



**HAL**  
open science

# Retail distribution and urban form : Street-based models for the French Riviera

Alessandro Araldi

► **To cite this version:**

Alessandro Araldi. Retail distribution and urban form : Street-based models for the French Riviera. Geography. Université Côte d'Azur, 2019. English. NNT : 2019AZUR2018 . tel-02467816

**HAL Id: tel-02467816**

**<https://theses.hal.science/tel-02467816>**

Submitted on 5 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

## DISTRIBUTION DES COMMERCES ET FORME URBAINE

MODÈLES ORIENTÉS-RUE POUR LA CÔTE-D'AZUR

**Alessandro ARALDI**

ESPACE-Etude des Structures, des Processus d'Adaptation et des Changements  
de l'Espace UMR 7300

Présentée en vue de l'obtention  
du grade de docteur en Géographie  
d'Université Côte d'Azur

Dirigée par :

Giovanni Fusco,  
Christine Voiron-Canicio

Soutenue le : 16 Septembre 2019

Devant le jury, composé de :

Cécile Tannier, Directrice de Recherche CNRS,  
Université de Franche-Comté  
Dominique Badariotti, Professeur,  
Université de Strasbourg  
Didier Josselin, Directeur de Recherche CNRS,  
Université Avignon  
Jean-Claude Thill, Professeur,  
University of North Carolina at Charlotte  
Meta Berghauser Pont, Professeure,  
Chalmers University of Technology



# **DISTRIBUTION DES COMMERCES ET FORME URBAINE**

**MODÈLES ORIENTÉS-RUE POUR LA CÔTE-D'AZUR**

## **RETAIL DISTRIBUTION AND URBAN FORM**

**STREET-BASED MODELS FOR THE FRENCH RIVIERA**

### **Jury :**

#### **Rapporteurs**

Cécile Tannier, Directrice de Recherche CNRS, Université de Franche-Comté

Dominique Badariotti, Professeur, Université de Strasbourg

#### **Examineurs**

Christine Voiron-Canicio, Professeure, Université Côte d'Azur

Didier Josselin, Directeur de Recherche CNRS, Université d'Avignon

Giovanni Fusco, Chargé de Recherche CNRS, Université Côte d'Azur

Jean-Claude Thill, Professeur, University of North Carolina at Charlotte

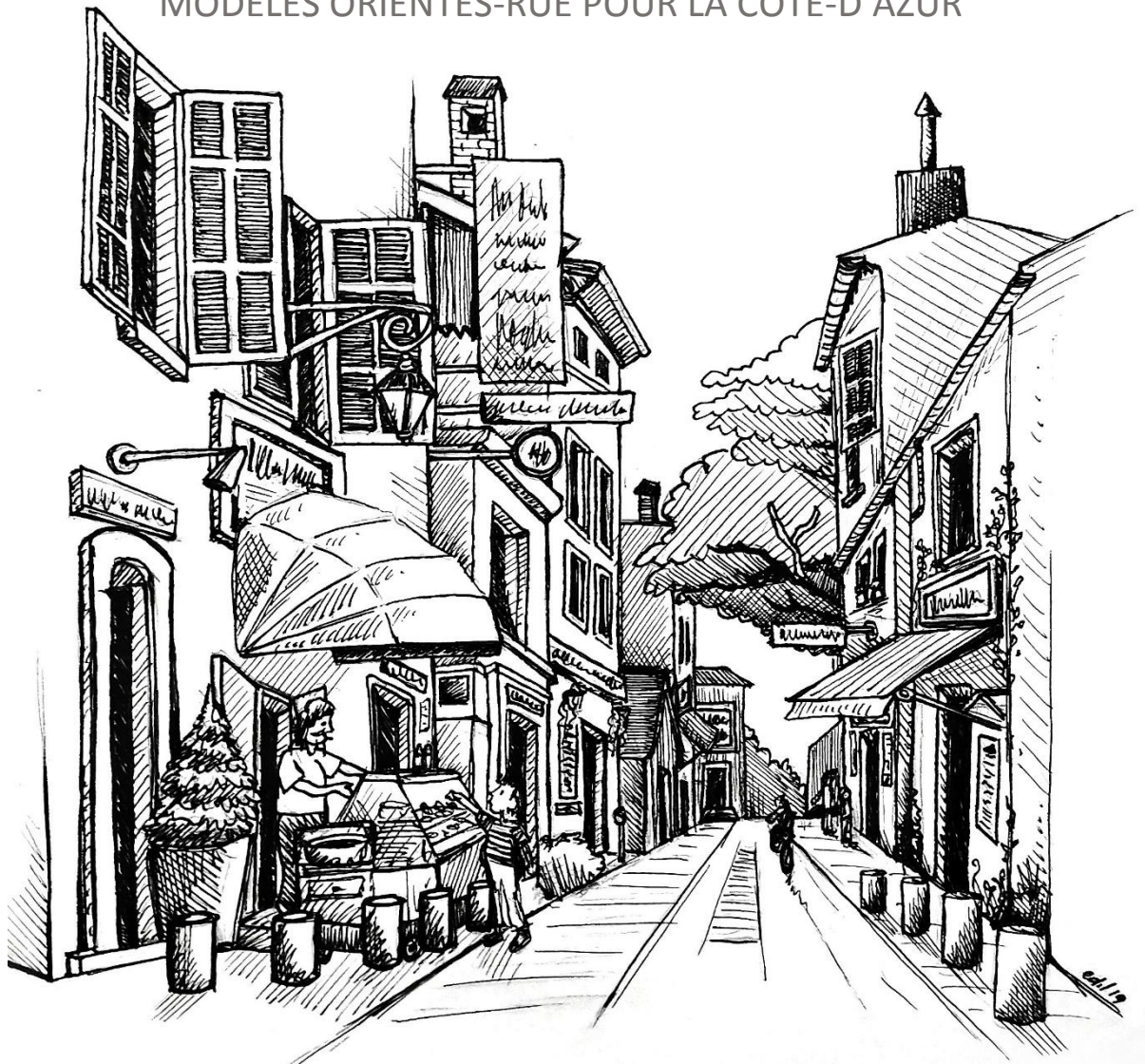
Meta Berghauser Pont, Professeure, Chalmers University of Technology

# RETAIL DISTRIBUTION AND URBAN FORM

STREET-BASED MODELS FOR THE FRENCH RIVIERA

DISTRIBUTION DES COMMERCES ET FORME URBAINE

MODÈLES ORIENTÉS-RUE POUR LA CÔTE-D'AZUR



Phd Dissertation in Geography - 2019

Alessandro ARALDI

# Abstract

This doctoral dissertation analyses and discusses the relationship between the spatial distribution of retail and urban form. More precisely, in this work, we focus on the spatial statistical relationships which occur between the localisation of small and average-sized stores and the physical properties of the urban form in the metropolitan area of the French Riviera.

The underlying hypothesis of this research is that the physical characteristics of the built-up landscape might influence how humans perceive and use urban space, and, ultimately, how stores are distributed and organised within cities.

In the last two decades, scholars have been increasingly investigating this relationship. Nonetheless, both retail and urban form characteristics are often reduced to the simple notions of store density and street-network configuration respectively.

Several aspects, such as the retail morpho-functional agglomeration typology, the geometrical characteristics of the streetscape and the contextual influence of the urban fabric are traditionally excluded from these analyses. These aspects should be even more important when studying highly heterogeneous metropolitan areas like the French Riviera, a combination of differently-sized cities and paradigmatic morphological regions: medieval centres, modern and contemporary planned areas, and suburban sprawl.

To overcome these limitations, computer-aided, theory-based protocols are accurately selected and developed in this dissertation, allowing for the extraction of quantitative measures of retail and urban form. In particular, starting from traditional theories of retail geography and urban morphology, two location-based network-constrained procedures are proposed and implemented, providing a fine-grained description of the urban and retail fabrics at the street-level. These methodologies are based on innovative combinations of geoprocessing and AI-based protocols (Bayesian Networks).

The statistical relationship between retail and urban morphological descriptors raise through the implementation of several statistical regression models. The decomposition of the study area in morphological subregions both at the meso- and macro- scale, combined with the implementation of penalised regression procedures, enables the identification of specific combinations of urban morphological characteristics and retail patterns.

In the case of the French Riviera, the outcomes of these models confirm the statistical significance of the relationship between street-network configurational properties and retail distribution. Nevertheless, the role of specific streetscape morphometric variables is demonstrated to be also a relevant aspect of the urban form when investigating the retail distribution. Finally, the morphological context both at the meso- and macro- scale is demonstrated to be a key factor in explaining the distribution of retail within a large urban area.

**Key Words:** Retail Distribution · Urban Form · Street-based perspective · Count Regression Modeling · French Riviera

# Résumé

Cette thèse de doctorat analyse et discute la relation entre la distribution spatiale des activités commerciales et la forme urbaine. Plus précisément, dans ce travail, nous nous intéressons aux relations statistiques spatiales entre la localisation des activités commerciales de petite et moyenne taille et les propriétés physiques de la forme urbaine dans la région métropolitaine de la Côte d'Azur.

L'hypothèse sous-jacente de cette recherche est que les caractéristiques physiques du paysage bâti pourraient influencer la façon dont les humains perçoivent et utilisent l'espace urbain et, en définitive, sur la répartition et l'organisation des activités commerciales dans les villes. Au cours des deux dernières décennies, les spécialistes ont étudié de plus en plus cette relation. Néanmoins, les caractéristiques du tissu commercial et du tissu urbain sont souvent réduites aux notions simples de densité de magasin et de configuration de réseau de rue.

Plusieurs aspects, tels que la typologie d'agglomération morpho-fonctionnelle du commerce de détail, les caractéristiques géométriques du paysage à l'échelle de la rue et l'influence contextuelle du tissu urbain sont traditionnellement exclus de ces analyses. Ces aspects devraient être encore plus importants lorsque l'on étudie des zones métropolitaines très hétérogènes comme la Côte d'Azur, où l'on retrouve une combinaison de villes de tailles différentes et de régions morphologiques paradigmatiques : centres médiévaux, zones planifiées modernes et contemporaines et étalement suburbain.

Pour surmonter ces limitations, des protocoles basés sur la théorie et assistés par ordinateur sont sélectionnés et développés dans cette thèse permettant l'extraction de mesures quantitatives de forme commerciale et urbaine. En particulier, partant des théories traditionnelles de la géographie du commerce et de la morphologie urbaine, deux procédures sont proposées et mises en œuvre, fournissant une description détaillée des tissus urbains et commerciaux à l'échelle de la rue. Ces méthodologies se basent sur des combinaisons innovantes de protocoles de géo-traitement et sur des approches d'IA (Réseaux Bayésiens).

Les relations statistiques entre les descripteurs morphologiques urbains et de distribution commerciale sont étudiées à travers la mise en œuvre de plusieurs modèles de régression statistique. La décomposition de la zone d'étude dans les sous-régions morphologiques à la fois à l'échelle méso et macro, ainsi que la mise en œuvre de procédures de régression pénalisées, permettent d'identifier des combinaisons spécifiques de caractéristiques morphologiques urbaines avec les caractéristiques de distribution spatiales du commerce.

Dans le cas de la Côte d'Azur, les résultats de ces modèles confirment la relation statistique entre les propriétés de configuration du réseau de rues et de la distribution spatiale du commerce. Néanmoins, il a été également démontré que le rôle de certaines variables morphométriques de paysage à l'échelle de la rue constitue également un aspect pertinent de la forme urbaine lors de l'enquête sur la distribution du commerce. Enfin, il a été démontré que le contexte morphologique à la fois à moyenne et à grande échelle est un facteur clé pour expliquer la répartition des petites et moyennes activités commerciales dans une grande région métropolitaine.

**Mots clés :** Distribution Spatiale du Commerce · Forme Urbaine · Analyses basées sur la rue · Modélisation par régression de comptage · Côte d'Azur

# Acknowledgements

It has been three and a half years since I started my PhD in geography. I never thought it would be such an incredible experience. This journey has given me the opportunity to enrich my personal knowledge and skills, but, more importantly, it has been invaluable for my personal growth. I've been able to explore and study a variety of subjects, as well as meet and exchange ideas with brilliant researchers and people from all over the world.

The first acknowledgement I would like to make is to everyone who helped me with this project, either directly or indirectly, with your kindness and support. Thank you for being there for me, no matter what capacity it was in; you helped me make it through.

Next, I'd like to thank my director Giovanni Fusco: our conversations have always helped fuel my studies, curiosity, energy, and you have challenged my ideas, creating a stimulating environment making me always eager to keep working and do the best I could. Thank you for your time and your patience, especially in this last part of the writing period. I sincerely admire your work and your dedication to the research activity; every discussion with you has always been an enriching moment. I hope to share with you many other collaborations in the future.

I am also grateful to my co-director Christine Voiron Canicio, who made it possible to integrate the ESPACE research laboratory to this project five years ago, and initiated me to the geographical and spatial analysis theories and methods. My former research director Giovanni Rabino offered me the chance to come to Nice and put me in contact with my current directors, but, more importantly, he is responsible for sparking my curiosity in spatial studies and urban systems.

I am grateful to the members of the jury Cécile Tannier, Dominique Badariotti, Didier Josselin, Jean-Claude Thill, and Meta Berghauser Pont who have promptly accepted to read and evaluate this manuscript. I am honoured and truly grateful for your interest in, and the time you dedicated, to this work.

I would also like to thank all the members of the ESPACE Laboratory, who were always open to discussing the research, and who supported me with encouragement. In particular, I am grateful to Diego Moreno, Matteo Caglioni, Karine Emsellem and Fabrice Decupigny. A huge thanks to Veronique Gibello, Sylvie Greve for their patience and time for all the administrative assistance.

My fellow travellers have been a source of inspiration: all ESPACE PhD and postdocs with whom I had the chance to collaborate, exchange ideas, knowledge and support. I am really thankful in particular to Alessandro, Joan, Alexandre, Jasiel, Irem, David, Matthieu, Diego, Kilian and TERENCE. I truly wish you all successful careers and I hope to work with you in the future on new and exciting projects.

This incredible journey would have not been possible without the support of the Local Chamber of Commerce who agreed to fund my research proposal and supported my project both with their personal expert knowledge and with their economical support. In particular,



I would like to thank Dr Jean Claude Zeitoun and the scientifically responsible Mr Jean Pinna who both believed in this research since the beginning, and who always appreciated my engagement in my work. A huge thanks to all the collaborators of the local Chamber of Commerce of Nice Côte d'Azur, and, more specifically, those from the Commercial Urbanism department and the Economic Observatory SIRIUS. In particular, Julien Puel, Emilie Collignon, and Sylvie Topor for your support from the very first day of my internship, and for your assistance with the CIFRE collaboration project. A special thanks to Youssra, who, despite the fact that she might still be trying to figure out what I did these past three years, always cheered me up and made me start my working days with a huge smile.

I am also grateful for the fruitful collaboration with several professors and researchers with whom I had the honour to collaborate. In particular, three visiting student periods have been possible thanks to Professor Jose Tiberio Hernandez Carolina Pinzon and Meili Vanegas from the Universidade de Los Andes; Professor Takashi Fuse and Joan Perez from the department of civil engineering from the University of Tokyo; and Professor Isabelle Thomas, Madeleine Guyot, and the members of the CORE laboratory from the Université Catholique de Louvain-la-Neuve.

Shorter, but of equal importance, was my exchange with Prof. Ana Clara Moura, Prof. Luc Anselin, and their research collaborators that give me important suggestions for the development of several parts of this dissertation. Thanks to all the professors, researchers, PhD and Post-Docs that I had the chance to meet, listen to, and discuss with throughout the conferences and summer schools I attended.

Beyond the academic work, this research project would not have been possible without the support of some beloved friends. A particular huge thank goes to Cyril Bousquet who cheered me up and encouraged me in the most difficult and stressful moments of this journey. A big thank you for your priceless help with the technical and informatics problems I encountered, for resolving bugs and scripts saving time, but most importantly, for being there to support my mental health more than once, even during thousands of kilometres of distance. A huge thanks also to Angelica Cofer who had the energy and the patience to proofread the whole text. Thanks to Edil Bessa who took the time to design the cover image with his personal and artistic touch. I will never be able to thank you enough.

I could not conclude these acknowledgements without mentioning my family: thanks to my parents, my brother Massimiliano and my sister Stefania who have always accepted and supported my decisions. A special and unique thanks to my grandparents and their unconditional love; I dedicate this work to you.

Finally, thanks to all my friends living in different parts of the world; in my home region, the Lake Como; and in my adopted city, Nice. Thanks to Cyril, Emily, Ruggero, Fred, Alain, Joan and Alessandro. A special thanks to Cansin and Saru: you recently came into my life with an explosion of joy and love and I look forward to sharing with you new, exciting and unforgettable moments.

# Publications and research communications

While developing this thesis, a parallel and continuous work of scientific publication has been undertaken. Some research outcomes presented in this dissertation have already been the subject of articles in international journals and/or presented in international conferences.

## **Papers in Peer-Reviewed International Journals:**

Araldi A, Fusco G (2019) "From the built environment along the street to the metropolitan region. Human scale approach in urban fabric analysis", *Environment and Planning B: Urban Analytics and City Science*, 46(7), 1243–1263.

<https://doi.org/10.1177/2399808319832612>

Perez J, Fusco G, Araldi A, Fuse T (2019) "Identifying building Typologies and their spatial patterns in the metropolitan areas of Marseille and Osaka", *Asia-Pacific Journal of Regional Science*, 1-25. DOI.org/10.1007/s41685-019-00127-6

Araldi A, Fusco G (2019) "Retail Fabric Assessment: describing retail patterns within urban space", *Cities*, 85(2019) 51-62. DOI.org/10.1016/j.cities.2018.11.025.

Araldi A, Fusco G (2018) "Describing retail distribution in urban space: local fractal indicators for point pattern analysis", *Revue Internationale de Geomatique*, 28(4), 505-532. DOI.org/10.3166/rig.2018.00063

Fusco G, Araldi A (2017) "Significant patterns in urban form. Spatial analysis of morphological indicators", *Revue Internationale de Geomatique*, vol.27, p. 455-479. DOI.org/10.3166/rig.2017.00045

## **Book Chapters:**

Araldi A, Fusco G (2018) "L'analisi delle forme urbane percepite dal pedone. Contributi dal geoprocessing", in F. Angelucci et al. (Eds.) *Smartness ed Healthiness per la transizione verso la resilienza. Orizzonti di ricerca interdisciplinare sulla città e il territorio*. Collana BETHA, 86.1.4, Franco Angeli, Milano, ISBN 8891760870, p.163-179.

## **Papers in Proceedings of Peer-Reviewed Conferences:**

Araldi A, Fusco G (2019) "Bridging configurational, streetscape and urban fabric analysis: the street as a keystone" *International Seminar Urban Form-ISUF 2019*, Nicosia, Cyprus, July 6<sup>th</sup>-8<sup>th</sup> 2019, (under review)

Araldi A, Fusco G (2018) "A new methodological approach for retail fabrics in the city." *International Conference on Spatial Analysis and Modeling*, 8th-9th sept. Tokyo. <http://ua.t.u-tokyo.ac.jp/sam2018/abstract/A-1-6>

Perez J, Fusco G, Araldi A, Fuse T (2018) "Building Typologies for Urban Fabric Classification: Osaka and Marseille Case Studies" *International Conference on Spatial Analysis and Modeling*, 8th-9th sept. Tokyo. <http://ua.t.u-tokyo.ac.jp/sam2018/abstract/A-1-5>

<sup>1</sup>Araldi A, Perez J, Fusco G, Fuse T (2018) "Multiple Fabric Assessment: Focus on Method Versatility and Flexibility", in: O. Gervasi et al. (eds) *Computational Science and Its*

---

<sup>1</sup> Best Paper Award 17thICCSA 2017 Trieste- Geo&mod

Applications – ICCSA2018. Proceedings, Part III, Lecture Notes in Computer Science, vol.10962, pp.251-267. Springer, Berlin, p. 251-267. DOI: 10.1007/978-3-319-95168-3\_17

Araldi A, Fusco G (2017) "Retail distribution in the city: analyzing point patterns in non-uniform network space." Spatial Analysis and GEomatics - SAGEO 2017, Rouen, France. Proceedings SAGEO Nov. 2017 <https://hal.archives-ouvertes.fr/hal-01643363>

Fusco G, Araldi A (2017) "The Nine Forms of the French Riviera: Classifying urban morphology from the pedestrian perspective", International Seminar on Urban Form- ISUF 2017, Valencia, Spain. DOI.org/10.4995/ISUF2017.2017.5219

Erin I, Fusco G, Cubukcu E, Araldi A, (2017) "Quantitative Methods of Urban Morphology in Urban Design and Environmental Psychology", International Seminar on Urban Form- ISUF 2017, Valencia, Spain DOI.org/10.4995/ISUF2017.2017.5732

<sup>2</sup>Araldi A, Fusco G (2017) "Decomposing and Recomposing Urban Fabric: The City from the Pedestrian Point of View", in O. Gervasi et al. (Eds.) Computational Science and Its Applications - ICCSA 2017, Trieste, Italy Proc. Part IV, LCNS Springer, vol. 10407, p. 365-376. DOI: [https://doi.org/10.1007/978-3-319-62401-3\\_27](https://doi.org/10.1007/978-3-319-62401-3_27)

Fusco G, Caglioni M, Araldi A (2017) "Street network morphology and retail locations. Application to the city of Nice (France)" European Colloquium on Theoretical and Quantitative Geography- ECTQG 2015 Bari, Italy. in *Plurimondi VIII, 17, 2015, 15-22* <http://plurimondi.poliba.it/index.php/Plurimondi/article/view/26>

Fusco G, Araldi A (2016) "Urban Fabrics with LINCS. Empirical Evidence from the French Riviera" Spatial Analysis and GEomatics - SAGEO 2016, Nice, France. Proceedings SAGEO Dec. 2017 p. 229-244 <https://hal.archives-ouvertes.fr/hal-01424734v1>

Araldi A, Fusco G (2016) "Urban Form from the Pedestrian Point of View: Spatial Patterns on a Street Network." International Conference on Innovation in Urban and Regional Planning- INPUT, Turin, Italy Sept. 2016, Proceedings INPUT 2016 pp. 32-37. <https://hal.archives-ouvertes.fr/hal-01417484/document>

### **Conference Presentations without Proceedings**

Araldi A, Fusco G, Venerandi A (2019) "Assessing the relationship between urban fabric, configuration of the street network and spatial distribution of retail" European Colloquium on Theoretical and Quantitative Geography- 5th-9th Sept ECTQG 2019 Luxembourg

Guyot M, Araldi A., Fusco G., Thomas I. (2019) "Urban street landscape and well-being in Brussels" European Colloquium on Theoretical and Quantitative Geography- 5th-9th Sept ECTQG 2019 Luxembourg

Guyot M, Araldi A., Fusco G., Thomas I. (2019) "Une typologie du paysage urbain bruxellois dans une perspective de recherche en santé" ThéoQuant, 6th-8th February Besançon. France

---

<sup>2</sup> Best Paper Award 18thICCSA 2018 Melbourne- Geo&mod

- Guyot M, Araldi A, Fusco G, Thomas I (2018) “Multiple Fabric Assessment: application to the case of Brussels”, American Association of Geographers – AAG 2018, New Orleans, United States
- Fusco G, Araldi A, Borgo S, Stufano R (2018) “Asymmetry along the way. Applications and conceptualization for morphology of urban fabric”, First European Asymmetry Symposium -FEAS 2018, Nice, France
- Araldi A, Fusco G. (2017) “Retail Activity and the City: Contribution from the Analysis of Urban Fabrics”, European Colloquium on Theoretical and Quantitative Geography-ECTQG 2017 York, United Kingdom
- Fusco G, Araldi A (2017) “Multiple Fabric Assessment: Exploring the Forms of a Metropolitan Area “, European Colloquium on Theoretical and Quantitative Geography- ECTQG 2017 York, United Kingdom

### **Scientific Reports**

- Genre-Grandpierre C, Fusco G., Balac M et al. (2015) “Vers de nouvelles dynamiques de localisation des ménages et des activités dans les territoires urbains pour découpler accessibilité et mobilité automobile.” [Research report] Subvention 11-MT6PREDITG06-5-CVS-053, UMR Espace, Université d'Avignon; UMR Espace, Université de Nice; ETH Zurich; Laboratoire d'informatique d'Avignon, Université d'Avignon; UNIVERSITE Nice Sophia Antipolis; LET. 2015, pp.186. <https://halshs.archives-ouvertes.fr/halshs-01708490>
- Araldi A (2016) “Morfologia urbana e localizzazione delle attività commerciali : analisi, modelli e visualizzazioni. Applicazioni all'area metropolitana di Nizza” Environmental Engineering MSc Dissertation , Politecnico di Milano <http://hdl.handle.net/10589/107485>

### **Workshop Presentations**

- Presentation at the Centre for Spatial Data Science, The University of Chicago 09/01/2018: “Urban form from the Pedestrian Point of View: Spatial Patterns on a street network” <https://spatial.uchicago.edu/content/2018>
- Presentation at CORE Laboratory, Université Catholique de Louvain 08/05/2018: “Fabrics from the Pedestrian Point of View” <https://uclouvain.be/en/research-institutes/immag/core/events/geography-seminar-alessandro-araldi.html>

# Contents

**Acknowledgements**

*Remerciements*

**Abstract**

*Résumé*

**Publications and research communications**

**Introduction ..... 5**

**Introduction .....12**

**Part 1 *Retail and the Physical City. Formulation of Research Questions*.....19**

**1. The Point of View of Retailing Geography..... 22**

1.1 Neoclassical approach..... 23

1.2 Behavioural approach ..... 30

1.3 Micro retail in intra urban space: combining theories ..... 34

1.4 Structural approach ..... 40

1.5 Contemporary approaches ..... 42

**2. Retail in the Study of Urban Form ..... 48**

2.1 Definition of Urban Form..... 48

2.2 Street-Network Configuration and Retail Distribution ..... 52

2.3 Beyond Network Configuration ..... 59

2.4 Beyond Retail Density ..... 67

2.5 Overview of Research Questions ..... 81

**3. Study Area and Data ..... 83**

3.1 Urban and economic context..... 83

3.2 The retail system..... 91

3.3 The scientific interest of the French Riviera case study ..... 96

3.4 Data ..... 97

**Part 2 *Measuring Retail Distribution in the City*.....105**

**4. Quantitative methods for micro retail location analysis ..... 108**

4.1 Spatial analysis approaches in retail geography ..... 108

4.2	Point pattern analysis .....	117
4.3	Observations .....	120
<b>5.</b>	<b>Retail Fabric Assessment: a new proposal .....</b>	<b>126</b>
5.1	Spatial Unit.....	127
5.2	Measuring Form and Functioning Patterns .....	128
5.3	The geometry indicator.....	133
5.4	Classification .....	144
5.5	Results.....	145
<b>Part 3</b>	<b><i>Measuring the Form of the Physical City .....</i></b>	<b>181</b>
<b>6.</b>	<b>Computer-Aided Approaches for the Analysis of the Urban Form.....</b>	<b>184</b>
6.1	Configurational approaches.....	185
6.2	Quantitative Typo-morphological approaches .....	191
6.3	Finding a common ground for configurational, typo-morphological and streetscape analyses .....	208
<b>7.</b>	<b>Measuring the street configuration .....</b>	<b>211</b>
<b>8.</b>	<b>Measuring the streetscape.....</b>	<b>225</b>
<b>9.</b>	<b>Urban fabrics from the pedestrian perspective .....</b>	<b>238</b>
9.1	Network Constrained Streetscape Patterns .....	239
9.2	Clustering morphological patterns .....	246
9.3	Results: the nine urban fabrics of the French Riviera.....	248
<b>Part 4</b>	<b><i>Models for Retail Distribution in Urban Space .....</i></b>	<b>285</b>
<b>10.</b>	<b>Retail distribution and Urban Form:.....</b>	<b>288</b>
10.1	Overall Framework: the street as a keystone.....	288
10.2	Statistical and modelling procedures: literature review. ....	291
10.3	Exploratory data analysis .....	296
<b>11.</b>	<b>How to model the relationship between retail and urban form? .....</b>	<b>304</b>
11.1	Modelling micro retail distribution: from linear to count regression approaches .....	304
11.2	Multicollinearity and feature selection. ....	316
11.3	How to introduce spatiality: the contextual variable .....	322
11.4	Goodness of fit measures .....	326
<b>12.</b>	<b>Models of Retail Distribution on the French Riviera.....</b>	<b>336</b>
12.1	Counting model selection .....	336
12.2	Variable selection for count regression models .....	348

8.1 A model for the typology of retail agglomerations .....	375
<b>Conclusions and Perspectives.....</b>	<b>385</b>
<b>Conclusions et Perspectives .....</b>	<b>393</b>
<b>References .....</b>	<b>401</b>
<b>LIST of ACRONYMS and ABBREVIATIONS.....</b>	<b>436</b>
<b>LIST of TABLES.....</b>	<b>438</b>
<b>LIST of FIGURES .....</b>	<b>439</b>
<b>APPENDIX 1.....</b>	<b>445</b>
<b>APPENDIX 2.....</b>	<b>464</b>
<b>APPENDIX 3.....</b>	<b>473</b>





# Introduction

This dissertation addresses the relationship between retail distribution and the city. The association between these two components has been widely discussed and demonstrated in the large and diversified scientific literature of the last century (Lemarchand 2009).

The city is a complex system composed of a large number of elements—or agents—of many different types that interact in different ways and are connected through countless relationships. Commercial activity is one of the main components of this complex system and the physical localisation of stores might be seen as its spatial and physical manifestation. Studying its spatial structure enables us to identify and disclose important information about the organisation and the functioning of urban spaces.

Retail geography is the scientific research domain devoted to the investigation, identification, and description of the retail localisation factors and how they influence the spatial distribution and organisation of stores. Since the localisation of stores and services is influenced by a complex combination of micro- and macro-scale social, economic, and environmental factors, several approaches have been developed and explored, each time focusing on the analysis of a specific subset of aspects. Despite this already rich and diversified scientific literature, a large number of studies is constantly being published on this subject.

The main underlying reason behind this constant, and, in certain aspects, growing interest in retail geography is that the spatial organisation of retail is subject to continuous modifications which raise fundamental questions about the future socioeconomic characteristics of our cities. Over time, patterns of stores in urban spaces are continuously under evolution, stimulated by different economic and technological transformations, the development of new retail formats, but also the spatial development of urban spaces. The commercial activity is, indeed, the urban function which is most reactive to societal and physical transformations of cities: studying the characteristics of retail distribution might reveal important information about the causes underlying these processes and their possible evolution.

Two main retail *revolutions* have already transformed the retail system and its spatial structure in the last century (Desse and Lestrade 2016): in the late 19<sup>th</sup> century, the public transport system has engendered the centralisation of retail activities with the birth of department stores. After WWII, a second revolution has been triggered by the proliferation of the car and the consequent development of large retail surfaces in the peripheral areas of cities. More recent technological advancements, such as the Internet and mobile phones (or, more broadly, Internet of things), have triggered a third commercial revolution transforming consumer behaviours and proposing new retail formats: to the traditional physical formats of the 20<sup>th</sup> century, a new generation of dematerialised stores and services are today proposed to costumers, allowing the buying experience to exist independently of spatial and temporal constraints. This undergoing transformation is at the origin of a global *seismic shift* (Dart 2018) which is expected to profoundly transform the retail system as we know it today.

Several contrasting projections of the future retail system have been depicted by researchers and practitioners; the common idea underlying these projections is that the current revolution, as the previous ones, should not be seen as a complete substitution of the old

retail formats with new ones, but rather as a new coexistence (Desse and Lestrade 2016) with a redefinition of the equilibrium of the whole system.

What we are observing in recent years is an overall reduction of small and average-sized stores with a different magnitude depending on both local and global characteristics of the urban system. While in the North American and UK context, the popular press describes the first consequences of the retail system transformations as “*apocalyptic*”<sup>3</sup>, in France, these signals of change seem less catastrophic but still significant: in the last 15 years, a growing vacation ratio has been observed in French cities, from 6.2 to 12.1%<sup>4</sup> with an acceleration in the last few years (Madry 2018).

The observed shrinking of retail distribution in cities is not only threatening local economies. Beyond the economic impact, it constitutes a major challenge for public policies which particularly aim for a balanced and sustainable development of cities and metropolitan areas. Moreover, the presence of stores is a generator of human and social interaction and its dematerialisation becomes a threat for the social life, vibrancy, cohesion and sense of place in a street, a neighbourhood, and ultimately affects the attractiveness of a whole city (Smith and Sparks 2000).

National governmental institutions, aware of these problems, have proposed several studies and policies for the revitalization of the retail sector in town centres: for instance, in France, the “Action Coeur de Ville”<sup>5</sup> (literally “City-Core Action”) established a set of guidelines for the support of reactivation and development of the retail system in 222 municipalities with a declining retail system. In the specific case of the French Riviera, for example, two mid-sized cities are concerned: Grasse and Vallauris. Two of the five strategical axes of this policy consider the specific role of accessibility of public spaces, the characteristics of urban form, and the role of the architectural heritage of cities. As evidenced by this example, a growing interest is associated with how cities are designed and how we can intervene and modify the urban form of cities in order to better enhance the social interaction and the economic durability.

Retail geography has traditionally been focusing on the economic, social, behavioural, and cultural context, while less consideration has been given to the morphological and design properties of the city. On the contrary, a growing interest in this subject might be observed in urban morphological studies in the last twenty years, which has produced proposals of new methods for investigating how the physical characteristics of the urban space might influence the retail spatial organisation. More specifically, the Movement Economy Theory explains how the spatial configuration of public spaces might influence the movement patterns and indirectly the spatial localisation of stores (Hillier 1996). Based on this theory, several empirical studies have demonstrated in the last twenty years how the spatial configurational properties of the street network represent a significant localisation factor of the store distribution.

---

<sup>3</sup> *Forbes* ‘Retail apocalypse: a look at what comes next and it isn’t pretty’ Danziger, P. N. (2017). *Business Insider*. ‘Here are the 18 biggest bankruptcies of the ,retail apocalypse ‘of 2017. Taylor, K. (2017).

<sup>4</sup> Source *Institut pour la ville et le commerce*, données Codata.

<sup>5</sup> <https://www.entreprises.gouv.fr/coeur-de-ville>

The general question underlying this strain of research is as follows: what role do urban morphologies play in the definition of the spatial distribution and structure of retail activities?

The interest in this specific question has been stimulating my research activities before undertaking this dissertation. While studying Civil and Environmental Engineering at the Polytechnical School of Milan with the specialisation in *Urban and Land Planning and Management*, I had the opportunity to take the course of *Urban Systems* given by the Professor Giovanni Rabino. Considering my particular interest in questions of spatial organisation and functioning of the urban system, Prof. Rabino suggested I finish my master's programme with an internship at the multi-site Laboratory of theoretical and quantitative geography ESPACE, at the Universities of Nice, Avignon and Aix-Marseille, directed then by Professor Christine Voiron-Canicio.

Between April 2014 and May 2015, I thus had the chance to work on the research project DAMA (*Vers de nouvelles dynamiques de localisation des ménages et des activités dans les territoires urbains pour découpler accessibilité et mobilité automobile*) directed by Giovanni Fusco in Nice and Cyrille Genre-Granpierre in Avignon. In the meantime, the double degree agreement between the Polytechnical School of Milan and the University Côte d'Azur (University of Nice Sophia-Antipolis back then) allowed me to participate in the MSc programme *Geoprospective, Planning and Sustainability of Territories*, and more precisely to its *Spatial Structures and Dynamics* research option. This gave me the opportunity to enrich my university background as an engineer with theoretical concepts from geography and with methodological skills in spatial analysis and modelling. In 2015 I graduated both in Environmental Engineering and in Geography with the project entitled: "*Urban morphology and retail localisation: analyses, models and visualisations. Applications to the metropolitan area of Nice*". In this master's thesis research, part of the aforementioned DAMA project, I had the chance to study and implement the Movement Economy Theory to the central area of the French Riviera around the city of Nice. In particular, I investigated the relationship between street-network configurational properties and retail distribution. By collaborating closely with DAMA's research director Giovanni Fusco, I further deepened my interest and knowledge in the theories of urban form and in the data analysis protocols of Bayesian Networks.

At the end of this research project, several unfulfilled questions were further stimulating my interest in this specific research field, such as the role of other local and contextual morphological variables of the urban form, the absence of stores in specific subspaces, and the spatial functional patterns of retailers.

The Presentation of the MSc Thesis research outcomes to Dr Jean Claude Zeitoun and his collaborators at the Commercial Urbanism department of the Local Chamber of Commerce of Nice Côte-d'Azur (CCINCA) made me recognise the practical meaning of these research questions in a planning context. In particular, the collaboration with Julien Puel, Emilie Collignon and Sylvie Topor (from the department of Commercial Urbanism) provided highly valuable expertise and support both in geo-marketing and retailing in general and in the

specificities of the retail system of the French Riviera. An internship of six months allowed me to produce a technical audit of the analytical tools already deployed in the Commercial Urbanism department for the assessment of the retail distribution and to propose a detailed action plan suggesting possible integrations of research-based methodological innovations with praxis-based expertise.

Based on these premises, a new research project has been developed between my research unit, ESPACE, and the CCINCA, and validated both by the academic and the private partners, as well as the National Authority of the Scientific Research ANRT, allowing me to launch a doctoral research program within a CIFRE framework (Conventions Industrielles de Formation par la REcherche<sup>6</sup>) in the month of March 2016. At ESPACE, my new research project was to contribute to the emerging actions “Urban complexity and Networks: Morphology, Modelling, Multiplexity and Metamorphosis” and “DAMIA - Geographical Big Data, Modelling and Artificial Intelligence” which would eventually become integral parts of the new 2018-2022 scientific programme of the research unit.

As a result of these particular scientific and personal trajectories, this dissertation is at the convergence point between several theoretical and applied research questions within an interdisciplinary context. Notions from retail geography and urban morphology are both implemented and developed with new analytical procedures from data mining and statistic modelling. Finally, the institutional framework between academia and a local agency fostered a research agenda aiming to answer fundamental issues in planning and management of urban space. The outline of the manuscript reflects this tension between specific knowledge domains (retail geography, urban morphology and planning) and more general theoretical and methodological issues.

The dissertation is organised in four parts.

**Part 1 discusses the overall research questions.** In **Chapter 1** we review the scientific literature of the works concerning the retail geography domain: the traditional distinction between neoclassical, behavioural, and structural approaches is presented. As in the past, when different transformations of society required the development of new research approaches, the ongoing processes observed in more recent years are stimulating the research on new contextual approaches. The urban form approach developed in this dissertation contributes to this wider research effort.

In **Chapter 2** we review and discuss how the study of the retail distribution has been approached in the urban morphology scientific literature. Since the notion of urban form is multifaceted, is interpreted differently depending on the context, and is used in several disciplines, we will first specify the meaning which we associate to it. Urban form is thus understood, in this dissertation, as the spatial organisation of the mainly man-made physical space of the city. Then, we discuss how urban studies and network configuration theories have launched the study of the quantitative analysis of the urban form and its relation with retail distribution. In the last 20 years, a growing number of works assessing the relationship between urban form and retail distribution have been implementing original concepts to

---

<sup>6</sup> <http://www.anrt.asso.fr/fr/cifre-7843>

different case studies. Nonetheless, few theoretical and methodological innovations have been put forward. Based on these works, we will propose our main research questions: we firstly discuss the importance of the retail system and how the description of the retail spatial organisation should not be limited to the simple enumeration of stores. Secondly, we discuss how urban form should not be limited to the street network configuration and why new analytical procedure should be developed for the quantitative description of the other aspects of urban form. These research questions and their associated methodological challenges are summarised at the end of the chapter.

**Chapter 3** describes the study area of the French Riviera and focuses on both on its urban and economic characteristics. Different from previous research, the study of a large metropolitan area at the fine-scale of the street will allow us to include eighty-eight municipalities within five major cities (Nice, Cannes, Antibes, Grasse and Menton), which, along with the principality of Monaco, represent a continuous and highly constrained coastal conurbation. This research area represents an interesting study region both for its urban morphological characteristics and for the variety of the retailscapes. The observed over-representation of small- and medium-sized retailers in the central cities (Schema Departemental d'Equipement Commercial des Alpes-Maritimes, SDEC 2005), is thus associated with a very early appearance of shopping malls in more peripheral areas (France's first modern shopping mall opened in Saint-Laurent-du-Var in 1969).

**The following Parts 2, 3 and 4** focus on new methodological protocols for the quantitative description of the retail distribution, the urban morphological characteristics, and the statistical modelling of their relationship, respectively. The three parts share a similar structure: after a review of existing analytical approaches in the literature, applications and new analytical procedures are proposed and implemented. Their results for the French Riviera are then presented and discussed.

Moreover, the separation of these three components of the research provides several advantages. Firstly, the three sections might be individually read and appreciated by readers having a more specific interest in retail geography, urban morphology, and statistical modelling, respectively. Secondly, higher flexibility and modularity is provided to the overall research architecture: each part/module might be individually studied and developed in future research developments. Finally, this separation allows for reducing possible tautologies in the morpho-functional analysis by previously drawing a clear line between the retail system and the urban morphological system before proposing their cross-analysis.

In **Chapter 4**, we thus review the quantitative analytical protocols developed for retail assessment with a specific focus on store (point) pattern analysis on networks. In **Chapter 5**, a new original protocol (RFA, Retail Fabric Assessment) is therefore described and implemented in our study area. RFA outcomes are presented and a descriptive model for the retail distribution of the French Riviera is finally proposed.

In **Chapter 6**, we review the computer-aided approaches for the analysis of the urban form and describe three main approaches traditionally developed in urban studies: the street network configuration, the urban streetscape, and the urban fabric analysis. For the first two,

several measures have already been proposed in the scientific literature and **Chapters 7 and 8** describe the methodological choices implemented in this dissertation. On the contrary, the analysis of the urban fabric from the street point of view represents a new approach which required us to develop an original procedure (MFA, Multiple Fabric Assessment), described in **Chapter 9**. MFA outcomes are presented and a descriptive model for the French Riviera urban fabric spatial organisation is finally proposed.

**Part 4** studies the statistical relationship between retail spatial distribution and urban form properties. We firstly review in **Chapter 10** the traditional statistical approaches implemented in the works preceding this dissertation. In **Chapter 11**, we describe the modelling procedure which has been implemented: count regression approaches. Specific attention is given to three main concerns of our application: the multicollinearity problem, the inclusion of contextual variables in regressive modelling, and a coherent methodological approach for the correct evaluation of the modelling outcomes. Finally, **Chapter 12** presents the outcomes of our modelling procedures applied to the French Riviera. The relationship between the spatial distribution of the retail system and the urban morphological properties of this specific metropolitan space is outlined and discussed.

A conclusive section overviews the main outcomes of this dissertation both from thematic and methodological points of view. Limitations and perspectives of future research are finally discussed.

# Introduction



Cette thèse porte sur l'étude de la relation entre la distribution des activités commerciales et la ville. L'association entre ces deux composants a été largement discutée et démontrée dans la vaste et diversifiée littérature scientifique du siècle dernier (Lemarchand 2009).

La ville est un système complexe composé d'un grand nombre d'éléments, ou d'agents, de nombreux et différents types, qui interagissent de diverses manières et sont liés dans d'innombrables relations. L'activité commerciale est l'une des composantes principales de ce système complexe et la localisation physique des magasins peut être considérée comme une manifestation physique et spatiale. L'étude de sa structure spatiale nous permet de révéler des informations importantes sur l'organisation et le fonctionnement des espaces urbains.

La géographie du commerce est le domaine de recherche scientifique consacré à la recherche, à l'identification et à la description des facteurs de localisation de la vente au détail et à la manière dont ils influencent la répartition spatiale et l'organisation des magasins. Étant donné que la localisation des magasins et des services est influencée par une combinaison complexe de facteurs sociaux, économiques et environnementaux à l'échelle micro et macro, plusieurs approches ont été développées et explorées, axées chaque fois sur l'analyse d'un sous-ensemble spécifique d'aspects. Malgré cette littérature scientifique déjà riche et diversifiée, de nombreuses études sont constamment publiées sur ce sujet.

La principale raison sous-jacente de cet intérêt constant et, sous certains aspects, croissant pour la géographie du commerce, est que l'organisation spatiale des magasins est soumise à des modifications continues qui soulèvent des questions fondamentales concernant les caractéristiques socio-économiques futures de nos villes. Au fil du temps, les types d'activités commerciales et leurs localisations dans les espaces urbains sont en constante évolution, stimulés par différentes transformations économiques et technologiques, le développement de nouveaux formats, et des espaces urbains. L'activité commerciale est en effet la fonction urbaine qui réagit le plus aux transformations sociales et sociétales des villes : l'étude des caractéristiques de la distribution de détail pourrait révéler des informations importantes sur les causes sous-jacentes de ces processus et sur leur évolution possible.

Deux grandes *révolutions* de la vente au détail ont déjà transformé le système et sa structure spatiale au siècle dernier (Desse et Lestrade 2016) : à la fin du XIX<sup>e</sup> siècle, les transports en commun ont engendré la centralisation des activités de vente au détail avec la naissance de grands magasins. Après la Seconde Guerre mondiale, une deuxième révolution a été déclenchée par la diffusion de la voiture et le développement consécutif de grandes surfaces de vente au détail dans les zones périphériques des villes. Des avancées technologiques plus récentes telles qu'Internet et les téléphones mobiles (ou, plus largement, l'Internet des objets) sont à l'origine d'une troisième révolution commerciale transformant les comportements des consommateurs et proposant de nouveaux formats de vente au détail: aux formats physiques traditionnels du XX<sup>e</sup> siècle, une nouvelle génération de magasins et de services dématérialisés est aujourd'hui proposée aux clients, permettant une expérience d'achat indépendante des contraintes spatiales et temporelles. Cette transformation en cours est à l'origine d'un *changement sismique* (Dart 2018) global qui devrait transformer profondément le système de vente au détail tel que nous le connaissons aujourd'hui.

Des chercheurs et des praticiens ont présenté plusieurs projections contrastées du futur système commercial. L'idée commune qui sous-tend ces projections est que la révolution actuelle, comme les précédentes, ne doit pas être vue comme une substitution complète des anciens formats de vente au détail par de nouveaux, mais plutôt comme une nouvelle coexistence (Desse et Lestrade 2016) avec une redéfinition de l'équilibre du système global.

Ce que nous observons au cours de ces dernières années est une réduction globale du nombre de magasins de petite et moyenne taille, d'une ampleur différente selon le système urbain, en fonction de caractéristiques locales et globales. Alors que, dans le contexte nord-américain et britannique, la presse populaire décrit les premières conséquences des transformations du système de vente au détail comme «*apocalyptiques*<sup>7</sup>» en France, ces signaux de changement semblent moins catastrophiques mais néanmoins significatifs: au cours des 15 dernières années, la vacance commerciale a progressivement augmenté dans les villes françaises, de 6,2 à 12,1%<sup>8</sup> avec une accélération au cours des dernières années (Madry 2018).

La contraction observée des activités de vente au détail dans les villes ne menace pas seulement les économies locales. Au-delà de l'impact économique, il constitue un enjeu majeur pour les politiques publiques visant notamment un développement équilibré et durable des villes et des métropoles. De plus, la présence du commerce de détail génère des interactions humaines et sociales et sa dématérialisation devient une menace pour la vie sociale, le dynamisme, la cohésion et le sens du lieu d'une rue, d'un quartier et affectant en définitive l'attractivité de toute une ville (Smith et Sparks 2000, autres).

Les institutions gouvernementales nationales, conscientes de ces problèmes, ont proposé plusieurs études et politiques visant à revitaliser le secteur de la vente au détail dans les centres-villes: par exemple, en France, l'Action cœur de ville<sup>9</sup> a établi de lignes directrices pour soutenir la réactivation et le développement du système de vente au détail dans 222 municipalités où le tissu commercial urbain est en déclin. Dans le cas particulier de la Côte d'Azur, par exemple, deux villes de taille moyenne sont concernées : Grasse et Vallauris. Deux des cinq axes stratégiques de cette politique considèrent le rôle spécifique de l'accessibilité des espaces publics, les caractéristiques de la forme urbaine et le rôle du patrimoine architectural des villes. Comme en témoigne cet exemple, un intérêt croissant est associé à la conception des villes et à la manière dont nous pouvons intervenir et modifier la forme urbaine des villes afin de mieux renforcer les interactions sociales et le développement économique durable.

La géographie de la vente au détail a traditionnellement été axée sur le contexte économique, social, comportemental et culturel, tandis que les propriétés morphologiques et de design urbain de la ville ont été moins prises en compte. Au contraire, un intérêt croissant pour ce sujet pourrait être observé dans les études morphologiques urbaines des vingt dernières années proposant de nouvelles méthodes pour étudier comment les caractéristiques physiques de l'espace urbain pourraient influencer l'organisation spatiale du commerce de

---

<sup>7</sup> Forbes 'Retail apocalypse: a look at what comes next and it isn't pretty' Danziger, P. N. (2017).

Business Insider. 'Here are the 18 biggest bankruptcies of the, retail apocalypse' of 2017. Taylor, K. (2017).

<sup>8</sup> Source Institut pour la ville et le commerce, données Codata.

<sup>9</sup> <https://www.entreprises.gouv.fr/coeur-de-ville>

détail. Plus précisément, la Théorie de l'Économie du Mouvement explique comment la configuration spatiale des espaces publics peut influencer les comportements et les déplacements et, indirectement, la localisation spatiale des magasins (Hillier 1996). Sur la base de cette théorie, plusieurs études empiriques ont démontré au cours des vingt dernières années que les propriétés de configuration spatiale du réseau de rues représentaient un facteur de localisation significatif de la distribution.

La question générale sous-jacente à cette partie spécifique de la recherche est de savoir quel rôle les caractéristiques morphologiques urbaines jouent-ils dans la définition de la distribution spatiale des commerces ?

L'intérêt pour cette question spécifique a stimulé mes activités de recherche avant même cette thèse. Durant mes études d'ingénieur civil et environnemental à l'École Polytechnique de Milan avec une spécialisation en *Urbanisme et Aménagement du Territoire*, j'ai eu la chance de suivre le cours de *Systèmes Urbains* dispensé par le professeur Giovanni Rabino. Compte tenu de mon intérêt particulier pour les questions d'organisation spatiale et de fonctionnement du système urbain, le professeur Rabino a suggéré la possibilité de terminer mon programme de maîtrise par un stage au laboratoire multi-site de géographie théorique et quantitative ESPACE, aux universités de Nice, Avignon et Aix-Marseille, dirigé par la professeur Christine Voiron-Canicio.

Ainsi, entre avril 2014 et mai 2015, j'ai eu la chance de travailler sur le projet de recherche DAMA (Vers les nouvelles dynamiques de localisation des environnements urbains pour découpler l'accessibilité et la mobilité automobile) dirigé par Giovanni Fusco à Nice et Cyrille Genre-Granpierre à Avignon. Entre-temps, l'accord de double diplôme conclu entre l'École Polytechnique de Milan et l'Université Côte d'Azur (à l'époque l'Université de Nice Sophia-Antipolis) m'a permis de participer au programme de maîtrise en sciences *Géoprospective, planification et durabilité des territoires*, et plus précisément à son option de recherche *Structures spatiales et dynamique*. Cela m'a donné l'occasion d'enrichir mon parcours universitaire d'origine en tant qu'ingénieur avec des concepts théoriques issus de la géographie et des compétences méthodologiques en analyse spatiale et en modélisation. En 2015, j'ai obtenu mon diplôme en génie de l'environnement et en géographie avec le projet intitulé « *Morphologie urbaine et localisation des activités commerciales : analyses, modèles et visualisations. Applications à l'agglomération de Nice* ». Dans le cadre de cette recherche faisant partie du projet DAMA susmentionné, j'ai eu l'occasion d'étudier et d'appliquer la Théorie de l'Économie du Mouvement dans la zone centrale de la Côte d'Azur, autour de la ville de Nice. En particulier, j'ai étudié la relation entre les propriétés de configuration de réseau de rue et la distribution de détail. En collaborant étroitement avec Giovanni Fusco, directeur de recherche de DAMA, j'ai approfondi mon intérêt et mes connaissances des théories de la forme urbaine et des protocoles d'analyse de données des Réseaux Bayésiens.

Au terme de ce projet de recherche, plusieurs questions non résolues ont encore stimulé mon intérêt pour ce domaine de recherche spécifique, telles que le rôle d'autres variables morphologiques locales et contextuelles de la forme urbaine, l'absence de magasins dans des sous-espaces spécifiques, les schémas spatio-fonctionnels des magasins.

La présentation des résultats de la thèse au Dr. Jean Claude Zeitoun et à ses collaborateurs du service Urbanisme commercial de la Chambre de commerce de Nice Côte-d'Azur (CCINCA) m'a permis d'apprécier le sens pratique de ces questions de recherche dans un contexte de planification. La collaboration avec Julien Puel, Émilie Collignon et Sylvie Topor (du Département de l'urbanisme commercial) a apporté une expertise et un soutien de grande valeur, tant dans le géomarketing et dans les spécificités du tissu commercial de la Côte d'Azur. Un stage de six mois m'a permis de réaliser un audit technique des outils d'analyse déjà déployés dans le département d'Urbanisme Commercial et d'élaborer un plan d'action détaillé proposant l'introduction de nouvelles compétences et l'emploi de nouvelles méthodes innovantes et issues de la recherche.

Sur la base de ces prémisses, un nouveau projet de recherche a été développé entre mon unité de recherche ESPACE et la CCINCA. Il a été validé à la fois par les partenaires académiques et privés et enfin par l'autorité nationale de la recherche scientifique ANRT permettant de lancer un programme recherche doctorale dans un cadre CIFRE (Conventions industrielles de formation par la réforme<sup>10</sup>) en mars 2016. Au laboratoire ESPACE, mon nouveau projet de recherche contribue aux actions émergentes « Complexité urbaine et réseaux : morphologie, modélisation, multiplicité et métamorphose » et « DAMIA - Donnée géographique massives : modélisation et intelligence artificielle », qui finiront par devenir partie intégrante du nouveau programme 2018-2022 de l'unité de recherche scientifique.

Fruit de ces trajectoires scientifiques et personnelles particulières, cette thèse est le point de convergence de plusieurs questions de recherche théorique et appliquée dans un contexte interdisciplinaire. Des notions issues de la géographie de la distribution et de la morphologie urbaine sont mises en œuvre et développées avec de nouvelles procédures analytiques issues de l'exploration de données et de la modélisation statistique. Enfin, le cadre institutionnel entre le monde universitaire et une structure administrative locale a favorisé un programme de recherche visant à répondre aux questions fondamentales en matière de planification et de gestion de l'espace urbain. Les grandes lignes du manuscrit reflètent cette tension entre des domaines de connaissances spécifiques (géographie de la vente au détail, morphologie urbaine et planification) et des questions théoriques et méthodologiques plus générales.

La thèse est organisée en quatre parties.

**La première partie aborde l'ensemble des questions de recherche.** Au **Chapitre 1**, nous passons en revue la littérature scientifique des travaux élaborés dans le domaine de la géographie du commerce : la distinction traditionnelle entre les approches néoclassique, comportementale et structurelle est présentée. Par le passé, différents changements de la société ont conduit à l'élaboration de nouvelles approches dans la recherche. De la même manière, les processus observés ces dernières années dirigent la recherche vers des approches plus contextuelles. L'approche de la forme urbaine développée dans cette thèse participe à cet effort de recherche plus large. La notion de forme urbaine étant une notion à multiples facettes, sa compréhension et son utilisation varient suivant les disciplines. Nous allons donc dans un premier temps préciser le sens que nous lui associons. La forme urbaine est comprise,

---

<sup>10</sup> <http://www.anrt.asso.fr/fr/cifre-7843>

dans cette thèse, comme l'organisation spatiale de l'espace physique essentiellement urbain de la ville. Nous verrons ensuite comment les études urbaines et les théories de configuration du réseau ont initié l'étude de l'analyse quantitative de la forme urbaine et de ses relations avec la distribution de détail. Au cours des 20 dernières années, un nombre croissant d'études évaluant la relation entre forme urbaine et distribution au détail ont appliqué la notion originale à différentes études de cas. Néanmoins, peu d'innovations théoriques et méthodologiques ont été proposées. Sur la base de ces travaux, nous proposerons nos questions de recherche principales: premièrement, nous discuterons de l'importance du système de vente au détail et du fait que la description de l'organisation spatiale de la vente au détail ne devrait pas se limiter à la simple énumération des magasins. Deuxièmement, nous examinons comment la forme urbaine ne devrait pas être limitée à la configuration du réseau de rues et pourquoi une nouvelle procédure analytique devrait être mise au point pour la description quantitative des autres aspects de la forme urbaine. Ces questions de recherche et les défis méthodologiques associés sont résumés dans la section finale de ce chapitre.

Le **Chapitre 3** décrit la zone d'étude de la Côte d'Azur axée à la fois sur ses caractéristiques urbaines et économiques. Contrairement aux recherches précédentes, l'étude d'une grande métropole à la fine échelle de la rue nous permettra d'inclure quatre-vingt-huit municipalités dont cinq grandes villes (Nice, Cannes, Antibes, Grasse et Menton) qui, avec la principauté de Monaco représentent une conurbation côtière continue et très contrainte. Ce domaine de recherche représente une région d'étude intéressante tant pour ses caractéristiques morphologiques urbaines que pour la variété des paysages de détail. La surreprésentation observée des petites et moyennes enseignes dans les villes centrales (Schéma Départemental d'Équipement Commercial des Alpes-Maritimes 2005) est donc associée à une apparition très précoce de centres commerciaux dans des zones plus périphériques (centre commercial ouvert à Saint-Laurent-du-Var en 1969).

**Les Parties 2, 3 et 4 suivantes** portent sur de nouveaux protocoles méthodologiques pour la description quantitative de la distribution du commerce de détail, les caractéristiques morphologiques urbaines et la modélisation statistique de leurs relations, respectivement. Les trois parties partagent une structure similaire : après un examen des approches analytiques existantes dans la littérature, des applications et de nouvelles procédures analytiques sont proposées et mises en œuvre. Leurs résultats pour la Côte d'Azur sont ensuite présentés et discutés.

De plus, la séparation de ces trois composantes de la recherche offre plusieurs avantages. Premièrement, les trois sections pourraient être lues et appréciées individuellement par les lecteurs ayant un intérêt plus spécifique dans la géographie du commerce, la morphologie urbaine et la modélisation statistique, respectivement. Deuxièmement, l'architecture globale de la recherche est dotée d'une plus grande flexibilité et d'une plus grande modularité : chaque partie / module peut être étudié et développé individuellement dans de futurs travaux de recherche. Enfin, cette séparation permet de réduire les éventuelles tautologies dans l'analyse morpho-fonctionnelle en traçant préalablement une ligne de démarcation nette entre le système commercial et le système morphologique urbain, avant de proposer leur analyse croisée.

Au **Chapitre 4**, nous passons donc en revue les protocoles analytiques quantitatifs mis au point pour l'évaluation de la vente au détail, en mettant l'accent sur l'analyse des modèles de points de vente sur les réseaux. Au **Chapitre 5**, un nouveau protocole (RFA, Retail Fabric Assessment) est donc décrit et mis en œuvre sur notre zone d'étude. Les résultats obtenus sont présentés et un modèle décrivant le tissu commercial de la Côte d'Azur est proposé.

Au **Chapitre 6**, nous passons en revue les approches assistées par ordinateur pour l'analyse de la forme urbaine décrivant trois approches principales développées traditionnellement dans les études urbaines : la configuration du réseau de rues, le design urbain et l'approche typomorphologique. Pour les deux premières, plusieurs procédés ont déjà été proposées dans la littérature scientifique et les **Chapitres 7 et 8** décrivent les choix méthodologiques mis en œuvre dans cette thèse. Au contraire, l'analyse du tissu urbain du point de vue de la rue représente une nouvelle approche qui nous a obligé à développer une nouvelle procédure (MFA, Multiple Fabric Assessment), décrite au **Chapitre 9**. Les résultats obtenus sont présentés et un modèle décrivant le tissu urbain de la Côte d'Azur est proposé.

La **Partie 4** étudie la relation statistique entre la distribution spatiale des activités commerciales et les propriétés de forme urbaine. Nous revenons tout d'abord au **Chapitre 10** sur les approches statistiques traditionnelles mises en œuvre dans les travaux précédant cette thèse. Au **Chapitre 11**, nous décrivons la procédure de modélisation mise en œuvre : approches de régression de comptage. Une attention particulière est accordée aux trois préoccupations principales de notre application : le problème de multicollinéarité, l'inclusion de variables contextuelles dans la modélisation par régression et une approche méthodologique cohérente pour l'évaluation correcte des résultats de la modélisation. Enfin, le **Chapitre 12** présente les résultats de nos procédures de modélisation appliquées à la Côte d'Azur. Les relations entre la distribution spatiale, les activités commerciales et les propriétés morphologiques urbaines de cet espace métropolitain sont décrites et discutées.

Une section concluante passe en revue les principaux résultats de cette thèse du point de vue thématique et méthodologique. Les limites et les perspectives des recherches futures sont enfin discutées.

## **Part 1**

### *Retail and the Physical City. Formulation of Research Questions*

*Retail geography is the scientific research domain devoted to the investigation, identification and description of the relationships between the retail system and the city.*

*Within this domain of research, different approaches have been developed in the last century. Three main lines of research are traditionally identified (Brown 1991): the neoclassical, the behavioural, and the structural studies. All these approaches share the same goal: to identify the locational factors allowing for the understanding of how and why retail is differently distributed in the geographic space.*

*As highlighted by Brown (1992), the same spatial distribution of stores in a given region might be coherently explained with every aforementioned approach, providing different insights about the importance of different locational factors. Moreover, these approaches are not mutually exclusive, and, on the contrary, they could suggest new aspects between one another.*

*Nonetheless, the importance of the physical properties of the urban space on retail, also named the 'morphological sense of commerce', have been poorly analysed in the traditional approaches of retail geography literature (Saraiva et al. 2019).*

*Conversely, in the field of urban morphology, the last two decades have seen growing attention on the study of the relationship between retail distribution and the urban form.*

*At the origin of this line of research, we find the relatively recent Movement Economy Theory (Hillier 1996). Within this theory, specific attention is given to the spatial properties of the physical urban space as a determining factor of the economic potential of urban spaces. More in particular, the configurational properties of the street network have been demonstrated to be an important locational aspect describing the spatial distribution of stores.*

*From the first implementation on the city of London (Hillier 1999), an increasing number of projects have been implementing and assessing Hillier's theory against case studies. However, underlying these works, several assumptions simplify the retail systems to simple measures of store density while the urban form is represented by the street-network and described by its configurational properties.*

*While retail geography has often overlooked the importance of the urban morphological properties of the urban space, on the contrary, urban morphology has been reducing the retail system to a description of store density.*

*In this part, we will go through these different approaches and set the theoretical background of this dissertation. Based on the original definition of the Movement Economy Theory, we will propose new research that would enrich its original version with aspects from both urban morphology and retail geography.*



*La géographie du commerce est le domaine de la recherche scientifique consacré à la recherche, à l'identification et à la description des relations entre le commerce et la ville.*

*Dans ce domaine de recherche, différentes approches ont été développées au cours du siècle dernier. Trois approches principales de recherche sont traditionnellement identifiées (Brown 1991): l'approche néoclassique, l'approche comportementale, et l'approche structurelle. Ces approches partagent le même objectif : identifier les facteurs de localisation permettant de comprendre comment et pourquoi les activités commerciales sont réparties différemment dans l'espace géographique.*

*Comme décrit par Brown (1992), la même distribution spatiale des magasins dans une région donnée pourrait être expliquée de manière cohérente avec chacune des approches susmentionnées, fournissant ainsi des informations différentes sur l'importance de facteurs liés à l'emplacement. De plus, ces approches ne sont pas mutuellement exclusives et, au contraire, elles pourraient suggérer de nouveaux aspects entre elles.*

*Néanmoins, l'importance des propriétés physiques de l'espace urbain sur la distribution des activités commerciales, également appelée « sens morphologique du commerce », n'a été que peu analysée dans les approches traditionnelles de la littérature couvrant la géographie du commerce (Saraiva et al. 2019).*

*À l'inverse, dans le domaine de la morphologie urbaine, l'étude de la relation entre la distribution des activités commerciales et la forme urbaine a suscité un intérêt croissant au cours des deux dernières décennies.*

*À l'origine de cette ligne de recherche, nous trouvons la théorie relativement récente de l'économie du mouvement (Hillier 1996). Dans cette théorie, une attention particulière est accordée aux propriétés spatiales de l'espace urbain physique en tant que facteur déterminant du potentiel économique des espaces urbains. Plus précisément, il a été démontré que les propriétés de configuration du réseau de rues constituaient un aspect de localisation important décrivant la répartition spatiale des magasins.*

*Depuis la première mise en œuvre pour la ville de Londres (Hillier 1999), un nombre croissant d'études ont appliqué et évalué la théorie de Hillier dans une multiplicité de cas d'études différentes. Cependant, sous-tendant ces travaux, plusieurs hypothèses simplifient les systèmes commerciaux en se basant sur de simples mesures de la densité de magasins, tandis que la forme urbaine est représentée par le réseau de rues et décrite par ses propriétés de configuration.*

*Dans cette partie, nous allons passer en revue ces différentes approches et définir le cadre théorique de cette thèse. À partir de la définition initiale de la théorie de l'économie du mouvement, nous proposerons de nouveaux questionnements de recherche afin d'enrichir cette définition originale en considérant des aspects issus à la fois de la morphologie urbaine et de la géographie de la vente au détail.*

# 1. The Point of View of Retailing Geography

An extraordinarily large and diverse theoretical literature of retailing has been developed in the last century. Retail studies are varied depending on the approach, disseminating a diverse mix of scientific domains such as geography, economy, psychology, sociology, urban planning, architecture and regional sciences, among others.

Wholly-comprehensive literature reviews have been regularly proposed throughout the last decades; however, the growing number of approaches developed from the contribution and combination of different scientific domains, as well as the complications brought about by recent methodological procedures made this task highly challenging. In the last decades, this work has been successfully undertaken by academics as Stephen Brown (1991,1992a,1992b,1996), Wrigley and Lowe (2002) in the English-speaking literature. Within the francophone community, where the geography of retailing attained its own intellectual independency <sup>11</sup>, interesting insights and extensive reviews have been more recently proposed, among others<sup>12</sup>, by Mérenne Schoumaker (2003), Wayens (2006), Lemarchand (2008), Madry (2016).

Based on the aforementioned author works, this first part of the literature review proposed aims to provide a general framework of the retail geography discipline. A threefold classification considering the main approaches developed along the last century has been firstly proposed and described by Brown (1991) and successively acknowledged in the same form by several authors (Wayens 2006, Lemarchand 2008, etc.).

**Neoclassical approach** is mainly constituted by scientific works and theories developed in the fields of economic geography, land use and land planning studies. This approach reached its apex in the second post-war, in particular in the decades between the '50s and late '60s when Central Place Theory, Spatial Interaction Theory, Bid Rent Theory, and the Principle of Minimum Differentiation dominated the academic debate in the field of the economic geography. Beneath these standing and well-defined pillars, are two hundred years of extensive antecedents and contributions from geographers and economists studying the relationship between city and commerce (Brown 1989, 1993; Dawson 1969; Madry 2018). As all dissertations dealing with retail distribution, a brief description of these four main theories

---

<sup>11</sup> As deeply reviewed by Lemarchand (2008) and Renard-Grandmontagne (2016) French geographers along the 21<sup>st</sup> century have dedicated a specific attention to the geography of retailing (English name from Wrigley and Lowe 2002) (French name: "*géographie du commerce*"). In the last fifty years, this discipline has been recognized as an independent scientific domain at the crossing point between geographic and economic studies; with the creation of a specific Commission within the French National Geographic Committee CNFG starting from 1977 and a specific research group at the National Center for the Scientific Research CNRS between 1984 and 2000 by Alain Metton. An important number of contributions is attributed to Belgian academics, equally involved in this specific scientific domain recently developing a more quantitative approach comparatively to French authors (Madry 2016). More recently the French-speaking retail geography community is gathered around the *Institut pour la Ville et le Commerce* <https://www.institut-ville-commerce.fr/>

<sup>12</sup> Colette Renard-Grandmontagne (2016) offers an overview about the main research axes and evolutions of 50 years of the French National Commission of the Retailing Geography

is inevitable; critics to their assumptions as well as changes both in society and in the retail industry have been stimulating the scientific debate and triggered the exploration of innovative approaches.

Starting from the '70s, **behavioural approaches** have been proposed as a response to the several criticisms related to the economic vision of individuals. It encompasses empirical and cognitive studies and it investigates the decisional process of consumers and the factors influencing the frequentation of retail locations.

Since the '90s, the complete maturity of the new format (large surfaces and store chains) caused a growing number of works investigating the spatial organisation and strategies of store chains at multiple scales. **Structural approach** investigates both the spatial organisation of these new formats and their impact on more traditional independent store distribution. The importance of these retail formats, beyond the district level, has engendered the reaction of the local and national administrations. As a consequence, a large number of studies investigated the legislative and administrative framework and their intervention on the retail industry and its spatial organisation.

These approaches should not be considered as a chronological succession where the development of one approach informed the others; they rather represent the prevalent approach in a given historical moment as a response to specific research questioning raised from the challenges imposed by societal and technological evolution.

With the entering in of the digital era, new technologies are revolutionising consumer behaviour and modifying the way we shop; the emergence of new retail formats such as the e- and m-commerces create interesting questions about the dematerialisation of the retail sector. The consequent reaction of retailing geography studies is a growing interest in the spatialization of retail: the analysis of different environmental, social and cultural conditions where the ongoing transformation of the retail sector is taking place. A fourth **contextual approach** is emerging from these recent works.

In this work, we focus more specifically on the urban morphological context; despite the recent trend of dematerialisation of the retail system, the traditional brick and mortar format is still resisting and it represents a fundamental component of public spaces, influencing their social and economic vitality and attractiveness.

But before discussing our specific approach let us briefly review these four aforementioned approaches which represent an important source of insight and part of the theoretical background to this dissertation.

## 1.1 Neoclassical approach

Despite the strong English-speaking dominance in the post-war planning and land value studies, the main theoretical pillars might also be found in the German literature (Carol 1960). The importance of Central Place Theory (Christaller 1933, Lösch 1945), Spatial Interaction Theory (Reilly 1931), Bid Rent Theory (Haig 1927) and the Principle of Minimum Differentiation (Hotelling 1929) have been widely acknowledged beyond the scientific community: private

retail organizations, as well as retail planning policies of numerous countries, have been profoundly influenced in their locational decision-making from these works (Dawson 1980).

**Central Place Theory** (CPT) has been attributed as the combined works of Christaller (1933) and Losch (1954), proposing planning theories about the spatial structure of the commercial space. CPT proposes an economic model able to explain the number, size, spatial threshold and final functional composition of commercial centres in a microeconomic free-trade market world with specifications about geographical conditions (Beavon 1977, Brown 1992). In order to address these questions, Christaller based his CPT on several strong assumptions: consumers are small, well-informed, rational, agents with similar purchasing powers aiming to maximise their utility function; providers of goods are small, rational agents aiming to maximise their profit. The geography as conceived in CPT considers a uniform space where agents are uniformly distributed: in this isotropic world, growing distances are positively related to transport costs and, therefore to the price of a good. Inversely, the demand for a good follows a distance decay function.

From these functions, two key-variables of CPT are defined: the range of a good and the threshold. The former is given by the maximal distance a consumer would travel to purchase a good while the latter represents the minimum demand necessary for a store to stay in business. The combination of these two variables is used by Christaller and Lössch to explain the well-known hexagonal pattern allowing an even repartition of market areas and equal distances between same-goods stores. Range and threshold values are influenced by the frequency at which goods are purchased. Higher frequencies correspond to smaller values of range and threshold and consequently to a market partition defined by smaller-sized hexagons, as in the case of daily-based goods; on the contrary, products rarely bought like furniture or jewellery correspond to a partition made by larger hexagons. Overlapping market areas of different sizes result in the hierarchical organisation of centres as illustrated in Fig.1, determining the functional composition of centres: higher hierarchical levels combine a variety of stores offering both high- and low-order goods while lower hierarchical levels offer only low-order goods to smaller marketing areas.

The spatial organisation of stores as described in CPT represents the optimal spatial repartition that would occur when the aforementioned strong assumptions are respected: this spatial optimum is obtained through the maximisation of the number of stores, minimising the consumer-based sizes of market areas.

Despite Christaller's criticisms of Kohl's theory (Kohl 1850) about the emphasis on the effects of traffic factors on retail distribution rather than its economic causes, the same author is aware of the presence of transportation inefficiency in CPT as originally conceived. Kohl's traffic principle, in which travel routes follow the most economical paths between centres, is not respected since lower hierarchical levels lie on the connections between two high hierarchical level cities. The same Christaller proposed a rearrangement of the spatial organisation of centres.

Fig.2. Die Ergänzungsgebiete im System der zentralen Orte.

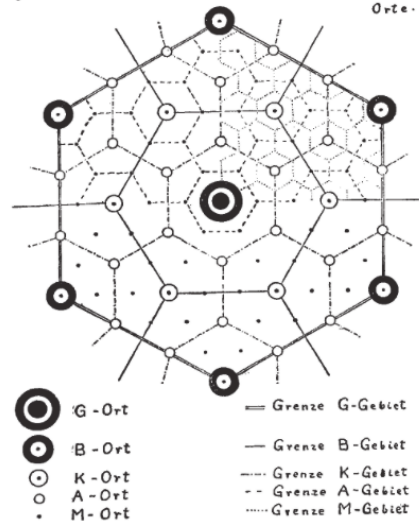


Figure 1 The illustration of nested hexagons in Christaller's central place theory (Christaller 1933, p. 71).

Kohl principle, as well as several CPT assumptions, have been stimulating a constant debate in the following generations' literature, reaching contrasting observations and conclusions. On one side, the strong simplifications of the economic and geographic world underlying the CPT have been considered as "too implausible to serve as a basis for empirical work" (Marshall 1978). Critics are concerned, for example, with the single-purpose shopping, the even distribution of identical consumers (Kivell et al. 1980), the "optimizing men" behaviours assumption (Golledge et al. 1966) and the "static-equilibrium" on which the representation of the spatial structure is built on, which does not measure the dynamics of retail evolutions (Brown 1993). On the other, it has been considered as "the most innovative and successful attempt to construct a fully realised theory of spatial structure" (Stevens 1985); an extensive literature has been developed from the original CPT formulation, each time relaxing or modifying each underlying assumption (Potter 1980, Berry 1963). All these variations tried to assess the validity of the CPT, recognising a hierarchical structure of places in a free market condition. Despite more or less recent criticisms, CPT has been recognized as one of the fundamental theories in all disciplines interested in the study of the spatial distribution of centres in cities and regions (Batty 1994; Fujita, Krugman, Vernables 1999).

The second theory of land planning is the **Bid Rent Theory** (BRT); this land use theory was first proposed by von Thünen (1826) to study agricultural land allocations of areas surrounding an isolated town and represents one of the main references of Christaller's CPT. In BRT, the economic function to be minimised is the total costs given by the production and transportation of food supplies to the town market (Fig.2.A). As for CPT, traditional assumptions of microeconomics underlie BRT. Here again, additional geography assumptions consider a homogenous space: transport is possible in all directions and transport costs increase with the distance from the town centre towards peripheral areas. Unplanned competition between agents (farmers and landowners) defines the distribution and allocation of land values and consequently, land uses are described by concentric patterns of increasing distances from the town centre.

Haig (1927) reformulated von Thünen's BRT in the case of land use distribution and firm spatial organisation. The principle on which Haig based his version of BRT is that settlements should be considered as an economic process: their form and spatial structure are both defined by demand and supply market forces. City centre represents the convergence point of the transport system where accessibility is maximal for all agents providing the maximal potential of the market space. Land users (commercial, industrial and residential agents) compete with one another for land close to the city centre, trying to maximise their profitability. In the long run, all urban sites are occupied by the activity capable of paying the highest rental. Firms and offices locate at the core of the city (Central Business District-CBD), residential areas and industries surround the CBD at intermediary distances and land uses which makes the less profitable use of a given site forced towards the most peripheral distances.

Alonso (1964) proposed further developments: maintaining the underlying traditional assumptions of BRT, the author outlined bid rent curves describing the willingness of individuals and firms to pay. The slope of these curves reflects the sensitivity of each land use in relation to accessibility. Similar to his predecessors, Alonso's BRT identifies a highly segregated core of offices and retail activities surrounded by rings of manufacturing, and residence (Fig.2.B). Two additional contributions of Alonso's BRT might be resumed as follows: firstly, land price and built-up density are described by a decay function from the CBD. Secondly, BRT is verified both horizontally and vertically: retailers located at the ground floor outbid services and offices at higher floor levels. The pattern defined by the BRT might be observed at the urban neighbourhood, but also at the individual street level as described by Villain (2011) and Wayens (2006).

The three versions of the models confirm a monocentric structure of land use and land value where the accessibility to the centre is simplified with the raw Euclidean distance. Alonso observes how considering street network distances, regulations zoning and natural/artificial barriers might alter the concentric spatial layout. Further developments readapted BRT to multicentre city patterns (i.e. Ladd and Wheaton 1991).

As observed by Brown (1993), BRT has been mainly applied and developed in studies related to land value and land density with specific attention to the residential sector. Fewer investigations have focused on the application of BRT to study the distribution of retailers: following the same concept of the BRT, Garner et al. (1966) and Scott (1970) defines bid rent curves for retail categories by describing their sensitivity to change in accessibility (Fig.2.C): commercial activities try to locate in the most central/accessible city position in order to maximize their catchment area, therefore, maximizing customer attraction. The rent they are willing to pay falls rapidly with the distance from the centre. Consequently, those retailers trading accessibility for lower rent are pushed to peripheral regions of the centres while on the contrary, high order retail functions are located in its core. Shallower bid rent curves describe the former retail categories while steep curves are associated with the latter. The same concentric arrangement is detected around community-level centres, department stores and specialised retail areas. While Scott's model is based on subjective estimations of bid rent curves with a more accurate description of retail categories, Garner et al.'s model

incorporates the concept of total sales volume as a decreasing function of the distance from the “peak intersection” and the concept of commercial centres influence. Different typologies of retail format (Centres, Ribbons and Specialized Areas) are integrated into Garner et al.’s BRT model and successively reviewed by Davies (1976).

**Spatial Interaction Theory (SIT)** relies on the hypothesis that consumers trade off the attractiveness of alternative shopping areas against the deterrent effect of the distance from the city centre (Clarkson et al. 1996). SIT finds its origins from the pioneering works of Reilly (1931), based on the analogy to the Newtonian law applied to retail spatial distribution. While consumers in CPT evaluate shopping area destinations as a function of their distance to

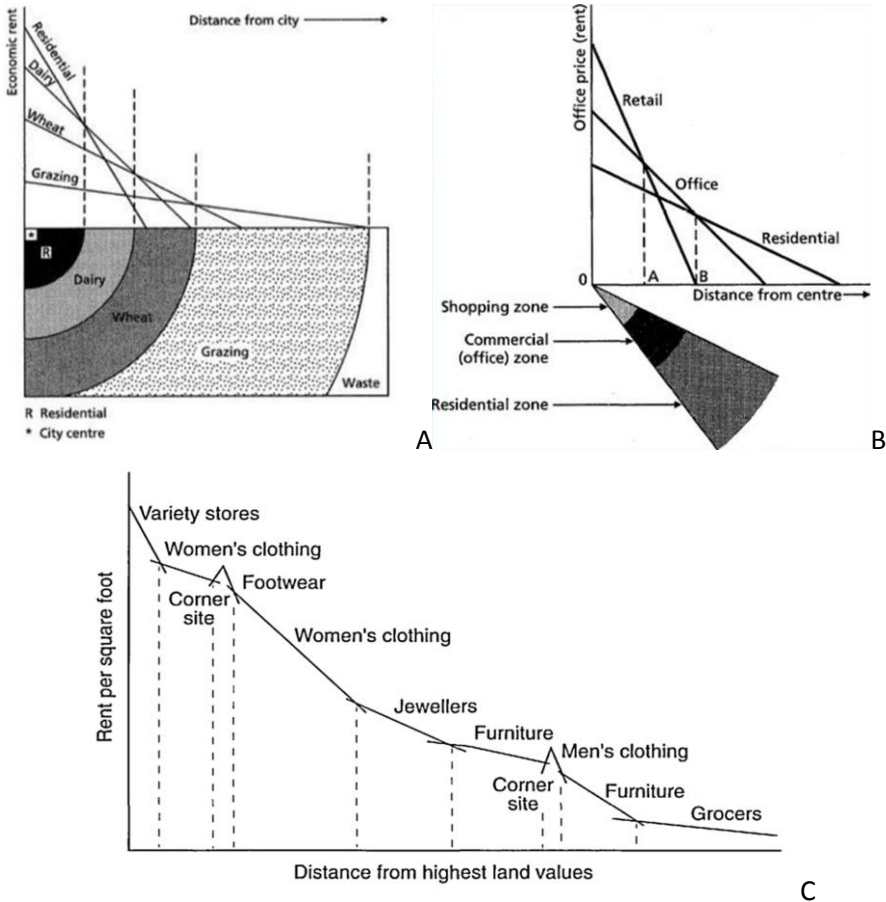


Figure 2 Illustration of the Bid Rent Theory: A) von Thünen 1927, B) Alonso 1964, C) Scott 1970

centres, consumer choice in SIT is derived from a more complex process involving weighing of alternatives. The author states that two places “draw trade from an intermediate city or town approximately in direct proportion to the population of the two cities and in inverse proportion to the distances of the two cities to the intermediate town” (Reilly 1931, p.9). The proportion of trade attributed to each city is defined as follow:

$$\frac{T_a}{T_b} = \frac{P_a}{P_b} \left( \frac{d_b}{d_a} \right)^2 \tag{1.1}$$

Where T is the trade measure of the two cities a and b, P the population and d the distance to each market area. Specific parameters and variables have been subject to further developments from the original Reilly's heuristic. Among them, a significant contribution came from Converse (1949) through the identification of the inertia factor reflecting the reluctance of consumers to move towards the retail location. From Reilly's theory, Converse develops the following measure for the determination of the Break Point (BP).

$$BP = \frac{D_{ab}}{1 + \sqrt{p_1/p_2}} \quad (1.2)$$

The application of the (1.2) allows the definition of the relative position between two places of size p1 and p2 (population or other weighting measures) spacing D at which the influence of the two cities is equal. This measure allows significant advances in trade area estimation and investigation (Goldstucker et al. 1978): trade areas might be outlined for each place allowing the identification of statistical variables of the surrounding socio-economic environment that could be associated with its performance (Ghosh and McLafferty 1987). It is interesting to highlight how in the case of equal place size (or weight) p, equivalent trade areas are described; this would coincide with the spatial partition defined by the Proximal Area Model with the well-known Thiessen tessellation procedure (Thiessen 1911).

The most important conceptual refinement is the contribution of Huff (1962, 1963) extending the original Reilly's laws of retail gravitation for multiple attributes and retail centres and the following work of Lakshmanan and Hansen (1966) address how areas can overlap. Continuous evolutions of Huff model (Ghosh and McLafferty 1987, Gonzalez-Benito et al. 2005, Li and Liu 2012, Roy and Thill 2003) allowed a clearer representation of the realistic interaction between retailers; for this reason, SIT is one of the most implemented approaches among academics and practitioners<sup>13</sup> (Brown 1992). Nonetheless, several issues have been highlighted considering the lack of tests to measure the closeness of fit (Kivell and Shaw 1980), the subjective calibration and variable indexing and the difficulties in defining measures of attraction and distance (Clarkson et al. 1996) but also conceptual problems relatively to its theoretical weaknesses (Jensen-Butler 1972). Recent evolutions in commercial habits might represent further potential limitations that need to be integrated into a revisited model (Friske and Choi 2013).

**Principle of minimum differentiation** PMD (Hotelling 1929), which is less known than CPT and SIT, represents the basis for micro-retail studies. With the same traditional assumptions of neoclassical theories, Hotelling demonstrates that two retailers selling identical products, if they can relocate, they tend to maximise their hinterland to maximise their profit. Consequently, the optimal location is the one seeing the two competitors close each other at the centre of the market area, converging to the same results of the CPT and BRT. This principle has been popularised by the monodimensional version of the two ice-cream sellers

---

<sup>13</sup> as a decision support model for site selection, and performance evaluation of a site as for example in the ESRI application Business Analyst or in the specific case of the local chamber of commerce it represent the underlying model implemented in the specifically conceived MagCalculator



on the beach (Losh 1954) even if the main focus of the original version of the PMD was the study of price stabilisation in a context of market competition (Brown 1989,1992).

As for the previous theories, the underlying assumptions assumed by the PMD are not encountered in real life and several critics have mainly considered the assumption about the well-informed agents (Browet 1997, Thill and Timmermans 1992); Eaton and Lipsey (1975) observed that, by relaxing these hypotheses, no reasons could be found to explain why firms should cluster. The implicit benefits of agglomerations are not directly explained by Hotelling's PMD. From a consumer perspective, store agglomerations provide higher variability of goods and values compared to the one that might be provided individually by a retailer (especially for high-order products, Brown 1989). This wider offer of merchandise helps the consumer in the activity of research and comparison of products, thus reducing the risks of "bad purchases" derived by the incomplete information of the market which is in contrast to the well-informed-agent assumption underlying PMD. Similarly, from the retailer point of view, possible damages from a higher competitive environment are more than balanced by a minimisation of the risk of a wrong location and a higher flow of consumers attracted by the overall agglomeration. Other factors might be at the origin of this clustering behaviour such as historical, urbanistic and legislative reasons (Wayens 2006).

Further developments have been partially rehabilitated PMD through i) the incorporation of agglomeration economies, positive externalities and cost-reducing benefits flowing from spatial propinquity (Fujita and Smith 1990, Thill and Thomas 1987). ii) relaxing the hypothesis of perfectly informed agents (retailers and consumers) (de Palma Ginsburgh and Thiesse 1987, Pascal and McCall 1980); iii) permitting consumers to minimise costs with multipurpose, multistep and comparison shopping behaviours (Fujita et al. 1988). Despite profound changes in its underlying assumption operated by these works, Hotelling PMD is still considered to be one of the main principles defining retail agglomerations.

Recent technological progress allowed the investigation of Hotelling's agglomerative hypothesis through the application of advanced statistical procedures. As we discuss in Part 2, a great variety of measures in the fields of spatial and economic statistic have been developed in the last few decades, aiming to identify and describe the clustering behaviour of firms. In particular, Okabe and Baddeley research groups have contributed to the understanding of point pattern analysis. In the context of retail agglomeration studies, while some authors might consider these great efforts as a support to the PMD (Okabe, Asami and Miki 1985, Rogers 1969, etc.), others could consider the continuous investigation of methodological procedures as a lack of general agreement about its validity.

Independently from the underlying reasons behind retail clusters, the agglomerative behaviour of stores represents the main focus of retailing studies: the spatial delineation of commercial centralities represents the most critical step influencing every spatial and economic analysis of retailing geography. As further discussed in the following sections, spatial statistics approaches provide valid tools which could contribute important insights only if used in combination with a specific-contextual knowledge of the phenomenon under study.

## 1.2 Behavioural approach

The theoretical models reviewed in the previous sections are based on “*positivist, neoclassical premises, which assumes essentially, that there is an identifiable order in the material world, that people are rational, utility-maximising decision-makers and that economic activity takes place in a freely competitive manner*” (Brown 1993, p.10). These theories are deductively obtained: given highly simplified assumptions, they predict patterns of retail activity. As observed by the same author: “*the close correspondence between these normative theories and empirical evidence might lead to infer that the latter proves the former*” (Brown 1993, p.13).

Nonetheless, the spatial patterns of retail observed in an urban space may result from a different combination of processes other than those economic principles of utility maximisation underpinning all the theories reviewed up to now (Scott 1970; Brown 1993,1994; Wayens 2006).

The strong underlying assumptions about agents responding exclusively to economic principles is a mechanical vision of the “optimising men” that similarly to the homogenous assumption of the space is meant to simplify the high complexity and variable dimensionality of the retail locational problem. Other approaches have been widely studied along the twentieth century in other disciplines of human and social sciences: starting from the ‘70s, as a reaction to the critics of neoclassical approaches, wider attention has been given by sociologists to the study of agent behaviour. The final goal of these studies is to provide a more realistic representation of human behaviour as an alternative to the simplistic definition given by economists (Hubbard 1978).

While neoclassical studies are characterized by a strong deductive approach, the mainly inductive approach of sociologists focus on the study of agents at the individual level with the consequent proliferation of studies on different spatial behaviours, perceptions and decisional processes. While the former is based on a theoretical framework, the latter adopt empirical methods based on observation-generalisation approaches.

Three subgroups of studies might be described (Brown 1992, Wayens 2006). Empirical-behavioural studies, mainly based on observations; cognitive-behavioural studies, combining both observations and cognitive aspects; a humanist approach, focusing on emotion and sentiment associated by consumers to market places. In each of these approaches, we may distinguish between demand and supply-side studies, as well as those works focusing on the perceived element or its influence on the perceiver.

The large amount and variety of **empirical-behavioural studies** finds its origins in the first pedestrian counts for store location purposes of Hurd (1903) and Brisco (1927). From direct (counts) or indirect (surveys) observations, shopper behaviours might be studied considering both the influence of individual factors describing consumers (socioeconomic environmental and temporal descriptors, demand-side) and the specific format of the retail centres (supply-side).

Following Brown (1992) and Wayens (2006), when considering the study of the impact of consumer factors on behaviour patterns, they highlight that:

- shopper behaviour is strongly dependent on time and day of the week and depending on consumer profile variables such as sex, age, socioeconomic level etc.
- the extent and the shopping areas visited might have different patterns depending on the locations where consumers live, work and their socio-professional job category; these characteristics influence the mobility capacity of each individual who can afford to buy products from higher or lower levels of the retail hierarchical structure (Davies 1969);

Similarly, in the findings of the first group of studies, the same authors resume that:

- the presence of magnet stores, which consist of large-sized stores within a commercial area, directly influences consumer circulation patterns. This phenomenon might be observed both in planned and unplanned centres; in the latter, the importance of this magnet force upon the surrounding spatial organisation might vary from neighbourhood to neighbourhood (Davies and Bennison 1978).
- customer/shopping interchange tends to occur between shops of similar/compatible or complementary trade when shop to shop movements are investigated (Jones and Simmons 1990)
- the importance of consumer entry/exit point influences the consequent circulation paths which in unplanned shopping centres are materialised by parking, bus stops, railways, and underground stops (Morris and Zisman 1962, Bennison and Davies 1977, Howard and Davies 1986)
- the frictional effect of distance: distances are always short, and consumers try to minimise the expenditure of effort (Zipf's principle 1949, Johnston 1973, Brown 1992)
- the type of environment and the space dedicated to pedestrians have a direct influence on both the behaviour and the customer composition (Mérenne Schoumaker 2003)

Multiple complications come with observation data needed in behavioural studies. Two approaches might be adopted: the first considers the description of individuals' behaviours and is based on survey approaches (diaries, interviews, street audit tools, GPS and accelerometers etc.), and it might raise issues of reliability and high costs in terms of both time and resources. The second analyses the pedestrian flow of people in a given place traditionally with mechanic or electronic counting methods. Through this approach, only the overall behaviour might be studied; important behavioural factors at the individual level might be overlooked. Moreover, although technical limitations are being overcome by new procedures such as Wi-Fi, surveillance cameras, mobile phones, GPS etc. strong constraints are still imposed by the juridical framework. The use of volunteering geographic information (VGI) or social media geographic information (SMGI) has recently been presented as a tool to introduce the human perception and uses within urban planning studies (Campagna et al.

2015, Massa and Campagna 2015, Floris and Campagna 2014 etc.). Nonetheless, several critics have been questioning the reliability of these alternative sources of data.<sup>14</sup>

**Cognitive-behavioural studies** integrate the cognitive dimension of consumer behaviours aiming to explain not only the spatial pattern deriving from individual and contextual factors but also the spatial choice of the consumer. The underlying concept of this approach is that the consumer choice is made upon an imperfect and subjective knowledge of the available opportunities of the market. Here again, we might distinguish between those works focusing on the individual level variables influencing the perception of a place and those analysing how different retail centres are perceived by individuals (i.e. Rushton 1969, Potter 1979, Mérenne Schoumaker 2003).

Within the cognitive approach, we might find those works initiated by Lynch a seminal study on the perception (the image) of city centres (Lynch 1960). Through the adoption of qualitative analysis based on mental maps and one-to-one surveys in US cities, Lynch identified five typologies of physical forms that inhabitants use to describe their individual perception of the city: paths, edges, districts, nodes and landmarks. Lynch highlighted how stores play the prominent role of landmarks in the mental representation of the city centre. However, while natural and historical landmarks are considered for their static location, stores acquire a dynamic connotation when considering their rhythmic layout along streets (Lynch 1960 p.97).

Empirical and cognitive behavioural studies allow the redefinition of the perfectly informed and optimiser "*homo oeconomicus*" with a more realistic description of the variability of human nature in regards to the retailing activity. These approaches, along with the one from the economic geography, share the research for simplification of the reality assuming different underlying hypotheses; statistical procedures are implemented to identify recurrences of spatial patterns supporting the process of generalisation. This modelling procedure is necessary to allow the comprehension and the formalisation of those concepts explaining the relationship between spatial patterns and the underlying human behaviours. As discussed in the previous section, the generalisation approach allowed the integration of findings of behavioural and cognitive-behavioural studies within the framework of classical theories, such as multipurpose and multistep shopping trips.

This research of generalisation is rejected by the humanist approach, instead prioritising the study of the profound meaning behind agents' actions more than the process itself. This peculiar approach is based on individual detailed interviews analysing the specific sensations and motivations behind a person's perception of the retail activity and place. Despite the small literature in the specific domain of retail studies, this approach might represent the point of departure from which a new version of human geography might be elaborated in order to answer to the need of individuality and respect of ethical values required by new generations of consumers (Dart 2018).

---

<sup>14</sup> A research activity carried out during the last three years has set up a data base gathering geo-located tweets of the study region from the streaming tweeter API. Starting from March 2017, more than 1.5 million tweets have been collected. Although this source of data has not been used, the conclusive part discuss how it might be integrated within the procedure developed in this dissertation, which could also be combined with the results of this research project.

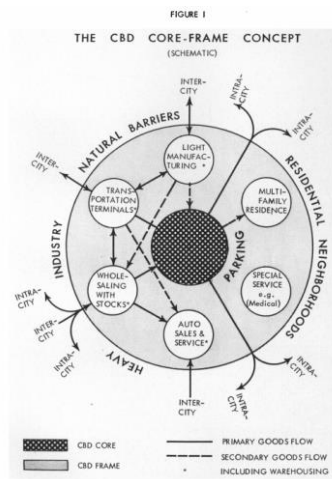
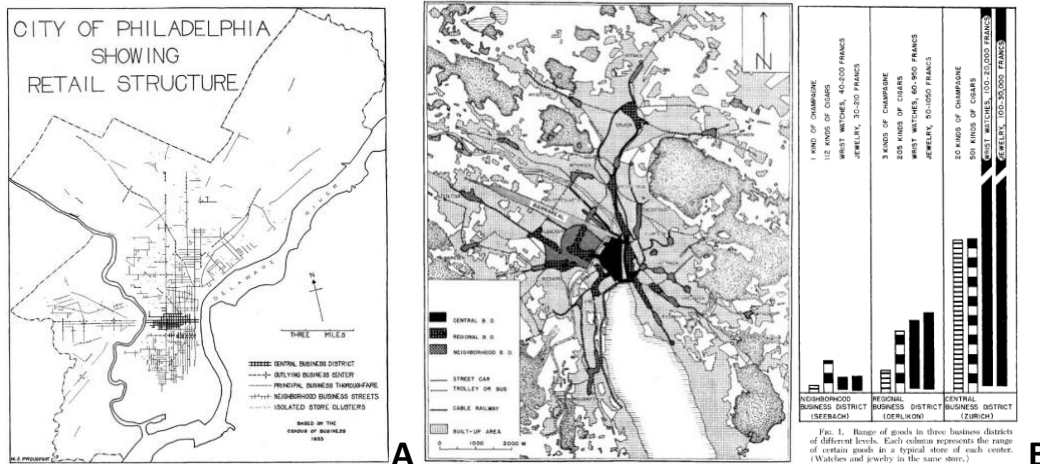


Figure 1 Application of the Central Place Theory to the intraurban space: A) Retail Structure of Philadelphia from Proudfoot (1937); B) Hierarchical retail structure of Zurich from Carol (1960) with the related variety of goods (champagne, cigars, watches and jewelry) observed in each center level. C) Core-Frame model of Horwood and Boyce (1959).

Future research in behavioural studies might consider the contribution of simulation analysis: advancements in computing efficiency and affordability have been catalysing and increasing the research activity in this field of research. The Agent-Based Model (ABM) approach is a novel perspective on how to simulate consumer behaviour. It allows a more realistic simulation of the behaviour of individuals and to simulate the observed patterns. Simple rules describing the micro-behaviour of consumers can simulate the complex emergent structure of retail distribution. In the last twenty years, the development of specific platforms for the (ABM) geographical studies have reinvigorated the interest of the academic in behavioural studies<sup>15</sup>. Four inputs are required by ABM: the physical environment through which agents interact, the definition of autonomous or heterogeneous decision making agents and the set of rules defining both the relationship between agents and environment and those between agents. In the retailing context, stores and consumers might be modelled in separates or combined simulations. The instructions defined by the analyst and based on empirical

<sup>15</sup> Two of the main ABM platforms used in geography studies: NetLogo and GAMA. The former has been firstly released in 1999 by Uri Wilensky at the CCL Northwestern University (Illinois, USA) handling raster data. The former, developed between 2007 and 2010 by a joint research program between French and Vietnamese universities.

observations and theories, enable the simulation of both the individual behaviour of agents and their mutual (social) interactions.

The implementation of ABM in behavioural retail geography might be considered under the following different strategies: it could be used as an alternative way of simulating the "homo oeconomicus" as described by neoclassical hypotheses or, as a way of overcoming these strong theoretical assumptions, implementing different observed behaviours (in reactive ABM). An alternative approach might integrate the perceptive and cognitive processes of agents (reactive-perceptive-and-cognitive ABM).

Nonetheless, empirical studies are still necessary to validate those behavioural rules injected in ABMs: the simple fact of being able to reproduce an empirically observed pattern is not enough to validate an ABM since different simulations with different rules might lead to the same outcomes (problems of equifinality, von Bertalanffy 1968).

Despite its advantages and potential use over the possible alternatives, ABM technique has yet to be extensively applied and investigated in retail geography studies.

### 1.3 Micro retail in intra urban space: combining theories

Before proceeding with the description of more recent approaches, it is fundamental to review works specifically focusing on micro-retail distribution within urban areas that are applying neoclassical theories and combining them with behavioural approaches.

Since their early conception at the beginning of the twentieth century, land use and planning theories have strongly influenced the research of models of retail distribution at the intra-urban level. Applications and adaptations of CPT and BRT to the study of retail activity distribution represented one of the most discussed approaches of retail locational studies.

Proudfoot (1937) firstly transposed the concepts of CPT to the detailed intra-urban retail distribution of several American cities (Fig.2.A). He proposed a first structural model representing the commercial distribution and organisation through the definition of five categories of retail districts: Central Business District (CBD), outlying business centres, principal business thoroughfare, neighbourhood business street and isolated store clusters. Following this example, several researchers applied the CPT to different English and German-speaking cities<sup>16</sup>: probably the most known in the European context is Carol's analysis of Zurich (Carol, 1960) where a hierarchical representation of commercial districts on four levels is outlined and described (Fig.2.B).

Combination Theories have also been developed from the consideration that CPT, BRT and PMD do not operate independently and concentration and zonation forces define together the spatial arrangement of land use. An early example in the case of US cities is the "core-frame" model by Horwood and Boyce (1959): two main areas of the city are defined following

---

<sup>16</sup> Carol's 1960 provides an exhaustive review of the German and English speaking literature before the 60s. Some of the authors he referenced are: Hassinger (Basel 1927), Struttgart (Innsbruck 1928), Christaller (south Germany 1933), Dickinson (East Anglia 1934), Proudfoot (Chicago 1938), William-Olson (Stockholm 1940), Harris (Salt Lake City 1940), Mayer (Chicago 1942), Carol (Karru, South Africa 1952), Brush (Winnconsin 1953), Spelt (South-Central Ontario 1955), Murphey, Vance and Epstein (1955) etc.

the concentric patterns of BRT, while PMD explains the coalescence of specific concentrations of retail categories within the core and/or the framework ring. Several authors have applied the core-frame model to several western countries,<sup>17</sup> always highlighting its descriptive and prescriptive value. Further developments have been proposed, integrating linear agglomerations (arterial) which have been described as the spatial manifestation of countervailing forces of agglomeration and dispersion (Berry 1963, Davies 1972).

While CPT represented the main approach in the study of retail centres within cities for almost thirty years, peculiar retail agglomerations were receiving a milder but growing consideration. During the '50s, the arrival and growth of chain stores, in conjunction with hypermarket shopping centres, were rapidly changing the retail industry. New retail forms were appearing along the major axes between and within urban areas as well as peripheral specialised clusters (Rolph 1935, Garrison 1950, Merry 1955)<sup>18</sup>. The increased use of the car reinforced the presence and extension of these new agglomerations, transforming the retail structure of cities ("*second revolution of the retail*", Desse and Lestrade 2016).

Motivated by these emerging features, B.J.L. Berry (1959) stated that despite the notion that CPT provides "*a logical and satisfactory explanatory scheme of the entire set of nucleated business centres [...], (CPT) is not adequate to explain stretches of highway-oriented or urban arterial businesses. Neither is it adequate to account for specialised functional areas, such as automobile row*" (Berry 1959).

Following these observations and the flood of empirical and conceptual refinement of the CPT during the early post-war ascendancy of social physics and spatial sciences (Brown 1995), Berry revisited the original CPT, proposing a new hierarchical model for the retail spatial organisation in urban spaces (Berry 1963,1967), integrating aspects of the land value theories, analysis of the pedestrian movement at a large scale, and observations on specialised concentration. The analysis of Chicago's business districts led this author to the identification of three morpho-functional centres:

- (i) A hierarchical structure of business centres characterised by high density, category-mixed agglomerations.
- (ii) Highway oriented ribbons identifying store concentrations with a linear layout alongside highways, routes, urban roads with or without a functional specialization (i.e. home supplies, furniture decoration etc.; motorcar industry-related etc.).
- (iii) Specialised functional concentrations of specific retail categories taking advantage from the customer comparative behaviour and/or from agglomeration economies.

---

<sup>17</sup> i.e. Newcastle-upon-tyne (White 1975), Stockport (Potter 1982), Belfast (Brown 1987), west London (Davies 1976)

<sup>18</sup> Unpublished PhD thesis reported by Berry 1958

Each of these retail forms is divided into sub-groups considering both size and complexity of retail agglomerations. While the first group follows the fourfold subdivision of Carol's classification, the other two represent the most interesting advance of Berry's model (Fig.3.A). Additional insights are provided by the author when analysing the internal hyper-local layout of retailer business centres as a consequence of specific economic and behavioural forces. More precisely, he describes the influence of the relative position of retailers on pedestrian movement and the concentric spatial organisation of retail categories (with clothing and jewellery at the centre) following the BRT. For the core of the city of Chicago (Fig.3.B), he finally suggests a spatial structure resulting from the combination of a concentric arrangement of the retail around the "best localisation" with sub-spaces of category and ethnical specialised local concentrations (Wayens 2006). Another seminal work is the study of the retail distribution in Boston and Detroit by Cohen and Lewis (1967) where the notions of retail *form* and *function* and their relationship are firstly proposed and investigated.

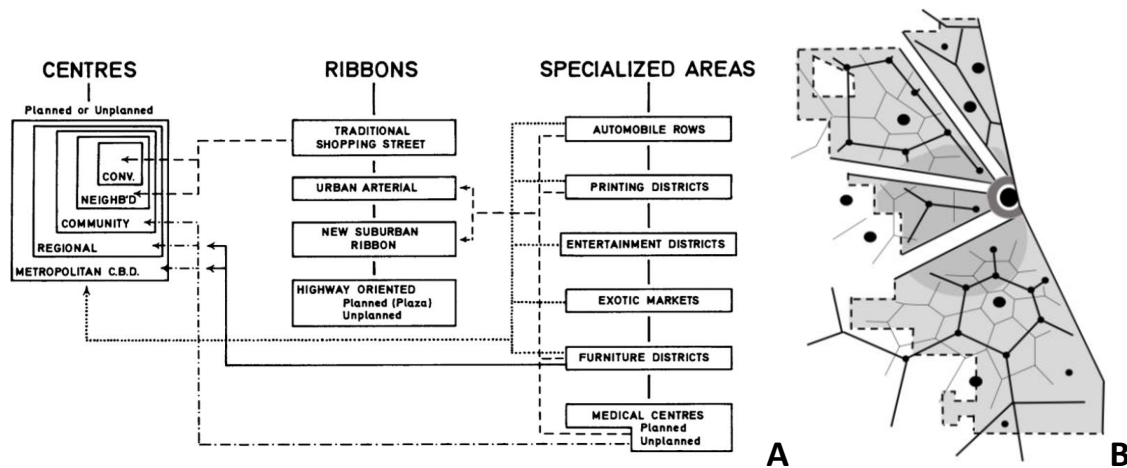


Figure 3 Berry's hierarchical model for retail system (A) and its application to the city of Chicago (B) (reproduction from the original cover of Berry's 1963 "Commercial structure and commercial blight").

A few years later, Davies (1972a, 1972b) outlines several differences of retail organisation between British and North American cities considering Berry's hierarchical model, specific observations of retail agglomerations in British cities of several authors,<sup>19</sup> and his personal study of the retail structure of Coventry. Some of his insights consider the higher dominance of the city centre in European cities within the hierarchal structure of nucleated agglomerations; a minor development of hierarchical lower levels is observed when compared to North American cities where the process of suburbanization, at that time, was more intense. Moreover, ribbon concentrations in the European context are associated to the historical urban development of the city, often inherited by the public transport system configuration; consequently, these agglomerations are identifiable more for their spatial form than for their specialisation. Further observations consider the impact of the population density and its socioeconomic characteristics on the retail hierarchical structure. Similar observations have been highlighted in the French literature by Beaujeu-Garnier (1977). While

<sup>19</sup> Davies 1972 refers to the description of specific retail configuration analysis realized: retail nucleations of London (Hartley 1961 ; Carruthers 1962); ribbons of Liverpool (Parker 1962); specialized areas of Glasgow (Diamond 1962) and of Cardiff (Carter and Rowley 1966)



both Berry's and Davies' studies focus on the specific relation between retail structure within the urban space privileging specific sociocultural contexts (of English-speaking cities), Beaujeu -Garnier follows the French academic tradition, exploring retail organisation within different countries, and providing a more general and qualitative description of the retail structure (Lemarchand 2008).

The analysis of the core area of the city represents one of Davie’s major contributions. Based on the land use studies of Burgess, Hoyt, Harris and Ullman, on the core-frame models of Horwood and Boyce(1959) and on Garner et al.’s model for the differentiation of retail activities within shopping centers (based on BRT), a specific structural model for the hyper center of European cities is proposed (Davies 1974) and applied to the core of the British city of Coventry (Fig.4).

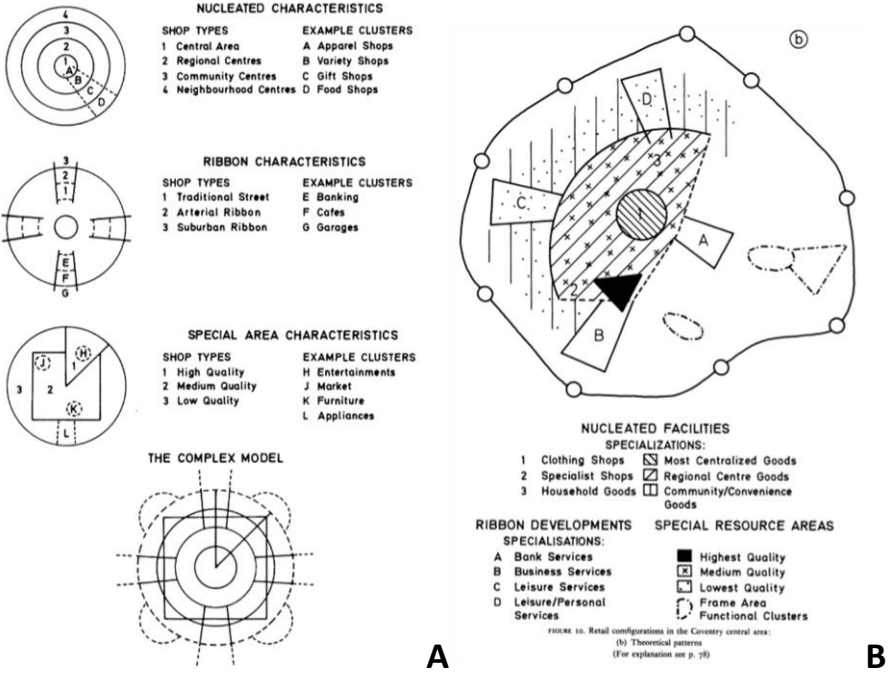


Figure 4 Davies’ Structural models of central are core retailing facilities (A), and its application to the Coventry central area (B) (Davies 1972)

For about 30 years, Berry’s and Davies’ models have been celebrated but also criticised due to their several underlying assumptions and implications (Dawson & Sparks, 1986). Despite evolutions in the socio-economic settings and retail formats, these models have been widely applied by researchers, town planners and retailers themselves (Brown, 1992; Brown, 1993).

Ever since the ‘80s, several studies identified specific issues with hierarchical models and more specifically with Berry’s model: the incomplete picture of spatial patterns of urban retailing (Jones & Simmons, 1990; Potter, 1982), the assumption of single-purpose shopping behavior (Dawson, 1983), single-retail organization (Dawson, 1983; Guy, 1984), and the new development of big box stores out-of-town (Davies, 1984) etc. Dawson and Kirby (1980) and Dawson and Sparks (1986) state the inadequateness of Berry's hierarchical model.

Consequently, several post- and quasi- hierarchical models have been proposed to address these criticisms.

While the majority of Berry’s model critiques and developments consider incongruences of the underlying hypothesis with the real functioning of the retail phenomena, Brown (1991) provides additional observations about weaknesses of Berry’s resulting classificatory framework. For example, difficulties of differentiation between form and function might lead to consider automobile row as both specialised and ribbon agglomerations; moreover, free-standing and ancillary stores are completely overlooked and several new facilities were modifying the retailing industry. Finally, but not less important, the radical changes of the underlying urban landscape should be considered: the well-defined pattern of towns and cities on which the new economic geography has based its main location theories is now faced with large metropolis and more complex urban structures.

Brown (1991) proposes a new taxonomy of retail agglomerations which represent an alternative solution to the traditional hierarchical model. Instead of readapting or reinterpreting existent models, Brown's classification is based on the independent identification and analysis of the two main aspects describing retail districts: the form of a shopping agglomeration and the function enclosed in its perimeter. As for previous models, the author recognises and redefines three types of form: cluster, linear and small isolated concentrations. The function of these agglomerations is then considered: general, specialised or ancillary functions are thus identified. The combination of these two factors generates a classification matrix composed of 9 groups of commercial districts, extended to 12 when considering planned/unplanned clusters (Fig.5). No applications of the post-hierarchical model have been proposed by its author. In the more recent work of Barata Salgueiro and Cachinho (2011), Brown’s taxonomy is discussed and applied to the city of Lisbon where examples of each retail agglomeration type are qualitatively presented.

		<b>FUNCTION</b>		
		GENERAL	SPECIALIST	ANCILLARY
<b>FORM</b>	CLUSTER (Unplanned)	Town and City Centres	Bright Lights District	Sandwich Bars in Financial District
	CLUSTER (Planned)	Megacentre	Speciality Shopping Centre	Shops in Airport Departure Lounge
	LINEAR	Traditional Arterial Route	Ethnic Shopping Street	Hamburger Alley
	ISOLATED	Corner Shop	Carpet Warehouse	Crush Bar in Opera House

Figure 5 Brown’s post-hierarchical model (1991)

Nonetheless, some limitations of this model have been identified by its author towards the end of his work. Firstly, it would be challenging to implement the identification of the 12 classes described by the post-hierarchical model in real case studies. Second, the model represents a static classification unable to recognise changes in the retail configuration. Finally, Brown highlights that despite its exclusion, “hierarchy still exists and will doubtless continue to exist” ((White, 1975), p.377). Several empirical studies have been applied and adapted CPT to intra-urban case studies (Brown 1993, Davies 1972, Potter 1982) and more recent works still recognise the hierarchical structure of the CPT when investigating the spatial organisation of retail in contemporary cities (Fleury, Mathian and Saint-Julien 2012, Saraiva 2017 etc.).

Nevertheless, two aspects should be highlighted. On the one hand, a constant need for more empirical quantitative analysis is continuously emphasised (Brown 1994, Davies 1972 etc.) in order to consider different retail agglomerations and categories within a common methodological framework.

The second aspect is a more theoretical problem which might underlie the CPT: Christaller was mainly concerned with the spatial structure of a hierarchical urban system offering goods and services in a rural space. In his theory, centres are point-like entities separated by rural areas. After the war with the first application of the neoclassical approaches, CPT theory was generalised to explain the spatial structure of hierarchized retail centres within the continuum of urban space. This transposition of the CPT theory to two different systems at two different scales has been accepted without a careful consideration of its pertinence for the new domain of research (Fusco 2018). Is it possible that the same notions and assumptions underlying the analysis of systems of cities and flows among them might also be implemented to explain the behaviour of individuals within a continuous urban space?

Almost the same energy that some academics invested in highlighting shortcomings and criticising the validity of neoclassical approaches, has been invested by other researchers in applying, revisiting and renovating the same theories (Brown 1994, Lemarchand 2008). The increasing availability in data availability and computational processing power allows new empirical works to integrate quantitative approaches and renovate classical theories (Wayens 2006 on Brussel, Fleury et al. 2012 on Paris, Wang et Jones 2002, etc.)

Independently from the theory under study, one common problem of retail studies is how to recognise and delineate the spatial extent of unplanned retail centres considering the different typologies observed and discussed in the literature of the XXI<sup>st</sup> century. Part 2 illustrates how several quantitative procedures from spatial statistics have been proposed in the last two decades. However, these approaches are mainly focused on the automatic detection of clusters, which might be considered as an overly simplistic measure compared to the more complex descriptions of retail agglomeration that has been described in the retail literature reviewed as far.

## 1.4 Structural approach

Neoclassical and behavioural approaches focus respectively on the economic forces and behavioural factors influencing the retail spatial organisation. Among the assumptions underlying neoclassical theories, the “optimising men” attracted the majority of academic critics delivering a vast number of studies starting from the ‘70s.

It has been since the ‘90s that “supply” side studies have gained more attention. This interest has been triggered by the multiplication and diversification of retail formats which have been transforming the retail landscape. While in the ‘60s we might have observed a prevalence of independent stores, less than 30 years later national and international firms with important financial dimensions control a large share of the retail sector.

As for behavioural studies where individual descriptors have been used to categorise and generalise behaviours depending on variables such as sex, age, professional category (among others) similarly, the different formats of retail might be identified by classifying retailers based on different variables: size, type of goods, type of store etc. (Guy 1998). Three retail formats can be easily identified: individual storekeepers, store chains, and planned retail centres.

Location choice process of small independent shopkeepers is based on a single-location optimisation problem, mainly influenced by exogenous factors. The spatial organisation of independent shopkeepers might be considered as the result of a spontaneous self-organised process of the retail system. On the contrary, the spatial organisation of store chains and retail spatial patterns observed in planned centres represent the result of a series of ex-ante calculations and planned choices.

While the former faces a complex problem of multi-location and multi-scale optimisation, the latter considers the maximisation of the overall retailers’ achievements in a given perimeter through a multi-location and multi-shop optimisation problem.

Higher levels of organisation of these retail formats correspond to larger market shares with higher financial impacts on social and political organisations, both at local and national scales. As a response, public authorities have long devised instruments and political measures to guide and exert control over the development and spatial organisation of retail.

Brown (1992) and Wayens (2006) named as structural approach those academic works related to the study of the spatial organisation of international retail networks/chains and the consequent dialogue with socioeconomic and political organisations. Within this approach, three thematics are distinguished: i) studies of localisation strategies of retail chains, ii) functioning and diffusion of planned retail centres and iii) influence of governmental and administrative regulations on retail locations.

Within the first group, the research activities have been focusing on the identification of those aspects to include in the multi-location optimisation process: spatial distribution of consumers, stores densities, relative distances between same-chain stores (to avoid overlapping catchment areas), and internationalisation (Cliquet 2002a, Cliquet 2002b). Spatial

strategies should also consider the planning approach of adopting potential competitor locations (Laulajainen 1987, Wayens 2006). In the same group, we might find the works investigating the impact of store chains on traditional shopkeeper location and distribution (Mérenne Schoumaker 1998, Jones and Doucet 2000, etc.)

A wide scientific literature has been studying planned retail centres considering the managerial strategy, their spatial location, and the control of environmental factors such as shape, size, aesthetical features, accessibility, commercial mix, etc. These studies have in common the same final goal: maximising the number of visits, sales, revenues of the overall planned centre (i.e. Jones and Simmons 1990b, Wrigley and Lowe 2002, Wayens 2006 etc).

The third approach firstly outlined by Brown (1992) investigates the socioeconomic and political aspects underlying the retail spatial distribution. Since its earlier works in the '90s, the main focus of the structural approach is on the interactions between stakeholders involved in the retailing process.

Recent studies with a specific focus on political influence on retail structure are for example those of Cliquet (2000) in France or Wang (Wang and Song 2008, Wang et al. 2013) in China, Bauer (Bauer et al. 2018, Bauer and Fusco 2016) in Germany, etc. In the specific case study of the French Riviera, Fabre (2008) investigated the role of the administrative and legislative tools for the management and development of retail at the departmental level.

Wayens (2006) further elaborates Brown's description of regulation factors influencing retail distribution. He describes: i) functioning regulations impacts on retail activity such as limitations on opening hours (Heinritz 2001, Karrholm 2017, Mulicek and Osman 2018 etc.), environmental legislation, different taxation in neighbouring regions (Kampshulte 2002), etc; ii) spatial regulations, potential effects of land use and urban limitations like distances from sensitive sites (Davies 1972), zoning areas, and the requests of specific procedures for the evaluation and authorization of specific retail formats.

The structural approach might be also studied in relation to recent evolutions of online retailing. E/m-commerce represents a fourth retail format that shares with store chain the need for a planned system of delivery points, namely pick-up-points. The network of delivery points frequently coincides with traditional retail formats. The spatial distribution of pick-up-points might (partially) represent the materialisation of virtual commerce. Studies on pick-up-points have focused on logistics and transportation issues known as "last kilometre/mile delivery" or the "last mile issue" (Dagher 1998; Punakivi 2001 etc.) with the final aim to optimise the localisation strategy of pick-up-points from a transportation and economic point of view. More recent works studied the retail distribution of automated lockers and in-store pick-up-points (i.e. Morganti 2014a, 2014b)<sup>20</sup>; other studies might consider drivers (Deprez 2017, Gasnier and Raveneau 2016).

---

<sup>20</sup> Some tests have been conducted studying the spatial repartition of pick-up-points: after collecting a complete database of 63 P-P within the study area we analysed socioeconomic characteristics of the surrounding environment. Since P-P represents the final point of delivery which should be positioned close to the recipient, small sized isochrones have been used to define P-P catchment areas. With the classification approach (Guy's 1998), we identified 11 typologies of environments served by pick up points ranging from metropolitan to town and suburban environments.

## 1.5 Contemporary approaches

A general agreement about the three main approaches described in the previous sections has been proved by several works of acknowledged academics within the micro-retail geography domain (Brown 1991, Wayens 2006, Lemarchand 2008).

Neoclassical, behavioural, and structural approaches should not be interpreted as a chronological sequence where the development of each approach replaces the previous one. Retail geography, as for most of the disciplines in social sciences, is composed of simultaneous coexistence of different analysis approaches. The evolution of the predominant approach in different chronological periods represents the scientific response to emerging questioning engendered by social and technological evolutions. In order to fully understand why these approaches have been blossoming in different periods, we should consider the historical and societal changes at the origin of the research in question.

The shift from a neoclassical approach to a behavioural one has been explained in the previous section as the need for a more realistic representation of human behaviour. Among the several assumptions underpinning economic theories, why has the one simplifying consumers' behaviour been the focus of numerous critics, debates and research developments? Three (highly interrelated) societal changes might have an impact: firstly, the entry of the woman in the job market with a consequent growth of dual-income households. Secondly, the market penetration of domestic appliances and all those saving-time products stimulating the development of leisure-related services and activities (for example, tourism). Thirdly, the market penetration of cars. The combination of higher mobility, free time and higher household incomes engendered the multiplication of consumer behaviour and consequently interrogations on the interaction between the functioning of a society and its retail system. Therefore, the growth of studies focusing on behavioural aspects is detected around the '70s.

The shift of the research focus from behavioural to structural approaches should also be understood as a consequence of the societal and technological context characterising the period between the end of the '80s and the '90s. The end of the Cold War resulted with the consequent acceleration of the expansion of western national retail companies to the international market; the reduction of production costs and customer prices resulted from the progressive delocalisation of the manufacture to developing countries combined with a growth in technological efficiency. Moreover, generational factors impacted western economies with the baby boomers reaching the peak spending ages. Within this macroeconomic context described by high levels of both demand and supply sides, large retailing organisations started their international conquest of the market with a spatially organised dissemination of stores. The need for a rational and efficient store network raised issues of the spatial distribution of the whole structure through an efficient choice of combined locations. Technological advancements in geographical information systems during the same period provided new tools as decisional support systems. The renovated interest of the spatial component is witnessed by Wrigley and Lowe manifesto where the independence of a "new retail geography" is stated (Wrigley and Lowe 1996). As discussed in the previous section, structural studies considered not only private company organisation and impacts but

also the influence of administrative policies: this aspect might be seen as a political response to the economic growth of retail companies which raised interrogation of public social responsibility (Clique 2000).

What, then, are the societal and technological challenges of the current period, which is stimulating further research on retail studies? A series of ongoing mutations are reported by Dart (2018) and Desse (2016): the generational composition of the market with the entry of millennials and a gradual shift of baby boomers to their pension age. This generational transformation brings a new attitude towards material goods with a greater emphasis on experience, meaning and wellbeing (this last one also triggered by the extension in life expectancy). Additionally, there is a higher propensity to collaboration and sharing economy; and the need to be understood as an individual with specific values, lifestyles and personal experiences. The hyper-fragmentation of the retail market is accelerated by new technologies of the digital age: among other things, the Internet, mobile phones, and 3D printers are causing the dematerialisation of the economy with the development of new retail formats named e/m-commerce.

Retail centralities as traditionally studied in retail geography have been shaped by the combination of centripetal and centrifugal forces stimulated respectively by public and private transport systems; the omnipresence of retail in space, and time allowed by m-commerce seem to weaken and ultimately dissolve retail centralities as were previously understood. As a consequence of this threat of dematerialisation, academics reacted with a renovated interest in retail geography, both in terms of cultural and physical components for a re-spatialization /re-territorialisation of commerce (Desse 2016, Madry 2016).

Recent works of retail geography academics are focusing on two main axes: on one side we find those works studying the cultural context influencing consumer spaces and practices (*consumm'acteurs* Lemarchand 2011,2008; the Consumerscape Culture Theory, Cachinho 2014; etc.), on the other, a transformation and reorganization of retail structure within urban spaces (at different scales): urban area requalification, sustainability concerns, social importance of public spaces, resilience of retail landscape, etc. (Cachinho 2014; Fernandes and Chamusca 2014; Desse 2016). This current research interest is the origin of new approaches to investigate the relationship between city and commerce. Madry (2016) and Renard-Grandmontagne (2016) review the most recent works: the latter detects a recent trend focusing on the physical city, exploring the need to rethink urban spaces and its morphology in relation with usages and practices of consumers and retail structure.

The recent trends in retail study can be seen as a contextual approach: the description and analysis of a retail landscape from both the urban and cultural context are what will be able to monitor the expected or ongoing retail seismic shift (Dart 2018).

In the following image (Fig.6), we propose a graphical summary of the evolution of the retail geography approaches in the academic context (right side) related to both socio-technological evolution and the developments of retail formats (left side). On the bottom part of the image we can observe the thematic shift of the different approaches: from the economic focus of

the neoclassical approach, to the growing attention on agents (consumers and retailers) and their organisation and, more recently, the spatial component both for its sociocultural and morphological point of view.

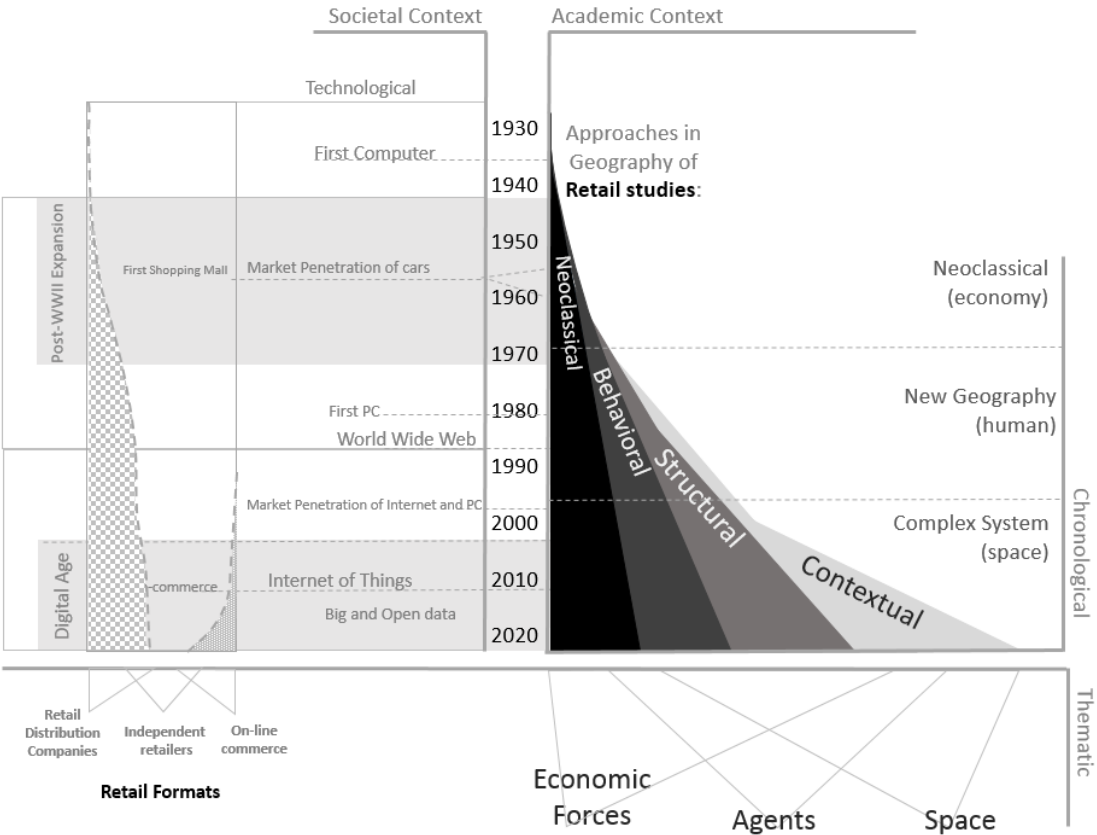


Figure 6 The chronological and thematic evolution of the retail geography approaches and the relative socio-technological context

As we have seen in the previous sections, when researchers are faced to the task of summarising this literature of retail geography they all face the same problem of a wide number of studies approaching this subject from different epistemological and historical backgrounds, which is further compounded by the multiplication of quantitative and qualitative methodological procedures. The panorama of the approaches outlined in the previous sections represents only one possible way to classify these works of retail geography, which is the most recognised among acknowledged authors of the domain (Brown 1991,1994; Lemarchand 2008; Wayens 2006; Desse and Lestrade 2016; Madry 2016; etc.).

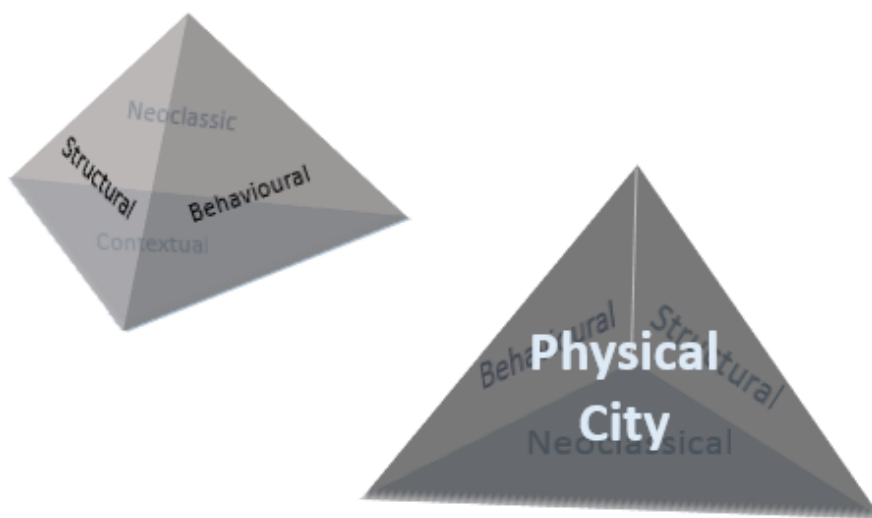
Four main approaches coexist in the study of retail distribution and have demonstrated their legitimacy, providing different insights about retail spatial organisation. Nonetheless, each of them focuses on specific aspects of the phenomena under study, trying to deconstruct the high complexity of the system; the unavoidable cost of this process is the simplification of those factors considered by the analyst as having marginal importance.

For example, neoclassical theories (Section 1.1) provided spatial models and principles focusing on the economic forces occurring between sellers and buyers. Strong assumptions about transport systems, urban morphology, human behaviour, and agent economic conditions are common premises of neoclassical theories. These assumptions have been



clearly defined in their original version, which has stimulated a wide scientific debate about their validity in the scientific literature of the following generations.

These approaches represent the point of view from which the analyst chooses to observe and understand a phenomenon, acquiring specific yet partial insights into the multifaceted reality. The retail geography literature review previously discussed might be then imagined as tetrahedron where each facet corresponds to a different approach: when the analyst sights are perpendicular to a specific facet, this surface acts as a lens, enabling and clarifying the view on the retail spatial organisation. In the meantime, the same chosen lens presents a partially or totally distorted vision of other sides. Every lens is a valid approach which allows describing a specific aspect of the reality and a complete vision of the phenomena under investigation is made possible only through a fragmented panorama given by the joint analysis of the tetrahedron (Fig.7).



*Figure 7 Schematic illustration of the different approaches of retail geography (left) and the specific approach chosen in this dissertation: the physical city (right)*

This tetrahedron is a graphical representation of the problem an analyst faces when trying to investigate a specific phenomenon where different possible approaches might be undertaken. Different classifications focusing and isolating other aspects are possible, such as the one of Mérenne Schoumaker (2003 p.187, identifying structural studies, retail fabrics, spatial behaviour of consumers, agent strategies and retailing policies).

Independently of the qualitative, quantitative, and epistemological approach, the final goal of retail geography is the study of the location factors and forces that are able to explain and model the spatial behaviour of retail distribution (Wayens 2006).

As previously highlighted, the contextual approach represents a specific approach which has stimulated the interest of retail geographers in the last years (Renard-Grandmontagne 2016): with this dissertation, we focus on the urban context, and, more precisely, we consider the spatial relation between retail and morphological organisation of urban space.

We believe that the spatial location of retailers, both at the individual and at the aggregated scale might be influenced by the physical properties of the surrounding urban space: the disposition of buildings along the street, the typology of urban fabric and the accessibility as

defined by the street networks might represent important factors influencing the attractiveness and thus the presence of retail activities in a given urban space. While these aspects have been recently investigated by urban morphologists, the urban context approach is still far from the mainstream literature of retail geography which is more focused on social, cultural and behavioural contextual aspects.

With this statement we do not mean that the retail geographer has never considered urban morphological aspects; on the contrary, a large number of references, observations and highlighted notes have been spread in the last 50 years' literature, representing a noticeable support for the development of this work. Academics constantly remember how we cannot separate the study of retail agglomeration and organisation from urban and transport geography (Berry 1967, Lemarchand 2008, Madry 2016). The importance of the urban morphological space has been partially included in retail geography studies and we can find several references to the built-up space and the influence on the retail organisation.

Parkins (1930), for example, focuses on how to measure in an objective way the intensity of land utilisation "to discover what physical conditions man was adjusting himself in locating the retail district"; in order to reach this goal, the author uses variables such as building height and build-up spaces coverage. A critical attention is given to the morphology of the site, its acclivity and the presence of natural barriers (Fig.8.A). These morphometric variables are implemented in a quantitative analysis of the urban space, with the ultimate goal of defining the CBD of Nashville, US.

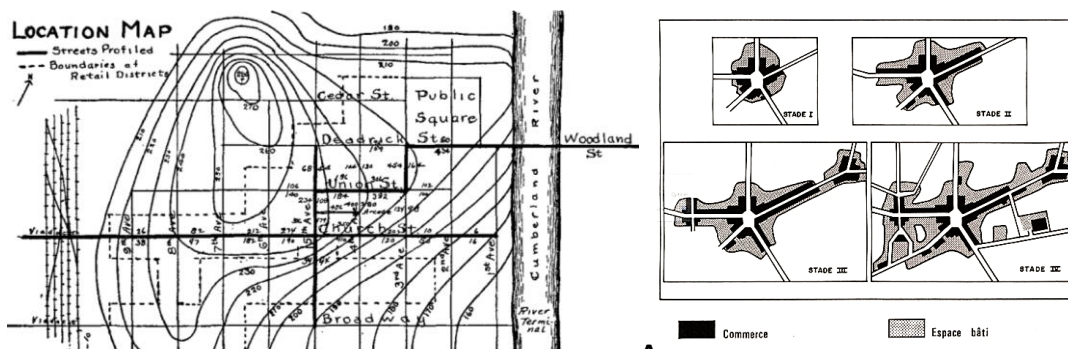


Figure 8: A) The retail center of Nashville, Tennessee (Parkins 1930); B) The phases of the urban development and retail distribution (Beaujeu -Garnier 1977)

When considering the work "La géographie du commerce" of Beaujeu-Garnier (1977), considered one of the main references of the retail geography in the French community, several considerations of the relation between the urban morphology and retail distribution are qualitatively discussed. The author uses the expression 'masses distribution' (p.207-208), referring to built-up spaces when describing the relation between urban form and retail distribution and how the street organization represent the backbone of their evolution (Fig.8.B): she describes how retail participated in the long process of anastomosis of squares in European city centres resulting in today's urban fabrics. The importance of the site morphology is also considered, for example, when describing the spatial distribution of the nine CBDs of Istanbul. Beaujeu-Garnier in 1977 suggested how the census of building properties would have represented an "excellent source of information for future studies".

The work on Le Mans, Carré and Rouleau (1974) goes further into the details of the urban form, describing quantitative and qualitative aspects of the interface between street and a building's facades, and they use these measures to describe the commercial dynamism of a street and in the urban space (with a focus on food retailers). Density, verticalization, aesthetical aspects of streets and other dimensional measures are evaluated in relation to retail agglomeration distribution.

In a more recent study cited by Lebrun (2002), among different factors influencing retail location, specific attention is given to urban morphology. For example, he highlights how retail continuity might be interrupted by the presence of "urban fractures" originated by both natural or artificial linear elements as transport infrastructures, rivers, parks, etc. (as observed in Parkins 1930). What is more interesting is that he notices that fractures might be defined not only by physical obstacles, but also by a more or less gradual transformation of the morphological organisation of urban fabric defined as "differentiation barriers". Citizen perception and pedestrian movement behaviour might be influenced by these demarcations between different urban fabrics, impacting retail diffusion in the same manner as physical obstacles. The same reason might be at the origin of different organisation and specialisations of retail agglomerations: it is common to observe historical centres and modern urban areas laying in two adjacent districts, both with a similar retail density but described by two different retail compositions (i.e. touristic and clothing, or touristic and business district etc.).



*Figure 9 Some morphological fractures identified by Lebrun 2002 in Reims(left) and Nancy (right)*

The importance of the building form is also acknowledged at a more intrinsic level: Lebrun discusses how the monumental approach used in the requalification of streets (i.e. Haussmann period) have a direct influence on the economic accessibility to certain buildings. Similarly, the presence of large historic administrative/religious buildings along squares and street edges is responsible for the expulsion of the commercial function from these urban spaces despite their importance and attractiveness. Because of these physical characteristics of the built-up spaces in given areas, retailers might not always find available surfaces adapted for their location. The urban space where large buildings are organised on their sides might "earn in emphasis, but they lose in polysemy" (Lebrun 2003, p.363).

Wayens (2006) discusses the historical extensions of the urban settlement and the relative proportion of ground-floor commercial surfaces: the author recognises eight urban growth stages (from the first and second medieval defensive walls, the incorporation of the surrounding hamlets and more recent residential developments), and they provide a fine description of the retail composition (42 retail categories).

Nonetheless, these aspects are spread in retail geography literature without a common framework. Moreover, several authors use both the morphological and functional notions of urban fabrics to describe the specific environment surrounding retail format and agglomerations (residential fabric, historic centre, CBD, peripheral, etc.). The selection of urban fabrics is used as a means to delimitate specific urban areas (generally described by homogenous morphological conditions), which allows a focus on the analysis and description of a specific phenomenon. Urban form is considered as a way to contextualise their works within which a specific process takes place, rather than considering its physical component as an active force which co-occurs in the retail location and organisation process.

The importance of the morphological space on retail named the 'morphological sense of commerce' in Saraiva (2013), is still overlooked in the retail geography literature (Saraiva 2019), which seems to be more focused on social and cultural contexts. On the contrary, in the field of urban morphology, the last two decades have seen growing attention on the study of the relationship between retail distribution and the urban form.

Before proceeding with the exploration and discussion of these works, it is fundamental to clarify some notions about urban form.

## 2. Retail in the Study of Urban Form

### 2.1 Definition of Urban Form

The definition of urban form is not unique and several interpretations and explanations might be found in the scientific literature, depending on the specific domain of research. The growing attention of urban morphologists (from different European schools) and urban geographers during the second half of the twentieth century is at the origin of a discrepancy on the notion of urban form, associating to this notion different meanings; consequently, several misunderstandings might arise when trying to converge concepts and ideas originated in these two domains of study.

Critics about the lack of a general consensus on the vocabulary, the epistemological weaknesses, and the absence of scientific rigour (Choay and Merlin 1988) stimulated an interesting debate in both urban morphology and urban geography highlighting how different connotations and interpretations might be associated with this specific expression. This variability is originated both from the polysemy associated with the notion itself and by the epistemological backgrounds of readers and of academic groups using this notion to describe and discuss specific scientific questionings.

The absence of a strict definition of the expression 'urban form' is also recognised in Raynaud (1999): the author proposes a semantic study of this nominal group, combined with the investigation of its uses within the scientific community literature. This semantic approach allows the author a *super partes* analysis of the expression urban form, going beyond the specific interpretation of disciplines and schools. After an interesting discussion about the reasons behind the semantic transformation originating the polysemy of the notion of urban

form, the author outlines its “spectrum of meanings” and finally proposes alternatives to enrich and specify the (French) vocabulary of urban-related studies.

In order to obtain the “*spectrum of meanings*”, Raynaud deconstructs the variability of the notion urban form and considers the uncertainty linked to each component of the nominal group urban form and the possible connotations associated to it.

He calls **internal variability** the different significations the two words might individually carry as well as the additional variability given by their combination. While the concept of urban on a first sight might be associated to the notion of a specific space with material characteristics and delimitations, at the same time, it could also be used to describe a subgroup of the population defined by a specific lifestyle with a particular social organisation and structure. Concerning the word form, the author illustrates eleven roots explaining its etymological origins. Three main uses might be associated with this word denoting respectively i) production process, ii) description and iii) perception of an object. The combination of these two words should be finally considered as a third component of the internal variability: the different interpretations of the nominal group urban form (which in French correspond both to *forme urbaine* or *forme de la ville*) are caused by a possessive or genitive relation between the word form and urban (the form within or generated by the city).

The **external variability** is generated by the different interpretations and traditions which are linked to a specific professional context and in the case of the expression urban form it might engender contradictory interpretations: normative/descriptive, causal position, physical/conceptual. Finally, an additional external aspect stimulating further variability resides in the choice to investigate a specific element (or a sub-group of elements) among the components constituting the object of study (in what might be defined as a synecdoche of urban study). Raynaud concludes his works observing how rather than imposing a unique definition to the notion of urban form, it would be better to use a more unambiguous vocabulary in order to reduce the variability and to clear out possible misinterpretations. Nine expressions are finally specified as alternative terminologies describing specific facets of the urban form with a lower degree of variability (i.e. social form, urban representation, urban model, urban plan, urban fabric, etc.).

The problem of the polysemy of the urban form discussed in Raynaud (1999), is further developed by Levy (2005). The author observes how the complexity of the notion of the urban form allows different conceptualisations; consequently, each specialised domain of urban-related research originated a large variety of exclusive and sector-specific definitions and analysis approaches. Therefore, there is the necessity for an epistemological revision of the notion of urban form able to overcome the traditional compartmentalised definitions: five registers of urban form are thus identified and described.

The urban form as **urban landscape** is defined as “*the urban space as visually perceived in its three-dimensionality and its plastic materiality*”, with a specific focus on the aesthetical, stylistic, cultural and historical aspect of the urban space. This first definition finds its origin in the works of Cullen (1961), Bacon (1965), Sitte (1889), Lynch (1960) and the picturesque approach of Pinon (1991).

The register of the **social form** is that which studies “*the occupation of the urban space by different social, demographical ethnical groups, household composition and the distribution of activities and functions within the city*”. For this definition, the references might be found in the works of Durkheim (1960), Jacobs (1961), Ledrut (1968), the School of Chicago and the French school of social morphologic studies (Halbwachs 1928) and more recently with the works of Roncayolo (1996, 2002), which investigates the link between social morphology and physical form of the city.

The **bioclimatic form** describes the urban space considering “*its environmental dimension, as a micro-climate, considering both its geographical variations in each neighbourhood and the diversity related to the types of urban fabrics (open/closed/semi-open), the orientation (solar-thermal), the site (water, relief and vegetation) [], the unequal repartition of the pollutants*”. Since the first works in the ‘80s-‘90s (i.e. Escourrou 1980; Newman and Kenworthy 1998; etc.) the attention to urban climatology, ecology, vegetation, etc. has known an important growth specifically in the last decade due to a higher sensibility to the environmental and ecological issues requiring a more sustainable way to think and plan urban spaces of the future.

The urban form as a study of **urban fabric** is “*the study of the interrelations between the elementary components: plots/streets/empty spaces/built-up spaces, which constitute every urban fabric [at different scales], in relation to the site, or focusing on the analysis of certain components.*” Two main goals are associated with the works interested in the urban fabric: the first aims to understand the dialectical relationships between building typologies within urban fabrics as well as the historical morphogenetic process generating the different forms observable in a given cultural context. The second is related to the acquisition of that knowledge which might be implemented in urban and architectural projects for the production and renovation of urban forms. As we further discuss in Section 11, this specific notion of urban form has been developed by the Italian (Muratori, Aymonino, Caniggia, Maffei), French (Depaule, Castex, Panerai, Borie, Denieul) and English (Conzen, Whitehand, Larkham) schools.

Finally, there is the urban form analysed through the **form of the urban layout** considering “*the geometrical form of the city plan (organic/geometrical plan; orthogonal/radio centric plan)*” in relation to its centralities (monocentric or polycentric city) with the works of Pinon (1994)

While Levy (2005) describes the Space Syntax as an attempt to renovate the study of the urban fabric with new computational approaches, Fusco (2018) highlights the need to distinguish and recognise the **configurational** as a sixth independent approach. The configurational approach finds its origins in the analysis of the complex networks firstly applied to the social science (Freeman 1977). First applications to the case of architecture and urban fabric are implemented in Kruger (1977) and Steadman (1983), Hillier’s seminal works on Space Syntax (Hillier and Hanson 1984, Hillier 1996, etc.), and more recently in the Multiple Centrality Assessment of Porta (2006, 2009 etc.) (further details will be discussed in Section 2.2 and Part 3). In his work, Fusco (2018) discusses the main differences between urban fabric and configurational register, justifying the consideration as two separated registers of the urban form; as we will further discuss in Part 3, while the former considers the geometrical

description of the urban form elementary components, configurational analysis quantifies the capacity of network configuration (street segments, visual axes, etc.) to structure movement and encounter patterns within urban space. This second approach describes the properties of form elements and considers their spatial relations established with all other form elements within the urban space (or within a shorter radius of analysis). The same separation between *configurational form* and *urban fabric* is also recognized in the urban morphology community known as respectively as the Conzenian and the Space Syntax approaches.

Beyond the specific definition of each register<sup>21</sup>, both Levy and Fusco highlight how the strong interrelation and interdependence of these different registers hinder the analysis of the urban form only through a specific register. The challenge for theoretical, empirical and applied works is the systemic analysis of these five/six registers in order to understand and describe the urban form as a whole.

Guided by the rich description of the complex and polymorph notion of urban form provided by the works of Reynaud, Levy and Fusco, it is now possible i) to better delineate the theoretical framework and goals underlying our research problem and ii) to reduce the variability associated with the expressions used in this dissertation.

First, it should be now clear how the retail geography literature might be considered as a part of the social form with a specific focus on the commercial function and its spatial organisation within the city. The different approaches reviewed in Section 1 show how scientific research has investigated different notions included in the internal variability associated with the word 'retail'. Indeed, following the example of Reynaud, we might observe how retail is associated with the process of selling goods as well as the place and the agents involved in the retailing activity<sup>22</sup>. In this dissertation, with the word retail we consider only the geographical location of its physical component and the relative functional typology; economic, social and administrative/policy are not analysed in this study.

Secondly, the form of the physical city is analysed in this work by combining the configurational and the urban fabric registers. In the methodological procedure discussed in Part 3, we propose to renew urban fabric analysis introducing the visual perceptive approach traditionally developed in the landscape form register.

Among the different approaches proposed for the analysis of the relationship between retail and urban physical space, street-network configuration attracted the most attention of urban morphologists. This next chapter will thus be dedicated to this prominent field of research.

---

<sup>21</sup> The identification of each register is mutable in the time as much as the urban form itself. The registers described by Levy and Fusco are related to the current meaning, uses and analytical approaches associated to the notion of urban form; evolutions are also possible, requiring redefinitions of the overall framework. We might for example imagine how the recent development of technologies could be the origin of a seventh register, the *virtual form* completing the description of the cities of the 21<sup>st</sup> century.

<sup>22</sup> While the use of the English words *retail* and *retailer* might help in the differentiation between the activity in the former case and place or agent in the latter, in French the word *commerce* carries both the activity and place notion while *commerçant* refers to the agent.

## 2.2 Street-Network Configuration and Retail Distribution

Starting from the '80s, both in retail and urban geography, growing attention has been given to the study of the city at the intraurban level; neoclassical approach models were considered too coarse when analysing the city at a higher resolution. While retail geographers focused on the role of retail as a structuring or structured element of urban space (Madry 2016), urban morphologists on their turn investigated the ability of the physical organisation to influence how cities work.

As observed by Alexander (1965) the socio-functional organisation of the city and its physical structure constitutes two profoundly intertwined characteristics of human settlements. The link between social functioning and space should not be reduced to a simple spatial representation of the former upon the latter: an important, dynamic relationship occurs between these two components, both of which are exposed to continuous modifications and restructures.

Within this context, a revolutionary contribution has been proposed by the research group of Bill Hillier at the University College London. With the seminal work *The Social Logic of Space* (Hillier and Hanson 1984), the authors developed a new descriptive theory of space: the Space Syntax (SSx). From the hypothesis that human societies use space as a key and necessary resource in organising themselves, the authors derived a fundamental relationship between the configuration of space in a city and the way it functions.

SSx theory develops the idea that physical boundaries strongly influence the visibility and perception of enclosed spaces. Focusing on those areas available for pedestrians and delimited by the interface with buildings, an innovative reconceptualisation of the complexity of the urban space is proposed. The urban open space is partitioned into convex polygons approximated with visual axial lines; the relationship established between these lines through their intersection enables the performance of configurational analysis, similar to graph analysis, previously elaborated by Freeman (1977) in the domain of complex social networks.

Configuration is defined by the same Hillier *“as, at least the relation between two spaces taking into account a third, and, most, as the relations among spaces in a complex taking into account all other spaces in the complex.”* For this reason, he concludes that *“Spatial configuration is thus a more complex idea than spatial relation, which needs to invoke no more than a pair of related spaces”* (Hillier 1996). This is an essential concept that Space Syntax (as well as other configurational analysis) derives from a systemic approach: the importance of an element is defined by its relative position within the whole system and not as a sum of the relations between all pair of elements.

SSx protocol provides a quantitative description of the spatial properties for each street of the urban space: the resulting set of indicators have been proved to be able to capture important aspects of social and behavioural phenomena. In particular, this method opened the way to a deeper understanding of the relationship between urban form and human movement.



Through street configuration: *“other things being equal, movement flows in different parts of a street network were systematically influenced by the spatial configuration of the network itself”* (Hillier et al. 1993).

The relationship between the street network structure and the movement patterns represent what Hillier calls the Principle of Natural Movement defined as *“the portion of the movement on each line that is determined by the structure of the urban grids itself rather than by the presence of specific attractors or magnets”* (Hillier 1999, p.120). Different studies on the city of London showed how the configurational properties of the network might explain the difference between the 55 and the 80% of the pedestrian movement (Hillier et al.1993, Hillier and Idia 2005; Hillier & Vaughan, 2007; Jiang 2009).

The axial representation of the urban space has been the subject of several considerations leading to the identification of alternative procedures, each of them based on a different interpretation of the human behaviour underlying movement patterns. While Jiang and Claramunt (2004), Rosvall et al. (2005), Figueiredo (2009) proposed to identify axial lines using named street networks, in order to consider the urban space as socially defined. Turner (2001b) proposes to measure angular distances between axial lines, instead of traditional topological measures, while Asami extended axial lines considering 3D topological changes (Asami et al. 2003).

A wide number of works have seen the implementation of SSx in transport planning studies in order to model and explain movement in urban space; this domain of study has been experiencing a growing interest to respond to both environmental and health issues with in the past two decades. Nonetheless, pedestrian movement should not be considered as the final goal of SSx theory; rather, it is an intermediary step to understanding the structure of a city and to model the way it functions.

Hillier, since his early works, noticed how streets with higher integration values were associated with larger volumes of human movement and, consequently, were becoming more attractive for socioeconomic functions. Retailers, for example, need to be visible and accessible by a large number of people in order to maximise their possibility of being visited and thus to sell their products and services. High values of configurational centrality might be considered as a location factor for several socioeconomic functions of a different nature; consequently, the phenomena of competition between functions for places with better urban space properties results in higher values of land-use intensity. Nonetheless, the indirect relation between configurational properties and land-use intensity as defined by SSx theory is not linear: when a higher presence of retail activities is observed along a street, more people are attracted to the same urban space, increasing its original attractiveness. A process of cumulative agglomeration of functions is thus described by what has been called the multiplier effect (Hiller 1999b) or circular causality (Omer 2016).

The Theory of the city as a Movement Economy (MET) might be summarized by the following statement: *“space organization in settlements first generates movement patterns, which then influence land-use choices and these, in turn, generate multiplier effects on movement with further feedback on land use choices”* (Hillier 1999, p 117) (Figure 10).



Figure 10 Graphical representation of Hillier's Movement Economy Theory

When compared to traditional interaction theories (neoclassical approaches of Section 1.2), a specific clarification should be highlighted. *“Spatial interaction modelling approaches consider the location of economic activities as an input variable with their importance and their relative co-localisation in the geographical space; from this information, they derive, as output, variables measuring interaction flows between the same urban functions. In a specular way, configurational approaches study the urban grid as an input variable and assess the distribution of movements and the location of the economic activities. In other words, what is considered an input data for the former, (the location of the economic activities), it becomes a result for the latter.”* (Cutini 2001, p.144).

Consequently, as someone might observe, a sort of morphological (or architectural) determinism seems to underlie this association between configuration and retail. In the specific case of SSx theory, this aspect has been discussed by Seamon (1994); nonetheless, the same observation might be extended to all configurational analysis as well as every approach using physical descriptors of urban space as factors influencing human behaviour and activities. Hillier dissociates himself from these deterministic ideas, specifying that *“spatial form [...] creates fields of probable--though not all possible--encounter and co-presence within which we live and move”* (Hillier 1989). Similarly, this dissertation focuses on the identification of those elements of urban morphology that could help in a better comprehension of the commercial potential. Studying the role of the urban morphology as a localisation factor of retail should not be misinterpreted as the research of those sufficient conditions which explain retail presence and agglomeration process in an oversimplified causal relationship.

The association between the configuration of the urban space and land-use intensity presented in the MET has been addressed in a large number of studies encompassing cities with different socio-cultural backgrounds as well as different economic-related observations: retail distribution, property and house value, office and retail rentals, land-use intensity, etc. Table 1-2 illustrates a synthetic view of these works. Starting from Hillier's seminal study of London (1999), for twenty years the scientific community of urban morphologists has investigated and applied the principles of the MET.

As we might observe in Table 1-2 (column *Conf.Approach*), although most of the studies investigating the association between economic intensity and urban form have been developed in a Space Syntax framework, alternative procedures of network configuration analysis have been independently investigated by the research groups of Sergio Porta (Porta et al. 2006a, 2006b) and John Peponis (Peponis 2008).

The former developed the Multiple Centrality Assessment (MCA), a specific set of measurements considering a graph representation of the topological street network. Similarly to SSx theories, MCA indicators have been proven to be correlated to land-use patterns: retail correlation in Bologna (Porta et al. 2009) and Barcelona (Porta et al. 2012), land-use intensity in Baton Rouge (Wang et al. 2011), and land-use intensity in Shenzhen (Wang et al. 2018). A similar approach has been developed by Sevtsuk with a specific focus on the study of retail distribution in Cambridge and Sommerville (Sevtsuk 2010, 2014). The latter proposed measures of centrality on a primal representation of the street network, considering metrical and directional measures (Peponis 2008) with application to retail frontage distribution in the city of Buenos Aires (Scoppa and Peponis 2015).

<b>Configurational analysis and economic variable distribution</b>				
<b>Authors</b>	<b>Year</b>	<b>Conf.App.</b>	<b>Phenomena</b>	<b>Space study</b>
Desyllas	2000	SSx	Office rent patterns	Berlin, Germany
Kim and Dong	2002	SSx	Land use intensity of office build.	Seoul, South Korea
Rui-lan and Xingi	2004	SSx	Properties and housing value	Jinan, China
Matthews and Turnbull	2007	SSx	Properties and housing value	Washington, USA
Min et al.	2007	SSx	Land Use	Seul, South Korea
Enstrom and Netzell	2008	SSx	Office and retail rental patterns	Stockholm, Sweden
Topçu et al.	2009	SSx	Properties and housing value	Konya, Turkey
Chiaradia et al.	2009	SSx	Residential Property Value	London, UK
Soesanti and Nakai	2010	SSx	CBD and market locations	Bandung, Indonesia
Wang Antipova and Porta	2011	MCA	Land Use Intensity	Baton Rouge, USA
Narvaez et al.	2012	SSx	Properties and housing value	Cardiff, UK
Law et al.	2013	SSx	Properties and housing value	London, UK
Netzell	2013	SSx	Office and retail rental	Stockholm, Sweden
Rui and Ban	2014	MCA	Land Uses	Stockholm, Sweden
Muldoon-Smith et al.	2015	SSx	Real estate	Leeds, UK
Gianninopoulou et al.	2016	SSx	Properties and housing value	Xanthi, Greece
Adebayo et al.	2019	SSx	Real Estate	York, UK

*Table 1: Literature review-part I*  
*E/NC-KDE: Euclidean/Network-Constrained Kernel Density Estimation; CDF: Cumulative Distribution Function- MLR/BLR: Multiple/Bivariate Linear Regression- ExpR: Exponential Regression - P-corr: Pearson Correlation; NBR: Negative Binomial Regression*

Configurational analysis and retail distribution						
Authors	Year	Config. Approach	Dependent variable	Phenomena	Space study	Analysis Method
Hillier	1999	SSx	N° Stores/street	Retail pattern	Camden London, UK	MLR
Cutini	2001	SSx	N° Stores/25m (100 streets)	Retail pattern	3 small-medim sized Italian towns	ExpR
Van Nes	2005	SSx	-	Retail aggl.	Amsterdam, Netherland	Visual
Sarma	2006	SSx	N° Stores/aggl.	Retail aggl.	New Delhi, India	BLR
Porta	2006	MCA	E-KDE: 100m-cells, BW 100-300)	Retail pattern	Bologna, Italy	P-Corr
Ortiz-Chao	2008	SSx	N° Stores/street	Retail (land use)	Mexico City, Mexico	CDF
Porta	2012	MCA	E-KDE(300mt BW) on a 10m-size cell raster	Retail Pattern	Barcelona, Spain	P-Corr
Tsou, Chen	2013	SSX	Retail density withn Traffic Zones	Retail Pattern	Taipei city, taiwan	MlogLR
Van Nes	2014	SSX	-	Retail Pattern	Pompeii, Ancient Rome	Visual
Wang et al.	2014	MCA	E-KDE (1.5km BW on 100m cell-side raster)	Retail pattern	ChangChun, China	P-corr
Sevtsuk	2014	MCA	Presence/Absence retail building level	Retail pattern	Cambridge and Sommerville, USA	MLR - Spatial Lag and Error
Sevtsuk	2010					
Cui and Han	2015	MCA	E-KDE (1.5 km BW, 100-m size cell)	Retail (POI)	Zhengzhou, China	P-corr
Omer and Goldblatt	2015	SSX	N° Build. with retail 50m street-buffer	Retail pattern (comm.build.)	8 Israeli Cities (3 types)	P-corr and MLR
Scoppa Peponis Scoppa	2013 2015	Primal approach metrics	Retail Frontage/ street length	Retail pattern (comm.parcels)	Buenos Aires, Argentina	BLR, PCA-MLR
Fusco and Araldi	2015	MCA	NC-KDE (100m)	Retail density	Nice, France	Bayesian Networks
Fusco et al.	2016					
Ye et al.	2017	SSX	N° Stores/ street-block	Catering businesses	Shenzhen, China	NBR
Lin et al.	2018	MCA	E-KDE (3.5 km BW 100m size cells)	Retail pattern (POI)	Guangzhou, China	
Wang, Xu and Guo	2018	MCA	E-KDE (1km-size cells)	Land Use Intensity	Shenzhen China	P-Corr
Cutini et al.	2018	SSx	N° Stores/ street (30 streets)	Retail pattern	Milan, Italy	Exp-Corr
Saraiva et al.	2019	SSX	E-KDE (20m-size cells)	Retail vacancy	4 medium-sized Portuguese cities	P-corr

Table 2: Literature review-part II  
E/NC-KDE: Euclidean/Network-Constrained Kernell Density Estimation; CDF: Cumulative Distribution Function- MLR/BLR  
Multiple/Bivariate Linear Regression- ExpR: Exponential Regression - P-corr: Pearson Correlation; NBR: Negative Binomial Regression

In Part 3, we provide a more detailed description and discussion about methodological alternatives for the implementation of configurational analysis. We ask the reader to temporarily overlook any difference between these approaches and focus the attention on their shared features. What we should take into account is that the significant number of studies in Table 1-2 using configurational approaches might be seen as a general agreement of the urban morphologist community about the validity of Hillier's MET: the analysis of urban space might explain the commercial attractiveness of streets where a higher density of stores is observed. In other words, (probably those of a retail geographer), among retail location factors, the configuration of urban space plays a non-negligible role.

The success of configurational analysis resides in its simplicity and robustness (Nes and Yamu 2017): despite few limitations and inconsistencies (Ratti 2004; Wang et al. 2007 etc.), these methods provide quantitative, and therefore objective, measures of the urban form, reducing the uncertainty of traditional qualitative approaches. These characteristics represent the fundamental premises which allowed the elaboration of specific tools for the automatic implementation of configurational analysis. DepthMap by Alasdair Turner (2001, 2004), and more recently Confeego (Gil et al. 2007), QuantumGIS plug-in (Gil 2016) facilitated the dissemination of SSx (and derivatives), while Urban Network Analysis' (UNA) extension for both ArcGIS and Rhino software has been provided by Sevtzuk and Meckonnen (2012) for automatic implementation of MCA; similar tools have been independently developed by Porta and Peponis research groups. The research group of Chiaradia at the University of Cardiff proposed sDNA (2012), allowing a wide number of street network descriptors. These tools, constantly under development, allow for the multiplication of configurational applications in several domains of studies: architecture, urban and transport planning, social science, and urban geography. A fast-growing number of works is continuously demonstrating how different configurational properties might be associated with different facets of the urban space behavioural and socioeconomic phenomena of urban spaces (van Nes and Yamu 2017).

Moreover, as an accelerating factor of this proliferation of studies, we might mention the increasing availability of large volumes of data, including location information such as mobile phone data, trajectories, social media data, etc. which are providing a finer vision of urban land uses and intensity (Wang et al. 2018). As we might observe in Table 1-2, recent years have seen an increasing usage of point of interest (POI) from open data platforms (i.e. Wang et al. 2018, Lin et al. 2018) both official (i.e. national statistic organisations) and unofficial sources (i.e. Open Street Map, Google POI).

The growing number of configurational studies on a large variety of social, cultural and historical backgrounds is far from being a simple repetition of Hillier's seminal work on London. Several aspects of urban morphology, as well as methodological and statistical advancement, are continuously proposed; a substantial number of works focuses on the description of methodological differences to obtain configurational properties achieved from further developments of the methodological procedure.

Nonetheless, when implementing configurational approaches, *“the interpretation of the results of the spatial analyses must be done in correlation with an understanding of the societal and human behaviour”* (van Nes and Yamu 2017, referring to SSx). This understanding should not be limited to the meaning of configurational indicators but extended to the specific economic or behavioural phenomena under study.

What we might observe when reviewing the works of urban morphologists in Tab 1 is the reasonable and justified attention to the analysis of morphological components, while retail distribution and its spatial organisation are overlooked. An unbalanced consideration of the first term of the relation "urban form-economic value", where the measures of economic intensity are considered as simple observations, spatially distributed as a direct consequence of urban configuration. This remark probably represents one of the reasons underlying those critiques relating architectural determinism and the limited consideration among retail geographers of MET caused by the extreme simplification of the retail system.

If the goal of the study is showing the importance of network centrality in influencing retail density (the traditional approach of urban morphologist), then those works in Table 1-2 represent the correct example to replicate. If the objective is studying the relationship between the urban fabric and retail organisation, the same works provide only a partial explanation of the phenomena, and further elements should be considered and integrated.

Price value, land use, office density, or, as in our case, retail distribution, have specific spatial structure and organisation which are differently related to the urban fabric and to the human movement. The natural movement generated by street configuration represents one of the main aspects of urban form associated with human activities, but further investigations are needed to provide relevant insights about how this relationship is verified in the case of retail distribution: MET should be better examined and tailored to the specific phenomena under study. We believe that the key to a deeper understanding of this relationship might be possible only through the analysis of the retail distribution as a system and not only as individual observations.

As we can perceive from this section, the urban morphological approach for the study of retail distribution within the city, despite its complete absence in literature reviews of retail geographers, is far from being overlooked. In the last twenty years, urban morphologists have provided a large number of empirical studies implementing configurational analysis as an alternative methodology for the study of the distribution of retail activities at the intraurban level. For this reason, MET represents the theoretical framework adopted in this dissertation, the starting point from which new developments are proposed and investigated in the following sections.

## 2.3 Beyond Network Configuration

Considering our specific interest in the retail distribution in urban spaces, those works in Table.1 which implement network configurational analysis associated with retail distribution have been highlighted. This subgroup of studies represents the specific scientific literature from which we highlight theoretical and methodological observations representing the basis of the specific developments presented in this dissertation.<sup>23</sup>

### **The Urban morphological context**

The main theoretical observations we are going to discuss might be traced back to the fundamental premises and generalisations underlying the Movement Economy Theory. As we cited in the previous section, Hillier states that “*other things being equal, movement flows in different parts of a street network were systematically influenced by the spatial configuration of the network itself*” (Hillier et al. 1993). The incipit “*other things being equal*” encompasses more information than what it seems to do.

Every model about the functioning of a specific process of the city is a simplified representation of the reality, suggesting a generalisation of the relationship between two (or more) components of the urban system. Modelling requires the inevitable reduction of the complexity of those characteristics which are considered (by the analysts) to play a peripheral role (the tetrahedron of Section 1.6). These simplifications stimulate the scientific discussion: relaxing and modifying the traditional premises allow to test and adapt the original models and theories with the twofold objective of verifying the model on a different urban context and of obtaining new insights into specific aspects involved in the model.

Similarly, MET focuses on the relationship between urban form, movement of individuals, and economic intensity relying on a series of assumptions which have received less consideration.

Netto (2016) identifies in SSx three forms of reductionism: i) the reduction of urban space to its syntactic properties, ii) the reduction of spatial practices to movement; iii) the reduction of social interaction to the physical presence in public space. While the first simplification concerns the physical space, the other two consider the human and its activities in the urban spaces (movement and interaction).

MET might be considered as the extension of SSx with the inclusion of the *Function* to the relation between *Form* and *Flows*. Therefore, not only SSx reductions highlighted by Netto are inherited in MET, but also a fourth simplification regarding the localisation of Functions (retail in our case study) is introduced.

Let us focus for a moment on the first two components, *Form* and *Flows*.

---

<sup>23</sup> In Table 2 We highlighted those works where the analysis of network configuration is correlated with retail distribution implementing quantitative and statistical procedures. We do not consider these studies to be implementing visual analysis because this simpler procedure does not allow to reproduce, replicate and compare the outcomes. For this reason, despite the great interest in many works of van Nes Akelelies, who proposed several interesting applications of SSx in relation to retail distribution (from Pompeii in the ancient Rome, to the Rijnland region in the Netherland, coffee shops in Berlin, etc.), these works are mainly descriptive, based on qualitative observations and they do not provide robust statistical models/outcomes.

The first assumption underlying MET reduces urban space to a succession of voids enclosed within walls, or, in other words, to the planar juxtaposition of built-up and non-built-up footprints. Almost every configurational analyses, independently by the specific procedure, share this same notion of the urban space as a network of interconnected empty spaces. The absence of the third dimension is only one and the most evident measure which has been highlighted by several authors (Asami et al. 2003; Netto et al. 2012, 2016; Ratti 2004). The rich description of the relative spatial arrangement of buildings, plots and streets is strongly overlooked<sup>24</sup> and simplified with simple polygonal lines. As we will develop in Section 14, the form of the urban fabrics and the morphological regions they define, following the classical schools of urban morphology, seem to be a fundamental description of the urban morphological context, overlooked by MET, for which urban form is synonymous with street-network configuration.

The second assumption concerns how human movement is related to the complex network of possible paths allowed by the urban maze. This second simplification represents, at the moment, the most discussed topic among urban morphologists: the combination of different graph representations of the urban space (primal, dual, named streets, axial lines, etc.) and measures (metrical, angular, topological, etc.) are constantly investigated by a growing number of studies (Ratti 2004; Porta 2006; Hillier and Ida 2005, Omer and Kaplan 2018; Serra and Hillier 2019; etc.). The goals underlying these studies consists in finding the best generalisation of human individuals' behaviour, allowing for the describing of the overall movement patterns within settlements<sup>25</sup>; the only way to prove the superiority of a specific methodology requires its validation through empirical data collected in pedestrian counts. The debate comparing and evaluating the goodness of a specific configurational procedure is far from being closed, and we expect in the next few years a growing number of studies supported by the diffusion of advanced tracking technologies will provide new sources of information (i.e. cameras, Wi-Fi, GPS tracking systems).

As previously announced, this dissertation does not discuss the validity of a specific configurational approach and that choice will be justified by considering the technical and conceptual conditions further explained in the following sections.

As we might observe, while the simplification of human behaviour in urban space has been studied and several network configurational procedures have been proposed as different ways to capture human movement, fewer considerations have been given to the effect of urban space simplification in MET.<sup>26</sup>

---

<sup>24</sup> Sevtsuk 2010, p.59 mentions also the land use as a missing factor

<sup>25</sup> Another research direction, oriented towards social and environmental psychology fields of study, might assume the existence of different individuals' cognitive profiles underlying the perception of urban space. As we have seen in section 1.2, retail studies developed behavioral approaches investigating the different perception of individuals' sub groups related to retail; similarly, we might think about different cognitive profiles underlying specific configurational approaches. Only their combination might result in the observed patterns of the overall movement of pedestrians. Probably the easiest example we can think about is the widely studied subject of "gender difference in spatial orientation" (i.e. Coluccia and Louse 2004). Van Nes and Nguyen (2009) represent a first step towards this direction where network configurational parameters are investigated with the final aim to identify gender differences in the use of space.

<sup>26</sup> As we will discuss further in Section 3, urban morphologists proposed several procedures combining configurational approaches with other morphometric procedure (Ye et al. 2017); nonetheless, these works overlay the outcomes of the latter



Several recent publications from the research group at the University of Tel Aviv guided by Itzhak Omer might be related to the discussion about the validity of these two assumptions (Lerman and Omer 2013, 2016; Omer and Zafrir-Reuven 2015; Omer, Rofè and Lerman 2015; Omer 2017; Omer et al. 2017; Omer and Kaplan 2017, 2019). Despite urban morphologists awareness about the importance of the '*network effect*' in the determination of pedestrian flows, Omer highlights how research still needs to better understand how this effect works (Omer and Kaplan 2018). In order to answer this question, two main research directions are investigated by Omer's group: the first one is aligned with the abovementioned trend of morphological studies, where different SSx axial representations of street network are studied and compared in relation to pedestrian volume observations (Omer and Jiang 2015; Omer et al. 2017; Omer and Kaplan 2018).

The second research direction considers comparative studies between settlements considering different planned/unplanned urban networks: this approach might show interesting variations and dissimilarities of the network effect on pedestrian and retail presence. SSx analysis shows a higher correlation with pedestrian movement in well meshed pre-modernist (self-organised) street patterns than in modernist (planned) morphologies (Stähle 2008, Netto 2012). Berghauser-Pont and Marcus (2015) observed different associations between pedestrian volumes and network properties when considering three districts of Stockholm (two in the centre and one in a semi central area) chosen for their noticeable differences in both street layout (measured with ABCD method, Marshal 2005) and density (measured with Spacematrix, Berghauser-Pont and Haupt 2010). Lerman and Omer (2016) apply a similar analysis to two sub-areas of Tel Aviv characterised by different urban design paradigms: significant variations are detected between traditional and contemporary neighbourhoods.

Similarly, when studying retail distribution Cutini (2010) observes different correlations between street configuration and store density in six sub-systems of Grosseto, Italy: higher values of correlation are found in five spontaneous sub-systems while a significantly low value is observed in a recent development area ('70s). Omer and Goldblatt (2015) focus on eight Israeli cities with three types of street patterns are under investigation: older cities, new towns and hybrid cities. While in older cities the relationships between self-organised network, pedestrian flows, and retail distribution corroborate the MET, the same cannot be asserted when considering new towns where a weaker network effect is observed. Retail activities in older cities tend to be more concentrated in a linear form along main roads originated from the city core, while a more dispersed presence of stores is detected in new towns.

This work represents one of the first of its kind responding to the recommendations by the urban morphology community of using a comparative approach which "*overcomes the problems of idiographic study in which particular findings fail to yield useful generalisations*" (Conzen et al. 2012).

The majority of the works studying retail distribution in relation to urban morphology (presented in Table 1-2) typically focus on central areas of larger cities (as observed by Van Nes and Yamu 2017, Saraiva 2019) with homogenous urban form characteristics and/or with a monocentric pattern of urbanisation. For example, in Peponis and Scoppa (2015), the study

---

analysis onto the former without a real comprehension of the interaction occurring between the two components under study.

area under investigation, the city of Buenos Aires is described as having a *highly regular layout* with an intense environment<sup>27</sup>. Similarly, Barcelona (in Porta et al. 2011) with Cerda's plan and Mexico City (in Ortiz-Chao 2008) with the typical Spanish colonial influence, the study areas are characterised by the dominance of an iron-grid network. Sevtzuk (2014) study on Somerville and Cambridge focuses on a homogenous urban space characterised by a *modestly dense urban environment* excluding both the dense historical centre of Boston and the outer suburbs.

When examining these works, some questions might arise in the reader's mind: it is not clear whether the outcomes from these empirical studies can be demonstrated in large, medium, and small-sized settlement, or whether they are peculiar to large cities. Similarly, we could not verify if and how these findings are only pertinent to planned or informal (self-organised) settlements, compact or diffused cities, etc. (Berghauer-Pont and Marcus 2015).

On the contrary, Omer's group work opens the way to the exploration of these aspects: the importance of different urban planning development contexts represents an additional variable participating in the definition of the relation between network configuration and retail distribution.

Similar observations about the possibility of having different locational behaviours, depending on the dimension of the city under study, underlies recent works of Saraiva on medium-sized Portuguese cities (2017, 2019). While Omer and Goldblatt's work compares different-sized and different-planned cities, Saraiva focuses on the same type of medium-sized cities, differentiating between city centres and peripheries.

What we might notice from these works is the growing interest and need to compare between typologies of neighbourhoods, cities and towns. This information plays an essential role in the definition of the network effect works on retail distribution; information which becomes even more evident and relevant when comparing different cities or when studying large metropolitan areas characterised by strong heterogeneities.

The problem here presented within the urban geography context and vocabulary might be clearly understood when its statistical description is provided. While a positive (or negative) correlation might describe the overall statistical relationship between two variables, the same analysis applied to a specific subset of the original dataset might show higher, lower, or even inverted correlations (Figure 11). The hypothesis is that the urban morphological context represents an important variable which might help in discriminating these different behaviours. As we will further discuss in Part 4, many statistical methodologies might be implemented in order to study if and how this specific hypothesis is verified. Nonetheless, in this dissertation, the problem is not only to verify but also how to identify those typologies of subspaces which might be considered as the key factor for a better interpretation of the phenomena under investigation.

---

<sup>27</sup> "blocks are, for the most part, of regular size and shape and choice nodes are found at regular intervals of slightly over 100 metres"; moreover, the regularity is also detected by a low number of dead ends, and a majority of 4-ways intersections

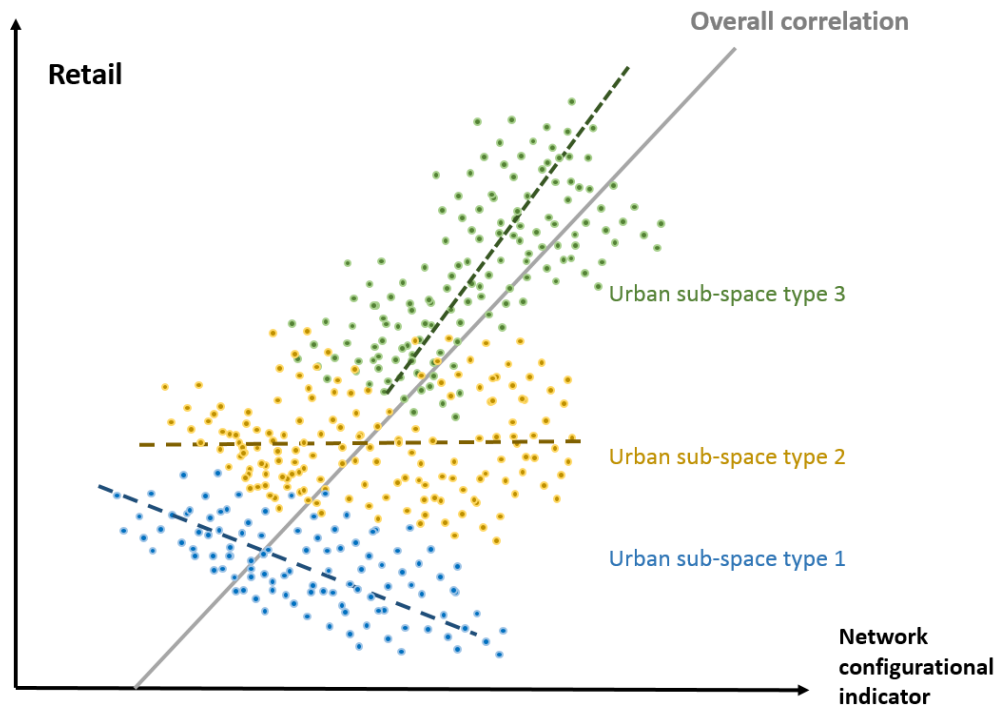


Figure 11 Urban morphological context, might outline different relationship between retail distribution and network configurational properties, traditionally hidden when the overall study area values are considered

These observations lead us to the first question which will be specifically addressed by this dissertation: **(1) how the contextual information of the urban form might help in better understanding the ‘network effect’ on retail distribution.** We believe that street network configuration properties represent an important locational factor for retail distribution; nonetheless, as tackled by Omer and Goldblatt (2016), this factor should be considered conditionally to the type of neighbourhoods, town and cities under study. Different from the works of Omer’s research group and Berghauser-Pont and Marcus (2015), where the authors propose a classification of urban (sub-) area types based on their knowledge of the settlements under study, our goal is to convert this expert-based information into a context variable which could be systematically inserted in the statistical analysis associating urban morphology and retail distribution. From this specific question, we investigate in Part 3 **(1.b) how to automatically identify urban morphological regions from the street point of view.** Only the identification of settlement sub-areas with specific characteristics might allow to: first, focus on network configuration properties reducing the possible bias induced by the heterogeneity of the urban morphological system; secondly, detect if and how the retail locational factor ‘network effect’ varies depending upon morphological subspaces.

In order to answer these questions, the comparative approach is essential: in this dissertation, the area under investigation considers a vast region encompassing different-sized cities, towns and villages where several morphological regions and urban fabrics might be detected (from unplanned medieval centres to planned grid-shaped and hybrid areas). The choice of studying a vast region within a common framework allows for the investigation of the combined influence of morphological and network configurational properties of the urban space on retail patterns.

### Detailed local urban morphometrics

In the previous section, we proposed to explore the combined effects of configurational properties with a contextual analysis of urban physical spaces: while the former focuses on the spatial organisation of empty spaces, the latter consists in the spatial pattern recognition of urban morphological characteristics, like urban fabrics and morphological regions (Scheer 2015).

Configurational and contextual analysis share a similar principle in the study of the urban physical space: spatial units are not evaluated individually for their internal properties, but relatively to the surrounding element position (configurational) or pattern of geometrical descriptors (context). When these approaches are implemented for the study of store distribution in the MET framework, the retail density (intensity, count, etc.) observed in a given spatial unit is explained through properties of the built-up environment obtained from the analysis of a more extended area including several neighbouring units.

For instance, if we would study the retail presence in the seven subspaces of Fig.12 using network centrality values and neighbourhood type as previously described, these urban morphological properties might be revealed only if considered by their relative position within a larger urban fragment.

Nonetheless, the intrinsic morphological characteristics observed within a spatial unit might also have a more direct impact on retail presence. The geometrical features observed in those same subspaces of Fig.12 might be considered as a hyperlocal environment for the retail location.

The relevance of these hyperlocal features of the built-up spaces become even more important to discriminate between urban spaces, having similar contextual and configurational characteristics. In the hypothetical case where the same contextual and configurational values describe the seven subspaces of Fig.12<sup>28</sup>, evident variations might still be detected in their intrinsic description. The geometrical descriptors of the layouts of urban features might be more or less appropriate to accommodate retailers, and several examples might be found in the scientific literature.

For instance, “continuous and well-defined building front can render streets more pleasant for pedestrians, while wider setbacks could degrade their access” (Sevtzuk 2010, p.55). Large ground floor surfaces might have a selective access effect on smaller retailers (Lebrun 2002); on the contrary, highest fragmentation of built-up surfaces might prevent stores with larger surface requirements to locate. These observations also surfaced from personal discussions with local administration experts: they explained to us how high fractioning conditions which we observe in attractive urban areas (i.e. European historical centres) represents one of the causes of the exclusion of stores with larger surfaces requirements, modifying the attractiveness of a given urban space<sup>29</sup>.

---

<sup>28</sup> Which, is not that far from the reality, since they all have been extracted by the compact area of Nice where the highest values of centrality are observed

<sup>29</sup> Larger surface requirement is not only imposed by the need to expose cumbersome products (electronic apparels, vehicles etc.) but also by policies and strategies of international brands: i.e. Apple stores in every city around the globe occupy large surfaces, above the necessity imposed by their “handy” goods; similar observation might be referred to with luxury goods.



Figure 12 Seven different urban landscapes extracted from the central area of Nice (Source: Google Maps 2019)

Sparse in the scientific literature, we might find several studies where intrinsic geometrical aspects of the urban subspaces have been correlated to retail distribution. In urban morphology studies, where the (street) block is traditionally considered as the minimal fragment of the urban space, we might find works from Yoshida and Omae (2005) investigating commercial street-blocks' morphological characteristics using LIDAR data in the central part of Tokyo, while Joosten and van Nes (2005) visually explore the co-occurrences of retail presence and block types in Berlin.

Several empirical studies have also been proposed by the behavioural approach in retail geography (Section 1.3) where the perception of detail urban morphological and urban design features are studied with the final goal to understand the human behaviour in retail centres.

Nonetheless, none of the morphological qualities of a single place could alone define a good store location; as observed by Sevtzuk (2010, p.55) only *“a cumulative collection of favourable spatial conditions might well inform the appropriateness of a place for selling goods of whatever type. The work of morphological analysis suggests that a location model that aims to capture configurational differences between various locations ought also to integrate a series of detailed place measures that go beyond graph theory”* (Sevtzuk 2010, p.55).

The combination of intrinsic and configurational properties represents a relatively recent field of research. As we might observe in Table 2, while the few years after Hillier's seminal study of London (1999) saw the implementation of MET in its original version, more recently a growing number of studies is investigating the importance of a combined evaluation of configurational properties and other urban morphological aspects.

In this direction, we found the recent work of Ye et al. (2017) combining configurational metrics with street block-based measures in the analysis of catering businesses' distribution (interpreted as a measure of the vitality of urban spaces) in Shenzhen, China. This work represents a specific application to retail (catering) which follows a series of studies from the same research group (Van Nes et al. 2012, Ye et van Nes 2013, etc.) where geoprocessing

procedures are suggested as possible solutions for the combination of SSx indicators with other block-based urban morphological descriptors<sup>30</sup>.

We share with these works the idea that urban morphological description of urban spaces represents an important aspect which should be considered together with configurational properties; instead of ambiguous modifications of traditional configurational approaches in order to integrate different morphological aspects within the same methodological framework, the combination of procedures highlighting different facets of the physical form should be preferred (Kropf 2009, 2017). This second approach would allow a comparative study between different components of the urban form outlining the relative and combined influence on urban phenomena.

Nonetheless, bridging these different aspects should not be considered as a simple superposition of miscellaneous procedures with different underlying notions and interpretations of spatial units and morphological measures. While configurational studies consider the space from the point of view of public spaces integrating human behaviour, block-based methods use streets only for the delimitation of spatial units, privileging an aerial vision of the urban space. For this reason, we believe the specific methodological procedure proposed by Ye et al. (2017) was not adapted when studying the retail distribution at the microscale level, since POI pattern, street network configuration, and block intrinsic descriptors are combined without a common human-behaviour-based perspective in the analysis of the urban form. Moreover, these combination procedures using both street-blocks and/or uniform grids (as in Van Nes et al. 2012, Ye et van Nes 2013, etc.) raise important questions about statistical robustness, due to the biases induced by the high level of manipulation and transformation of the original data.

Considering that the simple combination of existent methodologies is not sufficient for the analysis of the different aspects of the physical space, ***(2) how to bridge configurational contextual and local urban morphometrics in a common framework for the study of retail distribution.***

Surely, the GIS platform might help in the computational task. Nonetheless, we think that the first step towards the answer of this question is to define a clear and common conceptual interpretation of the urban space and the phenomena under study, in order to bridge different representations of the urban space. Several indications have been suggested in the scientific literature of the last two decades: Zhang and Kukadia (2005) recommend the use of a behavioural-based (scale) spatial unit. The importance of street spaces have been highlighted in (retail) geography (Fleury 2004), spatial statistical studies (Okabe 2012, Yamada and Thill 2004,2007, etc.), environmental psychology (Harvey 2016, 2017, etc.) and more recently Kropf (2009, 2017), which identifies in the street the possible element to bridge different urban aspects of urban morphology.

---

<sup>30</sup> The authors implement Spacematrix procedure (Berghauser Haupt and Pont 2010): a specific methodology for the identification of urban block types based on the combination of geometrical descriptors of building layout within blocks. Further descriptions and discussions about this procedure are provided in Section 6.

If the street might be considered as the best spatial unit partition in our case study, the following question naturally arises: **(2.b) how to provide a quantitative description of the built environment along a street.**

The description of the built environment at the intrinsic level of the street (substituting the more traditional block-based measures) would allow a better integration and combination of morphological measure in the study of human-based phenomena. As we further detail in Part 4, the answer to this question might partially arrive from the block-face analysis of Vialard (2013), and, more recently, from environmental psychology studies of Harvey et al. (2015), proposing several measures for the morphological description of streetscapes.

## 2.4 Beyond Retail Density

While in the previous section we discussed urban space and human behaviour simplifications, we focus in this section on some considerations about retail distribution. While retail geography provided essential insights about the spatial organisation of stores in urban spaces for seventy years, MET simplifies this notion to a simple measure of density. In this section, we show how MET completely ignores two main aspects of retail distribution despite their importance in the determination of the economic and social attractiveness and vivacity of urban spaces.

Regarding the three assumptions underlying MET concerning urban space, human behaviour, and retail locational process, this third simplification has not been discussed yet. Retailers are individually and uniformly considered (absence of differences and direct interactions among stores), and their behaviour responds to one single rule: high volumes of movement attract retail activities, and, in their turn, they generate additional movement in what we previously described as the multiplicative effect. In MET, the profit maximisation rationality of retailers is expressed through a utility function which is directly proportional to the natural movement observed in a given space. As for the previous simplifications, we might further detail this notion and propose alternative ways to investigate the spatial distribution of retailers in relation to urban space.

Retail geographers have extensively examined the spatial interaction between movement (named flow) and store location; the presence of different behaviours have been identified and discussed. Following the work of Lebrun (2002, pp. 337-347) we might outline a finer description of the relationship between retail and movement: the author describes a first broader distinction between **active** and **passive** interactions. In the first case, stores are both attracted by and generators of movement: they intercept a fraction of the pre-existent flows and they actively participate in the description of additional human movement within urban spaces. On the contrary, in the second case, stores are attracted by flows without stimulating any additional contribution to the pre-existent situation. This first classification might be further detailed.

Within active interaction we might identify **sympiotic interaction** which describes those stores attracted by a specific urban space where high volumes of movement are observed; once

located, they increase the pre-existent flow volume and consequently the attractiveness for new stores. Symbiotic interaction as described by Lebrun, might be considered the closest notion to Hillier's MET: microeconomic functional patterns are located on those streets described by the highest through-movement values at the global scale ("foreground global network", Hillier 2007), triggering the "multiplicative effect" process. Nonetheless, when observing retail patterns, we might notice that not all retailers are positioned along these main axes and some of them are observed in streets with lower flow volumes. In this case, "*the consideration of the observed flows [as a locational factor] seems to fade in favour of stronger confidence into the market activity [i.e. endogenous factors] [...]. Consequently, retail activities gain higher independence from the accessibility of places or at least they try to generate their own flows. In this case, we could talk about autonomy interaction between retail and flows*" (Lebrun 2002, p.341). The author identifies three sub-types of autonomy interaction triggered by three different, almost contradictory, causes. The first group, the **proximity autonomy**, describes those retailers located at the heart of a residential community and characterised by a small customer catchment area which provide the sufficient clients and thus incomes for its survival. These typologies of stores represent a service for the local community and their contribution to the local flows is limited. Within the second group, **influential autonomy** considers those stores with a high attractiveness effect, which is able to generate and modify the original flows; within this subgroup, we recognise highly specialised stores or concentrations of large specialised retailers. Finally, the **suffered autonomy** describes those stores that are located in more isolated places, forced by locational costs; in this specific case, the autonomy, which is originally a controlled phenomenon, is here an effect imposed by exogenous factors. We might find similarities of the suffered autonomy with BRT (Section 1.1) where retail activities with lower economic potential are forced in more peripheral spaces. Nonetheless, suffered autonomy might find its significance also in urban morphological factors: an emblematic example is represented by large Haussmannian boulevards where the planned *monumentalism* of the street frontage requires the realisation of buildings with large ground floor surfaces. Consequently, despite the high attractiveness of the main street, small retailers are pushed out to the surrounding areas.

Passive interaction describes those retailers considered in both their individual and agglomerated format, which do not have sufficient attractive influence on the surrounding spaces necessary to generate their own flows or to modify those already existent. We might identify **ancillary interaction** to be when stores are located along streets with high flow volumes, commonly generated by other urban functions which are specifically researched and used as a locational strategy in order to target a specific group of customers. Fast food restaurants located close to public transport hubs, or touristic attractions within highly frequented pedestrian strips represent the most evident example of ancillary retailers; despite their passivity, these activities presuppose a substantial economic investment, defined by the good potential location, and compensated by lower surface. Finally, **parasitism interaction** identifies those stores having the same passive interaction of ancillary stores, although their presence might also cause a modification upon the original flows; an example of this would be souvenir stores which redirect flows as originally outlined and reduce the attractiveness of the same space for the local community (Lebrun 2002, p.344).



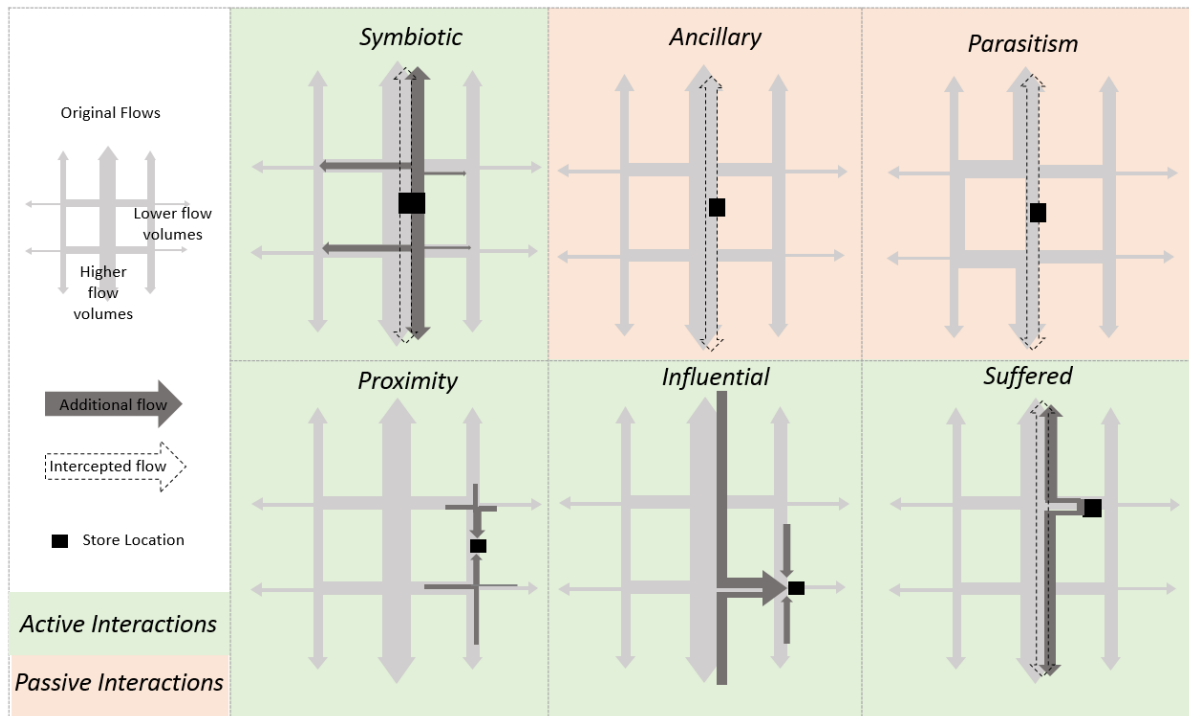


Figure 13 A graphical representation of the spatial interactions between retail and flows elaborated from Lebrun's description (2002).

Bearing in mind Lebrun's classification of the spatial interactions between flows and retail location, we might now reconsider Hillier's MET in order to better understand both its potential and limits.

Two observations should be highlighted: firstly, while MET considers only that part of the flows generated by the network configuration (the natural movement), retailing geography analyses the interaction between stores and flows considering the overall movement generated by the whole of urban functions (commercial and non-commercial). Secondly, it also accepts the possibility of (at least partial) autonomy of store location choice from the observed flows.

As we previously discussed, the combination of morphological contextual variables and network configurational properties discussed in the previous section (Section 2.3) might provide a better representation of the potential movement, reducing the discrepancy when compared to the observed flows<sup>31</sup>. In the hypothetical situation where natural movement is perfectly correlated with real flows, we might ignore the first of the two differences;

<sup>31</sup> Only the integrated study of the real distribution of urban functions with accessibility properties might provide a more precise description of the urban flows as we can observe in the urban transport planning literature. Nonetheless, we remind the readers that our goal is to study the relationships between urban morphology and retail distribution (form-functioning of the city): the distribution of urban functions to obtain a better estimation of flows with the ultimate goal of describing retail distribution will result in a tautological problem. This issue hinders all considerations and speculation about the importance of urban morphology and urban design as an influencing factor for the determination of urban spaces usages. The real challenge for urban morphologists is to provide a description of the possible flows allowed by urban morphological properties' ad descriptors, keeping in mind the important separation between form and function of urban components.

nonetheless, the assumption about retail behaviour as described by Hillier's MET would only fit the symbiotic interaction description of Lebrun, leaving the other five typologies unexplained.

On the other hand, we should also recognise that the finer description provided by Lebrun—that these forces do not perform individually and the spatial combination of different retail behaviours positioned on a same urban region might prevent the correct identification of each store behaviour. For instance, if we consider the well-known Oxford Street in London, how could we differentiate and model each store locational behaviour among the six typologies previously described? And, even if an experienced retail geographer might achieve this task, how could we replicate the same analysis for every store in London or an even larger region where different (almost unique) combinations of retailers are observed?

The assumption about retail locational behaviour in MET allows for a quantitative description of the overall distribution of retail density, summarising the six typologies of interactions described by Lebrun; this is made possible through the combination of the two concepts of attraction and multiplicative effect. Differently from what traditionally is described by Hillier's MET, we might underlie how, these two forces do not perform simultaneously: we could observe multiplicative effect without attraction to high volume flows, as in the proximity or influential interactions, and vice-versa attraction without multiplicative effect, as in ancillary and parasitism. This shortcut comes with an important cost: the loss of information about the complex reality of the interactions among stores with and through flows.

So far, we have compared the detailed classification provided by Lebrun and the simplification suggested by Hillier's MET. As is commonly known for every theory and model, even in this case we face an important trade-off between generalisation and detail of the information. However, we still have not found a solution that allows for the enrichment of the MET under the consideration of different retail locational behaviours.

Possible answers might be considered through the analysis and comparison of retail typologies (as in Porta 2012 and Wang 2014) or the selection of a specific retail category, traditionally catering businesses (as in Sevtsuk 2014 and Ye et al. 2017). The different correlations between retail pattern and network configurational properties described for different retail families might suggest the prevalence of a specific locational behaviour related to the type of service/good. Nonetheless, we believe that this direct equivalence is not plausible: a restaurant might be located in a central touristic area, in an ancillary position (business district area) or even in a more independent position within a semi-central residential neighbourhood. Similar observations might be highlighted for furniture and clothing stores which might be located both in exclusive hyper central districts and within more peripheral commercial ribbons. Within each retail category, we might identify almost all locational behaviour of Lebrun with different proportions.

Since the individual description of retail location behaviour seems hard to model, what we propose in the next section is to consider its aggregated form: the different locational behaviour previously outlined, as explained by the same Lebrun, might describe retail both at the individual level and in its agglomerated form.

### **Retail structure: agglomerations and organization**

Let us now consider how retail is measured and analysed in MET empirical works. As we might observe in Table 1-2, all studies using the original or a modified version of MET, measure retail distribution as a density, intensity, or count<sup>32</sup>: a monodimensional description of retail which cannot fully represent the multifaceted reality of retail. As observed by Greenberg (1995), the simple enumeration of stores and businesses within a street or a given urban space “does not tell the whole story”.

The same observations highlighted by urban morphologists about the reduction of the urban morphological complexity to a simple matter of built-up density, or the street network complexity to simpler measures of accessibility, should also be addressed in those works considering retail with a simplistic store count indicator.

Since our goal is to study the relationship between urban morphology and retail spatial organisation in urban space, we should not reduce the comprehension of retail to its simple measure of density.

Starting from the '80s, we began to observe a growing interest in urban morphology studies, surrounding the analysis of the complexity of built-up spaces at the intraurban level; similarly, retail geographers have produced an extensive literature with important insights about retail organisation and structure within settlements. As urban morphology studies have been exploring the spatial interactions among buildings, streets, and parcels, in order to better understand the urban morphological system, in the same way retailing geography has explored the spatial interaction of stores producing specific scientific literature about retail systems. The spatial and functional organisation of individual stores originates specific retail agglomerations structured and structuring the urban space.

What is then evident is the necessity to bridge the two scientific domains of retail geography and urban morphology which have been so far independently developed. As graphically represented in the following scheme, on one side we find retail geography studies focused on the recognition of retail patterns and their spatial organisation; within this domain, we might find several references and connections to urban morphological components, limited to visual qualitative analysis or simple measures of built-up density (as observed in Section 1.5). Symmetrically, on the other side, urban morphology studies focus on the recognition and description of urban structure; within this domain of research, retail is recognised to represent an important component of urban spaces but is frequently relegated to a simple measure of density or concentration (i.e. MET applications in Table 1-2). Bridging these two systems (as in the case of physical infrastructure) requires equal attention of the two components we are trying to connect, avoiding the impasse generated by an unbalanced consideration of the problem. Studying the relationship between urban morphology and retail distribution should uncover how the spatial structure of the two components under investigation are related.

---

<sup>32</sup> Sevtsuk PhD dissertation and the following 2014 paper represent the only works in Table 1-2 where retail is measured at the level of the building with a Boolean variable.

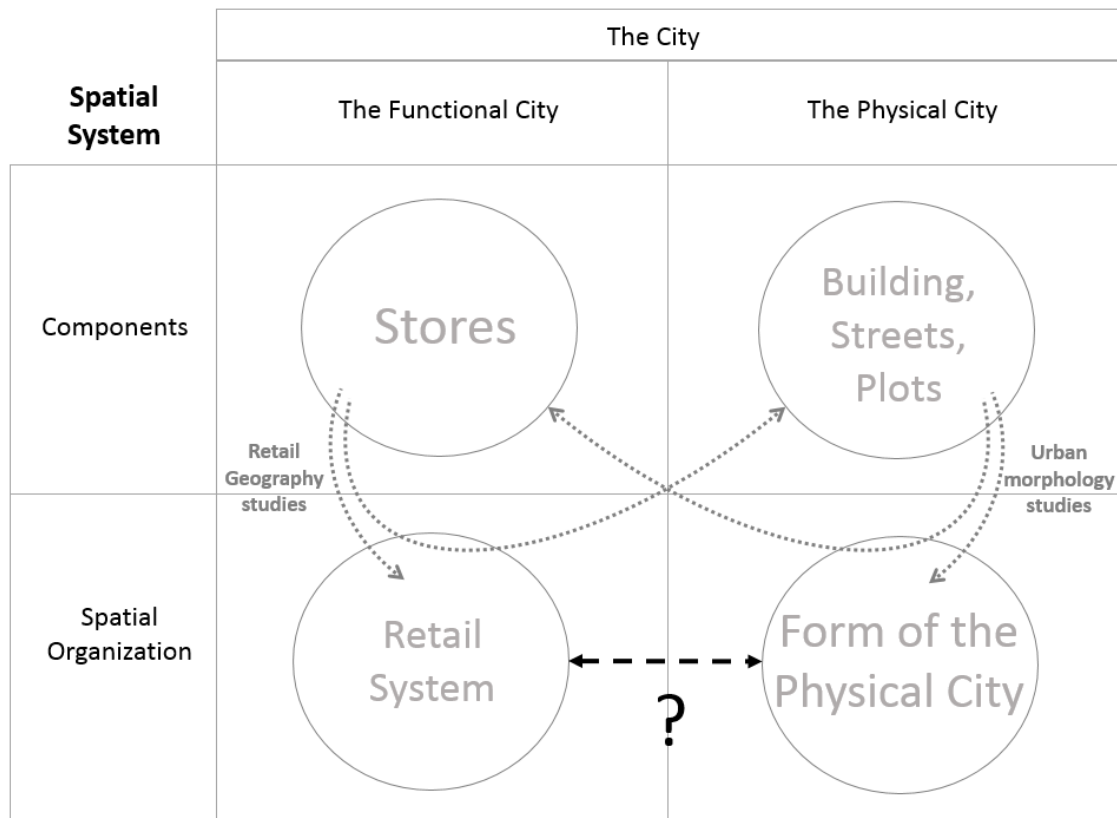


Figure 14 Graphical representation of the retail and physical systems, their elementary components

What we propose is a conceptual shift from the analysis of the individual elements composing the two spatial systems to the study of the relationships between their spatial structure (Fig.14).

This shift might help in answering the third question addressed by this dissertation: **(3) how urban morphological characteristics are related to the spatial organisation of retail agglomeration types.**

In the scientific literature, a lack of quantitative studies is noticeable in this specific direction, probably because of one fundamental reason: the nature of the objects which distinguish these two systems (Hermenault 2017). On one side the material notion associated with the physical city; on the other, retail activities could be conceived (and are often conceived by economists) as immaterial components, whose structure is described by interactions (mainly economic interactions) not easily measurable. While in the first case we might find support for a large number of quantitative works providing a solid background to our study, the nature of the second has hindered the development of a large consensus about quantitative procedures, and, consequently, qualitative approaches are traditionally preferred.

The spatial location of each store represents one means we dispose of for the materialisation of retail structure. Density or concentration represents one aspect, but not the only one, to describe retail agglomeration and its structure.

Differentiation *“is necessary to the retail concentration for two reasons: firstly, the need of demarcation in order to strengthen the legibility. The retail concentration should reach the*

*position of a recognised commercial entity. A commercial space not identifiable is a commercial space without attractiveness characteristics. Secondly, the need of differentiation among other retail concentrations in order to guarantee its specificity.” (Lebrun 2002 p. 379).*

Van Nes (2005) represents one of the few works we found which investigates the relationship between urban space properties and “typologies of shopping areas” in Amsterdam; in this study, the author visually describes the configurational properties (using SSx) concerning different types of retail agglomerations. Despite sharing the same final goals and interest of this study, as already noticed in other works from the same author, her analysis is mainly visual, realised through the traditional Space Syntax vision of the problematic; no statistical methodologies are implemented to demonstrate the objectivity of her conclusions and to allow the reproducibility of her study.

What is more, retail agglomeration types as intended by the author result from broad qualitative discrimination between clustered and linear retail agglomerations which poorly summarise the wide discussion provided by decades of retail geography studies (see Section above). The spatial identification of these two shopping area types is based on the expert knowledge of the author, which prevents any generalisation, comparison and implementation of modelling procedures (similarly to what was previously observed in the work of Omer and Goldblatt in the identification of city types).

As observed by DeLisle (2005, p.100) *“Although specialists who concentrate exclusively on the retail industry may be able to operate without an industry consensus on shopping center classification, the same cannot be said for those for whom retail is only a part of their focus.”*

It is thus evident how, in order to perform the analysis and modelling of the relationship between urban morphological characteristics and typologies of shopping areas, a robust method for the detection and definition of retailscape types is needed. In Part 2, we will try to answer the following question: **(3.b) how to automatically identify and describe the spatial patterns of retail activities (retailscape).**

The study of the relationship between built environment and retail agglomeration types is of a fundamental importance--not only for a better comprehension of the urban system but also for more practical purposes linked to planning studies and practices.

From the observations of Jacobs about the social uses of urban space and the following work of Hillier and Hanson on the social logic of space, urban designers, and morphologists, these researchers justify the specific attention placed upon the built environment of their work, stressing the importance of the possible effects on vivacity, attractiveness or vibrancy.

The hypothesis underlying point 3 and 3.b is that each type of retail agglomeration has specific morpho-functional characteristics, which directly impact the typology of social activity of the urban space where they are located, yet in different ways. Studying the combination of urban morphology and retail agglomeration types, it might better describe different types of vivacity, attractiveness and vibrancy of urban spaces.

The outcomes of these specific questions (as we will further discuss in the conclusive section) might help planners and local administrations in projects of requalification and reactivation of

urban areas: precise urban morphological characteristics might be better shaped when a specific social and retail environment is targeted, or conversely, which retail activity combination is better suited in a given neighbourhood.

### **The Survivorship Bias in retail studies**

As we introduced in the previous sections, quantitative analysis in both retail geography and urban morphology have been boosted by the fast development of geoprocessing techniques and statistical procedures combined with a growth in computer power and data availability (Aversa 2018).

The advantages brought about by these technological advancements have both reinvigorated spatial statistics and quantitative geography and opened the way to new research possibilities from the intersection with domains of research having a shorter tradition in quantitative spatial analysis. Nonetheless, the implementation of more sophisticated analytical procedures giving a higher robustness appearance might also be at the origin of vulnerabilities: methodological choices might unintentionally hide essential aspects about the process under study, leading to misleading conclusions. This work as well is not immune from this concern: we are aware that future scientific and technological advancement might unveil criticisms about our methodological choices and some possible development and improvements are already proposed in the last sections of each part.

In this section, we focus on the methodological aspects of previous works of Table 2, highlighting a specific weakness shared by many of these works and which we try to prevent in this specific dissertation.

Quantitative analysis based on robust statistical procedures have the advantage of providing support to traditional visual and qualitative examination; more important, they deliver to the analyst objective means to evaluate the goodness and robustness of specific outcomes, endorsing a specific research concept. The application of these statistical procedures might help in the reduction of critical interpretation bias which traditionally affects visual analysis; among these cognitive biases, the most recurrent is the confirmation bias, defined as the tendency of the researcher to look for, interpret, and focus on those elements which confirm a notion, a model, or a theory (Taleb 2007).

The doubts about the presence of confirmation bias might frequently come to one's mind when reading interesting but ambiguous works showing an apparent visual correlation of configuration properties (applying SSx configurational analysis) with social and economic phenomena. In these works, the attention of the reader is traditionally caught by appealing images where patterns of observed occurrences representing the phenomena under investigation (retail activities in our case) are superimposed on configurational indicator properties. These visual explorations of the data are frequently used as evidence to validate hypothesis about the link between social phenomena and urban morphology; moreover, these studies try to convince the reader by focusing the attention on specific examples which

might be hardly used for generalisation about the phenomena<sup>33</sup>. Oxford Street, Les Champs-Élysées or Friedrichstrasse might not be the most archetypal examples to confirm general theories and models about streets' economic/vivacity potential; they seem to us examples of outliers rather than representatives of general behaviours.

The graphical representation proposed in these visual analyses (namely exploratory data analysis EDA), is a common practice which GIS platforms made easier to implement; EDA is the starting point from which every analyst elaborates new ideas and hypotheses about a specific phenomenon. Nonetheless, it represents only the first step of the scientific investigation which requires a more robust statistical methodology to (un)validate these hypotheses and to avoid idiographic and subjective drifts of the analysis.

Even when considering more accurate and sophisticated procedures, cognitive bias might still affect the analysis. When reading works of Table 1-2, we might notice one common inclination of both visual and statistical analysis: the attention of the researcher (and the reader) is captured by those spaces where retail is detected. Only when its concentration/count/density is greater than zero, the characteristics of the urban space are investigated; the absence of retail is often ignored in the same way as missing data. Even in the discussion of our objectives in the previous sections, the reader may notice our specific focus on the presence of retail, the factor triggering its presence, its multiplicative effect or, again, the typology of retail agglomerations. The absence of stores in urban spaces has not been considered in any study as an integral part of the process defining the retailing spatial organisation.

This statistical bias is similar to what several disciplines have recognised as the survivorship bias: epidemiology, history, finance; they have all demonstrated the tendency of statistic studies to draw conclusions considering a subset of “successful” individuals which are not representative of the overall population. Authors such as Taleb (2001) expertly explained the phenomena in his book “Fooled by Randomness”, discussing the example of the survivorship bias in the evaluation of someone’s career performance when considering the sub-sample of successful individuals (spatially segregated in specific neighbourhoods or cities) without considering the entirety of the original population. Another well-known example is of the famous mathematician Abraham Wald: during the Second World War, analysis of the distribution of holes on damaged aircrafts were used in order to highlight those part of the planes requiring strengthening. Wald noticed how these studies were considering only the damages of aircrafts returning from the military conflict, while those dispersed were not included in the assessment.

In our case study, we might find a similar correspondence to the survivorship bias: the concept of “successful” or “survived” subset of data is automatically associated with those spatial units with a presence of retail different from zero, ignoring those spaces where no stores are observed. Consequently, statistics analysis revealing significant associations with

---

<sup>33</sup> Sevtsuk highlights how data in SSx analysis seems to be conceived “without a clear hypothesis about a given property of urban form and starting with ad-hoc data putting the cart before the horse” [...] “a methodologically more rigorous approach for investigating the relationship between the urban form and social processes that take place therein is urgently required” (Sevtsuk 2010, p.63)

morphological and configurational properties are not representative of the overall space units under investigation.

The emphasis on a specific subset of spatial units might be explained by the high heterogeneity characterising the micro-retail distribution in urban spaces (as well as most of the socioeconomic observations) which becomes even more evident when using fine-grained measures. In Omer and Goldblatt (2015), retail presence covers between 11 and 38% of the axial lines of the eight cities under study; in Barcelona, between 14 and 40% of the grid cells are described by a retail density higher than zero, depending on the retail activity family under consideration (Porta 2012), while 38% of plots have a commercial use in Buenos Aires (Scoppa 2013; Peponis and Scoppa 2015). These ratios might also vary depending on the size of the study area: in our study area, retail is detected only on 12% of the streets segment.

Selecting and analysing spatial units based on a specific criterion is not a bias in itself; bias might arise in the interpretation of its results. This selection has been frequently implemented (i.e. Porta et al. 2009, 2012, Wang 2014; etc.) and it might be considered as a legitimate choice of the authors, under the condition that interpretations and conclusions of results highlight this background restriction underlying the statistical analysis. Consequently, the same statistical relationships might or might not be verified when extending the analysis onto the whole dataset.

For instance, let us consider the hypothetical case where the vegetation cover index along a street is positively correlated with retail density, with the background condition that at least one store is observed. If this correlation is interpreted as a general relationship extended to all streets, this would lead the reader to a misconception of the phenomena: the analysis did not consider all those streets bordered with trees that we might find in suburbs, and, even worse, one could consequently expect more retailers on streets crossing woods and natural spaces than in the city centre!

Beyond this example, considering the case of tree coverage ratio, the same problem might be less evident but still present when studying network configurational properties (both in statistical and visual analysis). We might observe streets with the same network property values, but only those hosting retail are analysed for correlation studies. Urban form and configuration properties might be considered to be a necessary but not sufficient condition for retail presence.

These conditions are of primary importance, especially when considering that urban studies are frequently utilised to provide useful information to urban planners about those morphological characteristics supporting retail presence and the socioeconomic phenomena related to it (vivacity, etc.).

Few works have tackled this specific aspect on the absence of retail activities: Omer and Goldblatt (2015), when correlating retail presence and network configurational indicators, compare values obtained on the overall study area to those obtained on the subset of streets with retail presence different from zero. Sevtsuk (2014), using buildings as spatial units, applies multiple regression analysis on the subset composed by buildings with commercial functions and those located within a radius of 100 metres. Saraiva et al. (2019), when studying



the vacancy ratio in urban spaces, highlights a similar issue: the authors notice how retail studies traditionally focus on store presence instead of considering the overall surfaces including both active and inactive commercial surfaces.

Retail absence has never been fully integrated into this family of studies, often overlooked or totally ignored. Since more than 60% of the urban space (streets, building or plots) do not contain retail activities, its exclusion from the statistical analysis is not only a possible source of bias but also a significant loss of knowledge: for this reason, in this dissertation we propose specific measures which are able to integrate the analysis of retail absence in the overall models, and implement specific analytical procedures. **(4) What are the characteristics of the urban morphology inhibiting retail presence and development?** We believe that the absence of retail should be considered just predictive and informational as its presence; methodological choices for the analysis and modelling of retail distribution needs to consider both its density and its absence within a common framework. The absence of retail is at the core of the fourth question addressed by this dissertation.

The idea underlying this approach is that two processes might generate the distribution of retail activity: the first is responsible for its absence or presence, while the second explains its density. This dual nature of the process of retail distribution finds its analogies to what is discussed in the previous section (5): the presence/absence might be associated to the concepts of attraction, while density/intensity to the multiplicative effect. This dual description allows, for instance, to investigate if specific combinations of urban morphological parameters are associated with the presence/absence of retail while other different parameters are correlated with its density. From this observation arises the following question: **(4.b) retail attraction and multiplicative effect might be triggered by different combinations of urban morphological characteristics.**

In order to investigate the specific aspect of absence in retail distribution, it is essential that the analyst avoids those data manipulations leading to the manifestation of the survivorship bias. This bias might not be caused only from a direct selection of the subset. Other methodological procedures might hide the absence of retail; for instance, all those smoothing and interpolation procedures transform retail point patterns from a discrete distribution to a surface of continuous values. Kernel Density Estimation (KDE) is a non-parametric estimation method which is traditionally used to obtain the probability function of a variable; a well-known smoothing technique which would have been used in several works among those in Table 1-2 (Porta et al. 2009,2011; Wang et al. 2014, Saraiva et al. 2019). Similarly, Inverse Distance Weight (IDW) or Kriging methods (Ordinary Kriging Method OKM) have been implemented in Gianninopoulou et al. 2016 for residential house prices. Through these procedures, it is possible to estimate the value of a density function in a given space using the information about the surrounding observed points.

The first criticisms surrounding the use of these methods is that the anisotropy of the urban space is not considered: differently from environmental issues (as pollution or noise dispersion), when analysing phenomena related to the human movement (accidents, stores etc.), the network-constrained nature of the space should not be ignored (Batty 2005). Its importance (as we further discuss in the next sections) has been demonstrated in several

spatial statistical studies analysing point patterns on networks (Yamada and Thill 2004,2007; as well as several works from Okabe research team in Tokyo). In order to obviate this specific limitation, Network Constrained KDE (NC-KDE), Kriging and IDW have been implemented by Okabe's research group (2006, SANET Okabe et al. 2009, Okabe and Sugihara 2012 etc.), while a Network-Based KDE (NetKDE) has been studied in Timothée et al. (2010). Applications of these recent procedures to the specific case of retail distribution might be found in works of Borruso (2008), Timothée et al. (2010), Fusco et al. (2017) etc.

Nonetheless, perhaps even more accurate, network-based estimation and interpolations should still be considered to be smoothing methods which are used to provide "*a better representation of the phenomenon*" (Thimotee et al. 2010). As highlighted by Scoppa (2013, p.63), "*by using KDE, the highly detailed representation of both centrality properties and [more important] retail data is smoothed across the bandwidths areas*" and consequently "*there is a risk of overstating the relationship between measures of street connectivity and retail*".

Moreover, streets where stores are not originally observed might be associated with estimated values of density higher than zero, therefore raising the survivorship bias problem. If a specific morphological reason explains the retail absence of a given subspace, the implementation of KDE hinders the identification of this precise information.

What we should remind the reader is that KDE, IDW, and Kriging methodologies have been originally developed in the geo-statistic domain of studies with the underlying goal to estimate values of an unknown continuous function (precipitation, mineral concentrations etc.) in those subspaces where we do not dispose of direct observations. The estimated surface of the precipitations, for instance, represents a "better representation" of the one offered by partial and punctual observations (sample data).

When studying retail distribution, independently by stores' physical representations, we are dealing with discrete distributions precisely known for every point of our space study<sup>34</sup>. It is then evident how retail distribution does not share the same underlying premises and issues for which these methodologies have been developed. How could retail distribution be better represented by an estimated smoothed surface than the real location of each store itself?

The appropriate use of KDE might be observed in studies aiming to investigate, for example, the spatial influence of retailers on a specific phenomenon such as accessibility, mobility, etc. On the contrary, if the goal is to explain the observed distribution (like in our specific case study), the use of smoothing, or interpolation approaches, produces an artefact characterised by high autocorrelation values artificially introduced (by construction) and assigned to spaces where retail is naturally absent. It should be now more evident to the reader how these procedures contribute to the survivorship bias effect, precluding the study of the precise aspects of the real distribution of retail, of which includes its absence.

---

<sup>34</sup> Localization bias is always present with different degrees of importance; KDE might be used in order to spatially represent the uncertainty relative to the retail geo-localization process. In this work, attentive analysis and corrections have been applied to the original data in order to minimize the localization bias. For this reason, we assume that the retail locations are (almost) perfectly known and they do not need specific smoothing or interpolation procedures to estimate their presence.

A second important aspect leading to survivorship bias is related to the choice of the spatial unit. The partition of the study area in subspaces is a necessary procedure to provide a local quantitative measure about the variables under investigation and to study its spatial variation. Specific assumptions about the phenomena under study or a specific goal might justify the choice of a specific partition or the use of exogenous partitions.

Independently by the underlying theoretical assumptions motivating the spatial unit choice, specific attention should be given to how the spatial partition might be at the origin of survivorship bias. From the seminal work of Gehlke and Biehl (1934), a wide number of studies have shown how spatial units aggregating data at different scales can affect the analysis results. This problem is known under the name of Modifiable Areal Unit Problem (MAUP; Openshaw 1984, Openshaw and Albanides 1999). Spatial units are usually arbitrarily determined and modifiable, resulting in the generation of “artificial spatial patterns” (Heywood 1998). Two main strategies have been proposed in order to reduce the effects of the MAUP: when studying human-related phenomena, Zhang and Kukadia (2005) suggest the use of behavioural-based scales, while Holt et al. (1996) recommend decreasing the scale of aggregation. In Part 4, we will discuss further statistical criticisms relatively to the spatial partition choice arising from the heterogeneity of the spatial pattern of features under study (heteroscedasticity problems).

It seems to us that among the different discussions about the effects of MAUP, the spatial distribution of empty areas has been overlooked: as previously discussed, both absence and density represent two aspects describing the spatial pattern of retail (and the same might be argued about crimes, accidents and other discrete human-related phenomena etc.). When choosing large spatial units, alike to what is observed in the case of smoothing and interpolation, the precise and sharp information about retail absence is more or less diluted by the size of the partition: the majority of spatial units assume values that although very small, they are still higher than zero. By reducing the presence of zeroes, we will not be able to distinguish and evaluate any difference between the two processes. On the contrary, when small-sized spatial units are chosen, the distinction between absence/presence and density values is observable.

We discussed in this section the fourth and last point which we investigate in this dissertation: the analysis of both absence and density in the study of retail distribution. Several thematic and methodological procedures have traditionally hidden this concept both in qualitative (cognitive confirmation bias) and quantitative assessments. The data manipulation procedures might not only produce a distortion of the real correlation (as traditionally observed in MAUP discussion), but more importantly, it inhibits the study of those factors associated with the absence of the retail distribution.

The solution is not unique. Together with the two traditional criteria aforementioned (behavioural-based and small scale spatial unit), we propose in this work a third condition: the

inclusion of units with zero occurrences in order to analyse the hypothesis of a dual process underlying the spatial distribution under study.

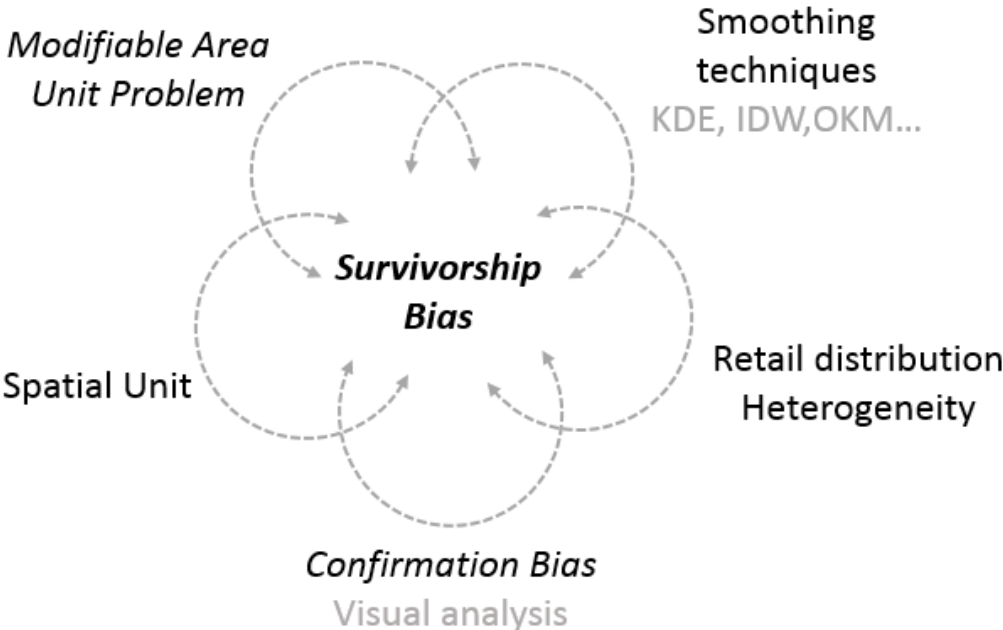


Figure 15 Possible causes at the origin of the Survivorship Bias

## 2.5 Overview of Research Questions

This dissertation finds its theoretical origin in the Movement Economy Theory. This relatively recent theory, elaborated in the urban morphological domain of study, is based on the specific application of graph theory to the study of the urban form. The simplicity and robustness characterising configurational analysis enabled a wide number of implementations. The recent growth of applied research in this direction has also been fostered by the combined effect of geographical data availability and the elaboration of automatic tools and GIS extensions. Despite the theoretical closeness of MET within urban morphology studies and the notions developed in retail geography, these two disciplines have hardly dialogued, and, even less, converged into a common framework.

In the two previous chapters, we have explored the premises of such convergence: notions from retail geography might be integrated into the MET enriching the original theory with new and more specific research questions. Similarly, recent developments in urban morphology studies might renovate the original MET, leading to the investigation of the role of different aspects of the form of the physical city.

Two main considerations have been specifically discussed in the previous chapters from which three specific research questions might be outlined.

***(1) Can the contextual information of the urban form help in better understanding the 'network effect' on retail distribution? If yes, how?***

The hypothesis underlying this first question is that the influence of the street network configuration, despite being generally valid in general, might take a different course of action in differently shaped urban spaces.

***(2) Beyond the network effect, are there other characteristics of the urban form inhibiting retail presence and development? If yes, which ones?***

The hypothesis underlying this second question is that other properties of the urban physical space, beyond or in combination with street network configuration, might also have an important role in the spatial distribution of presence and absence of retail.

***(3) Can specific urban morphological characteristics be linked to the spatial distribution of retail agglomeration types? If yes, which characteristics play the most important role?***

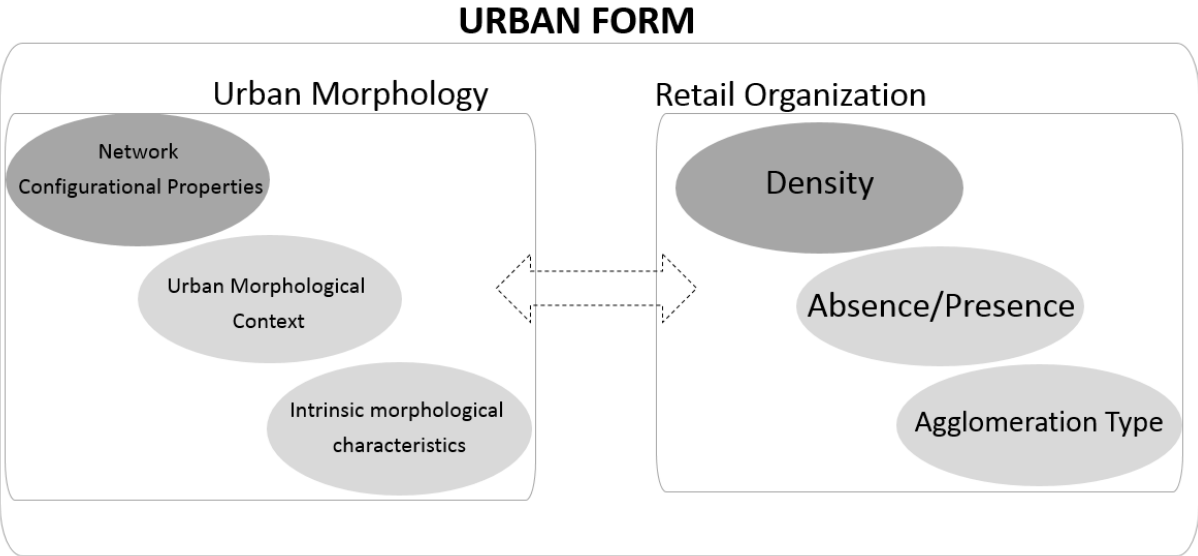
We would like to put forward a working hypothesis for this research question: The properties of the urban physical space (not limited to the street network configuration, if the previous hypotheses hold) might also have an important role in favouring the presence of specific types of retail agglomeration. In other words, we suggest that urban morphological characteristics might not only influence retail presence/absence and density, but also the typology of its functional agglomeration.

In order to answer these questions, it becomes clear how bridging concepts from retail geography and urban morphology should also give birth to common analytical approaches. For this reason, we decided to focus our attention on the street and on how retail and urban morphology can be characterized from the street point of view.

This renovated methodological approach responds to the recent suggestion of researchers to focus on the street as the spatial element of the analysis which we find both in geography and urban design/morphology studies (Gourdon 2001, Fleury 2004). Moreover, the analysis of retail and of the urban form from the street point of view might allow, in future research, to further integrate a third important aspect in the analysis of the relationship between retail and urban form, which is the socio/behavioural component.

Consequently, together with the aforementioned thematic questions, further methodological challenges to be tackled are discussed in this dissertation:

- (A) How can morpho-functional typologies of retail activities be automatically identified and described from the street point of view in a large metropolitan area?
- (B) How can we provide a quantitative description of the built environment along a street in a large metropolitan area? How can urban fabrics and morphological regions be automatically identified and characterised from the street point of view?
- (C) What approaches are most appropriate to tackle model and variable selection in studying the relationship between urban form and retail distribution?



*Figure 16 The study of the relationship between the retail spatial organisation and urban morphology: while the Movement Economy Theory (Hillier 1999) focuses on the relation between network configuration and retail density (dark grey), in this dissertation we also consider additional descriptors of both retail and urban morphological system (light grey)*

The next chapter will let the reader appreciate how these three research questions take all their relevance in the case study of our dissertation, the French Riviera metropolitan area.

## 3. Study Area and Data

### 3.1 Urban and economic context

The urban space of the French Riviera is the study area under investigation in this dissertation. The name of this coastal region, in southern France, has been attributed by the English aristocracy of the XIX<sup>th</sup> century, as an extension of the name Riviera traditionally associated with the neighbouring Italian coast. This area has also been known since 1887 as La Côte d'Azur, from Stephane Liegeard book's title where the author describes the coastal landscape from Marseilles to Genoa.

Despite representing a worldwide renewed touristic attraction since the period of "*La belle époque*", the French Riviera is not defined by specific geographical limits. While the national border on its eastern side is in a general agreement, the western limit is more ambiguous and it traditionally varies between Frejus or Cassis where a net change of the natural landscape (Calanques National Park) defines the beginning of the agglomeration of Marseille, and its Côte Bleu.

In our study, we focus on the most central part of the French Riviera: the coastal strip enclosed between the geomorphological regions of the maritime Alps (East) and the natural park of the Esterel Massif (West). Similarly, the northern part of this region is delimited by important protected natural landscapes: the Plateaux de Calern et Caussols, the Gorges de Mescla et Vesubie and the National park of Mercantour. Surrounded by a complex morphological space, the urban settlements of the French Riviera are limited on a marginal, linear and narrow strip between the mountains and the sea.

Within this corridor of about 60x30 km, the elevation variability is still relevant, particularly in the eastern side of Nice where it can reach 1000 metres within the first 20 km: villages such as Eze or La Turbie are located in less than 500 metres from the coastline and at about 700 metres above the sea level. From East to West, we observe a gradual transformation of the landscape from the Alpine to the Provençal regions with a gradual decrease of the slopes.

Our analysis considers the 88 municipalities of the littoral and sublittoral part of the Alpes Maritimes Department, within the southern French region Provence Alpes Côte d'Azur (PACA); completing the coastline, the independent Principality of Monaco.

Six main centres define the urban polarities of the Cote d'Azur. From West to East we find: Cannes-Grasse-Antibes conurbation counting respectively 74.2, 51 and 73.8 thousand inhabitants in their central cities<sup>35</sup>. Nice, with its 343 thousand inhabitants represents the largest municipality of the French Riviera and the administrative centre of the department. The enclave of Monaco and the border city of Menton have respectively 38 and 28 thousand inhabitants. Spread around these main centres, 295 thousand people find their home in

---

<sup>35</sup> In 2018, the city of Cannes has become the second most populated city of the French Riviera overtaking Antibes which had been the most populous city throughout the last fifty years (source:INSEE)

smaller cities, villages and hamlets. With a total of more than 1 million inhabitants, this places the French Riviera at the seventh most populated conurbation areas in France.

With the exclusion of Grasse, the other five centralities are located along the coastline, with the other 11 municipalities gathering 65% of the overall departmental population. When including the sublittoral band (*proche arrière-pays*) 33% of the departmental surface encompasses almost 97% of the population. A strong heterogeneity of the settlement distribution explains the exclusion from our analysis of the mountainous countryside (*haut arrière-pays*).

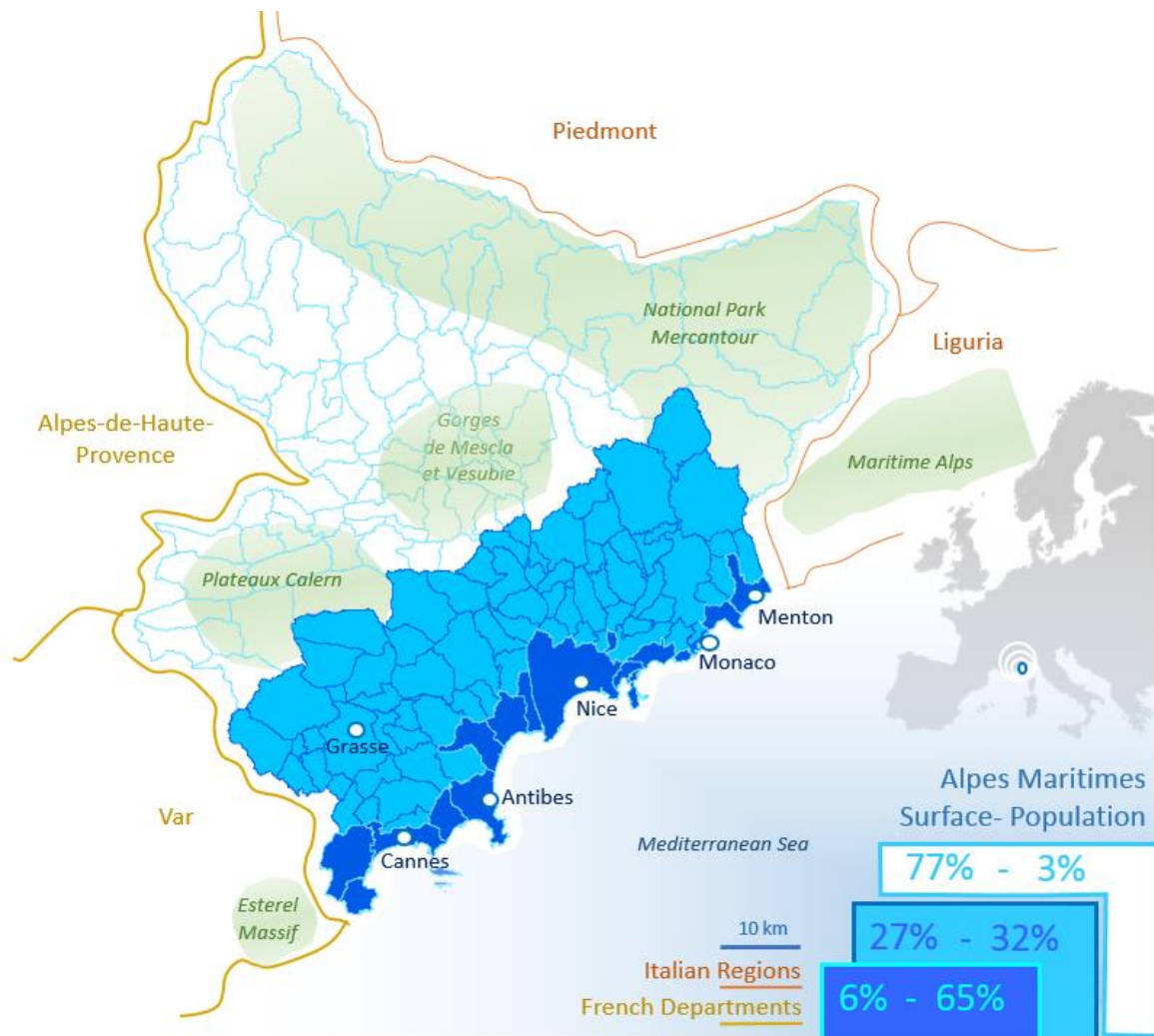


Figure 17 The administrative and natural context surrounding the Alpes Maritimes Department and the repartition of the population in the littoral, sublittoral and inland sectors.

To better understand the peculiarity of this space and the actual organisation of the settlement, it is fundamental to understand the evolution of the socio-economic system of this peculiar space.

While some archaeological finds show a human presence in this region since the Palaeolithic period, evidence of the first settlements of this region date back to between the VI and III B.C. After the foundation of Marseilles by the *Phoceans*, a Greek population coming from a region nearby the actual Izmir (Turkey), a network of smaller settlements was created along the coast



connecting the new colonies to the eastern Mediterranean basin through Corsica. *Antipolis*, *Nikaia*, *Monoikos*, the actual Antibes, Nice and Monaco are chosen for their strategic site morphology. A few centuries later, the Romans arrive and take control of the region, coming from the landside; new settlements are organised along the *Via Julia* and the *Via Aurelia*, two important routes connecting Genoa to Aix-en Provence and Arles for the control of the northern regions of the Empire. New settlements are founded as *Forum Julii* (Frejus) and *Cemeneleum* (the actual hill of Cimiez in Nice). Julio Cesar established the Varnum Flumen as the natural delimitation between the two regions of the *Gallia Cisalpina* and the *Gallia Narbonensis* (North Italy and South France, Franco 1987); between these two regions, Augustus defines in 13a.c. the *Alpae Maritimae* military district, corresponding roughly to the actual department of the Alpes Maritimes.

After the fall of the Roman Empire, until the XIX century, this region is characterised by a history of borders. In different ages, the western side is inhabited by Provençal populations, the Provençal County, the Burgundy Kingdom, the French Kingdom and the French Empire; while on its eastern side are the Italian Kingdom, the Genoa Republic, the Savoy Duchy and the Kingdom of Piedmont and Sardinia. In the middle we find the contended Nice County and the Principality of Monaco.

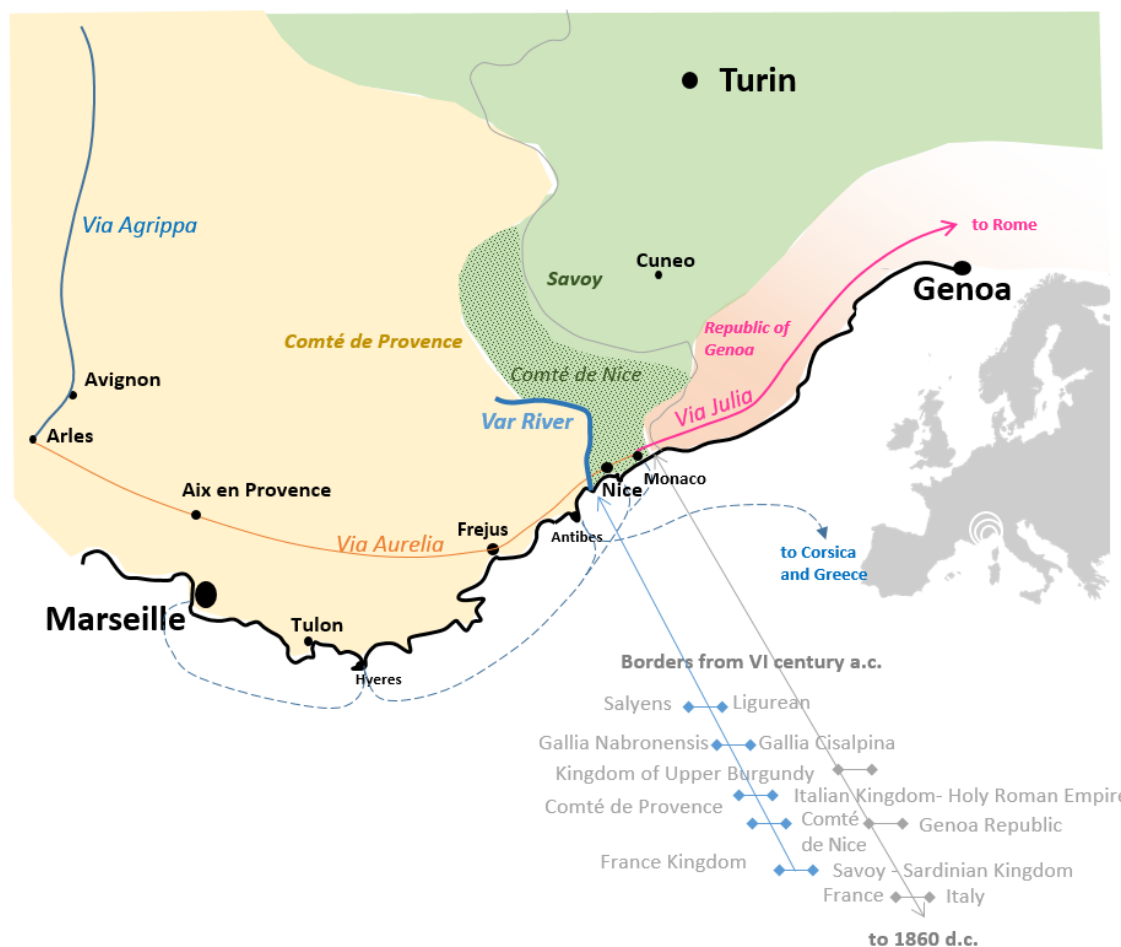


Figure 18 The French Riviera along the history: a region of borders and passageway

A region of natural and political borders, but also a region of passageway: while the ports represented the intermediary stops between Marseille and Corsica on the way to Greece, the Roman routes were connecting northern provinces of the Roman Empire. These same routes, during the medieval period, became a corridor for pilgrims heading to the Vatican; several monasteries and convents are distributed along the coast. For instance, the village of Saint-Laurent du Var finds its origin in a monastery succouring and assisting pilgrims in crossing the Var river. Between the XVII and the XVIII centuries, the same path has been crossed by intellectuals and aristocrats from northern Europe on their Grand Tour<sup>36</sup>, searching for the roots of western civilisation of the ancient world.

Beyond the physical characteristics of the coastline, it was an essential carrefour of European connections, yet underwent a scarcity of natural resources, its morphology and the unstable political conditions, hindered for centuries its economic and demographic development. Settlements remained limited for centuries in their protective walls of military, religious, and commercial centres.

The urban development of the French Riviera is a relatively recent phenomenon (Dauphiné 1990, Boyer 2002). The fragmentation and constant fluctuation of political delimitations of this coastal region finally reached its stability only in 1860 with the annexation to France of Nice County and of the municipalities of Roquebrune and Menton. For the first time in history, this space can benefit from a uniform and organically connected environment (Fusco 2003). However, the factor which triggered the economic and demographic development of this region is the construction of the railway lines: in the second half of the XIX century, the small littoral settlements of the French Riviera are finally connected to both the French and Italian national railway systems. The accessibility to a region of exceptional natural landscapes and climatic conditions make these spaces a touristic attraction for the European elite, stimulating a first urban extension outside of the historical centres: along the railway lines, those spaces with a favourable morphological condition allowed for the first planned, dense and compact urban growth.

After the first World War, important demographic transformations stimulated the urban expansion: the rural exodus from the rural countryside (Dauphiné 1990), the Italian and Russian immigration<sup>37</sup> (Ellis 1988), the repatriation of the Pied-Noirs (between the '50s and the '60s), and the demographic explosion of the Trente Gleurieuse (1945-1973).

This rapid demographic growth, combined with the proliferation of the car, had a strong influence on the urban development of the French Riviera. While at the end of the XIX century the first urban growth was characterised by planned developments mainly focused on those planar spaces along the railway line, the second urban expansion is defined by an unplanned process, spread on hilly areas of the peripheries of the city centres. Moreover, the inauguration of the motorway in 1960 (Autoroute A8) extended the dispersed single-houses urban fabric in the spaces between cities and villages of the sublittoral spaces (moyen pays), especially on the less sloped areas at the western side of the Var river (Fusco 2003).

---

<sup>36</sup> In 1766 the Scottish Poet Tobias Smolett described Nice in his "Travels through France and Italy".

<sup>37</sup> In 1936 Nice population counts 20% Italians and 18% Russian (Ellis 1988)

Between 1975 and 1990, the French Riviera gains 250 thousand inhabitants (+19% vs. +7% of the national scale). This is an unbalanced demographic growth within the sector Cannes-Grasse Antibes (+30,5%), as a consequence of a new urban project, the technology park of Sophia-Antipolis (see Text-Box 3, Section 9.3).

The urban growth and the higher individual mobility participate together in a process of spatial welding of the coastal settlements. While in 1975, 90% of the population were working on the same municipality of residence, between 1975 and 1990, the inhabitants start to ignore their municipality limits in their travelling behaviours, representing the emergence of one only coastal conurbation (Decoupigny and Fusco 2009, Fusco and Scarella 2013). As a consequence of this increased mobility, there flows out a new wave of urban expansion, particularly in the Grasse and Nice outskirts, where we find a prevalence of recently constructed villas and residential complexes (after the '90s).

Over 150 years, the French Riviera transformed its urban landscape: from scattered cities and villages to an almost continuous coastal conurbation as a consequence of an extraordinary population growth (+133% in the period 1861-1946 and +114.5% between 1947 and 1995, compared to a national level of, respectively, 8.5% and 39.6%).

The exceptional natural landscape and the temperate weather represent the main attractiveness of the French Riviera; since the *Belle Epoque*, tourism represents the primary economic resource. From an elite phenomenon<sup>38</sup> at the end of the XIX century to the 11 million of tourist of 2018, the French Riviera became the second French destination after Paris. Today tourism on the French Riviera represents 15% of its economy (compared to 7,4% at the national level) and 1% of the global touristic economy.

The development of infrastructure stimulated a continuous growth of attractiveness at the national and international scale, bringing to the French Riviera substantial exogenous investments for infrastructures and stimulating the local economy (Dagorne 1995). After almost fifty years of the arrival of the railway system, two airports were built in Nice (1929) and Cannes-Mandelieu (1937), both at less than six kilometres from the city centres.

Nice airport represents today's main international gateway of the French Riviera which (literally) transformed the coastline. Almost 14 million passengers make this airport the third most crowded of France (after those of Paris), and this number is estimated to reach 18 million with the recent extension of its Terminal 2. The proximity to the city centre and the recent urban expansion includes this infrastructure in the continuum of the metropolitan area. A new urban plan (see Text-Box 3, Section 9.3) is reconsidering its strategic position within a more integrated urban transport system; before the end of 2019, the airport will be connected with two new tramway lines to the logistic platform at the north of the Var valley and to the east with Port Lympia, the port of the city of Nice.

---

<sup>38</sup> In 1869, the historian Brunel wrote "The city of Nice became, with Naples, Rome and Florence one of the salon d'hiver of the European high society. The Russian colony is full of life and entertainment: [] they organise balls, soirées, garden-parties with plain air dances, in Nice as well as in the villas bought or rented in the surrounding areas"

Similarly, the port system modified the original coastline and today represents an important stimulus for the French Riviera Economy: 4.299 km<sup>2</sup> of ports organised in twenty-nine maritime and two fluvial centres, with an additional two ports in Monaco. Two main functions are associated with the French Riviera port system: Marinas-Yachting (with a total of 16 thousand yacht-parking lots) and Cruise-Ferrying. The Blue Economy is at the origin of 2,2 thousand commercial activities and generates about 2,3 billion euros of revenue. Cruisers and ferries count 900 thousand passengers yearly, connecting the ports of the French Riviera to a more extensive system of 298 ports spread across the coastal arch between Marseille and Tuscany (including Corsica and Sardinia). Ports are essential not only for their economic relevance but also for their architectural and urbanistic point of view. The genesis of littoral cities, towns, and fisherman villages found in ports the structuring element of settlements and still today they represent a stimulus for the local economy and urban planning such as the pharaonic extensions toward the sea in Monaco.

The important abundance of tourists might find in the French Riviera a wide offer of entertainment activities. Events and festivals of national and international importance are organised along the French Riviera and wisely distributed throughout the year to partially reduce the traditional seasonal fluctuation of coastal cities and improve the overall attractiveness. The most known examples are the Film Festival of Cannes (since 1946), the Monaco Grand Prix (since 1929), and the Nice Carnival (since 1875). This last event, organised in the same period of the Fete du Citron of Menton and Mimosa in Mandelieu, in 2018 attracted 300 thousand tourists, bringing about 30 million euros of revenue for the local commerce. Other international rendezvous are the Yacht Show in Monaco and the MIPIM in Cannes, an international real estate property industry event. Furthermore, on offer are a large number of cultural and leisure activities, including one hundred museums (art, botanic, history and even perfume) and 15 ski stations. Despite this continuous effervescence, a large variability in tourist affluence is still observed: from a minimum of 50 thousand in December and a peak of almost 700 thousand the 15 of August (65% of the department population).

Tourism is not essential only for the economic and social vivacity of the French Riviera, but it also responsible for the growth of these urban spaces. 25% of the housing stock is estimated to be secondary homes (2005), of which 40/50 thousand units are owned by an international population (50% Italians, 13,5% UK, 10% Scandinavian, etc.); moreover, 50% of the touristic arrivals are vehicle-based (highway A8). These numbers highlight how tourism plays a significant role in shaping the urban settlements as well as placing certain pressures upon the French Riviera; this is seen as a high demand for secondary housing (and relative infrastructure) characterised by a non-continuous use throughout the year. The recent phenomena of sharing economy, which is particularly affecting touristic coastal destinations (i.e. in Barcelona in Gutierrez et al. 2017, etc.), is also growing on the French Riviera. These new practices are stimulating the real estate market while they are perceived as a threat from the traditional hotel systems; however, official numbers are not yet available<sup>39</sup>.

The prevalence of secondary homes, a growth observed after the second war, is associated with another peculiarity of this region: the higher concentrations of senior citizens. Owners at

---

<sup>39</sup> Further research about this topic is under study both at the local CCINCA and at the laboratory ESPACE.

their retirement age have been deciding to make their secondary homes become their primary habitation, raising the average age of the French Riviera. In 1994, 28% of the population was more than 60 years old, compared to 16% at the national scale. Today we observe 30% at the department scale (24% at the national level), reaching values higher than 35% in specific urban sectors, (i.e. centre of Antibes, Cannes and Menton), and some peaks of 50% (eastern hills of Cannes). The demographic projections for the department show a growing trend of senior proportions (the population of 75+ is expected to double in 2030), and a consequent reduction of proportions of younger generations. The prevalence of seniors is detected not only for residents but also among tourists (an estimated average of 42 years old in 2017), which is correlated to longer permanence and with higher economical expenditure.

This characteristic has a strong influence on the local economy of the French Riviera: reduced mobility requires shorter distances between residential areas and retail distribution, a higher presence of health services (doctors, pharmacies etc.), a higher frequentation of shopping centres and higher inertia to new online retail formats. Elderlies are recognised to represent a new valuable resource for a local economy of a society, and recent studies on the *silver economy*<sup>40</sup> have been stimulating research at the local, national and European level. Leisure and health activities are the main activities and services related to the silver economy.

Health-related activities and environmental sustainability represent the two central notions at the centre of the design of a new urban development project located along the Var Valley, currently under development. The Eco-Valley project is a project labelled in 2008 by the French government as an Operation of National Interest (“Opération d’Intérêt National”), covering an area of 10 thousand hectares which is transforming this peripheral area from an urban fringe to a new urban centrality. This project is expected to become a national and international model of sustainable development; it represents only the latest of a series of large urban development projects that since the ‘70s have been planned and realised along the French Riviera, aiming to develop the technology industry with innovative research centres and industrial parks. With more than 2 thousand enterprises and start-ups employing more than 36 thousand people, Sophia-Antipolis represents one of five largest French technology parks where national and international firms work in collaboration with research and academic institutions. Further details about these specific areas and their urban morphological characteristics are further discussed in Section 9.3.

In this section, we briefly described the evolution of the economic and demographic forces which have shaped the urban settlements of the French Riviera. A highly complex and dynamic region whose development has been caused by and made for tourism and technology activities (with a recent growing interest to Blue, Silver and Green Economy). Such forces have shaped a large variety of different urban fabrics coexisting in the same highly constrained

---

<sup>40</sup> “Oxford Economics definition defines Silver Economy as the sum of all economic activity that serve the needs of people aged 50 and over, including the products and services they purchase directly and the further economic activity this spending generates. Thus Silver Economy encompasses a unique cross-section of economic activities related to production, consumption and trade of goods and services relevant for older people, both public and private, and including direct and indirect effects.” (EU final report on Silver Economy 2018 ) <https://ec.europa.eu/digital-single-market/en/news/silver-economy-study-how-stimulate-economy-hundreds-millions-euros-year>

coastal conurbation. A dual nature of the French Riviera settlement can be described as both as a continuous concrete wall (Dagorne 1995) and as a large mosaic of different urban spaces (Fusco 2003).

The main characteristics of the structure of our study area might be summarised as follows:

- 1) The municipality of Nice represents the largest centre of the French Riviera conurbation, including several internal polarities disposed on the more planar spaces and residential neighbourhoods in the more sloped areas. On the left bank of the Var River, Saint Laurent du Var and Cagnes-sur-mer complete the metropolitan area of Nice. The Var Valley is the subject of undergoing urban transformation, transforming this area into a new metropolitan polarity.
- 2) The cities of Cannes, Antibes, Monaco and Menton, aligned along the coast, represent the main centralities from which the touristic and residential development of the French Riviera blossom. In the smaller cities of Grasse and Vence, located in the western inland sector, have mainly the centre of residential developments.
- 3) Villages and hamlets are spread around larger polarities, affirming the hierarchical description of the centres. Their historical centres (dating back to the medieval period) are surrounded by vast residential areas, according to the morphological properties of the site.
- 4) New industrial areas and technological parks have been stimulating the development of particular forms of settlement, as in the case of Sophia Antipolis, Carros- Gattières, Cannes-La-Bocca and more recently the lower Var Valley.
- 5) All these differently-sized centres are interconnected by a pervasive discontinuous and car-dependent residential urban fabric (Fusco 2003, 2006)

In the scientific literature, several works further describe the natural, historical and demographic characterisation of this particular region, as well as its economy and transportation system. Nonetheless, we might observe the absence of studies focusing on the urban morphological characteristics in more detail.

The city of Nice represents the only exception: the works of Philippe Graff (2010, 2013) provide an outstanding detailed description of the architectural and urban evolution of the city. These two important references allow us to verify and interpret the results of our morphological analysis. Nice has also been the object of quantitative urban morphological analysis in the doctoral research of Marie Sevenet (2013): in this work, the author elaborates measures of volume and surface density with the final goal being to investigate the environmental and social sustainability evolution. In Part 4, the works of Graff and Sevenet will be further discussed.

## 3.2 The retail system

While in the previous section we focus on the overall characteristics and socioeconomic forces of our study area, let us now focus on the descriptions of the retailing sector.<sup>41</sup>

Retail on the French Riviera was responsible in 2016 for 10% of the working force (-1% from 2005) and about 26% of the departmental economy<sup>42</sup> (+1.5% from 2005) with an estimated revenue of 15.8 billion Euros. The total retail surface of the French Riviera includes 1,6 million m<sup>2</sup>, which corresponds to a density per capita of 1.4 m<sup>2</sup>/hab., higher than the 1,2 m<sup>2</sup>/hab. observed at the national scale. If we add to the local population the number of potential customers, which is estimated in 225 thousand tourists (in equivalent full time<sup>43</sup>), the density is closer to the national level with a value of 1.26 m<sup>2</sup>/hab.

Considering the conventional separation between small and average-large stores (<>300 m<sup>2</sup>), the former group accounts for about 13 thousand units over a surface of 487 920 m<sup>2</sup>, while the latter includes 971 stores for a total of 1 036 066 m<sup>2</sup>. Among them, 35 stores exceed 2000 m<sup>2</sup>.

Within the period between 2005-2016, the retail sector registered a growth of +4% for the small surfaces and 13% (20% in terms of surfaces) for the >300 m<sup>2</sup> stores (including the global economic recession of 2008). Seventy per cent of the commercial surface is found in 8 municipalities (Nice, Antibes, Cannes, Grasse, Mandelieu, Villeneuve-Loubet, Cagnes-sur-Mer et Saint-Laurent-du-Var), seven of which are disposed along the coast. The most widespread retail sectors, in order of prevalence, are: furniture, food, vehicles, clothes and lastly (but with the highest growth in the period 2005-2015) leisure activities.

From the spatial distribution in Fig.20, we highlight three main patterns: firstly, the western side of the French Riviera offers a higher presence of commercial surfaces compared to the eastern side. Secondly, the distribution of retail surfaces is spatially correlated to two aspects: the size of the compact urban centres and the access to the highway. This second aspect is at the origin of the dual nature of retail centralities in larger cities: Antibes and Cannes might be recognised from their urban centre along the coast and the peripheral polarity along highway access points. Similarly, Nice shows a large central polarity and three further poles located to its eastern and western highway access points. Thirdly, we might notice retail surfaces distributed along three ramifications of the departmental routes, connecting the highway with the cities of Grasse, Vence and the upper Var Valley.

---

<sup>41</sup> All the information provided in the following sections have been gathered from several documents from the local Chamber of Commerce CCINCA; More specifically, we refer to the following five reports: *Panorama de l'équipement commercial des Alpes-Maritimes* (2016), *Politique d'urbanisme Commercial* (2015, internal document), ADAAM (2012, internal document), *Schema Departemental d'Equipement Commercial des Alpes-Maritimes* (2005), *Document Intermédiaire de concertation* (1994, not public). Principality of Monaco is not discussed in these documents.

<sup>42</sup> 6,4% at the national scale (INSEE).

<sup>43</sup> Equivalent full time measures consider the number of nights spent by tourists hosted by the hotel sector (1 equivalent full time = 365 tourist nights) to which is added a coefficient counting for secondary home owners and friends/relatives hosted by local residents (Tourism Observatory).

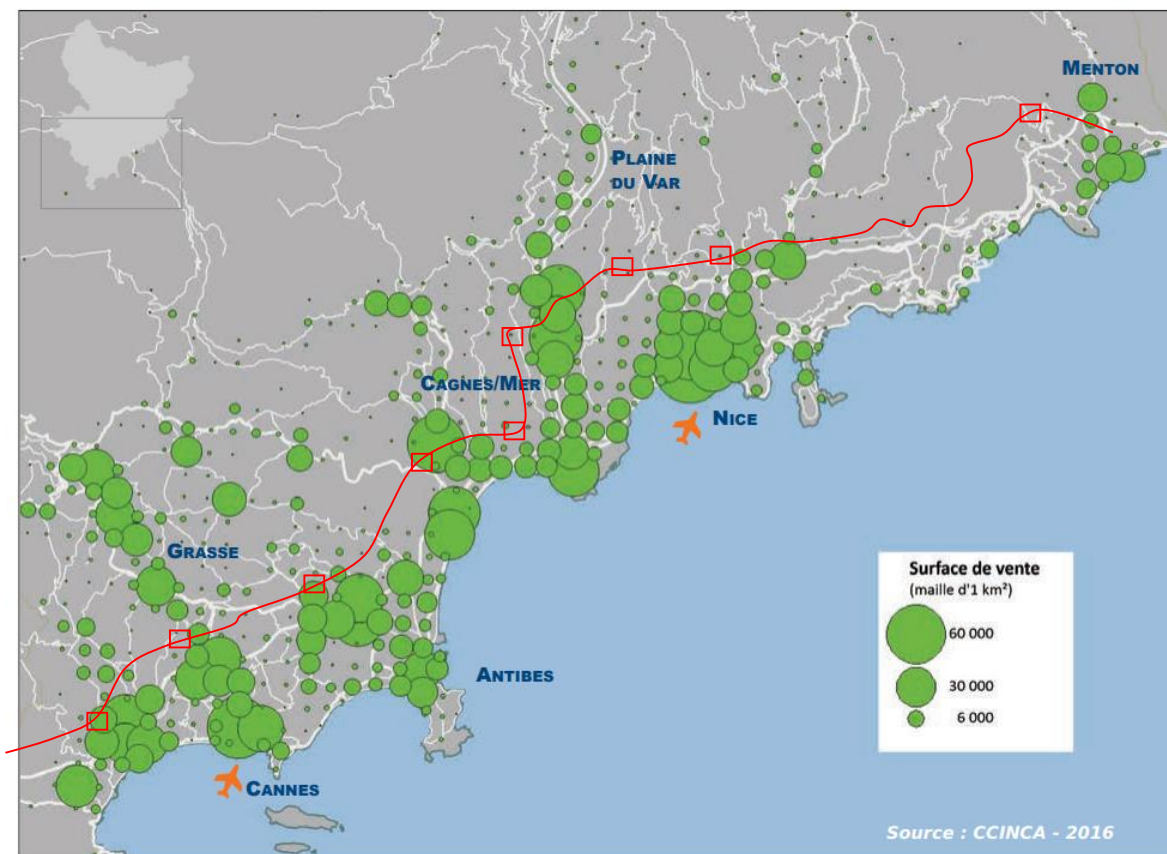


Figure 19 Repartition of the retail surface in our study area; in red, the national highway and its accesses (elaboration from CCINCA 2016).

These observations on the spatial organisation of retail are highly influenced by the unit of measure traditionally implemented in documents of the local administrations: the retail surface. Nonetheless, this information is precisely known only for average and large-sized retailers (>300m<sup>2</sup>) while smaller surfaces are approximated to 40 m<sup>2</sup>; this spatial representation might therefore lead to ambiguous interpretations. When observing the maps in Fig.19 we should remember that: i) 1% of department stores offer more than 64% of the commercial surface and ii) more than half of the retail economy is shared by few large national and international retail companies. Considering this information, the spatial patterns of Fig.19 might be probably associated with the repartition of the population expenditure or other economic variables, more than the social and economic attractiveness of urban spaces.

The extensive documentation produced in the last fifty years by the local Chamber of Commerce (among which some that are referenced in this section) proposes many reports about the retail surface distribution with specific attention to larger retail formats. In order to understand the reasons behind this uneven attention, we should remember that in France the “Commercial Urbanism” (Urbanisme Commerciale) originated in the late ‘60s with the specific intention of national and local administrations to: i) promote the development of commercial services within new peripheral urban areas (and more specifically collective housing and large complexes, Desse 2013); ii) regulate the opening of the large retail surfaces.

Starting from 1969, a series of national regulations have been guiding the expansion and distribution of large surfaces for almost fifty years (Desse 2013). During the ‘80s, with the



decentralisation of the French government, the control of these large surfaces was transferred to local administrations. Consequently, the peripheral urban areas were affected by an important growth of large surfaces. More specifically, non-alimentary surfaces opened in the proximity hypermarkets (Péron 2004), intensifying and diversifying the retail offer around these peripheral centralities; the easier vehicle accessibility and the presence of large car parking areas attracted a growing volume of costumers beyond the local scale and threatened the more traditional urban centralities. In the '90s, the effects of the decentralisation became more visible, requiring the elaboration of immediate intervention with new planning strategies and models for the retail organisation.

This same process has also affected the French Riviera; local authorities, such as the local Chamber of Commerce denounce a growing number of large retailing activities in the interstitial spaces between urban settlements, undermining the retail fabric of the city centres. As stated in the diffuse urban milieu, *“the retail urbanism becomes the result of the juxtaposition of a multitude of individual decisions frequently concentrated in the proximity of highway accesses, without a strategical vision and causing irreparable errors”* (Diagnostic commercial 1994). In this same document, the responsibility of this dysfunction is associated with the absence of cooperation and consultation among municipalities, engendering competition between projects and between hosting municipalities.

A specific concern of local authorities is about the low aesthetic and qualitative characteristics of these built-up spaces (*“architecture approximative [...] indigne de la Côte d’Azur”*), positioned at the entrance of the cities and affecting their attractiveness (Diagnostic commercial 1994, SDEC 2005). The absence of coordinated planning of large retail surfaces on the French Riviera is under scrutiny, and a stronger harmonisation and consulting process is being demanded in order to operate a selective control on the projects. In 2000, a new modification in the urban planning regulation (loi SRU) proposes a new tool *“Schémas de cohérence territoriale”* (Scot), which considers the development of the urban planning at the intercommunal scale over a period of ten years, and plans to coordinate the decision at the municipal level with a common vision (Local Urban Plan-PLU). Within the last ten years, only two of the five SCOT of the French Riviera<sup>44</sup> have been finally approved. More recent evolutions of the French and European legislative framework are widely discussed by the French retail geography community (Desse 2013; Madry 2016, 2018; etc.)

The political and administrative reaction to changes at the turning of the century were not only focused on the regulation, renovation, and reorganisation of these peripheral areas. The national debate and the following activities of local administrations also considered and intervened upon the urban development of more central areas for the revitalisation of the cores of cities, and villages.

In the last twenty years, the centres of larger cities on the French Riviera have been the subject of urban renovation projects in order to improve urban quality and accessibility. Anchor stores

---

<sup>44</sup> In the Alpes Maritimes department five SCOT are thus defined: SCOT Ouest for the western region of Cannes; CASA for the municipalities of the Antibes basin; (from west to east) Nice Côte d’Azur; Pays des Paillons and Riviera Française et Roya corresponding relatively to the littoral and sublittoral bands of the eastern region of the department.

have been strategically positioned within these spaces to stimulate their attractiveness and competitiveness. Large firms participated in the "reconquest" of the city centre with a wide range of retail formats responding to the increasing need of proximity: food stores have developed a store network with several smaller surfaces, as well as widening opening hours and eat-in services. Moreover, the omnichannel strategy of these national food firms has seen the development of *shopping drives* services both in the centre and in the outskirts of larger cities.

An important role in the orientation of the urban renovation project has been played by "Artisan and Retailer Associations": the number of associations has seen an important growth in the last fifteen years, and today 79 associations are distributed in the study area. Twenty-three of them gather retailers of five coastal municipalities at the eastern side of the French Riviera (such as Menton, Beaulieu-sur-Mer, Saint Jean Cap Ferrat, etc.) and eighteen sub-littoral municipalities (such as Vence, Tourrette-sur-Loup, La Gaude, Sospel, etc.). Fifty-six intra-urban organisations gather retailers of specific high streets and neighbourhoods in Nice (22), Cagnes-sur-Mer and Saint Laurent-du-Var (14), Antibes (2), Cannes (6) and Grasse (1).

The creation of these associations might be seen as a collective response of shopkeepers to a common local issue (at the district scale), such as the modification of the local urban environment (i.e. modification in the circulation, the construction of the new tramway line, etc.) or to a more general loss in attractiveness. The perimeter of these associations does not always describe a coherent morpho-functional agglomeration (as described in the previous sections), but they rather correspond to the spatial extent of those evolving urban subspaces (evolution towards a requalification or degradation process).

The renovated attention to peripheral and central polarities has consequently stimulated the rarefaction of the commercial presence in intermediary spaces. The commercial pauperisation affecting large urban districts, such as semi-peripheral residential neighbourhoods and social housing estates, might partially find an explanation in their urban morphological characteristics. The scarcity of commercial areas in these spaces forces inhabitants to increase their dependence on motorised transport.

We might consider another peculiarity of the retail system of our study area: retail vacancy is affecting urban spaces at different rates. In 2016, a specific national report of the Ministry of Territory Cohesion<sup>45</sup> analysed French cities (between 10 and 100 thousand inhabitants), with the purpose of identifying and describing those spaces where the vacancy evolved at higher rates. Based on this study, specific financial means have been deployed to help local authorities inverting the vacancy trend (Action Coeur de Ville). Among the cities with a shrinking retail presence, two cases have been identified in the French Riviera: Grasse and Vallauris. Following this report, the local Chamber of Commerce organised a new activity of observation and georeferencing of retail vacancy starting from the most vulnerable spaces in the inland municipalities. This dataset is still under development, and in the next few months, we will have access to a first complete picture of the retail vacancy in the department. Nonetheless, first estimations on partial data show that, despite a general and constant

---

<sup>45</sup> La revitalisation commerciale des centres-villes  
([https://www.economie.gouv.fr/files/files/PDF/Rapport\\_RevitalisationcentresvillesVdef\\_octobre2016.pdf](https://www.economie.gouv.fr/files/files/PDF/Rapport_RevitalisationcentresvillesVdef_octobre2016.pdf))

growth of the retail activities at the department level (+1%), an average retail vacancy of 9% might be observed (5% in the hyper-centres, and 25% in semi-peripheral sectors).

The brief description of the evolution of the retail sector of the French Riviera here presented is of paramount importance to understand its actual configuration. Two considerations should be highlighted.

The first is that administrative documents about retail distribution are characterized by an unbalanced preoccupation which privileges the analysis of larger surfaces and the development of planned centralities; undoubtedly, these urban projects might profoundly impact the equilibrium of entire cities and their realisation involve important economic resources from different stakeholders. Nonetheless, large surfaces, especially those in the peripheral areas, do not contribute to the vivacity and attractiveness of public open spaces. These aspects are traditionally associated with the presence of smaller stores and their disposition in public spaces, streets, and squares. This dissertation, as previously discussed, gives a larger attention to small and medium-sized stores and services and how they combine in different urban areas. As previously stated, behind the study of retail distribution in the city, we mask the broader goal of understanding the different typology of social life which we might observe in urban spaces. The Local Chamber of Commerce explicitly funded this research activity with the specific goal to provide an alternative analysis of the retail system going beyond their traditional approach and to focus on smaller retail surfaces, which are often overlooked.

The second observation is about the "inertia of the retail system": we should consider that a temporal lag elapses between the moment in which a criticism/trend is observed, the reaction of administrative and economic organisations, and the implementation of possible solutions. This lag might take up to several years. Since the data used in this dissertation is relative to 2016-17, it is crucial to consider that our outcomes and observations should be evaluated considering the ongoing process of the retail system in the period previous to this specific year. The analysis proposed in this dissertation does not respond to the questions about today's trends or future evolutions of the retail system; our focus is on the analysis of the relationship between urban morphology and retail distribution at a given point in time. Nonetheless, in the conclusion part of this dissertation, we describe the current evolution retail sector observed in our study area and how the methodological developments presented in this work might be deployed as a tool for the detection of temporal changes and prospective analysis for the retail system.

We might summarise the information about the retail system as follows:

- 1) Despite negative global trends observed in North American and UK contexts and the catastrophic scenarios described by the recent economic literature, the retail sector of the French Riviera shows a positive evolution with an uninterrupted and continuous growth of both number and surface of retail activities in the last decades.
- 2) Larger settlements including Nice, Cannes, Antibes and Menton gather important retail concentrations in both central and peripheral urban areas (in the proximity of the roadway). In these coastal municipalities more than 70% of the retail activities are concentrated.

3) About 20% of the retail offer is distributed among smaller municipalities in the western sublittoral sector, while villages of the eastern area only account for a limited offer of small surfaces (large and small retailers account for about, respectively, 2% and 3% of the total departmental number)

4) The positive trends at the departmental level mask an internal spatial heterogeneity of the retail system health: while larger coastal cities show a strong and complex retail organisation, the sub-littoral spaces have a more fragile environment witnessed by the recent governmental measures in Grasse and Vallauris (in the western sector) and by the creation of associations gathering retailers of smaller municipalities.

5) The main sources provided by local authorities have a specific (intentional) focus on large retail surfaces and planned retail centres, required by the specific local and national regulative framework of the last fifty years. Smaller retail activities are overlooked, and spatial representation at coarse levels are not able to provide a fine-grain description of the small retail agglomerations.

### 3.3 The scientific interest of the French Riviera case study

Several important characteristics described in the previous sections about both settlement and retail make the French Riviera an interesting subject of research. Considering the scientific context and goals of this dissertation, we might highlight those peculiarities of our study area which are expected to provide an alternative and complementary perspective to those works listed in Table 1-2.

- Differently from traditional studies (of both urban morphology and retail geography) where large cities with a monocentric pattern of urban intensity, our study area covers a wide conurbation originated by the welding process of several coastal cities which generates a continuous **polycentric urban settlement**.
- Cities of different size and villages are here analysed with a common methodological framework, allowing for a comparative study between **settlements of different sizes**.
- This large area allows investigating a coherent urban space where strong **differences in retail strength conditions** are observed: while retail in coastal cities benefits from a growing tourism, other realities have been recently described as having a more fragile system.
- The study area is of a **strong urban morphological heterogeneity**, which includes the following: historical medieval centres, planned and regular compact of the XIX century, new towns, urban fringes, sprawl of buildings and individual homes and large natural areas.

## 3.4 Data

Two sources of data have been used in this dissertation. Retailer information has been provided by the local Chamber of Commerce of Nice Cote-d'Azur (CCINCA), while from the National Institute of Geographical and Forest Information (IGN, *Institut national de l'information géographique et forestière*) we used several geographic databases (mainly BD TOPO®, 2017) to describe urban form.

While IGN data is freely available for every French research institution, access to official retail data was possible only through specific bilateral agreements. Nonetheless, starting from 2017, this information is now publicly available thanks to the implementation of the European policies about open data accessibility.

The information about retail has been always recorded by the CCINCA to inventory and analyse the economic health and evolution of the French Riviera; this information informs which economic measures and indicators are elaborated with the ultimate goal to support and guide the CCINCA strategies and actions. Starting from 1995 with the development of the Informatics System Department, an informatics database has been created, allowing for a systematic and precise account of this valuable information. While this database was initially created for administrative purposes and economic studies, GIS technologies have more recently extended its possible applications; cartographic representation and spatial analysis have been further developed and integrated into sectorial studies.

From the entire historical data, we selected about 50 thousand retail and service businesses active at the 1<sup>st</sup> of January 2017. We included every business selling goods and services located at the ground floor of every building along the edges of urban public space (streets, squares, public gardens). Services such as banks, insurance offices, hotels, real estate agencies, leisure and health services, etc. have also been included in our study. As previously discussed, services participate in the definition of the street social and economic attractiveness as much as stores selling material goods, and, furthermore, they compete for the same building frontage.

Among the several fields describing each business contained in the CCINCA database, three measurements are used in this dissertation: 1) the categorical classification following the international standard<sup>46</sup>; 2) the detailed measure of the commercial surfaces for average and large stores (higher than 300m<sup>2</sup>), whereas a default value of 40 m<sup>2</sup> is attributed to the smaller retailers; 3) Address information.

The address information allowed us to geocode the database and provide a spatial representation of the phenomena under study. This process has been realised through the National Open Addresses Database (*Base d'Adresses Nationale Ouverte* – BANO)<sup>47</sup>; since 2016, the French government has offered a geocoding service based on an address database constituted by the joint activity of IGN OSM France and the French National Post. In terms of the geocoding results, the BANO geocodation tool associates a score describing the

---

<sup>46</sup> Since 2008, the harmonization of the French classification NAF (Nomenclature d'activités Française), the European Statistical Classification NACE (Nomenclature statistique des Activités économiques dans la Communauté Européenne), and the United Nation classification ISIC (International Standard Industrial Classification of All Economic Activities) allows a direct correspondence of businesses classifications.

<sup>47</sup> <https://www.data.gouv.fr/fr/datasets/base-d-adresses-nationale-ouverte-bano/>

localisation precision in four levels: null, municipality, street, and house number. From our original dataset:

i) 7% of the data presented missing information, or they fell outside of our study area; this information has been excluded from our analysis.

ii) 2% of the information was geo-localised at the municipality level and 13% at the street level: the problem at the origin of these mis-localisations is often a result of incomplete address information in the CCINCA database (missing civic number, misspelt street names, names of isolated hamlets, etc.). A manual correction was carried out when the retail activity real address was available on other online sources.

ii) 78% of the data was correctly located at the house number level.

From the selection and manual correction of the original information, we obtained a final dataset of 45726 stores (44924 small and 802 large) distributed on 33221 locations (several activities might be found at the same address, especially in the city centre), 82%<sup>48</sup> with a precision at the civic number and 18% at the street level.

This cleaned and georeferenced information represents the starting dataset from which the analyses of the following chapters are carried on. In the following graphical representation (Fig.21), we observe the categorical classification of retail stores (overall dataset) and large stores' surfaces (>300m<sup>2</sup>).

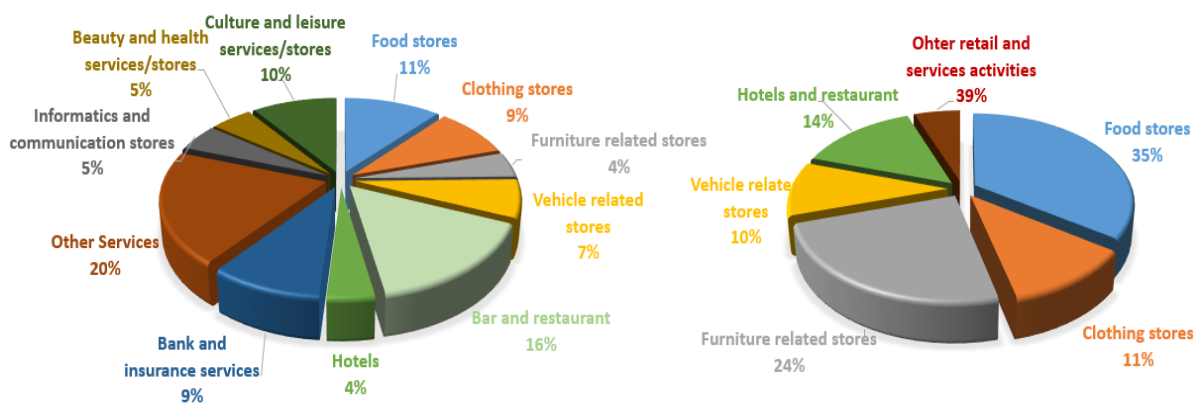


Figure 20 Repartition of the retail activities categories of the French Riviera: retail number (left) and retail large surfaces (>300m<sup>2</sup>, right).

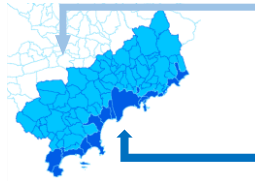
In the same study area, from the BDTOPO dataset we dispose of:

- 317 thousand buildings described by their footprint, height and a broad classification in three groups (remarkable, industrial, other);
- 113 thousand streets segments centrelines (motorways and pedestrian paths included) for a total length of 12.900 km. Each street segment is described by the number of lanes (from which we might estimate the correspondent width).
- 489 thousand parcels.
- Digital terrain model (DTM) at the resolution of 10 metres from the BDALTI (IGN).

Five layers of geographical information contain all the information here presented and from which all geoprocessing and statistical analysis of this dissertation have been realised.

<sup>48</sup> The same precision level found in Pavlis, M., Dolega, L., & Singleton, A. (2018); see Section 8.

In Table 3 we provide a first synthetic spatial description of the retail, street network and building data using a broad spatial repartition of the French Riviera. As previously introduced the study area counts 88 municipalities, extended above 1500 square kilometres. These municipalities are divided in littoral and sub-littoral (extended over respectively 20 and 80 per cent of the study area). Sub-littoral municipalities are further divided in Eastern (34%) and Western (47%) sectors using the Var River as a reference for evident physical and historical reasons previously described.



		Littoral 16 municipalities	Sublittoral		Study Area	
			Sublittoral 71 municipalities	Sublittoral East (33)		Sublittoral West (38)
Surface		18,7%	81,3%	34,1%	47,2%	100%
Network	Number	40440	62935	16847	46088	103375
	Length [km]	4124,2	8746,893	3282,63	5464,263	12871,093
	%	32,04%	67,96%	25,50%	42,45%	100%
	Density [km/km <sup>2</sup> ]	15,33	7,49	6,70	8,06	8,96
Average [m]		101,98	138,98	194,85	118,56	124,51
Building	Number	135439	178721	51338	127383	314160
	Surface [km <sup>2</sup> ]	22,712	20,615	4,673	15,942	43,327
	%	52,42%	47,58%	10,79%	36,79%	100%
	Density [m <sup>2</sup> /km <sup>2</sup> ]	84,40	17,65	9,54	23,51	30,16
Average		167,69	115,35	91,02	125,15	137,91
Retail	Number	35024	10702	1750	8952	45726
	%	76,60%	23,40%	3,83%	19,58%	100%
	N/Length[km]	8,49	1,22	0,53	1,64	3,55
	Number (>300)	604	198	19	179	802

Table 3 Repartition of Network Buildings and Retail in the three main geographical sectors of the French Riviera

The numbers shown in Table 3 confirm the panorama described in the previous sections: a hyper-concentrated urban development along the coast, less dense in the western sector, and urban underdevelopment in the eastern sloped sector. More than 70% of the retail is distributed in the littoral sector where we find 8.49 stores each kilometre of streets. While the remaining 23% is distributed mostly in the western sector with a density of 1.22 stores per kilometre of streets.

Since in this dissertation we are going to analyse the spatial distribution of stores considering their distribution in the urban spaces, we might propose a first visual representation of stores using the street network as a spatial unit of the analysis. Stores might be both represented as a count describing the number of stores along each street segment (Fig.22-upper part); this spatial representation might provide a biased representation of the spatial distribution of

retail since the underlying spatial unit might assume different sizes. In order to account for this dimension, we might divide the number of stores by the street length in order to obtain the retail density along each street segment (Fig.22-lower part). This second image provides the author with a first look at the simple spatial distribution of the phenomena which we are going to study and model in this dissertation. From these maps, we might easily delineate the most important urban centres of the French Riviera. Nonetheless, as previously discussed in Section 2.4, the simple enumeration of stores and businesses within a street or a given urban space *“does not tell the whole story”* (Greenberg 1995). Since our goal is to understand the spatial organisation of retail and its relationship with the urban form, in the first step of our analysis, we will discuss how to identify and describe quantitative descriptors of the retail spatial patterns from a simple point pattern representation.



### Retail distribution on the French Riviera: street count and street density

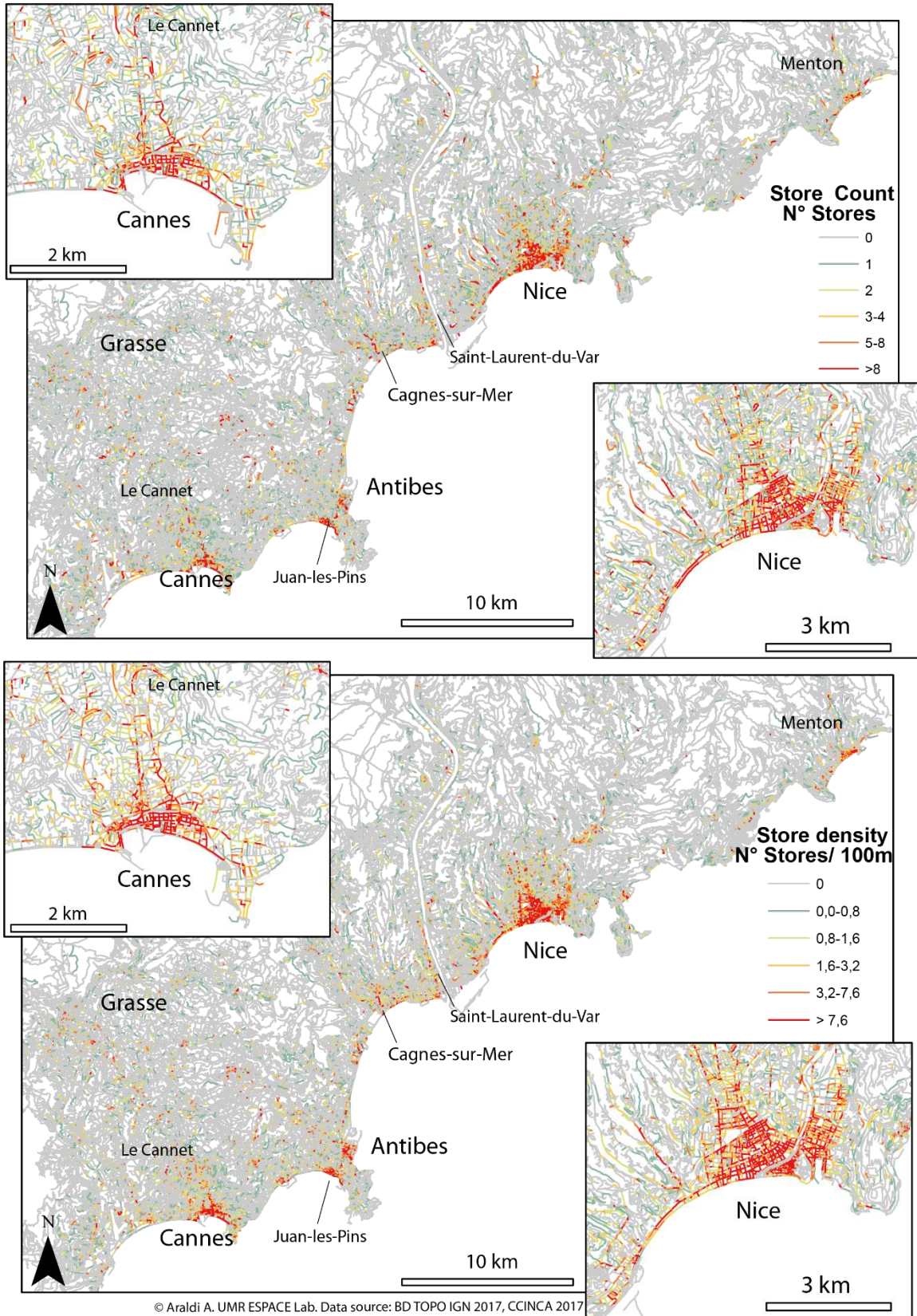


Figure 21 Spatial distribution of small and average-sized retail. On the left: street segment retail count; On the right: street segment retail density.

*In this first part, we presented the overall theoretical framework underlying this dissertation.*

*In Chapter 1, we focused on the scientific literature elaborated in the retail geography domain proposing the traditional distinction between neoclassical, behavioural, and structural approaches. More recent technological and societal transformations have required the development of new approaches for the study of the retail geography. Among contemporary works, we might detect a growing interest in the influence of the spatial societal and cultural context on the retail system. Nonetheless, the physical properties of the urban spaces are frequently overlooked.*

*In Chapter 2, we reviewed and discussed how the study of the retail distribution has been approached in the urban morphology scientific literature.*

*Firstly, in Section 2.1, once we explored the different notions associated to the term urban form, we then defined the specific notion investigated in this dissertation: urban form as the spatial organisation of the physical space of the city.*

*In Section 2.2, we discussed how urban morphology studies have approached the study of the retail distribution with quantitative approaches based on network configuration theories. The Hillier's Movement Economy Theory was described; based on this theory, new developments are proposed in this dissertation.*

*Section 2.3 and 2.4 discuss the simplifications underlying the MET: firstly, we focused on the simplification of the urban morphological system reduced to the configurational properties of its street-network. The importance of the morphological context both at the intrinsic level of the street and at larger scales is discussed. Secondly, we focused on the simplification of the MET regarding the distribution of stores. The retail system is traditionally reduced to a simple enumeration of stores, measured with street-based density measures. Based on these observations, we finally defined in Section 2.6 the thematic and methodological research question developed in this dissertation.*

*In Chapter 3, we described the study area of the French Riviera, a complex urban metropolitan region composed of eighty-eight municipalities within five major cities (Nice, Cannes, Antibes, Grasse and Menton). In Section 3.1, we described the urban and economical characteristics underlying our study area. Based on the technical reports produced by local authorities, we also proposed in Section 3.2 an overview of the French Riviera retail organisation. Finally, the data on which all the analysis of this dissertation is based upon was described in Section 3.4.*

*In order to better understand the spatial distribution of stores, in the following part, we start our quantitative analysis focusing on the study of the retail system of the French Riviera, investigating how quantitative approaches might be implemented for a better description and comprehension of the retail spatial organisation.*

*Dans cette première partie, nous avons présenté le cadre théorique général sur lequel s'appuient les différents développements de cette thèse.*

*Au Chapitre 1, nous nous sommes concentrés sur la littérature scientifique élaborée dans le domaine de la géographie du commerce, proposant la distinction traditionnelle entre les approches néoclassique, comportementale et structurelle. Les transformations technologiques et sociétales plus récentes ont nécessité l'élaboration de nouvelles approches pour l'étude de la géographie du commerce. Parmi les œuvres contemporaines, on peut déceler un intérêt croissant pour l'influence du contexte sociétal et culturel spatial sur les commerces. Néanmoins, les propriétés physiques des espaces urbains sont souvent négligées.*

*Au Chapitre 2, nous avons examiné et discuté de la manière dont l'étude de la distribution au détail a été abordée dans la littérature scientifique de la morphologie urbaine.*

*Premièrement, dans la Section 2.1, après avoir exploré les différentes notions associées au terme forme urbaine, nous avons défini la notion spécifique étudiée dans cette thèse: la forme urbaine en tant qu'organisation spatiale de l'espace physique de la ville.*

*Dans la Section 2.2, nous avons discuté de la manière dont les recherches dans le domaine de la morphologie urbaine ont abordé l'étude de la distribution des commerces de détail avec des approches quantitatives basées sur les théories de configuration de réseau. La théorie de Hillier sur l'économie du mouvement a été présentée. Sur la base de cette théorie, de nouveaux développements sont proposés dans cette thèse.*

*Les Sections 2.3 et 2.4 traitent les simplifications sous-jacentes à la théorie de l'économie du mouvement: en premier lieu, nous nous sommes concentrés sur la simplification du système morphologique, aux propriétés de configuration de son réseau de rues. L'importance du contexte morphologique, à la fois au niveau intrinsèque de la rue et à plus grande échelle est discutée. Nous nous sommes ensuite concentrés sur la simplification du MET en ce qui concerne la distribution des magasins. Le système de vente au détail est traditionnellement réduit à un simple recensement des magasins, mesuré à l'aide de mesures de densité basées sur la rue. A partir de ces observations, nous avons finalement défini à la section 2.6 la question de recherche thématique et méthodologique développée dans cette thèse.*

*Au Chapitre 3, nous avons décrit la zone d'étude de la Côte d'Azur, une métropole urbaine complexe composée de quatre-vingt-huit municipalités réparties dans cinq grandes villes (Nice, Cannes, Antibes, Grasse et Menton). Dans la Section 3.1, nous avons décrit les caractéristiques urbaines et économiques de notre zone d'étude. Sur la base des rapports techniques produits par les autorités locales, nous avons également proposé à la Section 3.2 un aperçu de l'organisation du commerce sur la Côte d'Azur. Enfin, les données sur lesquelles est basée toute l'analyse de cette thèse ont été décrites dans la Section 3.4.*

*Afin de mieux comprendre la répartition spatiale des magasins, nous entamerons ensuite notre analyse quantitative axée sur l'étude du commerce de détail de la Côte d'Azur, en examinant comment les approches quantitatives pourraient être mises en œuvre pour une meilleure description et compréhension de l'organisation spatiale du commerce.*



**Part 2**  
*Measuring Retail Distribution  
in the City*

*Retail in the city represents one of the main urban functions stimulating the social attractiveness of public spaces (Jane Jacobs 196). Its spatial organisation is influenced by and it, in turn, influences the social activity and the movement of people from the single streets to entire neighbourhoods and cities; understanding how stores are distributed might reveal information about the functionality of an urban system.*

*The analysis of the spatial and functional arrangement of stores might reveal the nature of the surrounding urban area, the social and economic composition of the neighbouring population, and the main functional activity, etc. It is clear, then, how the simple enumeration of stores and businesses along a street, or within a given urban space, “does not tell the whole story” (Greenberg 1995); their morphological and functional characteristics should also be considered.*

*Size, concentration, functional mix, and geometrical layout of stores in the urban spaces are only some of those aspects which should be combined for the definition of agglomeration typologies. As reviewed in the first chapter, each of these descriptors might play important roles in the definition of individual retail agglomeration typologies, as well as for the formulation of descriptive models for the overall retail system.*

*The retail agglomeration typologies and structure described by both hierarchical and post-hierarchical models studied in retail geography by respectively Berry (1959) and Brown (1991) are both based on the analysis of two groups of variables describing form and functional characteristics of store patterns. Nonetheless, within traditional approaches, these two components are traditionally evaluated separately.*

*Moreover, recent empirical works from spatial statistics have also demonstrated the importance of the integration of street-based analytical procedures for the description of point patterns on network constrained environment (Okabe, etc.). This specific aspect is also traditionally overlooked in quantitative analysis developed in retail geography studies privileging administrative partitions of the space. Conversely, street-based analysis of the retail system becomes one of the aspects of primordial importance for our dissertation.*

*Based on these considerations, Part 2 is dedicated to the presentation and implementation of a new computer-aided procedure able to support the analyst work of identification and description of retail patterns on the French Riviera metropolitan area.*

*Le commerce de détail dans une ville représente l'une des principales fonctions urbaines stimulant l'attractivité sociale des espaces publics (Jane Jacobs, 1960). Son organisation spatiale est influencée, et influence également à son tour l'activité sociale et le mouvement des personnes dans chaque rue, quartiers et plus largement la ville toute entière. Comprendre comment les magasins sont distribués peut révéler des informations sur la fonctionnalité d'un système urbain.*

*L'analyse de la disposition spatiale et fonctionnelle des magasins pourrait révéler la nature de la zone urbaine environnante, la composition sociale et économique de la population avoisinante, l'activité fonctionnelle principale, etc. Il est donc évident que le simple dénombrement des magasins situées le long d'une rue ou dans un espace urbain donné «ne racontent pas toute l'histoire» (Greenberg, 1995); leurs caractéristiques morphologiques et fonctionnelles doivent également être prises en compte.*

*La taille, la concentration, la mixité fonctionnelle et la disposition géométrique des magasins dans les espaces urbains ne sont que quelques-uns des aspects qui devraient être combinés pour la définition des typologies d'agglomération. Comme indiqué dans le premier chapitre, chacun de ces descripteurs pourrait jouer un rôle important dans la définition des typologies d'agglomération de magasins, ainsi que dans la formulation de modèles descriptifs pour l'ensemble du système de vente au détail.*

*Les typologies et la structure des agglomérations de détail décrites à la fois par Berry (1959) et Brown (1991) respectivement dans les modèles hiérarchiques et post-hiérarchiques, reposent toutes deux sur l'analyse de deux groupes de variables décrivant la forme et les caractéristiques fonctionnelles des modèles de magasins (retail patterns). Néanmoins, dans les approches traditionnelles, ces deux composantes sont habituellement évaluées séparément.*

*De plus, des travaux empiriques récents dans le domaine de la statistique spatiale ont également démontré l'importance de l'intégration de procédures analytiques basées sur le réseau pour la description de modèles de points (Okabe 2012, etc.). Cet aspect spécifique est également traditionnellement négligé dans les analyses quantitatives développées dans les études de géographie du commerce privilégiant une partition administrative de l'espace. Au contraire, l'analyse du système de vente au détail dans la rue revêt une importance primordiale pour notre recherche.*

*Sur la base de ces considérations, la Partie 2 est consacrée à la présentation et à la mise en œuvre d'une nouvelle procédure assistée par ordinateur capable de supporter le travail d'analyse et d'identification des caractéristiques spatiales de la distribution du commerce de la région métropolitaine de la Côte d'Azur.*

## 4. Quantitative methods for micro-retail location analysis

### 4.1 Spatial analysis approaches in retail geography

Between the 30s and the '50s, neoclassical geographical theories have been studied and applied at the intraurban level (Section 1.1). From a methodological point of view, these studies were mainly based on *subjective approaches*, with broad nominal categories assigned to retail agglomerations (i.e. Proudfoot 1937 in North American cities) or on *univariate methods*, employing measures of centre-size at different scales (i.e. Guy 1976 in Reading, UK).

The seminal studies of Berry (1959, 1963) proposed innovative theoretical and procedural developments in the analysis of the retail spatial structure: unlike previous works, the author proposes a bivariate method where retail agglomerations are identified and described through the combined analysis of the number of establishments and their functional prevalence. The introduction of functional measures with the traditional size, the well-known hierarchical model (Fig.3) which influenced and still influences the retail geography of the last decades. As we might remember from Section 1.1, Berry's model suggests *morpho-functional* centres described by:

- i) A hierarchical structure of business centres characterised by high density, category-mixed agglomerations.
- ii) Highway-oriented ribbons identifying store concentrations with a linear layout alongside highways, routes, urban roads with or without a functional specialisation (i.e. home supplies, furniture decoration, etc.; motorcar industry related, etc.
- iii) Specialized functional concentrations of specific retail categories taking advantage of the customer comparative behaviour and/or from agglomeration economies.

In the following twenty years, despite its theoretical importance, Berry's model has been little applied and discussed in empirical studies both in the original North American and in the British context as highlighted by **Davies (1976, p. 124)** and **Potter (1981)**.

Based on these observations, these two authors develop specific analysis in the British context for the city of Coventry (1974,1976, Fig.22.A) and Stockport (1983, Fig.22.B). In these works, Berry's notion of morpho-functional agglomerations is emphasised with a detailed multivariate analysis considering both form and function attributes. A common three-step procedure underpins these two works:

- A. Firstly, the identification of the most relevant agglomerations based on size measurement: 42 centres in Coventry and "72 separate retailing conformations" in Stockport (respectively with a minimum of ten and four stores).
- B. Secondly, agglomerations are described by an extensive set of functional characteristics (respectively 67 and 69 retail functions); factor analysis is implemented in order to reduce the large number of variables under study.
- C. Thirdly, classification of the agglomerations (Ward grouping procedure) and description of the retail structure.



Similarly, in the French-speaking literature **Beaujeu-Garnier** in **1977** proposed a comparable quantitative procedure to analyse the Parisian Atlas of 1967 (*Atlas de Paris et de la Région Parisienne*). A specific survey in the French capital streets provided an accurate measure of the retail distribution at the street level. Based on values of retail concentration at the street-level (only spatial units with a density higher than 3.6 stores per 100 meters of street, considering an average value of 2.6 stores/100 m), a first selection of street blocks (*ilôt*) is carried out and subsequently, a three-step procedure (similar to the one previously described) is implemented. Few methodological differences might be identified: while in the British case studies spatial agglomerations were selected by the authors using counting thresholds, in Paris the author uses street density thresholds applied on administrative units. From the 1056 *ilôts* retained, several explorative analyses describe different aspects of the Parisian retail organisation. In a second phase of the analysis, a selection of twenty-one most significant indicators is carried out, encompassing both retail-related variables and socioeconomic descriptors of the administrative unit. From a factorial analysis of these indicators, eight components are described and included in a classificatory analysis; seven typologies of retail agglomerations are identified. A hierarchical organisation is finally recognised and described (Fig.22.C).

Of these three seminal studies, only one considers the geometry of retail agglomerations: while the works on Paris and Stockport focus only on functional indicators in the analysis of the retail structure of Coventry, Davies consider both form and function of the retail centres. The author proposes two analyses: the first is equally applied to every agglomeration while in the second ribbons and nucleated concentration are separately evaluated. The goal of this specific dual analysis is to investigate functional differences between the two morphological configurations. Nonetheless, the distinction between ribbons and nucleated agglomeration is based only on the expert knowledge of the author.

As highlighted in the work of Beaujeu-Garnier, several works investigated retail spatial organisation, combining retail descriptors with socioeconomic information about the population distribution (Guy 1977,1983; etc.) driven by the growing attention to social studies (behavioural approach, Section 1.2). These works traditionally recognise retailer patterns using both demand and supply-side data: for example, population distribution, range quality, prices and distances between stores and households are combined in the same analysis aiming for a broader description of the *social form* of cities. As presented in the previous part, this dissertation investigates the spatial organisation of retail and its relationship with urban form; hence, specific consideration is given only to the spatial and functional distribution of stores, excluding additional socioeconomic data.

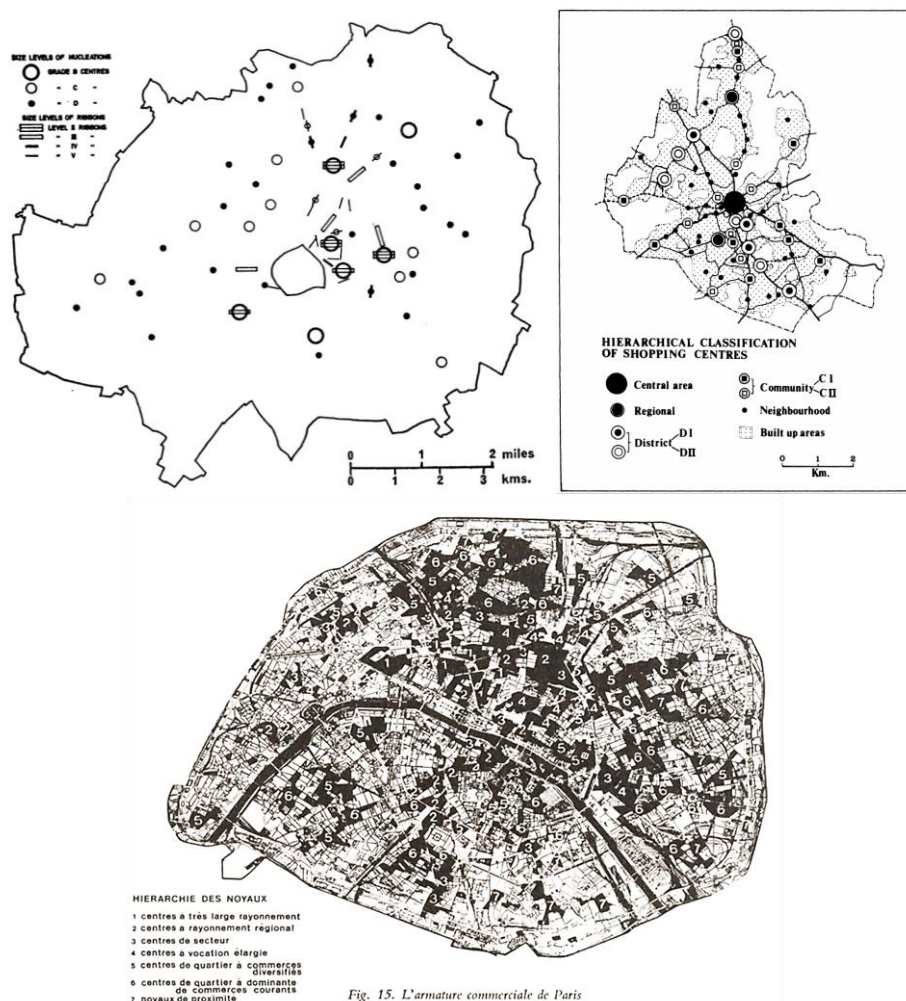


Fig. 15. L'armature commerciale de Paris

Figure 22 Retail agglomeration classification outcomes in the three seminal works of: a) Davies (1976), Coventry; b) Potter(1981), Stockport; c) Beaujeu-Garnier (1977), Paris.

In the four decades following the first intraurban retail study of Proudfoot (1937), the diffusion of computers and computer-aided procedures allowed the analysis of larger volumes of data and the implementation of more robust quantitative procedures. These significant methodological developments allowed new analytical approaches able to detect and describe retail spatial organisation with a more accurate description of agglomerations, beyond simple size/count measures.

As observed by Potter and Beaumont (1984), “classifications of urban retailing areas have tended to parallel the development of classifications in other fields, i.e., starting with subjective typologies based primarily on “eyeballing” and culminating in ostensibly more objective and increasingly sophisticated quantitative methods.” (p.353). In the same work, the authors describe six methodological approaches which have been developing over time: together with the subjective, the univariate, and the bivariate method, they also identify the growing interest in the functional index, the trait complex, and the multivariate approaches. They observe how “these approaches become increasingly quantitative, sophisticated and, [], objective, but as a direct consequence, also increasingly demanding with respect to data collection” (p.353).

Concerned by this last observation, the authors replicate the analysis of Stockport implementing a parsimonious multivariate classification. Starting from the same 42 agglomerations, they analyse seven form measures such as size, frontage quality, fabric-gap, distance from the city centre, road location, vacancy ratio and shape and only four functional indicators (food shops, furnishing and household, motor-related and professional service activities). The Ward classification on these 11 variables, showed an 82 per cent level of agreement with the previous 69 functional variable analysis showing how a set of carefully chosen variables *“might be just as efficacious in producing much the same classification”* (Potter and Beaumont 1984, p.354).

What is more relevant about the subset of variables proposed by Potter and Beaumont is that they *“can be assessed from map evidence plus cursory field inventories of centres, i.e., without full field surveys of the functions and character of each shop. Further, these variables together cover all of the aspects of retail area structural differentiation.”* Consequently, these variables *“can be assembled with relative ease and at low cost by planners, marketing specialists and retailing geographers, either by means of cursory field surveys and/or from published data sources, thereby further contributing to the parsimony of the method. The technique also has the advantage of drawing attention to the need for the examination of a balanced array of retail area variables at a time when many retail geographers have been over-concerned with the single attribute of centre size”* (Potter and Beaumont 1984, p.356).

These multivariate approaches from the late seventies are still applied in more recent years by several authors investigating the retail structure of urban spaces. Despite a few procedural innovations, the same three-step methodological approaches might be still identified in more recent works. Following the British approach with a multivariate analysis based on retail agglomerations, we might find for instance the work of Wayens (2006) on Brussels, while the work of Beaujeu-Garnier from the French-speaking literature has been recently implemented in the work of Delage and Fleury (2011) and Fleury et al. (2012).

**Wayens (2006)** proposes an in-depth study of the retail system in the city of Brussels. This work is based on several data sources (such as Cadaster, Belgian National Institute of Statistic, Real-estate agencies databases, etc.) and of a precise survey of almost 200 thousand buildings. Two spatial units are proposed for the analysis of retail structure: a first group of analytical analysis is based on a fine-sized grid of 33 x 33 metres (and a second of 100 metres) designed for spatial statistical analysis and the elimination of size effect or multiscale properties. A second group is based on a point pattern representation of retailers: the author provides a specific definition of retail agglomerations, namely, *nodules* (as defined by the Institut de Gestion de l'Environnement et d'Aménagement du Territoire; IGEAT, ULB). Nodules are defined as those agglomerations of at least three stores located within a 100 metres radius (3rd nearest neighbourhood within 100 metres), subsequently extended to every store located at less than 50 metres. Those retailers excluded from these procedures have been considered as isolated. From the original nodules buffers of 20 and 25 metres have been further applied. Since this procedure might generate large agglomerations encompassing different retail centralities, a final phase of fragmentation is necessary: *nodules* with more

than 1km of extent have been subdivided, considering the internal presence of specific polarities (see Wayens and Grimmaud 2004).

The *nodules* are used as spatial units for a typological analysis of the retail system with a set of variables going beyond the simple functional description: measures of retail category densities are evaluated together with estimated influence (*rayonnement*), luxury shops (*standing*) assessments of vitrines, number of employees etc. Factor analysis is then implemented to describe the main characteristics of the retail distribution in Brussel. In order to summarise the spatial distribution, 10 elaborated indicators are finally used in a weighted HAC (only on nodules containing more than ten stores); a final structure of eleven types of nodules is defined.

**Delage and Fleury (2011)** analyse retail distribution of Paris Intramuros from two datasets with different aggregation levels: point pattern (a dataset of two thousand points) and administrative units (918 iris). Two different procedures are therefore implemented to provide a combined analysis of the retail structure of the French capital. The point pattern dataset is analysed with spatial statistic approaches: in order to identify retail agglomerations a nearest neighbour hierarchical clustering (Nnh) routine is performed using CrimeStat: the authors implement a first EDA exploring different research radii between 100 and 600 metres. The stability of the cluster number is identified at the threshold distance of 350 metres. Once defined the perimeters of these clusters, functional indicators are studied and used in a hierarchical agglomerative clustering (HAC) procedure. This first analysis provides two levels of retail agglomerations respectively described by 2 and 3 sublevels defined by their functional composition.

The second database provides a density of 167 retail categories, reduced by the authors into 32 families, at the administrative unit level (IRIS). These density measures are used to produce four indicators (Total density, Variety, Prevalent density category and Local relative density) as input for a hierarchical agglomerative clustering analysis (HAC). From this second classification, the authors describe three levels of retail spaces composed respectively by 5, 2 and 1 sublevels. This second analytical approach has been further developed in **Fleury, Mathian Saint-Julien (2012)**.

Finally, the superposition of the results of these two classifications allows the authors to provide evidence about the hierarchical structure of retail in Paris: and a description of the spatial structure is therefore provided.

The works on Brussel and Paris represent two examples of more recent quantitative analysis of the retail structure, issues of the French-speaking retail geography research.

Other recent works have also been investigating the retail distribution and its evolution: for instance, we might mention the work of Axenov (1997) where retail organisation in Saint Petersburg is analysed (using 57 retail categories) along with the description of the socioeconomic and transport accessibility with the final goal to provide a graphic model for the retail organisation in Russian cities. Wang and Jones (2002) describe the retail structure

of Beijing using only retail variables such as number, category, location, sales surface and number of employees. The authors summarise the spatial organisation of retail distribution with a hierarchical structure composed of three first-order nodes and six second-order nodes; nonetheless, this analysis is based on an expert-based knowledge of the city without any quantitative and reproducible procedure. Similarly, in a second work on Shanghai, Wang and Zhang (2005), the spatial distribution and evolution are qualitatively described. Thurstain-Goodwin and Gong (2005) analyse the historical evolution of the retail distribution in London using two datasets (1971 and 2000); in this work, the administrative partition provides the spatial units on which statistical descriptors are measured. Similarly, quantitative analysis of retail at the national scale traditionally privilege administrative partitions for which socioeconomic data is regularly provided by national authorities (Grimmeau, Wayens and Hanson 2011, Madry 2016 etc.).

A primary observation we want to bring to the attention of the reader is that all the works reviewed so far share a similar methodological format where the statistical analysis are always subordinated to a previous definition or detection of a specific partition. Three spatial partitions might be identified:

- i **Exogenous partitions**, traditionally administrative units (in Beaujeu-Garnier in 1977, Thurstain-Goodwin and Gong 2005, Delage and Fleury 2011, etc.).
- ii **Endogenous partitions**, where agglomerations are defined by the author on an expert-based knowledge of the space study (as in Davies 1976, Potter 1981, etc.) or, more recently, by specific procedures such as thresholds or clustering procedures (as in Wayens 2006 and Fleury et al. 2012).
- iii **Neutral partitions**, such as regular grids with more or less coarse cell size; this specific approach is traditionally described as a possible solution to overcome statistical biases induced by the MAUP or the multiscale behaviour of the retail distribution of the previous spatial partitions methods (Wayens 2006).

Independently from the specific spatial unit choice, empirical studies in retail geography literature have been focusing on spatially aggregated retail locational data (administrative units, agglomerations, regular grids, etc.); only few works have been proposing studies of micro-retail patterns.

In the '90s, almost 30 years after the seminal work of Berry, Brown (1992, 1993,1994) highlighted how micro-retail studies was a relatively underdeveloped field of research necessitating more attention. Similarly, a few years later, Sadahiro (2000) states how "*though there are some exceptions such as microscopic studies of store arrangement pattern (Getis and Getis, 1968; Brown, 1992), analysis of urban retail structure has been mainly based on the spatially aggregated locational data.*"

Ha also observes that "*The use of aggregated data, however, is often problematic in retail analysis. This is, in short, because there do not exist spatial units used for data aggregation that are defined in a reasonable, consistent and objective manner and suitable for analysing the spatial structure of retailing. Administrative units and lattices impose unnatural*

*boundaries to the distribution of retail stores and thus sometimes yield strange results. Shopping areas still cannot be defined objectively, though numerous definitions have been proposed in the literature (Brown, 1992). It is obviously desirable in retail analysis to use the locational data of stores without aggregation.” (Sadahiro 2000, p.237).*

Hence, new methodological approaches with individual locational data of retail stores are required.

**Sadahiro (2000)** proposes a surface-based method for the investigation of the retail structure using a Probability Density Function (PDF) method. This approach overcomes the problem of the choice of the spatial unit with a transformation of the store point pattern in probability surfaces. The author estimates the PDF of 78 store categories, employing a kernel density estimation method, defined by a two-dimensional Gaussian distribution. These functions are summarised with several indicators describing the overall spatial patterns of every retail category and their degree of agglomeration. More interesting is the procedure developed to obtain the description of the overall retail structure: the first step consists of measuring the Jeffrey-Matsushita distance, a dissimilarity measure between PDF. These measures are successively used in the unweighted pair-group clustering method (UPGMA) using arithmetic averages. This methodological procedure allows classifying 78 PDF in 6 location types. Nonetheless, this classification does not consider the local spatial distribution of these classes. In order to study the spatial organisation of the retail structure, Sadahiro proposes a second implementation of the same UPGMA procedure on a lattice grid (of 300m), obtaining seven cells typologies.

Some observations should be highlighted. Firstly, the goal of this work to reduce the possible biases induced by the traditional spatial partitions using a point pattern-based procedure is achieved only in the first a-spatial retail structure description. The second classification for a spatial description of the retail structure is still implemented through the adoption of a regular grid partition of the space. This approach obliges the author to face the traditional problem of the choice of the most suitable cell size (different dimensions are explored). Secondly, for the same reasons discussed in Section 2.4, PDF represents a similar approach to other smoothing techniques (KDE, IDW, etc.) which might hide important hyperlocal characteristics of retail agglomerations. More specifically, two aspects of the retail agglomeration form are overlooked: the geometrical layout of retailer agglomerations (ribbon, nucleated etc.) and the absence of retailers (which represent one of the four main goals of our dissertation).

A more recent work investigating micro-retail distribution (point pattern representation) is discussed by **Hidalgo and Castañer (2015)**; the authors implement a recommender system to identify missing amenities considering current specialisation patterns within neighbourhoods. An ad hoc clustering algorithm is here implemented starting from a distance-based accessibility indicator. After the identification of amenity agglomerations in the first part of the algorithm, functional patterns and similarities among agglomerations are explored; this procedure is implemented over 1.26 million amenities (POI) of 47 US cities. Similarly to traditional studies previously reviewed, even in this work the analysis of form and functions characteristics of amenities agglomerations is not simultaneously considered. The spatial point repartition might influence (biases) the results of the subsequent implementation of

functional descriptors. Although the goal of this specific work does not correspond to those investigated in our dissertation, is still interesting to notice how recent studies are developing new procedures for community detection based on point pattern representation of amenities.

As we might observe, starting from 2000, retail geography studies include point pattern analysis, (such as the works of Sadahiro, Wayens, Hidalgo and Fleury, etc.) with the consequent development of more accurate procedures for the detection and delineation of retail clusters.

From these works, two further observations should be highlighted.

Firstly, despite the importance of a combined evaluation of form and function descriptors as traditionally described in the celebrated Berry and Browns models, empirical studies consider these aspects individually and in a subsequential order. Secondly, the geometrical layout is traditionally overlooked: we still do not find in the scientific literature methodological approaches integrating the description of the agglomeration shape. In more recent works, the form description of retail agglomerations has been reduced only to size/density measurements.

If we look back to the first seminal works previously mentioned, the geometry is everything but a neglected component: Davies provided a separate evaluation of ribbons and nucleated agglomerations while in the work of Potter and Beaumont (1984) we find the *Shape* indicator represented by a dummy variable (Nucleated centres = 1 and Ribbon and intermediate - shaped areas = 0) and manually attributed to 71 centres.

Analytical procedures for the automatic description of retail agglomeration shape seems to be missing in the scientific literature. The only exception might be found in the work of **Vazquez Parras (2011)** on the Wallonia region.

In this work, a first step of the analysis identifies retail agglomerations from a point pattern (similarly to Wayens 2006 with a quantitative definition of nodules). Secondly, a set of indicators is associated to each agglomeration describing their internal geometrical layout and functional characteristics; after the identification of three main factors through a Principal Component Analysis (PCA), a k-means classification is finally implemented helping the author to reorganise the 225 agglomerations in 10 typologies.

The real innovative aspect of this work is the development of the geoprocessing procedure for the introduction of a new geometrical indicator. The author proposes a specific procedure able to differentiate between compact and elongated shaped point agglomerations. Firstly, he implements the minimal convex elliptical-hull containing each agglomeration of stores. Secondly, he evaluates the ratio between the two axes of the ellipses: while values close to 1 describe a compact shape, elongated layouts are represented by values closer to 0. Despite the innovative aspect introduced by Vazquez, the same author highlights several limitations which make this indicator far from what is observed in the reality: the number and the density of points within an agglomeration might strongly influence and bias this measure. The same authors outline several inconsistencies when measuring specific spatial configurations. For instance, i) a retail strip which is disposed along a curved street might be described by values

close to 1 (compactness), ii) hybrid forms of retail agglomeration (such as a park or a square within a linear distribution or a linear ribbon rooted in a nucleated centre, etc.). iii) “artificial alignments” of the overall distribution of streets which in reality are differently orientated. The author suggests further developments should consider the improvement of this indicator, allowing for a more reliable detection of point layout.

As we highlighted in this section, only a small number of recent attempts in retail geography have been investigating micro-retail pattern agglomeration organisation through a point pattern representation of stores.

The reasons of our interest to point patterns is twofold: firstly, as previously explained in Part 1, we are looking for a fine-grained description of retail agglomerations within urban space to ultimately allow the cross-analysis with detailed urban morphological descriptors. Secondly, this interest is supported by the recent trend in point pattern analysis in social sciences which seem to have yet to be fully integrated into retail geography studies.

When considering the last thirty years, the development of information system (IS) databases and more recently geographic information system (GIS) databases allowed local authorities to implement a systematic follow up of the retail activities. Thurstain-Goodwin and Gong (2005) observe for example that it is only starting from 2000 that the city of London disposes of exhaustive quantitative store data; similarly, in Barcelona and Paris respectively in 2000 and 2001, local administrations provided regular data collection of the retail on offer for touristic purposes (Wayens 2006). In our case study (as mentioned in Section 7.4) since 1994, an IS database allows a regular and detailed data collection and has been transformed into a GIS database only more recently. This trend is nowadays accelerated by voluntary geography information platforms such as OSM (as previously discussed in Section 7.4) and by open-source data policies which provide free access to accurate retail data location: in France, starting from 2017, retail data has been made publicly available from the national statistical institute (INSEE).

A vast number of studies from several disciplines are investigating different aspects of retail point patterns with innovative approaches; on the contrary, point pattern procedures have not been fully integrated into retail geography studies (or limited to point cluster detection).

Moreover, our interest in micro-retail might find an additional explanation: the overall methodological approach developed in this dissertation tries to integrate a behavioural-oriented approach in the analysis of retail and urban form analysis. In order to develop this aspect, we should include a street network perspective in the analysis of the retail distribution. Nonetheless, in every empirical work reviewed so far, distances between stores (or agglomerations) are only evaluated with traditional Euclidean approaches. On the contrary, within the scientific statistic literature of the last ten to twenty years, growing attention has been given to the study of point patterns on street networks.



## 4.2 Point pattern analysis

While in the previous section, we reviewed works from the retail geography domain of research, retail point pattern analyses have also been investigated in other disciplines, such as spatial statistic and spatial econometrics. In this domain of research, the notion of *retail* is traditionally extended to several complementary aspects of the retailing sector: from the location of retail chains (multi-location problem) to the organisation of the supply distribution chain but also the location of industrial sites. The development of large international firms in the '90s raised several questions about the spatial organisation and optimisation of the retailing chain location, from an economic and logistical point of view (see structural approach Section 1.4). Therefore, when looking for retail point pattern statistical studies in this discipline we could encounter a wider notion of retail which is not only focused on store patterns in urban spaces, like in the works reviewed in the previous section.

Nonetheless, as observed by Zhao et al. (2018), two main aspects differentiate store patterns in urban space from other retailing studies: "*First, retailers in retail clusters do not [directly] cooperate and support each other like suppliers in supplier clusters do. Second, retail clusters directly serve the end consumer market, for which product valuation uncertainty is an important factor*" (Zhao et al. 2018). For these reasons, a specific sub-literature in spatial economic statistics has been investigating the specific distribution of store patterns.

Among these studies, a further distinction should be made: a first group of works have been focusing on modelling the statistical and economic process generating the observed patterns of stores with the final goal to support empirical observations of retail geography studies. In the last decades, we might mention the works of Eaton and Lipsey (1979), Dudgey (1990), Fisher and Harrington (1996), Gehrig (1998), Konishi (2005), Takahashi (2013), etc. A second group of spatial statistic studies have been focusing on the analysis and description of agglomerative properties of point patterns.

Within this second group, one last separation has been traditionally described by Besag and Newell (1991) and Lawson (2006) between:

- (i) **tests for cluster existence or measurement of its degree** where we find procedure traditionally developed for planar point patterns, such as nearest-neighbour distance (Clark and Evans 1954), K-function (Ripley 1976, 1981), M-function (Marcon and Puech 2006) and several others (e.g. Holgate 1965; Stephens 1974; Rao et al. 1979).
- (ii) **test for the spatial delimitation of clusters**, such as the multivariate methods developed from the late '80s in Geographical Analysis Machine (Openshaw et al. 1987), Besag and Newell's test (Besag and Newell 1991), SPLANCS (Rowlingson and Diggle 1993), a spatial scan statistic (Kulldorff and Nagarwalla 1995), etc.

The interest in network-based studies of the last twenty years (previously described in urban morphology studies Tab.1-2, Section 2.2) has also influenced the research in point pattern spatial statistics. In particular, in social sciences beyond spatial economics, the importance of the network constrained-shape of the urban space has been acknowledged to play an essential role in the detection of point patterns observations such as car accidents, crimes and

retail/firms distribution. Consequently, a recent growing number of studies have renovated traditional methodological procedures for the analysis of retail agglomerative behaviour introducing a network constrained space.

Early attempts to study point pattern analysis on networks might be found in Clark (1956); nonetheless, it is only starting from 1995 that Okabe and his research team at The Tokyo University have been developing specific procedures for point pattern analysis (Okabe et al. 1995; Okabe and Funamoto 2000; Okabe and Yamada 2001; Sato and Okabe 2005; Okabe 2009; Okabe and Sugihara 2012; etc.). In 2006, Okabe's group released SANET<sup>49</sup>, a specific tool conceived for the implementation of a large number of point pattern analysis on networks. This tool represents the outcome of the research activities of Okabe's group between 1998 and 2003, and the stimulus for contemporary researcher projects in urban studies.

In the last 15 years, point patterns on networks have been also investigated by several other authors: important references are the works of Flahaut et al. (2003), Borruso (2005, 2008) Yamada and Thill (2004,2007), Xie and Yan (2008, 2013), the research group of Baddeley (Baddeley et al. 2015<sup>50</sup>), etc.

Nevertheless, these recent works on point pattern analysis on networks focus on the methodological development of robust statistical procedures rather than the analysis of a specific phenomenon. Applications to real-world case studies are provided as a demonstration of the methodological validity of the proposed procedure. Only in last few years have these tools been implemented in more domain-oriented studies; in the specific case of store point pattern analysis, some applications should be mentioned.

When considering *agglomerative behaviour description*, we might mention the work of **Garrocho-Rangel, Alvarez-Lobato and Chavez (2013)** where K- and M- functions are implemented on 4.3 thousand stores on a 23.4 km street network in Toluca, Mexico. Both Euclidean and network constrained spaces are studied and their results compared. The same procedure is then applied considering twelve distinct retail functions in order to describe the different agglomerative behaviour associated with each retail category. Similarly, network-constrained and category-based weighted point pattern analysis are implemented in Nanjing, China in **Rui et al. (2015)**. The authors apply NetKDE and network constrained K-functions on a set of 4550 stores of one national and 16 international firms (national: Suguo, international: Carrefour, Metro, Auchan, etc.). **Saraiva and Pinho (2017)** in a broader study of the retail distribution in four Portuguese average-sized cities implement network constrained nearest neighbourhood.

What we should, however, keep in mind is that the results of K and M (or derivatives) and NN statistics are represented by distance-based functions, illustrating the overall clustering behaviour of the point pattern, by comparing it to the complete spatial randomness hypothesis. These three recent studies, similarly to what was previously highlighted in the work of Sadahiro (2000), describe the overall structure or agglomeration behaviour without a local description of the retail pattern. Only in Saraiva and Pinho do we find a separate

---

<sup>49</sup> <http://sanet.csis.u-tokyo.ac.jp/>

<sup>50</sup> New libraries have recently released the analysis of point pattern analysis on network in R environment.

assessment of central and peripheral areas which describe the agglomeration behaviour in two different urban contexts.

On the contrary, our research aims to identify local agglomerations behaviours which might help us in understanding the overall spatial organisation in our study area.

A network-constrained version of the M function is also at the origin of the work proposed by Jensen (2006): the author suggests a local appropriateness indicator for each retail activity in a specific location, considering store presence within its surrounding area. This indicator is based on quantitative measures of local spatial attractive/repulsive behaviours between retail categories, based on the M function. Jensen's local adaptation of the M statistic is a first attempt to identify local over/under-representation of retail activities. A local network-constrained version of K-function has been further discussed in Yamada and Thill (2007); nonetheless, we have not found any applications of the specific study of retail patterns.

Within *cluster delimitation* statistical procedures, despite the impressive development of planar methods, fewer attentions have been given to network constrained cluster detection. Exceptions might be found in the work of Shiode and Shiode (2009), where the authors propose an extension of the Planar Variable-Distance Clumping Method (PL-VCM, Okabe and Funamoto 2000) to a network constrained space (NT-VCM) and a specific test on Tokyo's restaurant distribution is used to assess the new procedure.

Finally, when considering the specific applications of clustering methods to store patterns, innovative methodological procedures have been recently investigated at the Geographic Data Science Lab at the Liverpool University by the recent works of **Dolega, Pavlis and Singleton (2016)**, **Pavlis, Dolega and Singleton (2018)**. Network-constrained cluster detection is overlooked in these works, which mainly focus on the study of density-based spatial clustering algorithms applied to retail distribution.

Their approach is *“strictly spatial; utilizing only the locations of the retail units. This is an appropriate approach for the identification of retail agglomerations where the extent of the clusters is determined by spatial discontinuities in unit distribution”* (Dolega et al. 2016). Algorithms for point pattern clustering identification such as k-means, quality threshold, kernel density estimation, random walk and density-based spatial clustering of application with noise (DBSCAN, Ester et al. 1996) are applied and compared. This last procedure provides the best performances when applied to a large dataset.

DBSCAN represents an interesting and statistically robust exploratory method for point pattern clustering identification. The spatial arrangement of a given set of points at different scales is compared to random distributions, in order to detect and extract clustering information pertaining only to anomalies, which means identifying agglomerations that are unlikely to happen by chance.

The identification of retail clusters with the DBSCAN procedure discussed in Pavlis, Dolega and Singleton (2018) is successively used as the spatial unit partition for a retail agglomeration classification in the most recent work of Dolega et al. (2019) of shopping and consumption

spaces in Great Britain, using both non-commercial socioeconomic data. This last study represents the most recent work in retail geography study at the time of writing this dissertation (which has been published a few months after the release of the methodology discussed in this section, Araldi and Fusco 2019). For this reason, it would not be correct to present the Dolega et al. (2019) work as part of the methodological background of this dissertation. We decided to provide a review of this work in the conclusion section of this part (Section 5.5), providing a brief comparative analysis of the results of the two methodologies.

### 4.3 Observations

In the two previous sections, a large number of works have been reviewed; advantages, innovative methods and shortcomings have been individually highlighted considering the specific goals of this dissertation. Nonetheless, to better understand this heterogeneous literature, it is imperative to reorganise these works from a methodological point of view in order to provide a clear vision of the methodological challenges and differences proposed by our specific approach. In the following Table 4, we firstly recapitulate the works: in the first part, we might find retail geography studies on store spatial organisation. Applications to stores distribution of point pattern analysis approaches complete the second part of the table.

Authors	Year	Data	City	Method	Spatial unit	Results
<b>From Retail Geography Studies</b>						
Beaujeu-Garnier	1977	115k O	Paris, France	Factor Analysis + Classification	1056 Admin.unit	7 classes
Potter	1981	72 aggl.	Stockport, UK	Correlation+ PCA+ Wald Clustering	72 centers	11 classes
Potter and Beaumont	1984	71 aggl.			71 centres (7 from and 7+4 form-function factors)	11 classes
Axenov	1997	4433(1988 ) 16561 (1996)	St. Petersburg Russia	Expert base analysis		Graphical /geographical representation
Sadahiro	2000	78 retail families	Yokohama, Japan	PDF (78 families) + UPGMA	Overall PDF	6 classes
					Grid (300m)	7 classes
Wang and Jones	2002	11712	Beijing; China	Expert based analysis	Admin.unit	Graphical /geographical representation
Thurstain-Goodwin and Gong	2005	-	London, UK	Evolution statistics (1971-2000) by districts	Admin.unit	Statistics at the district level
Wayens	2006	24.5k O	Brussels, Belgium	Nnh + HAC	3 <sup>rd</sup> Nnh within 100 mt and corrections	11 classes (5 types)
				Correlations	Regular grids (33 and 100 m size)	
Delage and	2011 and 2012	2320 O (retail chains)	Paris, France	Euclidean Nnh hierarchical clustering + HAC	63 Clusters defined by Nnh (450m and 10 O min.)	8 classes (three density levels)

Fleury and Fleury, Mathian Saint-Julien		56.6k O		HAC	919 Admin.units	5 classes (2 levels)
Vasquez	2011	20k O	Wallonia, Belgium	k-means on eight geometrical indicators	Agglomerations	Identification of 10 agglomerations types
Hildago and Castaner	2015	1.26 million amenities	47 US cities	Accessibility based Ad-hoc algorithm	Agglomerations	Functional co-occurrences
Dolega et al.	2019	3110 aggl.	Great Britain	DBSCAN + k-means classification	Agglomerations	5 Groups divided in 15 subgroups
Araldi and Fusco	2019	33k O	French Riviera, France	Point patterns + Bayesian Classification	Point Pattern on Networks	18 groups
<b>From Retail Point Pattern Analysis on Networks Studies</b>						
<b>Tests for cluster existence or measurement of its degree</b>						
Jensen	2006			Network <b>M-function</b>	Point pattern	Local functional agglomeration description
Garrocho-Rangel Alvarez-Lobato and Chavez	2013	4.3 O, 23.4 km S	Toluca, Mexico	Network and Euclidean <b>K-function</b>	Point pattern	Distance-based Functions of the overall space study
Rui et al.	2015	466 O	Nanjing, China	Network-constrained, category-based weighted point pattern analysis <b>K-function</b>	Point pattern (Bandwidth 500m and 1500 m for convenience stores and hypermarkets)	Distance-based Functions of the overall space study
Saraiva and Pinho	2017	3.9k O	Four Portuguese Towns	Network <b>k-NN</b>	Point pattern	Distance-based Functions (for central and peripheral areas)
<b>From Retail Point Pattern cluster detection</b>						
<b>Identification of the actual spatial delimitation of clusters</b>						
Dolega Pavlis and Singleton	2016	437k O	Eight representative location across GB	k-means, KDE, QT, random walk and DBSCAN	Point pattern	Cluster Detection
Pavlis, Dolega Singleton	2018			DBSCAN	Point pattern	Cluster Detection

Table 4 Methodological Literature Review of Micro-Retail Studies in Economic Geography and Point pattern analysis

As the reader might observe from this table, the common goal underlying the selected works from retail geography is the identification and description of the spatial organisation of retail agglomerations. This information is traditionally obtained as a multivariate classification of retail agglomerations (column Results Tab.4). Although few methodological variations, all these works share the same three-step procedure previously described, representing the research design underlying these studies (Fig.23). The only exception is represented by Sadahiro’s work where the study of the functional patterns through PDF is implemented before the identification of the spatial unit (inversion of Step A and Step B). Even the most recent work of Dolega et al. (2019) (as we further discuss in the conclusion section) replicates a more sophisticated version of this three-step procedure, based on retail agglomerations identified with DBSCAN approach (Pavlis et al. 2018).

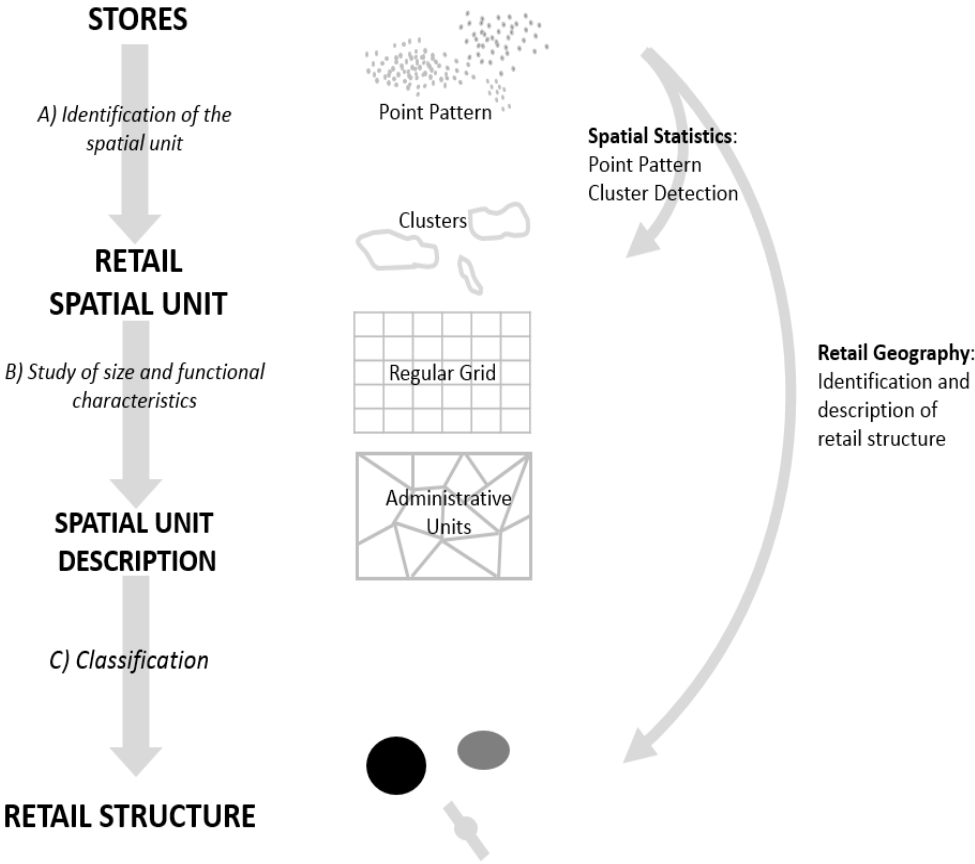


Figure 23 The three-step procedure traditionally implemented in Retail Geography studies for the identification of retail structure at the intra-urban scale

From the works in Table 4, let us now discuss the possible pitfalls associated with each level of this three-step protocol.

**Step A)** The first observation we should discuss is about the spatial unit underlying these works. As previously mentioned, in retail geography studies, three spatial unit typologies are traditionally considered: regular grids, administrative units and retail clusters. Each one presents specific peculiarities: the regular grid is suggested as a neutral spatial unit when the analyst aims to reduce the statistical biases such as MAUP and multiscale spatial organisation. Administrative units present the important advantage to share the same spatial repartition with socioeconomic information provided by public institutions allowing a straightforward

implementation of cross-analysis with other authoritative sources of data. Nonetheless, both grids and administrative units might be considered as exogenous partitions since they are traditionally obtained from residential population distribution criteria which do not always fit the retail spatial organisation (residential and commercial functions do not always coincide in urban spaces). On the contrary, partitions based on retail clusters have the advantage to be directly determined from the observed store spatial distribution, hence, an endogenous definition of spatial units. Moreover, they could be defined through robust statistical procedures from point pattern statistical analysis and they provide a finer definition of individual agglomeration borders. These last features might represent an important advantage when the goal of the study, like in this dissertation, is the identification of retail locational factors based on spatial co-occurrences.

When considering the final goal of our dissertation (the relationship between retail distribution and the physical city), another group of spatial units might be mentioned. As we previously discussed (Section 3), in urban morphology studies, elementary components of the physical urban form are traditionally used as spatial units for retail measurements: streets are the predominant spatial units for obvious theoretical and methodological reasons (MET, Tables 1-2), but also buildings (Sevtzuk 2010, 2014) and street-blocks (Yoshida and Omae 2005, Ye et al. 2017, etc.).

At this point in the dissertation, the reader might think that the natural solution for our case study could be the adaptation of the traditional three-step procedure of retail geography to street segment units (for example as illustrated in the data description, Fig.23 Section 3.4).

Nevertheless, it is of paramount importance to highlight how the significance of street-based analysis, acknowledged by both retail geographers and urban morphologists, not be limited to a simple street-based partition.

Moreover, a street-based partition (as well as every other tessellation based on other urban morphological elements) should be considered as an exogenous partition (similarly to administrative units) with specific pitfalls which might prevent its implementation. Firstly, discussions and criticisms might arise from the simple definition of streets elements: as for configurational studies (further described in Part 3), the outcomes of the analysis might vary when considering named streets, visual axes, street segments, etc. Secondly, the application of the traditional three-step procedure to streets would result in a taxonomy of streets which does not directly address the question about the retail agglomerations identification and organisation which we clearly stated at the beginning of this chapter. The fragmentation of street segments might influence the resulting street taxonomy, and a specific weighing approach should be introduced in order to compensate for the difference in retail densities observed on each element.

As we might see from this discussion, the choice of the spatial unit for the study of retail agglomeration is everything except an easy task for the analyst; the specific position of this dissertation between retail geography and urban morphology might increment the possible criticisms surrounding the choice of the spatial unit definition. Consequently, in this work, we decided to develop a new procedure entirely based on a point pattern representation of

stores, without a preconceived exogenous partition, integrating a street-based perception of store distribution.

**Step B)** The second step of defining the retail agglomeration description might also be at the origin of numerous discussions.

As observed in Potter and Beaumont (1984), we should be able to describe the retail agglomeration with a limited number of indicators: while the first multivariate analyses of the late '70s were considering more than 60 functional variables (as in Potter, Davies and Beaujeu-Garnier, etc.), Potter and Beaumont demonstrate how the classification of 11 form and function variables shows an 82 per cent level of agreement with the previous classification made with 69 (only) functional measures on the same study area (Stockport).

The high number of variables might also be explained by the implementation of measures not limited to the form-function descriptors but also including other socioeconomic aspects from other data sources or specific surveys (as in Axenov 1997, Wayens 2006, Dolega et al. 2019). Although these last sources of data might provide insights about the specific socio-cultural context of the case study under investigation, their implementation might strongly influence the classification results with the risk of an idiosyncratic description, which could limit its generalisation and its use in comparative and cross-analysis studies.

For these reasons, we decided to focus on the fundamental notion of Berry's hierarchical model, which considers the definition of retail agglomerations from a combined assessment of form and function features only. These two aspects are also acknowledged by the post-hierarchical model of Brown (1991): the author still identifies in form and function the two main components from which a combinatory matrix of 9 families of retail agglomeration is derived (extended to 12 when planned centres are included, Fig.5 Section 1.3). In this second model, the equal importance of the two factors is emphasised.

Nonetheless, in all empirical works previously reviewed, two shortcomings have been highlighted. Firstly, the evaluation of form and function is never simultaneous: spatial units are delineated and subsequently described by functional characteristics. Secondly, the form descriptors are reduced to size or density measures, while the geometrical layout of stores is traditionally overlooked. Independently of the more or less sophisticated procedure which might be implemented in order to consider this specific aspect, these two aspects are strictly related to the previous step (Step A) of spatial unit definition.

- I. The use of exogenous spatial units (both administrative or morphological) prevents any form of shape detection. In some works, this limitation is imposed by the data source resolution (Fleury et al. 2012). The same observation might be associated with smoothing procedures based on the only notion of density distribution (such as the PDF procedure of Sadahiro 2000).
- II. The detection of agglomerations with manual, distance-based rules or clustering methods results in a separated evaluation of form and functional aspects in two different moments of the analysis where the form (Step A) might influence the latter (Step B) (Hildago and Castaner 2015, Vazquez Parras 2011, Dolega et al. 2019).



III. The fine-grained point pattern representation of stores is of paramount importance for the description of retail agglomeration shape; nonetheless, as we further discuss in Section 9.3, cluster detection tests (both planar and network-based) always compare observed patterns to random distributions in order to corroborate the simple hypothesis of concentration/dispersion, while the geometrical layout of points is traditionally overlooked.

Some of the problems in Step A and B) might find possible solutions when considering a point pattern distribution as the spatial unit of the analysis. The innovative approaches proposed in point pattern analysis statistics (the second part of Table 3) might represent the support for the development of a new procedure.

**Step C)** Represent the key aspect differentiating statistical studies from retail geography studies. This third step allows academics to go beyond the simple description of individual agglomerative behaviours and the identification of the overall structure of retail is thus possible.

Nonetheless, when considering the third step of classification, another important observation should be outlined: the geographical/spatial analysis of the phenomena is limited only to the spatial partitioning (Step A).

Space is almost absent in the second phase of the analysis (Step B), which describes the internal form and functioning characteristics of each unit; similarly, the final classification (Step C) is implemented considering only the statistical distribution of the individual variables.

In order to better understand the implications of this traditional approach, let us consider for a moment the example of the classification of the vegetation of several islands of an archipelago. A classification based on the individual variables describing the trees' species proportions of each island is independent of the location that each island takes within the archipelago configuration. This means that only the statistical distribution of the variables is considered while the spatial distribution is overlooked. Shuffling the spatial distribution of the islands would result in the same classification. Nonetheless, from a different spatial organisation of the archipelago, a different description of the vegetation would be associated with each island. With this observation, we wanted to highlight how the spatial distribution of the units under analysis strongly influence the observed patterns; specific procedures should be implemented in order to analyse both the statistical and spatial distribution of the population under study.

Some authors, conscious of this problem, introduced a specific variable measuring the distance from the city centre of each spatial unit (Potter 1981); a measure which might generate another long discussion about the most accurate procedure to define the city centre, and more importantly, another source of possible biases. A task which is even more problematic when considering a complex study area such as our polycentric and coastal conurbation.

As highlighted in this section, an impressive number of methodological and theoretical debates might consider the validity of every single step of the traditional procedures

implemented in retail geography. The challenge we face in this Part of the dissertation is the need to find a common methodological framework able to integrate and solve (if not all, at least the majority of) the observations and shortcomings previously described. A profound reform of traditional approaches is thus needed.

In short, we are looking for a new methodology producing an automated delimitation of retail agglomerations typologies:

- I. without any preconceived or exogenous partition of space or expert-based recognition of retail agglomerations;
- II. able to associate both functional and form characteristics within an acknowledged theoretical framework;
- III. able to consider the geometrical layout of retail agglomerations with an objective and automatic procedure;
- IV. able to consider the spatial relative disposition of individual stores (without a problematic definition of the city centre).

Together with these objectives, some methodological constraints are included:

- V. network constrained distances to respect the nature of retail distribution within the city; minimal and standardised information on retailers, required to guarantee the reproducibility and adaptability to other case studies.

## 5. Retail Fabric Assessment: a new proposal

Retail Fabric Assessment (RFA) is a new methodological procedure for the identification of retail agglomeration typologies and the analysis of their spatial organisation. The three-step protocol of the traditional multivariate analysis of retail geography is here profoundly renewed with a modern geoprocessing approach and the combination of geostatistic procedures.

RFA can be summed up as follows:

6. Defining a network constrained catchment area around each event and collecting retail activity information.
7. Identification of significant spatial *morpho-functional* patterns describing different aspects of the retail context surrounding each store.
8. Clustering of relative event-based patterns through Bayesian methods in order to identify and characterize retail fabrics (*retailsapes*).

## 5.1 Spatial Unit

To overcome most of the problems associated with the delineation of the spatial partition, RFA operates a conceptual inversion of the notion of Spatial Unit. Traditional studies define a specific tessellation of the study area with a subsequent evaluation of the internal morpho-functional characteristics of each unit (intrinsic description); the chosen partition influences the whole procedure and the classification results.

RFA, on the contrary, is based on a point pattern representation of stores on a street network where each point is evaluated for its relative position within its surrounding context (extrinsic description), similarly to what happens when applying moving average methods on a shifting temporal/spatial subset or with local measures of spatial association (Anselin 1995, etc.).

We consider a floating network-based catchment area of radius  $r$  and centred on the precise location of each store. This specific approach not only integrates the network-based distances as suggested by more recent statistical works, but it might also allow a behavioural-based scale (Zhang and Kukadia, 2005) in the analysis of the retail distribution.

This specific approach might be associated with the notion of street-based retailscape defined as the spatial organisation of stores as perceived by the pedestrian exploring the retail offer distributed along the street network.

Within the network-based catchment area, the number of stores for each commercial typology is collected and evaluated. This procedure is the starting point of many statistical methods (i.e. Ripley's  $K$  and derivatives etc.) exploring the spatial distribution of points at increasing distances  $r$ . Similarly, we might also mention a similar approach implemented in local fractal studies (such as local functional sandbox counting method in Kanevski and Pereira 2017; see further details in Text-Box 1). While these works investigate the scale-dependence of point clustering behaviour, RFA focuses on specific scales of analysis, relying on the scientific literature of urban studies related to pedestrian movement.

The collection of the number of point for each retail category is implemented at three selected scales depending on the specific indicator (further detailed in the next section). We refer to them as street scale ( $r=150$  m), neighbourhood scale ( $r=300$  m) and district scale ( $r=600$  m). The choice of these three measures is based on traditional urban studies related to walkability and retail distribution where an average distance between 300 and 450 metres is traditionally considered. From the works previously reviewed and collected in Tab.4, we might also observe the how specific studies issued from retail geography literature usually explore radii of similar distances: between 100 and 600 metres in Delage and Fleury (2011) and Fleury et al. (2012), 100-500 metres in Sadahiro (2000), 60-500 metres in Hildago and Castaner (2015) etc.

Twenty-two retail categories are retained and further grouped into twelve families: food stores, clothing stores, furniture related stores, vehicle related stores, bars and restaurants, hotels, bank and insurance services, real-estate agencies, informatics and telecommunication (IT) stores, beauty and health services/stores, reparation services, culture and leisure services/stores. The choice of these twelve families is inherited by a specific functional grouping (which is based on the five-digit NAF codes, with a direct correspondence to the

international ISIC) adopted by the Commercial Urbanism department of the local Chamber of Commerce (CCINCA). Several advantages are associated with this 12-fold typology grouping: firstly, it might be considered as a reasonable differentiation of retail functional categories, between a too concise classification (such as the 4 classes of Potter and Beaumont 1984) and the excessively fine-grained and redundant description (i.e. 69 classes in Potter 1982, 57 Axenov 1997, 78 Sadahiro 2000, etc.). Secondly, this specific retail categorization has been assessed and employed by local administration with an expert knowledge of the urban and economic ecosystem under analysis: for example, the separation of real estate agencies (traditionally grouped with financial services) is suggested here because of the relevant presence of a secondary home market on the French Riviera (Section 3.1). Similarly, for the touristic nature of our study area, hotels are separated from bars and restaurants. Thirdly, the results of this study might be more easily understood by local practitioners and promptly applied implementation in official assessment documents.

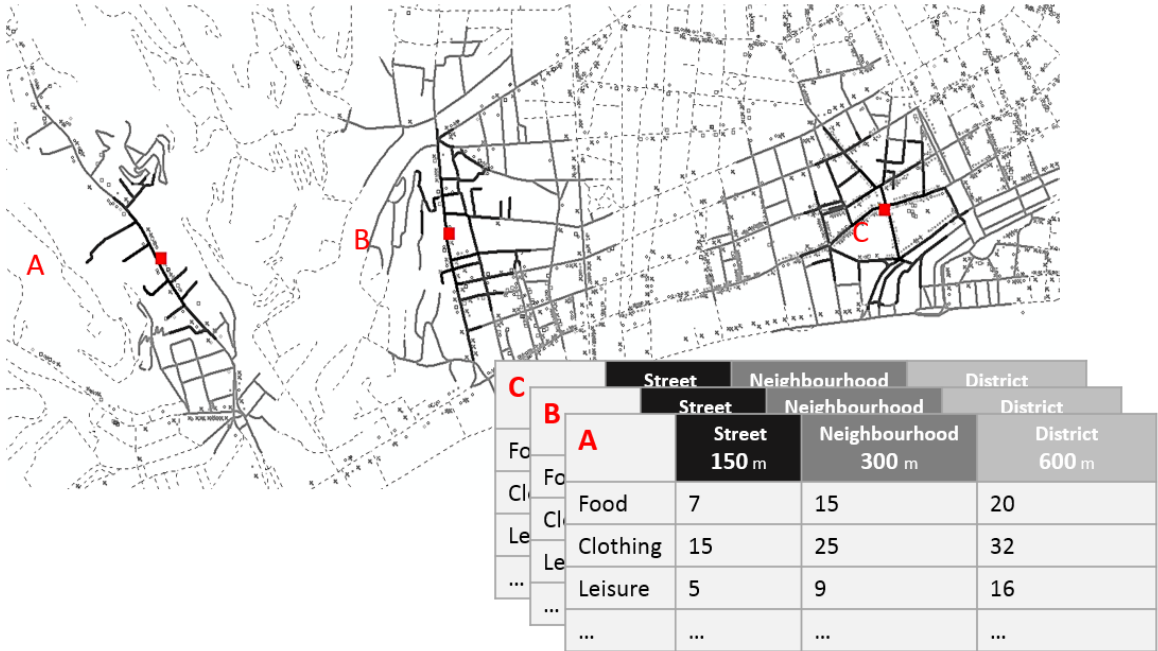


Figure 24 The network-based catchment area around three stores (A, B, C) and their relative number of store families at three different scales (street, neighbourhood and district)

## 5.2 Measuring Form and Functioning Patterns

As previously described, form and function represent the two most recurrent features measured in retail geography empirical studies. This section presents how these two aspects are conceptualised and implemented through quantitative algorithms based on statistical and geoprocessing methods. From the raw data collected in the previous step where the number of stores for different retail categories has been automatically collected at different scales, significant indicators need to be extracted. Both statistical and spatial distribution should consider the local distribution of stores compared to the overall values observed in the study area. The final goal of this second step is to associate each store with a set of indicators which allows the describing of the local *morphofunctional* patterns within which each store is included.

In order to identify the functional characteristics of a retail agglomeration, the number of stores for each retail family is now analysed. Three indicators describe the local **Functioning** patterns of a retail distribution:

- **Retail category variety** is described by the number of retail types present within the neighbourhood distance. This indicator represents the amplitude of the commercial offer enabling a multi-purpose shopping behaviour. Lower values of variety might represent both specialised areas and small convenience/street-corner developments. For this specific indicator, the 12-typologies have been separately considered. The row indicator has been classified using a 20th-percentile discretisation method.<sup>51</sup>
- **Retail category prevalence** aims at identifying significant patterns of specific function concentrations. In seminal works of Potter, Davies and Beaujeu-Garnier, the prevalence is traditionally implemented as the ratio of the number of a specific retail category compared to the overall number of stores for each spatial unit. The same notion has been further developed in more recent works in spatial economic studies considering a point pattern distribution of industries (Marcon Puech 2009,2017; Lang et al. 2016 etc.). Similarly, Jensen (2006) implemented a local version of the M index (Duranton and Overman 2005) specifically conceived for the detection of attraction/repulsion between pairwise retail categories. Considering these works, we derive a simplified procedure for the measurement of the local over/under-representation of each retail category: we firstly implement the ratio between the number of each specific store category and the overall retail number observed within a neighbourhood distance. The simple use of this local ratio might lead to a distorted representation of the phenomenon caused by the higher/lower presence of specific categories in the overall study areas and attractive/repulsive behaviour of different retail categories, etc.<sup>52</sup> To ascertain the statistical significance of this indicator, a double one-tailed binomial test is then applied with a confidence level of 0.95<sup>53</sup>. The results of the application of this methodology to each group of the 12 retail typologies identify statistical High, Low, and Not Significant patterns of functional prevalence (Fig.25). The patterns of retail category prevalence represent a relative measure based on both the (local) spatial and statistic distribution of the simple prevalence indicator.

---

<sup>51</sup> Alternative more sophisticated univariate discretization approaches such as k-means or genetic algorithm have also been tested with similar resulting intervals. Consequently, the percentiles method has been preferred for its reproducibility and easier interpretability.

<sup>52</sup> For instance, let us consider two stores A and B surrounded by 2 and 20 stores of which 50% belong to the food category. The prevalence ratio in the two cases should be differently evaluated: only in B we might detect a significant overrepresentation pattern while in A it is more likely to be observed by chance. Furthermore, if we consider that on average the ratio of food stores is observed with 0.5 of frequency, it is then clear that even in B the value is not significantly higher or lower than the average.

<sup>53</sup> As discussed in the scientific literature, a Monte Carlo randomization of the points around each centre of the analysis would be the best way to evaluate the correct local statistical significance of the "prevalence" computed within our methodology. Our case study would then require the implementation of one thousand Monte Carlo randomizations around 33 thousand centres of analysis.

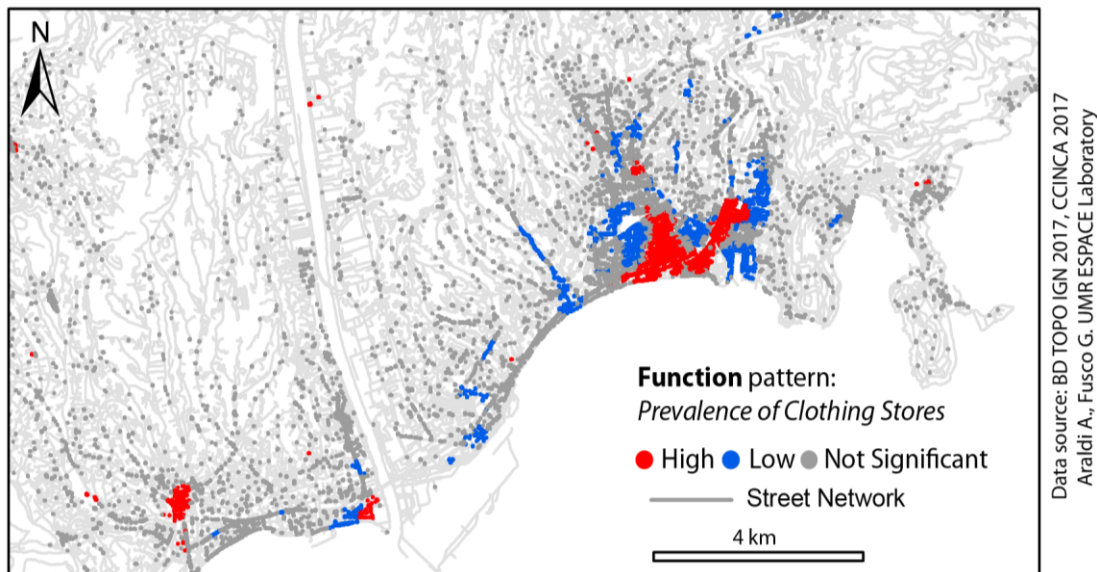


Figure 25 Example of Function Pattern applied to clothing store distribution (Source: Araldi and Fusco 2018)

- **Anchor Store** category prevalence. This group of indicators are introduced in the analysis to consider the presence of anchor stores (AS) in the retail ecosystem. The importance of these magnet/attractor stores (Brown 1994) requires a specific group of indicators to quantify their influence on smaller retailers and also those “out-of-town” store configurations, missing in traditional studies. The statistical analysis of category prevalence applied above for small- and average-sized retail surfaces (<300 m<sup>2</sup>), is reiterated here considering surface information. Since the spatial influence of anchor stores covers wider spaces compared to small retailers, the prevalence measure is thus implemented at the district level. Six typologies of retail activities are analysed (Anchor Store category prevalence in Tab.5)

The local Form of retail activities is described by the following indicators:

- **Size** indicator is defined in the seminal work of Potter and Beaumont (1984) as the “total number of establishments” observed in a specific agglomeration. Similarly, for each store-point, we compute the total number of retailers  $N(r)$  within a given radius  $r$ . A segmentation based on quantiles is applied to the statistical distribution of  $N(r)$  on the overall study area. This discretisation informs us of the relative dimension of the local retail offer. The procedure is applied to the three selected scales. For example, in Fig.26, the patterns identified illustrate the hierarchical importance of centres at the district scale.

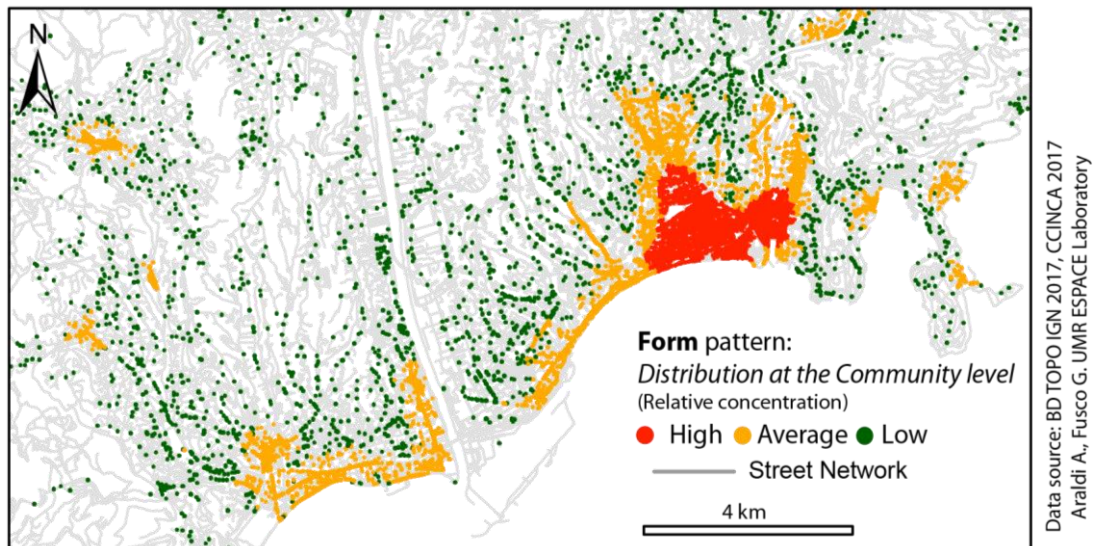


Figure 26 Example of Form Pattern: distribution variable at the district level (Source: Araldi and Fusco 2018)

- **Concentration** indicators consider the ratio of the store size at two different scales: street/neighbourhood levels and neighbourhood/district levels. This indicator discloses the information on the aggregation/isolation of the *retailscape* at a given level compared to the larger one. We can thus detect isolated stores and high concentrations surrounded by empty/lower concentrations (High-Low distributions), lower retail presence close to high concentrations (Low-High) and homogenous distribution on both scales (High-High and Low-Low). This triple description is obtained by applying the same discretization method used for the previous indicator.
- **Geometry** While functioning, size and concentration variables have been readapted and easily implemented on the new definition of the point-based network-constrained spatial unit, the geometry indicator requires specific considerations. The challenge we faced at this point of the research was to find a quantitative automatic and robust procedure able to assign to each point a numerical or categorical variable describing the geometrical layout of the surrounding store layout. A challenging task since the majority of point pattern statistic procedures (as previously highlighted) focus on the agglomeration detection/description (sometimes including multi-scale approaches, as in DBSCAN) but always overlook the geometrical layout. The reasons for this absence might be found in the limited interest of this aspect to retail geography: only in this context does the geometrical layout imply a specific connotation (ribbons, nucleated, distributed, etc.). On the contrary, within other disciplines and phenomena, this same aspect might be considered as irrelevant (i.e. traffic accidents, crime, etc.). In the next section (Section 5.3), we provide the theoretical and methodological description of a new methodology which has been specifically conceived to answer to this task (and if the reader is already aware of this procedure published in Araldi and Fusco 2019, we suggest to skip the next chapter). Here, in order to complete the overall picture of the indicators, we might summarise this methodology as follows: based on similarities between local point pattern statistical approaches (previously described) and recent measures of Local Fractal Dimensions (LFD, Taud and Parrot 2005, King 2014, Bruno and Costa 2011, Kanevski and

Pereira 2017) a specific Network Constrained LFD has been developed. This procedure implements the traditional *functional sandbox counting method* developed in fractal studies (Tél et al. 1989) at the three scales mentioned earlier. At the street scale, NC-LFD identifies fragmented or linear continuity of stores. At the neighbourhood and district scales, three geometrical layouts can be described: pervasive distribution, when the whole local sub-network is equally filled by stores. Linear distribution describes the presence of a commercial ribbon feature of retail compared to the overall distribution of the point subset. Finally, scattered store layouts are small and fragmented concentrations or isolated stores.

The next table (Tab.5) shows a condensed representation of the 27 indicators implemented in RFA procedure: when considering retail categories and the different spatial scales, each store is described by 27 categorical values describing local morphofunctional patterns.

Aspect	Indicator	Pattern	Name	
Functioning	Variety	Number of retail categories within 300 m	High / average / low retail variety at the neighbourhood level	
	Retail category prevalence	Ratio between the n° of stores of category(i) and total n° of stores within 300 metres	Prevalence of food stores	Rich300
			Prevalence of clothing stores	Food
			Prevalence of furniture related stores	Clothing
			Prevalence of vehicle-related stores	Furniture
			Prevalence of bar and restaurant	Vehicle
			Prevalence of hotels	BarRest.
			Prevalence of bank and insurance services	Hotels
			Prevalence of real estate agencies	Finance
			Prevalence of informatics and communication stores	RealEstate
			Prevalence of beauty and health services/stores	IT
	Prevalence of repairation services	Beauty		
	Prevalence of culture and leisure services/stores	Reparation		
	Anchor store category prevalence	Ratio between the surface of anchor stores of category(i) and total surface of anchor stores within 600 metres	Prevalence of food anchor stores	Leisure
			Prevalence of clothing anchor stores	Food_AS
Prevalence of furniture related anchor stores			Clothing_AS	
Prevalence of vehicle-related anchor stores			Furniture_AS	
Prevalence of home hardware anchor stores			Vehicle_AS	
Form	Size	Total number of stores within 150 metres	Relative high/average/low store presence at the street level	
		Total number of stores within 300 metres	Relative high/average/low store presence at the neighbourhood level	
		Total number of stores within 600 metres	Relative high/average/low store presence at the district level	
	Concentration	Ratio n. stores 150 m / n. stores 300 m	H-L, L-H, concentration between street and neighbourhood levels	R150
		Ratio n. stores 300 m / n. stores 600 m	H-L, L-H, concentration between neighbourhood and district levels	R300
	Geometry	Spatial layout of stores within 150 m	Fragmented-Continuous-Not Significant stores lay out at the street level	R600
		Spatial layout of stores within 300 m	Scattered-Linear-Pervasive-Not Significant stores lay out at the neighbourhood level	C150300
		Spatial layout of stores within 600 m	Scattered-Linear-Pervasive- Not Significant stores lay out at the district level	C300600

Table 5 Form and Function measures implemented in Retail Fabric Assessment methodology.



## 5.3 The geometry indicator

The increasing awareness of the importance of spatial relations in every discipline leads to the creation of ad hoc solutions and consequently to a large variety of methods for spatial pattern recognition (Perry et al. 2002). Each of them studies a particular aspect of spatial relations, which is necessary for the comprehension of the spatial behaviour of the phenomenon under investigation. Dale et al. (2002) analysed a large number of methods arising from several disciplines in order to identify similarities from both mathematical and conceptual points of view. They concluded that despite the partial redundancy of these methods, their combination is required for a better comprehension of different aspects of spatial structures. This remark legitimates the need to explore new approaches in point pattern analysis and to couple them with more traditional ones, with the aim of identifying new aspects of spatial distributions of events. Here, we will try to understand if a new approach, strongly built on well-established techniques, could lead us to better characterise and understand retail point geometrical distributions in urban space. As suggested by Perry et al. (2006) “*spatial statistics have most utility when used alongside a process-based or mechanistic investigation of the underlying processes driving the spatial patterns observed, whether experimental or model-based*”.

As previously introduced, in this dissertation we specifically focus on tests for the detection of clusters (Besag and Newell 1991) and, more in particular, on recent network-constrained extensions of these approaches. The implementation process described in these statistical procedures is also the starting point of other analyses performed in urban studies, based on fractal and configurational approaches, which lead to indicators describing different aspects of urban space complexity.

### **TEXT-BOX. 1: Counting methods for point pattern analysis**

#### **Point Patterns in Spatial Statistics**

A first synthetic descriptor of a point distribution is represented by its density/intensity, calculated as the ratio between the number of points and the size of the domain space. When analysing a network constrained distribution, the denominator is represented by the length of the analysed street network. This procedure gives us general information about the whole study area; applying the same ratio to subspaces, we can investigate how density and intensity vary in space. Local measures of density/intensity on networks can be obtained with several methods e.g. Kernel density (Okabe and Sugihara 2012), moving window on networks (Steenberghen et al. 2010), etc.

If we want to consider the reciprocal position of points, the implementation of second-order methods is necessary: nearest neighbour (or n-th NN), pair correlation function, Ripley's K and its derived functions. These methods are characterised by the study of the relative distribution of points around each event providing a distance-based function summarising the overall point pattern.

While NN, n-th NN, pair correlation functions study points with specific spatial conditions, Ripley's K function analyses patterns of all points at different scales. This function is obtained

considering the mean of the count of events within disks centred on each point; this count is divided by the overall density in the study region. Increasing the radius, we obtain a function describing the point pattern at different scales. In order to evaluate the statistical significance, K function considers as reference  $n$  realisations of a Poisson spatial process with known probability using a Monte Carlo sampling scheme. The definition of this process and its properties are widely discussed in the literature (ex. Stoyan and Penttinen 2000). The purpose of the  $n$  random realisations is to obtain significant upper and lower bounds for different radii of analysis: if the empirical K function of the point distribution falls inside this envelope, the hypothesis of randomness is satisfied. If the empirical value is higher than the upper bound for a given radius, our point pattern presents significant spatial clustering behaviour at this scale of analysis; if it falls below the lower one, we can infer the presence of regularity in our distribution.

Ripley's K is defined as a cumulative function; if the focus of the study is the variation of concentration at different radii, probability density functions  $g$  is implemented in the same way, considering rings instead of disks. When Ripley's K ( $g$ ) function is applied on a point pattern on a network, the traditional planar approach needs to be revisited: the spatial relation between points are distorted because of the network spatial constraints which reduce the spatial domain where events occur. In this case, the procedure presented above needs to consider network-based subspaces instead of Euclidean disks/rings (Okabe and Yamada 2001, Yamada and Thill 2004). A local version of the function K on networks has been proposed by Getis (1984; Getis and Franklin 1987) while its network constrained version has been developed by Yamada and Thill (K-LINCS, 2004).

In the research field of spatial economics, further developments of K functions consider how to evaluate the relative spatial concentration of specific categories of industrial sites: Cumulative Kd and Kd density function in Duranton and Overman (2005), M cumulative and m density functions Marcon and Puech (2009, 2017). The hypothesis underlying these works is that the set of existing sites represents the set of all existing and possible location for any plants. Consequently, the randomisation process of statistical significance tests of concentrations consists of reshuffling the relative mark values.

We should, however, keep in mind that the results of the K statistics (or its derivatives) are a function of the distance illustrating the overall clustering behaviour of the point pattern, by comparing it to the complete spatial randomness hypothesis. On the contrary, our research aims to identify local form descriptors of point distributions able to detect morphological properties and related variations along the street network.

What we have to retain from this discussion is the radial-centred, network-based approach used by all cumulative K function and derivatives, to count events at different distances and to inform us of the distribution of the points around any selected event. As for K-LINCS, we are interested in identifying local measures of clustering. In the next section, we will see how this approach is adopted in other methods, namely fractal analysis and network configuration analysis.

## **Fractal studies**

When analysing the spatial distribution of a human settlement, its social and economic functioning, independently from the scale of observation, we can observe apparently irregular distributions. Geographers, sociologists, and economists have tried to analyse these kinds of distributions (Frankhauser 1998). Under their apparent irregularity, a more complex spatial ordering has often been identified. Self-organisation phenomena have been associated with these complex forms and distributions (Tannier and Pumain 2005).

Fractals have been demonstrated to be a valid approach to the analysis of these irregular distributions. The impossibility of finding a regular pattern within the usual Euclidean geometry of discrete dimension (namely, Topological Dimension of the point =0, the curve =1, the plane =2, the volume =3) does not mean that regularities cannot be detected in a space possessing a Fractal Dimension (FD, also called Hausdorff-Besicovitch Dimension; Hausdorff 1919, Besicovitch 1964). Fractal analysis tries first to find a single parameter describing a repeating pattern at different scales, capable of explaining the apparent irregularity of a geometrical feature spatial distribution. In some cases, breaking points in the fractal dimension can be found over the scale range. The purpose of the analysis is then to find a simple set of parameters able to explain the complexity of the non-homogenous spatial organisation. Starting from Mandelbrot (1977), researchers widely apply the fractal approach in different domains, starting from physics and cosmology. When analysing urban systems and agglomerations, specific morphological parameters have been studied in the last 30 years (Fotheringham et al. 1989, Goodchild and Mark 1987, Batty and Longley 1994, White and Engelen 1993, Frankhauser 1994b, 1998. Salingaros 2003, etc.).

Based on the extensive scientific literature of urban fractal forms, we will focus our discussion on a particular formulation of fractal analysis. In this domain, analysis can have a global or local approach. In the first case, we can find grid, dilation, correlation and box (sandbox) counting methods (Mandelbrot 1983, Serra and Karas 1997, Tél et al. 1989, Frankhauser 1994b). These methods are traditionally implemented to assess fractal or multifractal analysis: a unique value is obtained (FD), used for describing the regularity of the whole distribution. More interesting for our purposes are the local methods. Local Fractal Dimension (LFD, Kanevski and Pereira 2017) can be obtained through the application of the box-counting method, radial analysis and curves of scaling behaviour have been applied in several domains of research to identify shape complexity variation, image texture classification and geometric composition. To name a few, measurement of surface roughness in geology (Taud and Parrot 2005), human cerebral cortex in medicine (King 2014), leaves classification in botany (Bruno and Costa 2001) etc. LFD may allow a more specific analysis, sensitive to complex changes in non-uniform distributions, avoiding problems related to spatial stationarity and inhomogeneity. As in the case of 2D/3D image texture analysis (magnetic resonance of brain, digital elevation models, satellite images, etc.), the application of local fractal measures to retail distributions may allow the identification and classification of retail textures.

Recently, Kanevski and Pereira (2017) applied the LFD approach to the case of forest fires in Portugal. In their work, the authors implement for the first time a local version of the sandbox counting method (Tél et al. 1989), namely functional sandbox counting method, to analyse local spatial patterns of clustering and compare the observed distribution to a randomly generated pattern on the validity domain. Both traditional and functional sandbox counting methods share the same four-step procedure: definition of circles centered in each and all events with increasing radii  $r$  (i); counting of the number of events  $N$  in each circle (ii); plotting the functional dependence of  $N(r)$  versus the radius length (iii); estimation of the slope on  $\log[N(R)]$  versus  $\log[R]$ . While the global approach is performed on the average  $N(R)$  in order to obtain a global measure, the local approach is applied individually to each event in order to obtain a local fractal dimension.

In urban studies, the reticular constrain domain of socio-economic phenomena suggests the application of the network-based approach (as seen in K functions in the previous chapter). This method has been studied in transportation field and in accessibility measurements (Gouyet 1992; Genre-Grandpierre 1999; Frankauser et al. 2007): starting from a given point on a network, these researchers analyse the amount of building surface within a network-based service area at different reticular lengths.

The application of the local sandbox counting method to the case of retail distribution on the street network shares some similarities with the statistical method of K function. They both consider a specific point as centre of the analysis, and, for different radii  $r$ , they count the number of elements falling inside each network-constrained disk (service area). The fractal law that we obtain through this analysis is given by a regression line on a scatterplot with logarithmic axes, plotting the radius of calculus  $r$  and the corresponding number of elements  $N(r)$ :

$$N(r) = \beta r^D \quad (5.1)$$

$$\log N(r) = \log(\beta) + D_r \log(r) \quad (5.2)$$

From (5.2), we obtain the LFD  $D_r$ ; this parameter resumes the trans-scalar spatial organisation of the geometrical pattern around the centre of analysis. As suggested by Frankhauser (1998, p. 217), this local method is a detailed approach in studying variation of local spatial patterns *“by comparing the results obtained for different counting centres”*; additionally, he described this approach as an instrument for research for the analysis of empirical structures through their comparison and classification.

As we can see, this procedure is equivalent to the one described above in Ripley’s K function in a network space (K-LINCS). The counting method with disks of increasing radii is applied by fractal analysis to identify the parameter synthetically describing the apparently uneven distribution. Like the K function, we need to apply the Local Fractal Dimension on a Network Constrained Domain (NC-LFD): the network approach preserves the same structure of equation (5.2).

### Street Network Configurational Studies

The same principle of non-homogeneity in the spatial distribution of urban, social and economic patterns discussed before is at the centre of network analysis. Based on the seminal work of Bavelas (1948) on CPT in social networks, several indicators of network properties have been developed (Leavitt 1951, Shimbel 1953, Shaw 1954, Freeman 1977, etc.). This approach has been later applied to urban systems and street networks.

As introduced in Section 3, within the study of the relationships between urban network configurational properties and the localization of social and economic activities within a city, *The social logic of space* (Hillier Hanson 1984) and *Space is the machine* (Hillier 1999) are unanimously considered the milestones of this research field and, more particularly, of the analytical methods of Space Syntax (SSx). Porta et al. (2006) identified two approaches of configurational analysis based on two different representation of the network: the dual approach applied in space syntax studies, where street segments are modelled as nodes and intersections as arcs (Hillier et Hanson 1984; Hillier 1996, 1999); and the primal approach, base of Multiple Centrality Assessment (MCA), with an opposite representation (Batty 2005; Porta et al. 2006; Sevtsuk 2010). This approach has been widely applied in the last few years to different urban phenomena in order to study their highly heterogeneous spatial distribution: property price, office rent, land-use intensity, and others (Tab.1-2).

Similarly to statistical and fractal analysis, we now focus our attention on a particular indicator within configurational analyses. When we count the number of nodes (links, destinations, etc.) attainable from a certain location, we are studying an accessibility measure which is a function of the connectivity of a point within the network: Reach in MCA, Node Count in SSx. Their formulation is of interest when they are calculated locally within a given radius:

$$Reach(r)_i = \sum d[i, j] \leq r(j) \quad (5.3)$$

Their implementation procedure is--once again--the same we encountered in Ripley's K function and local fractal analysis. Configurational analysis, unlike fractal analysis, compares the centrality value in the space at a selected scale (given radius r). Some studies try to compare the relation of these properties at different scales or at different times (Barthelemy 2011). Unlike K function and fractal radial analysis, though, configurational indicators were conceived in the first place for network analysis and do not need any particular adaptation to network environments.

When observing the procedures developed in statistical, fractal and configurational approaches, we might observe a substantial methodological similarity. Local measures of spatial patterns are based on the counting method applied to disks obtained by increasing radii around each point of the analysis; while the configurational analysis focuses its attention on a particular radius  $r$ , the fractal approach and K function analyse the evolution of the count with increasing size of  $r$ . Moreover, fractal analysis summarises this evolution with the fractal dimension  $D_r$ .

From this observation, a new procedure is here developed for the development of local measures for the detection of the geometrical layout of stores, similarly to the Local Fractal Dimension (LFD) developed by Kanevski and Pereira (2017).

LFD is based on the local implementation of the functional sandbox counting method<sup>54</sup>, at every point of analysis, with different radii within the range  $[1:r_{\max}]$ . Differently from Kanevski and Pereira (2017), this method is applied here on a street network and the resulting NC-LFD is used for a straightforward classification of the geometrical configuration of points depending on the closest topological dimension.

Functional sandbox method approach is implemented to the pattern of stores on networks: for each point and for each radius  $r$ , we calculate the number of events  $N_r$  at network distance less than  $r$  from the counting centre. From equation (5.2) the parameter  $\alpha = D_r$  is obtained for different  $r$ -max; this value might vary between 0 and 2. The application of this fractal approach to the point pattern gives an indication of the local proportional factor. When  $\alpha \rightarrow 2$  the number of points grows almost as in a two-dimensional homogeneous plane and many more event points are at larger distances than at smaller ones. When  $\alpha \approx 1$ , the distribution of points is aligned on a line<sup>55</sup>. When  $\alpha \rightarrow 0$  faraway points are so rare that the centre of the analysis is almost isolated.  $\alpha$  is relatively low even when the point of analysis is within a dense point cluster, whose dimension is much smaller than  $r$ . In this case, the cluster, and not the single point, is isolated from the rest of the point distribution.

The study of the distribution of points in space is characterized by higher irregularity and discontinuities. This might result in strong fractal transitions at different radii. For instance, if we consider a small scale where  $r \rightarrow 0$ , then  $\alpha \rightarrow 0$  (the Euclidean dimension of a point is 0). By growing the distance  $r$ , the value of  $\alpha$  could persist equal to 0 until other points are encountered. Depending on the spatial concentration and density at different  $r$ ,  $\alpha$  could increase assuming higher values (up to 2). The study of the variation of  $\alpha(r)$  with the local approach could reveal the relative scale at which the point distribution changes locally: these variations are called fractal transitions (Nottale 1993, Forriez et al. 2010)<sup>56</sup>. The coefficient of determination  $R^2(r)$  obtained from the (5.2) is used here as a significance filter: when this parameter is lower than a certain threshold,  $\alpha$  is not representative of the local fractal

---

<sup>54</sup> Sandbox counting method has been often applied to networks for fractal and multifractal characterization (i.e. Liu et al 2015). In this paper we apply the *functional* version as in Kanevski and Pereira (2017). Moreover, the centroids of the implemented sandboxes correspond to the events of our point pattern analyzed, instead of the edges of the network.

<sup>55</sup> If  $D_r = 1$ , the point process is described by a Poisson process of intensity equal to the ratio of the number of points on the line on which occurrences are aligned (for the demonstration see Stoyan et al. 1987 and Stoyan 1994)

<sup>56</sup> The precise study of the transition distances goes beyond the scope of this work; more insight and methodological procedures are discussed in Nottale (1993) and Forriez et al. (2010).

dimension at the given distance  $r$ . This results from the presence of a local fractal transition within distance  $r$ .

Before analysing a real case study, we apply the aforementioned procedure to a simple model in order to better understand the peculiarity of the NC-LFD  $\alpha$ . Three hundred points are uniformly distributed on some links of a regular meshed grid, specifically selected in order to recreate three different ideal geometrical patterns (Fig.27a).

Firstly, the Global K function is calculated. Fig.28.b informs us about the presence of a statistically significant clustering behaviour: indeed, the observed function lies outside the envelope generated by one thousand Monte Carlo random simulations (5% significance). As previously discussed, Ripley's K function describes the overall point pattern distribution. This graphic does not allow a local description of the clustering behaviour around each point.

Second, we apply the local K function: Fig.27c-e show the observed functions for the three points  $i, j, k$  extracted from the centre of the three simulated configurations. These graphics illustrate the distance at which the statistical significant agglomeration is detected. For example, considering Fig.27b, we deduce the presence of a cluster within a small distance (<400m) from point  $i$  and a second significant agglomeration at higher distances (>1000m). Similarly, Fig.27c and Fig.27d show the relative distance of the simulated structures from, respectively,  $j$  and  $k$ . Although local K function permits a finer spatial description of the agglomeration behaviour of a point pattern, nothing could be said about its geometrical arrangement in space.

NC-LFD is then applied considering  $1 \leq r \leq 400$  metres. The resulting local fractal dimensions range from 0 to 2 (Fig.27). As we can observe in Fig.27b-d, the local fractal dimension  $\alpha$  corresponds to the slope of the local K function within a given radius in a log-log plot (Kanevski and Pereira 2017). The segmentation of this value based on fractal theory allows the description of the geometrical layout of our point pattern: two "point" concentrations, a "linear" structure and last a "planar-like" distribution that covers uniformly a subset of the grid.

Since all the cumulative functions are implemented on networks, this method is able to identify linear structures along networks. The fractal dimension  $\alpha$  would vary depending on the disposition of the underlying network; from 0, perfectly linear to 2 when the street layout strongly deviates from a straight line and covers uniformly the planar space. Nonetheless, as is well-known from the fractal literature, the fractal dimension of a network is described by a measure between 1 and 2, traditionally close to 1.6 and as similarly supported by the test in Fig.27.

Let us implement NC-LFD to our real case study where the store point pattern distribution on the street network is considered.

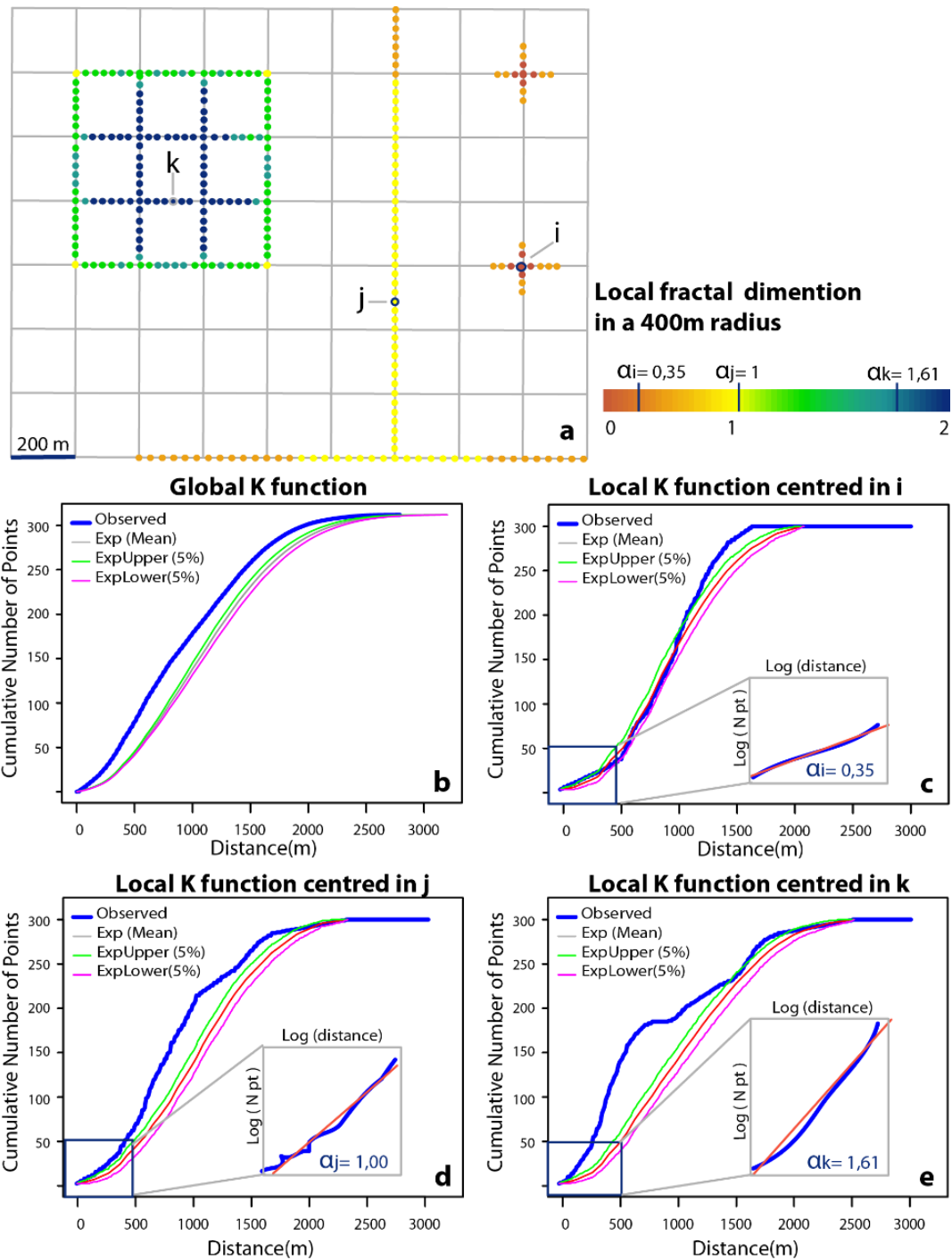


Figure 27 NC-LFD (a), Global K (b) and Local K (c-e) for three different theoretical layouts of retailers disposed over a regular meshed grid.(source: Araldi and Fusco 2019)

The value  $\alpha$  is implemented for each point considering street, network and district scales (respectively, 150-300 and 600 metres), with steps of 25 metres. In Fig.28, we can visualise the spatial distribution of  $\alpha$  (a) and  $R^2$ (b): the parameter  $\alpha$  varies between 0, when the retail-point is isolated, and 1.3 when the filling of network space by other retailers around the point is higher.

The local fractal dimension is used to study and to classify surfaces and image textures in other disciplines. In the case of retail distribution, the interpretation of the local fractal dimension, like in its original conception (Husdorff 1919), could lead to a segmentation of the values  $\alpha$ .



Considering this parameter as a geometrical description of the local point pattern, we can thus identify three typologies of retail textures or retail fabrics:

- when  $\alpha$  is close to zero, scattered retail fabric: this distribution is characterised by individual fragmented activities or small concentrations dispersed along the network;
- when  $\alpha$  is close to one, linear retail fabric: this distribution is represented by an alignment of retail points along a subset of streets which constitute in their turn an almost linear pattern.
- when  $\alpha$  is higher than one, pervasive retail fabric: this class is defined by a continuous penetration of retail presence along the whole urban street network.

The coefficient  $R^2$  informs us about the goodness-of-fit of these three patterns and it can be used to filter those points where the spatial layout cannot be well characterised with our method due to the presence of strong heterogeneities (Fig.28b).

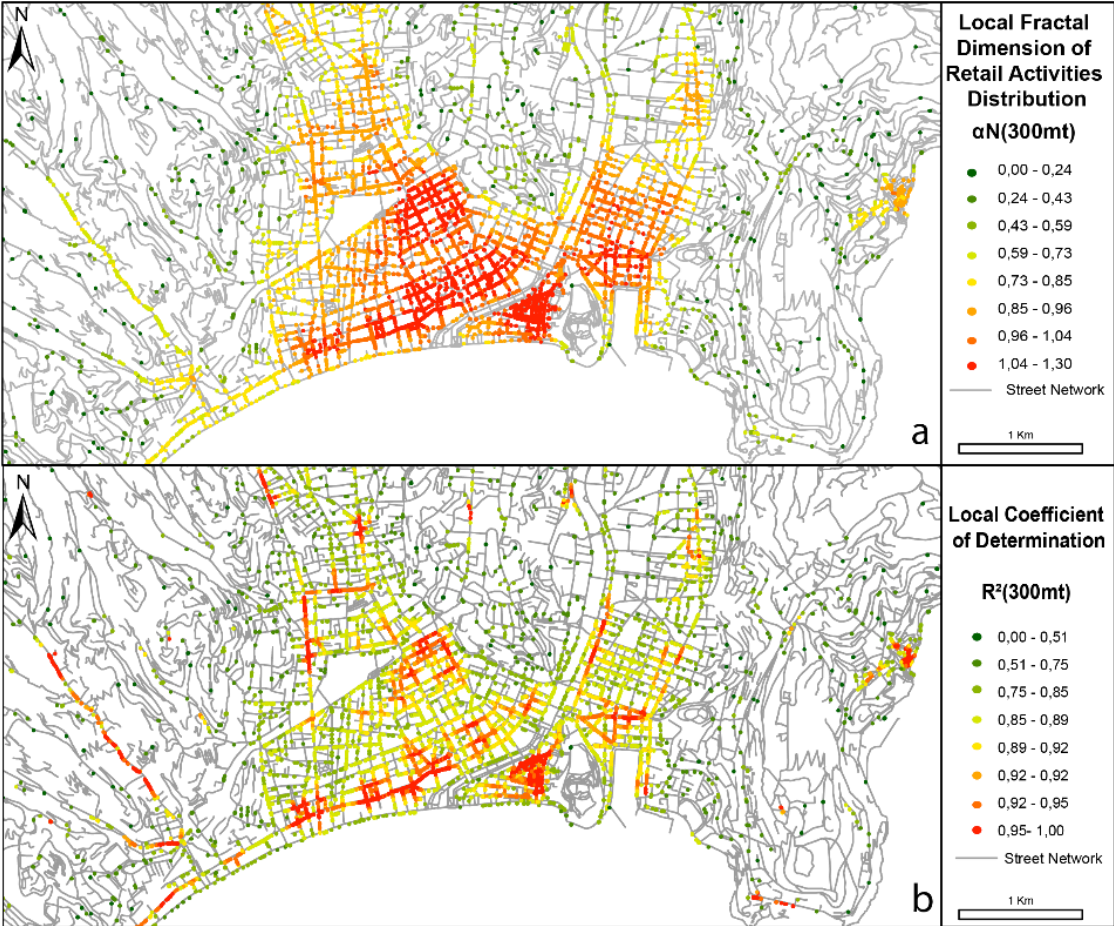


Figure 28 Retail activity fractal dimension, coefficient of determination and local relative concentration, calculated in a 300 m radius. (Source: Araldi and Fusco 2019)

The interpretation of the three different patterns depends in turn on the scale of the analysis (defined by the radius of calculus  $r$ ). The segmentation of retail activities at different scales is represented in Fig.30. All five images represent the classification with the same goodness-of-fit ( $R^2 \geq 0,85$ ) and with the same class repartition: scattered retail fabric [0-0.8], linear retail fabric [0.80-1.20], pervasive retail fabric (1.20-2).

At the street scale of 150 metres (Fig.29a) only two classes are identified: the uninterrupted linear presence of retail activities along some streets (linear retail fabric), opposed to others where gaps among retailers produce a scattered retail fabric. Both at the neighbourhood and district scales, the three classes of retail fabrics have different meanings: the scattered retail fabric at mesoscales of 300-600 metres is made up of isolated concentrations of retail activities. These small concentrations are typical of small villages scattered in the countryside (Fig.29b). Linear retail fabrics occur for example in residential or mixed-land-use areas where retail activities are concentrated along main axes, composing commercial strips. Linear configurations do not characterise exclusively suburban and peripheral areas: old cities like Antibes and Menton present one or more commercial axes within the downtown area, which represent the backbone of their commercial activity (Fig.30). Pervasive retail fabrics cover well-meshed street networks with a continuous presence of retail activities, typical of urban central and historical areas. The pervasiveness of retail presence in the street network depends on the scale of analysis: discontinuities of a few street blocks with fewer commercial activities are for example overlooked already at 600-meter scale.

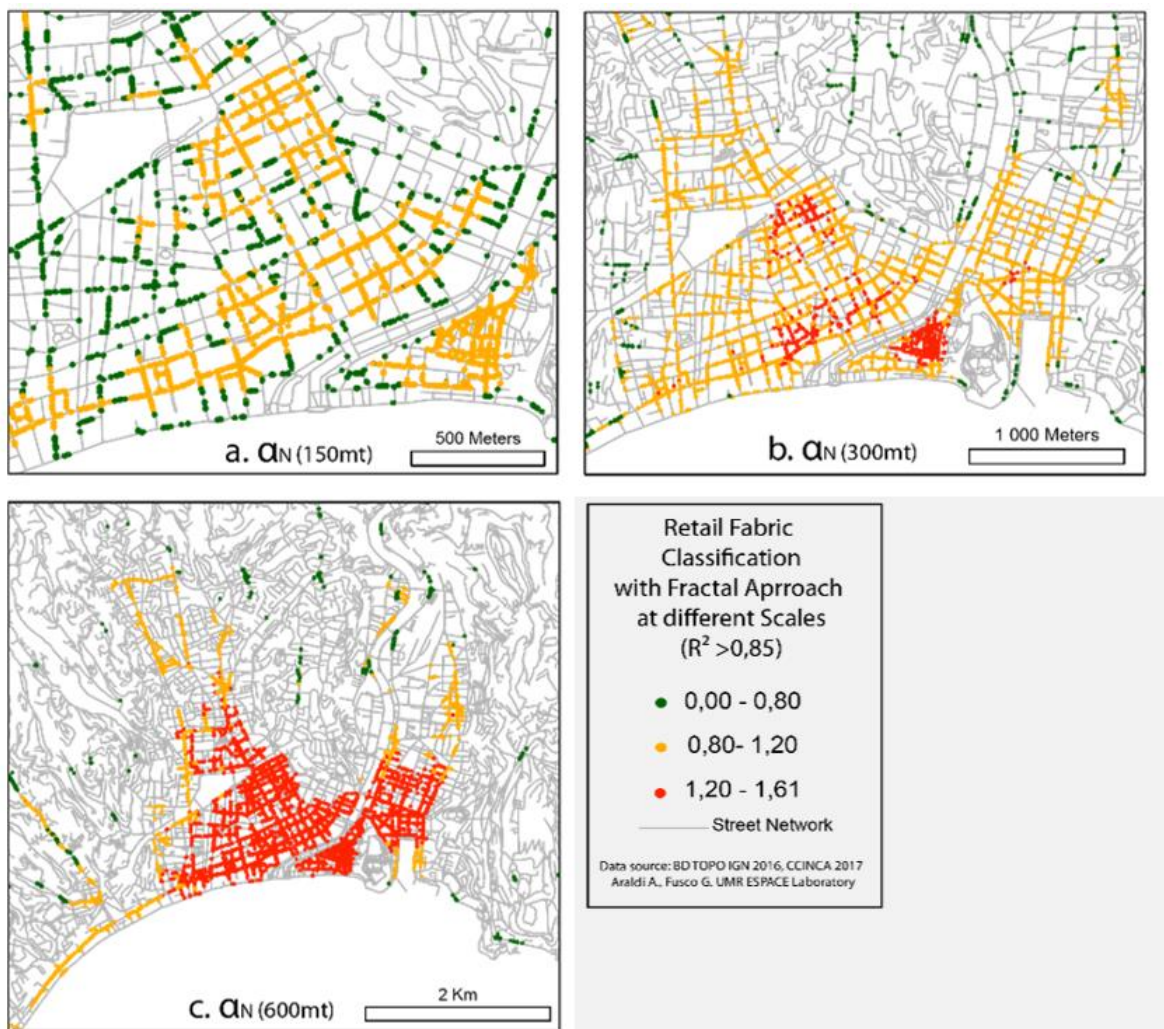


Figure 29 Classification of the retail fabric at different scales. (Source: Araldi and Fusco 2019).



Figure 30 Retail fabric classification in Menton (left) and Antibes (right). (Source: Araldi and Fusco 2019)

The computational limits given by the large data-set size used in this study do not allow the implementation of K functions (or derivatives) for comparative analysis. Nevertheless, in Fig.31, it is possible to visually compare the results of the G-LINCS method applied to retail densities on networks and the fractal approach here proposed (at the same scale of 600 metres). G-LINCS method defines High, Low and Non-Significant concentration segments: the first group includes 83% of what is here defined as pervasive retail fabric and 60% of the linear retail fabric. Non-significant clusters cover 23% of the retail pervasive class and 77% of those regions where the fractal approach is still able to identify linear and scattered retail fabrics (the highlighted section in the north of Nice is a good example of a linear retail fabric within a relatively well-meshed street network, which remains undetected by the geostatistical approach). A no-retail area is detected by the G-LINCS Low cluster. Despite the fact that these two methods do not share the same objective (verifying spatial randomness of retail distribution vs. characterising geometrical patterns of retail distribution), it seems to us that the fractal approach represents a better-suited alternative in the specific case of retail geometry pattern classification.

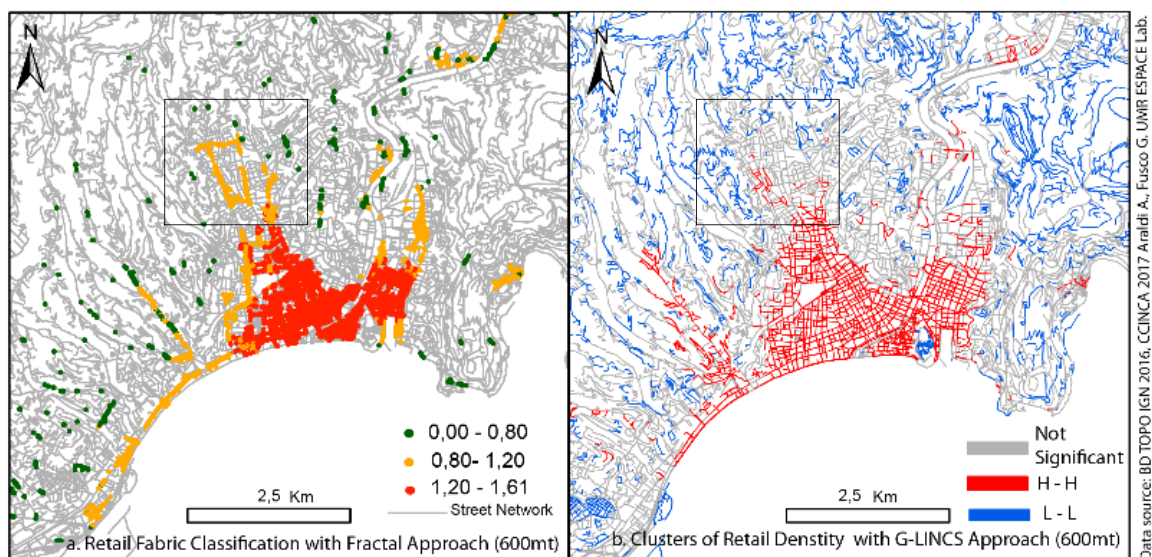


Figure 31 Retail fabric classification through NC-LFD (a) and G-LINCS (b) (Source: Araldi and Fusco 2019)

## 5.4 Classification

In order to identify retail fabrics, we still have to combine the results of the geo-statistical and morphology patterns; the third and last step consists of studying how these patterns overlap to yield specific combinations. The different retailscapes which might characterise streets and neighbourhoods of a settlement might be defined as those urban spaces where persistent co-occurrence of given morpho-functional characteristics might be observed. At the same time, a few key characteristics could identify a given retail fabric even in the presence of a certain variability of other indicators. For example, the geometrical indicators might help in the identification of a specific agglomeration type despite the high variability and mix of other functional descriptors; on the contrary, the overrepresentation of one specific functional indicator might be responsible for the identification of specialised retail fabrics.

Clustering approaches aiming to achieve high intra-cluster homogeneity based on variance minimisation of the values of all the indicators (like k-means, SOM and geo-SOM) are not well fitted to this task.

The Naïf Bayes Classifier (NBC) is applied here.

The Bayesian inference method is based on conditional probabilities (Duda 2001); this method tries to minimise the probability of error in a decision, by associating to each individual the most probable outcome. This methodology is based on the Bayes' Theorem and it allows each spatial unit  $PBi$  to be described with a set of probabilities of belonging to each Retail Fabric class  $RF_j$  considering its description given by the set of attributes  $A_{1i}, A_{2i}, \dots, A_{ni}$  (morphological patterns).

$$p_{PBi}(RF_j | A_{1i}, A_{2i}, \dots, A_{ni}) = \frac{\prod_{k=1}^n p(A_{ki} | RF_j) p(RF_j)}{p(A_{1i}, A_{2i}, \dots, A_{ni})} \quad (5.4)$$

with the constraint

$$\sum_j p_{PBi}(RF_j) = 1 \quad (5.5)$$

Two main reasons led to the decision to use these formulas: first, Bayesian clustering allows the identification of groups of individuals sharing a few key characteristics among the set of attributes  $A_1, A_2, \dots, A_n$ , which could vary from one group to another (as discussed in Fusco 2016). Secondly, Bayesian probabilities can address the uncertain assignment of observations: hybrid retail agglomerations (both from the form and functional point of view) could have non-null probabilities of belonging to different retail fabrics. As observed by Mackaness and Chaudhry (2011), Bayesian clustering methods in retailing studies might be a good option considering the fuzzy and incomplete knowledge of the domain.

The careful theory-driven feature selection of our analysis justifies the use of a simple naïf Bayesian classifier on the 27 morpho-functional indicators. Differently by the traditional approaches, in our methodology, the clustering procedure is applied on categorical values resulting from geostatistical analysis and morphological-based segmentation, instead of raw indicator values describing the internal characteristics of each spatial unit.

For 25 random seeds, 1000-step random walks explored the solution space to obtain optimal Bayesian clustering of our spatial units. The search constraints were a minimum cluster content of 2% of the spatial units, a minimum average probability of assignment of units to each cluster of 0.9 and a maximum of 25 clusters. The evaluation of the optimal classification is defined by minimising Akaike's Information Criterion (AIC, Akaike 1974) which combines log-likelihood of the clustering solution given the data and penalisation for the increasing number of clusters. The best AIC score obtained corresponded to an 18-cluster solution. A successive refining phase of the 18-cluster solution was carried out by imposing this fixed number of clusters to 6 independent random walks. The best result shows a contingency table fit score of 67.77%. By construction, the 18 clusters are spatially coherent retail fabrics.

The relatively important number of classes identified by the Bayesian clustering is a consequence of the variability within the set of indicators and their internal distribution. A correct number of classes is not determinable a priori. As observed by DeLisle (2005) a "dynamic tension between too few and too many classes" exists when applying classifications to retail activities. The applied Bayesian protocol is an attempt to use unambiguous, meaningful and measurable metrics to obtain an acceptable compromise. As we further discuss in the conclusion section, the recent work of Dolega et al. (2019) has similarly identified 15 retail sub-classes grouped into 5 macro categories.

The main result of the Bayesian classification is represented by the assignment of each individual store to its most probable cluster. In addition, a rich set of probabilistic information is provided to the analyst to identify and describe each class and the role which each indicator has played in the classification procedure. When projected in the geographic space, the interpretation of these outcomes offers a synthetic model of retailscapes. In the next section, we explore these outcomes in order to provide an attentive description of each class.

## 5.5 Results

The Bayesian Network (BN) approach has several advantages compared to other classification methods: beyond merely producing predictions/classification, BN classification helps the analyst in generating a deeper understanding of the underlying problem domain (Conrady and Jouffe 2013). The most relevant feature of BN classification is the possibility to precisely quantify parameters of *Mutual Information* describing the importance of individual variables (or combination of a subset of variables) for the definition of the overall classification as well as for each individual class. A brief explanation of this measure is here provided.

## TEXT-BOX. 2: Mutual Information

While correlation indicators are traditionally provided in multivariate statistics, similarly, in BN probabilistic environment the specific notion of Mutual Information (I) is implemented. While the former is a measure of the linear distance between two random variables, the latter is a measure of the distance between two probability distributions and has its roots in information theory. Both these two approaches express the degree of relationships between the distributions of two discrete variables X and Y with probability mass function  $p(x)$  and  $p(y)$ , and joint probability mass function  $p(x,y)$ .

The linear correlation (such as Pearson's correlation) is based on the measure of covariance  $Cov(X, Y)$  obtained as:

$$Cov(X, Y) = E(XY) - E(X)E(Y) = \sum_{xy} p(x, y)xy - (\sum_x p(x)x) \cdot (\sum_y p(y)y) \quad (5.6)$$

$$Cov(X, Y) = \sum_{xy} [p(x, y) - p(x)p(y)]xy \quad (5.7)$$

Similarly, the Mutual Information is implemented as:

$$I(X, Y) = E \left( \ln \frac{p(x,y)}{p(x)p(y)} \right) = \sum_{x,y} p(x, y) [\ln p(x, y) - \ln p(x)p(y)] \quad (5.8)$$

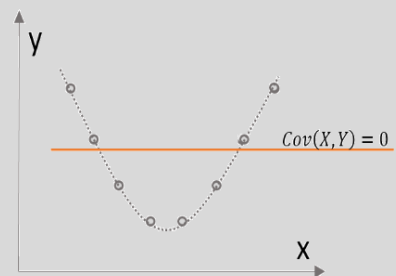
Also described as the difference between the marginal entropy of X,  $H(X)$  and its conditional entropy given Y,  $H(X|Y)$ :

$$I(X, Y) = H(X) - H(X|Y) \quad (5.9)$$

where the entropy  $H(\cdot)$  is obtained as follow:

$$H(X) = -\sum_x P(x) \log P(x) \quad (5.10)$$

The formulae in 4 and 5 show how both  $Cov(X, Y)$  and  $I(X, Y)$  provide two similar measures of distance from the independence expressed as the distance between the joint Probability Mass Function (PMF) and the product of the marginal PMFs. The main advantage of  $I(X, Y)$  is that we do not assume a linear correlation between two variables: on the contrary, when  $Cov(X, Y)$  describes a correlation value equal to zero, the two variables might be still stochastically dependent.



The Mutual Information might be also interpreted as the amount of information of Y which we would gain on average by observing X as a predictive variable; this average information gain would reflect the predictive importance of X. In our classification, for example,  $I(R300, RF)$  should be interpreted as the amount of information of the Retail Fabric classification which is predictable only through the knowledge of variable R300.

Firstly, the projection of our 27 variables in the mutual information space (Fig.33) provides a visual representation of the degree of similarity of the information contained in each variable. The absence of superposition among variables indicates few redundancies: nonetheless, the heterogeneous distribution highlights some differences among form and functional indicators. While indicators of small surface prevalence and distribution are well spread, denoting specific dissimilar patterns, two groups of indicators show a relatively higher redundancy: anchor store prevalence indicators on one side while, on the other, we find distribution, geometry, and variety indicators. These two groups of indicators also have the highest mutual information with the cluster variable: in Table 6 we reported the mutual information that variables share with the classification when individually considered. Size, variability, clothing, large surface prevalence, and the geometry at the neighbourhood level represent the five most determinant variables.

Despite unsupervised BN learning methods being able to handle redundant variables without any previous correlation reduction (Zhang 2004), the use of the redundant variables can conceal the role they play in the clustering analyses. From this point of view, mutual information with the target variable alone is of little use: if two variables show both high mutual information with the cluster variable, they could also share mutual information between them and eliminating one of them could have little impact on the final clustering result. A way to evaluate the parsimony of the clustering model is to force the Bayesian inference of the cluster values using only a subset of the original variables. We did this in a 10-fold cross-validation scheme, which also allows for the assessment of overfitting problems in the clustering model.

Using the same learning algorithm as before, we randomly divided our database of 33 thousand stores in 10 folds of 3300 stores, using nine of them for each test as the learning set to produce the clustering model and the remaining one as a test set to assess the clustering accuracy of this reduced model (we assume the clustering result of the complete model to be the reference). The clustering accuracy of the reduced models is assessed by forcing the Bayesian inference to use only 10 variables (Bayesian algorithms can identify the combination of variables producing the most accurate predictions). This means that for each record and for each model, a different combination of 10 most informative indicators might be used to best predict the cluster variable. The overall evaluation of the 10 collectively most informative variables selected by the cross-validation procedure identifies the indicators in Table 5 in descending order of frequency of observation. This procedure of 10-fold cross-validation, with inference of the cluster values limited to the use of the 10 most informative variables, shows an overall accuracy of 94.9%.

While in Tab.5-left we might observe the rank of the values of the mutual information of each variable individually assessed, the Tab.5-right shows the frequency of observations of the 27 variables implemented in the 10-fold cross-validation with limited observations used in the inference procedure. In this second column, since the variables have been collectively evaluated, higher positions are occupied by those variables showing less redundancy, and, therefore, a stronger power of identification and differentiation among classes. Although the two values are not directly comparable, we might observe the relative positioning of the two variable rankings.

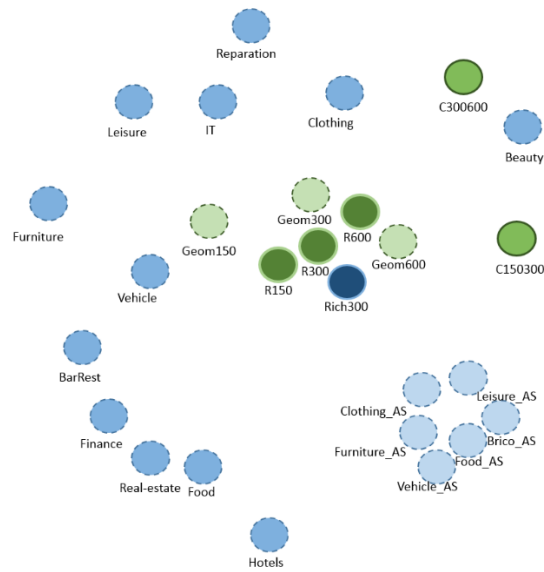


Figure 32 Projection in mutual information space of the variables (form-green, function-blue)

Indicator Name	Mutual Information (%)	Frequency of Observation (%)	Indicator Name
R300	35,38	100,000	R300
Rich300	33,09	100,000	Clothing_AS
Clothing_AS	32,88	79,001	R150
R600	32,31	73,559	R600
R150	30,98	63,613	Food
Geom600	30,20	62,892	Real_estate
Leisure_AS	29,51	61,668	Rich300
Brico_AS	27,28	60,513	Geom600
Vehicle_AS	27,21	48,330	C300600
Food_AS	26,84	43,445	Clothing
Furniture_AS	26,60	35,759	C150300
Geom300	23,70	31,587	BarRest
Vehicle	16,95	29,496	Vehicle
Geom150	16,57	27,492	Geom300
Clothing	14,66	26,791	Geom150
Food	14,14	25,309	Beauty
RealEstate	13,97	17,409	IT
C150300	13,35	16,919	Finance
BarRest	11,53	14,903	Leisure_AS
IT	11,19	10,698	Leisure
C300600	10,73	10,055	Food_AS
Finance	10,56	9,665	Reparation
Hotels	7,71	8,246	Hotel
Beauty	7,38	3,624	Furniture_AS
Reparation	7,18	3,502	Furniture
Leisure	6,64	0,828	Brico_AS
Furniture	3,21	0,171	Vehicle_AS

Table 6 RFA Indicator mutual information with cluster variable (left) and frequency of observation among the 10 collectively most informative variables (right)



What we might observe is a swap in the relative importance between large and small surface functional prevalence indicators. This might be due to the spatial clustering of anchor stores which result in similar overlapping patterns, determinant only for a smaller number of retail classes (as already described in our study area as an important characteristic of those interstitial peripheral areas of the French Riviera conurbation, see Section 3.2). What the higher position of food-stores and real estate agencies in the right column shows is that these variables add essential information to the clustering that was not integrated in the many redundant variables preceding food-stores and real estate agencies in terms of mutual information with the cluster variable. We see here all the interest of the Bayesian approach: in a variance-minimising approach (as k-means, SOM, etc.), the contribution of these variables would have been overlooked and the clustering results would have reflected the role played by the group of redundant variables.

The clustering outcomes of stores we firstly observe to the projection of the optimal classification in the mutual information space (Fig.33): the size of each dot is proportional to the number of stores attributed to each group, while their spacing highlights the relative information similarity. 76% of the stores under analysis is divided almost equally grouped in 17 classes between a minimum of 2.1% and a maximum of 8.2%; one class gather almost 24% of stores and represents the more isolated stores and services and therefore which could not be considered as integrated within a specific retail agglomeration type.

After the attentive analysis and description of each class that is further described in the second part of this chapter, we might anticipate a few main features which might help the reader in the apprehension of the classification.

Starting from the upper-left part of Fig.33, we observe those retail fabrics (RF 1, 2, 3) which are usually found within the core of larger cities. Within these areas, the high values of retail density and pervasiveness allow the coalescence of different subspaces with specific specialisation (i.e. clothing, service, touristic, etc.); functional indicators become of paramount importance for the distinction between different RF classes. From the core centre, moving clockwise, we gradually shift towards semi-central areas (RF 4, 5) and specialised peripheries (RF 6). In the middle of this upper part of the mutual information space, a specific class (RF 7) might be observed: its central position is indicative of the co-occurrences of all characteristics of the surrounding classes. Indeed, this specific retail fabric is found in those smaller cities where form and functional characteristics collapse in a single centre where functional specialisations, if present, are only perceivable in few streets but not in a significant extension which might be recognised as a specific retail agglomeration (district). We proceed with the clockwise exploration of the mutual information space, and from peripheral agglomeration we find isolated stores. In turn, closing the circle, there appear small isolated retail agglomerations (RF 8, 9) which might be observed for example in small villages. While in the upper part off Fig.33 we find those stores counting in their surrounding areas the presence of large retail stores, in the lower section, on the contrary, we find those stores located further away from large retail surfaces.

While the projection of the RF in the mutual information space allows a brief overview of the overall classification, a more detailed description of each class is still needed; the BN classification provides four main outcomes expressed in probabilistic terms supporting this specific task.

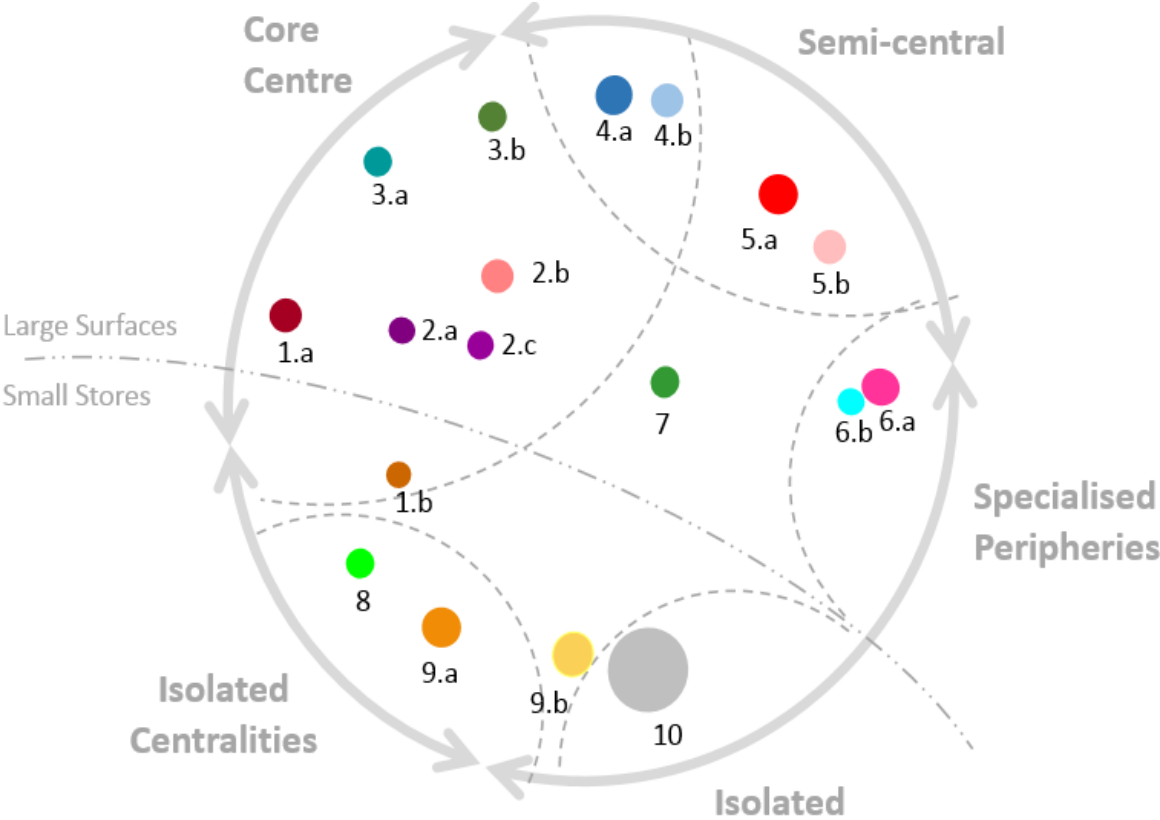


Figure 33 Projection in mutual information space of the clusters

The same quantitative description of the role of each variable in the overall classification might also be obtained for each class: this information is collected in Table 7 where each cell informs us about the targeted mutual information  $I(\text{Indicator}, \text{Class})$ . These values might be interpreted as follows: when considering, for instance,  $I(R300, RF1.a)$ , we might say that by knowing the variable R300, we have a 25% probability of correctly identifying if that store belongs to RF 1.a. These values are disposed in the same descending order defined by the indicators' mutual information for the overall classification; this specific disposition allows to easily identify which subgroup of indicators is the most correlated to each RF. Nonetheless, as previously observed, the Mutual Information evaluates indicator significance individually without considering the possibility of their combined effect.

Store Classification	100,0	4,7	2,5	5,0	2,1	2,5	3,5	3,2	6,0	4,5	7,3	5,0	6,2	2,1	3,2	3,1	7,0	8,2	23,9
Indicators	1.a	1.b	2.a	2.b	2.c	3.a	3.b	4.a	4.b	5.a	5.b	6.a	6.b	7	8	9.a	9.b	10	
R300	35,4	25,0	19,5	25,4	20,2	21,2	23,0	22,5	26,9	16,3	32,1	30,1	26,3	12,4	18,1	16,2	39,6	15,6	56,5
Rich300	33,1	15,2	13,6	18,0	17,5	20,7	22,6	22,2	23,6	15,7	15,9	20,6	24,8	15,4	15,9	11,2	24,7	34,3	76,4
Clothing_AS	32,9	15,2	17,6	46,5	42,6	19,1	33,8	13,8	22,5	16,2	25,8	21,2	19,9	12,9	20,3	18,5	23,6	24,8	43,8
R600	32,3	19,3	16,2	29,2	23,1	20,9	26,4	25,8	30,0	16,9	30,7	21,3	18,1	8,9	25,9	16,2	25,9	11,8	52,7
R150	31,0	26,9	8,7	25,7	23,7	21,3	25,9	19,4	20,5	23,1	26,0	14,6	13,8	12,4	22,6	9,8	29,5	12,1	56,6
Geom600	30,2	18,1	8,5	32,4	25,6	14,3	28,3	25,0	20,9	10,9	37,2	7,0	18,4	7,3	22,3	20,0	23,2	15,6	48,2
Leisure_AS	29,5	17,1	17,4	9,5	37,7	16,2	36,9	19,3	23,7	16,2	21,9	18,5	16,8	13,4	14,8	18,3	23,3	24,5	43,2
Brico_AS	27,3	17,1	17,6	16,2	12,5	18,2	15,3	16,1	23,7	15,1	17,4	15,5	16,5	12,3	16,3	18,5	23,5	24,8	43,7
Vehicle_AS	27,2	15,2	17,6	18,9	15,6	16,0	14,3	15,6	17,9	15,4	17,4	16,0	16,5	18,1	17,2	18,5	23,5	24,8	43,8
Food_AS	26,8	16,3	17,2	15,3	29,4	25,0	11,0	13,5	13,8	15,8	17,5	15,2	14,9	20,9	15,9	18,1	23,0	24,3	42,8
Furniture_AS	26,6	16,6	17,2	16,8	26,7	16,4	13,3	16,2	17,1	14,0	16,2	14,7	16,8	12,8	16,1	18,1	23,0	24,3	42,7
Geom300	23,7	19,0	9,0	23,4	49,4	17,3	16,8	15,0	18,6	2,2	20,2	16,7	18,1	7,0	18,6	13,3	2,6	18,9	34,3
Vehicle	16,9	28,1	11,6	28,6	22,7	23,6	25,9	17,3	0,7	1,6	7,0	6,6	12,9	8,0	3,4	10,0	2,1	7,1	23,0
Geom150	16,6	20,3	3,9	18,7	21,7	18,2	19,5	12,4	14,4	7,7	0,9	9,1	11,2	6,6	18,3	7,1	6,6	11,5	20,8
Clothing	14,7	29,0	2,6	39,7	31,2	4,3	15,9	9,4	1,9	0,9	6,4	6,7	7,3	0,2	4,2	5,5	7,0	6,0	19,6
Food	14,1	19,1	4,3	50,5	38,1	2,5	1,9	10,9	13,0	0,7	2,3	1,6	3,3	16,6	20,1	1,5	0,2	1,5	19,6
RealEstate	14,0	32,2	11,5	12,5	38,2	14,0	6,8	14,6	25,2	1,4	0,7	0,6	3,4	1,5	15,8	3,0	0,4	1,3	20,3
C150300	13,4	9,0	5,9	12,5	13,7	7,8	11,9	10,7	10,5	25,9	8,8	4,4	3,6	5,2	7,8	16,4	5,1	4,4	25,2
BarRest	11,5	38,2	8,0	14,5	22,0	8,1	1,7	12,0	10,3	0,5	2,7	1,6	4,4	11,2	5,9	11,8	0,1	0,5	16,1
IT	11,2	39,4	17,6	19,4	31,6	5,2	1,2	1,1	1,0	1,4	0,5	3,6	7,1	8,4	3,7	0,4	1,6	8,4	12,8
C300600	10,7	11,2	3,0	14,0	16,6	11,2	10,8	17,8	11,9	9,0	3,9	2,9	0,4	10,2	23,1	20,3	2,0	3,0	7,0
Finance	10,6	42,1	8,6	6,1	38,5	8,2	1,1	12,8	10,2	1,0	1,8	2,1	4,2	7,5	2,3	7,1	0,8	2,1	10,7
Hotels	7,7	23,5	2,5	5,4	2,4	3,8	18,5	2,4	9,5	0,6	0,5	1,3	5,9	1,8	8,9	1,7	1,4	4,3	9,3
Beauty	7,4	25,4	16,6	1,6	9,9	4,1	6,8	18,1	3,5	0,4	1,5	0,7	3,2	1,2	4,1	3,2	0,7	0,6	10,2
Reparation	7,2	19,3	0,9	26,4	21,8	5,5	0,1	0,7	3,0	1,1	3,1	2,7	3,0	2,2	2,5	3,3	3,1	2,6	5,6
Leisure	6,6	24,3	3,6	1,4	47,7	7,4	2,7	3,8	1,8	1,4	1,5	1,5	3,0	2,3	0,6	1,6	0,9	3,2	5,7
Furniture	3,2	10,5	3,4	3,6	1,9	3,3	0,8	1,9	0,1	0,1	0,1	0,2	0,8	10,9	0,4	1,8	0,3	2,0	5,3

Table 7 RFA Mutual Information of each Retail Fabric

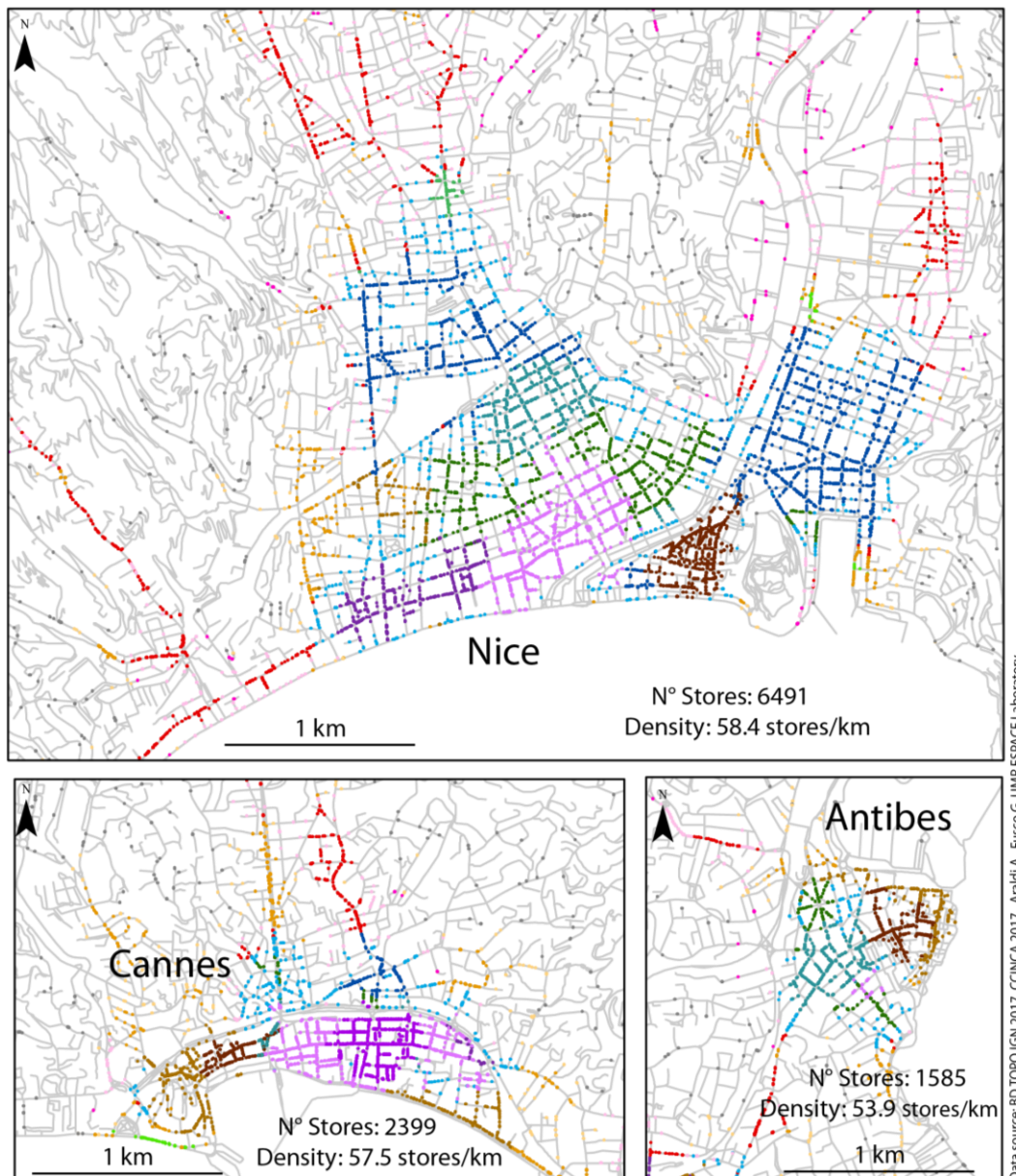
Finally, the fourth and last piece of information which is provided by the BN classification is the individual probabilistic description of each store: while traditional deterministic approaches assign only one class to each spatial unit, BN associates with each store a specific probability of being associated to the 18 RF (with the condition in 5.5). Together, with the most probable RF, we might also observe other retail classes which might also be associated with the same store allowing to evaluate the uncertainty of the classification results. This information acquires an essential meaning when projected in the geographical space: stores with a higher uncertainty between two or more classes might identify those specific spaces defined by a hybrid character or located in an intermediary position between two well-defined agglomerations or, again, those urban spaces under transformation.

Based on these rich quantitative and probabilistic outcomes, we provide in the following section a qualitative description of each Retail Fabric with some examples and images allowing the reader to better apprehend the different *retailsapes* which we might observe on the French Riviera.

### **Retail Fabrics description: Core Centres of Large-sized Cities**

We start our description from those Retail Fabrics which might be found in those urban spaces of the French Riviera conurbation which show characteristics of core centrality (RF1-4). Four core-centres are identified along the coastline gathering 37% of the overall number of stores of the study area: in Nice, we find the larger agglomeration (with more than 6 thousand stores), followed by Cannes (2.3k), Antibes with Juan-les-Pins(1.9k) and Menton (0.8k)<sup>57</sup>. Despite their differences in size and geographical extents, what these areas have in common is the high density and pervasiveness of store distribution (Fig.34).

<sup>57</sup> We remind the reader the absence of Monaco in our dataset for which we only know the overall number of stores (801 in 2017, source: Institut Monégasque de la Statistique et des Etudes Economiques, or IMSEE).



Data source: BD TOPO IGN 2017, CCINCA 2017, Araldi A., Fusco G. UMR ESPACE Laboratory

**Retail Fabrics of the French Riviera**

- |  |  |  |                         |
|--|--|--|-------------------------|
| <ul style="list-style-type: none"> <li>● <b>RF1.a</b> : Pervasive touristic retail fabric (culture and leisure oriented)</li> <li>● <b>RF1.b</b> : Pervasive touristic retail fabric (small restaurants and clothing stores)</li> <li>● <b>RF2.a</b> : Pervasive continuous retail fabric (clothing district)</li> <li>● <b>RF2.b</b> : Pervasive continuous retail fabric (clothes, real estate, dep. stores)-Cannes</li> <li>● <b>RF2.c</b> : Pervasive continuous retail fabric (clothes and real estate agencies)-Nice</li> <li>● <b>RF3.b</b> : Pervasive continuous service-oriented retail fabric with anchor stores</li> <li>● <b>RF3.a</b> : Pervasive continuous mixed retail fabric with anchor stores</li> </ul> | <ul style="list-style-type: none"> <li>● <b>RF4.a</b> : Pervasive continuous residential retail fabric</li> <li>● <b>RF4.b</b> : Pervasive fragmented residential-ancillary retail fabric</li> <li>● <b>RF5.a</b> : Residential ribbons with large store presence</li> <li>● <b>RF5.b</b> : Scattered out-of-RF5a-ribbon stores</li> <li>● <b>RF6.b</b> : Specialized Nucleated retails with food anchor stores</li> <li>● <b>RF6a</b> : Specialized nucleated or vehicle-oriented ribbon, (furniture-clothing-leisure)</li> <li>● <b>RF7</b> : Nucleated, mixed retail fabric (small, medium-sized cities)</li> </ul> | <ul style="list-style-type: none"> <li>● <b>RF8</b> : Nucleated isolated touristic oriented retail fabric</li> <li>● <b>RF9.a</b> : Residential fragmented ribbons, small stores /services</li> <li>● <b>RF9.b</b> : Scattered out-of-RF-9.a ribbon stores</li> <li>● <b>RF10</b> : Isolated stores</li> </ul> | <p>— Street Network</p> |
|--|--|--|-------------------------|

RF in Core-centers

RF in semi-peripheral and suburban areas

Figure 34 The core-centers of Nice, Cannes, Antibes (point-pattern representation of RFA results)

The highest density values are found in these spaces: 59 stores per kilometre in Nice, 57 in Cannes and about 55 in Antibes with Juan-les-Pin and Menton<sup>58</sup> (a significant difference when compared to the average value of 8.35 for the littoral municipalities in Tab.3, Section 3.4). This spatial form characteristic represents the necessary condition that allows for functional clustering of stores. The coalescence of nucleated agglomerations within these cities might find its causes in historic, economic, social, or urban characteristics. In these urban areas, nine retail fabrics are identified by RFA, grouped in four families (touristic, clothing, services and residential oriented agglomerations) and are differently distributed in our four core-centres.

Two Retail Fabrics identify touristic and leisure-oriented agglomerations.

**RF1.a** gathers 4.7%, of the overall store volume. In these spaces, we find the second highest values of size variables (an average of 128 -360 -828 stores in the surrounding 150-300-600 metres from each store). What most defines these centres is the strong under-representation of bank and insurance, real estate, IT and vehicle-related stores. The high fragmentation of the building ground floor enables the presence of a wide number of small commercial activities with an overrepresentation of bars and restaurants, clothing stores and specialised food stores (traditional local food, butcheries, etc.). When observing the geometry indicators, we might describe a linear continuum of stores at the street level, with a pervasive organisation observed already at the district level allowed by the underlying well-meshed street grid. Clothing and Furniture large stores are both under and over-represented, meaning that these stores might be scattered in these agglomerations. Low values of concentration (C300600) describes how these agglomerations are not isolated but always integrated within the urban core of large cities; nonetheless, when comparing the values of this specific indicator to the others RF found in core-centres, C300600 shows relative higher values. This difference explains how RF1 are not located at the geometrical centre of the core areas, rather in a relatively less commercial integrated space (as observed when projected in the geographical space, Fig.35).

RF1.a is found in the four historical downtowns where the picturesque Mediterranean urban landscape, dating back to the late medieval period, represents an important touristic attraction. As we might perceive from the pictures of Fig.36, the economic and social activity we find in the urban areas might be described as vibrant, especially during evening and weekend periods, influenced strongly by the touristic seasonality oscillation.

The only exception is the centre of Juan-les-Pins: as we will further describe in Part 4, the historical evolution of this urban settlement is strictly interwoven with the urban development of the close centre of Antibes. In the late XIX<sup>th</sup> century, after the inclusion of the Nice County to France, Antibes started a long and important phase of urban development for its conversion from a military to residential settlement (further details in Text-Box 3 in Section 9.3). In the meantime, the development of the first tourism of the Belle Époque in the Côte-d'Azur found

---

<sup>58</sup> These densities are obtained as the ratio between the number of stores and the street total length contained within the minimal network-constrained convex hull created around RF1-4 types stores (Figure.36)

in Juan-les-Pins the closest attraction to the city of Antibes (less than 1.5 km between the two centres). This peculiar urban evolution is at the origin of the double polarity observed in the conurbation of Antibes and Juan-les-Pins.

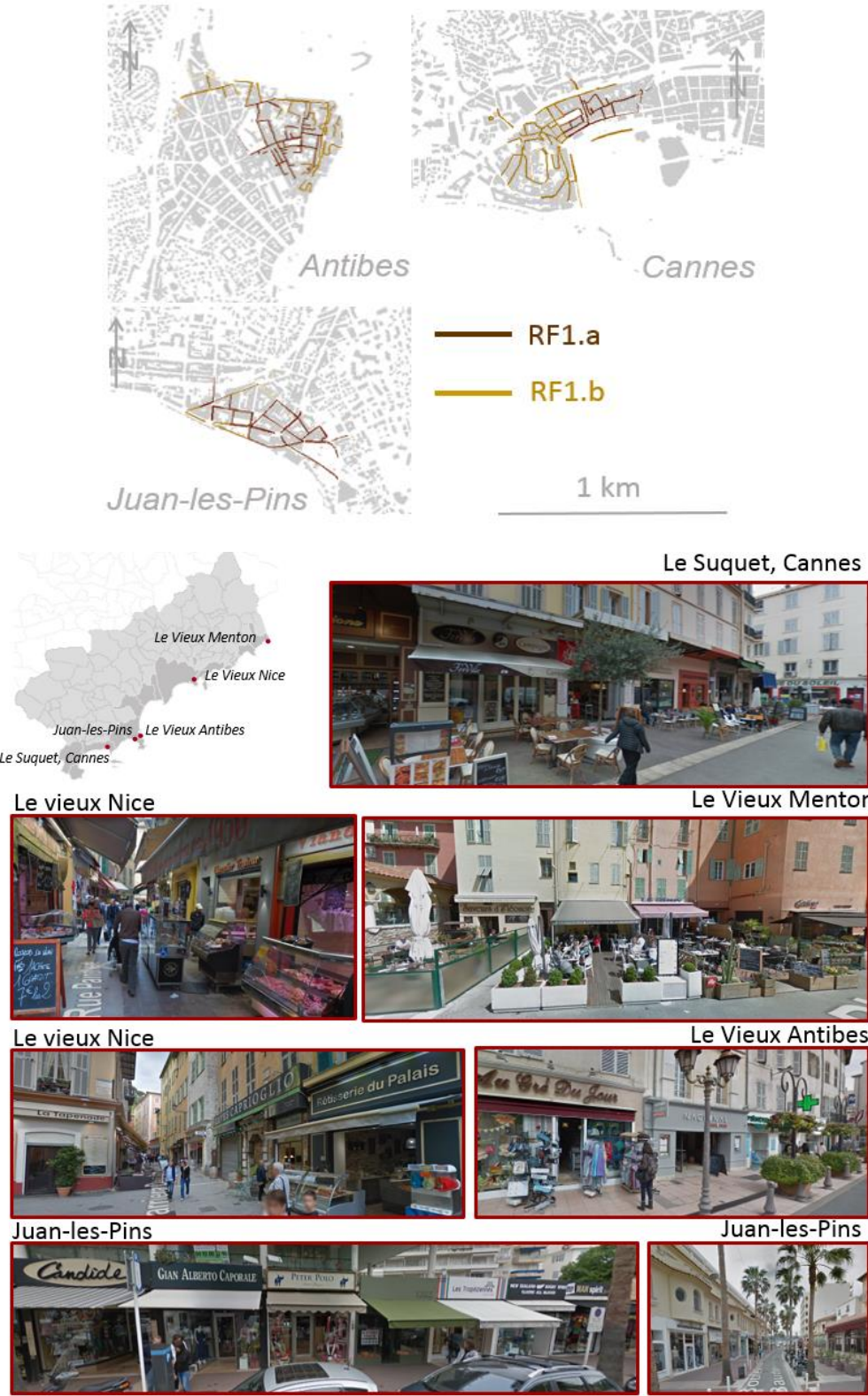


Figure 35 RF1 Touristic-oriented agglomerations in Antibes, Cannes and Juan-les-Pins (above); street view of RF1 (Source: Google Street View 2019)

A second touristic retail fabric is detected in Cannes, Antibes, and Juan-les-Pins in RF1.b. This second typology shows the same main characteristics previously described for RF1.a, with few significant variations. From a functional point of view, we might notice a higher distance from large anchor stores, a lower presence of small food stores, and a relatively higher presence of restaurants, bars, and real estate agencies. When observing differences in form indicators, it is possible to detect a higher fragmentation of the store linearity at the street level and a more isolated position of the overall agglomeration: as we might observe in Fig.35, RF1.b is found in edge areas of the core-centres. These characteristics explain the relative positioning of this class in the mutual information space closer to isolated centralities without large anchor stores.

The 9.6% of the stores are gathered in **RF2** describing clothing and leisure-oriented retailscapes; three sub-groups of retail fabrics identify different variations of these agglomerations. As for every retail fabric identified within core-centres, even RF2 is characterised by a pervasive distribution of stores at the district and neighbourhood scale; hyperlocal variations of store distributions might result in a more or less fragmented linear continuity. While form descriptors show strong similarities among these retail agglomeration typologies, functional variables highlight significant differences in over/underrepresentation and co-occurrences of specific stores categories.

When projected in the geographic space, RF2 classes identify the emblematic urban areas of Le Carré d'Or (The Golden Square) in Nice and the Centre-Ville (City centre) of Cannes. Two internationally well-known neighbourhoods for their historical palaces of the decades La Belle Epoque such as Le Negresco, Palais de la Méditerranée in Nice and Le majestique, InterContinental Carlton in Cannes. Palaces hosting luxury hotels and Casinos and attracting wealthy customers. The same retail class is also observed in few streets of Menton between the neighbourhoods Hotel de Ville-Bioves and Casino (Fig.36). Only few streets in Antibes show a clothing prevalence: the same urban historical reasons explaining the duplication of the touristic-oriented retail fabric in Juan-les-Pins (further details in Part 3), might be at the origin of the absence of an important clothing/leisure-oriented district in this cities.

RF2.a (4.99%) is defined by an important underrepresentation of small food stores and an overrepresentation of both small and large surface clothing stores (national and international firms), restaurants and bars, retail estate and hotels. When compared to the spatially contiguous RF1 of historical city centres, these distinctive features are even more emphasised; in particular, via the relatively higher distribution of clothing anchor stores. Contradictorily, although RF2.a shows a higher presence of bars and restaurant compared to the overall study area, lower values are found when compared to historical centres. Similarly, large food stores are generally underrepresented and scattered only in marginal subspaces.

When observing the urban areas identified by this specific class in more detail, what we might observe is a hyperlocal specialisation of individual streets juxtaposed within the same neighbourhoods. In Nice, for example, we find clothing hyper-specialisation in streets such as Rue Paradis, Rue Longchamps, Rue Alphonse Karr, Av. Jean Medecin (southern part),

coexisting and intersecting restaurant and bar concentration along the pedestrian island of Rue Massena and the lateral Rue Massenet, Place Grimaldi etc. The same juxtaposition is found in Cannes: while Rue d'Antibes is mainly clothing-oriented, bars and restaurants are concentrated along the pedestrian Rue Hoche or at the edges of Place General de Gaulle.

A specific class is detected in Cannes: RF2.b (2.54%) identify the core area of the Centre Ville neighbourhood where, compared to RF2.a, a significantly higher concentration of financial services, real estate and large retail surfaces is observed. While in Cannes, financial and real-estate agencies are concentrated at the hearth of RF2, in Nice, this specialisation is shifted toward the eastern side of the Carré d'Or and identified by RF2.c (2.1%). This specific variation is detected along Rue de France, which represents the non-pedestrian prolongation of Rue Massena; compared to RF2.a and b, a more fragmented retail continuity describes these urban areas. From a functional point of view, streets lose the mono-specialisation of both clothing or restaurant for a more mixed combination of small stores; nonetheless, together with the small clothing store prevalence, a higher number of financial services and real estate agencies might be found. When analysing large surfaces, clothing stores are substituted by food anchor stores; some car and furniture luxury shops complete the retail offer. The same category is identified in few streets in the central area of Menton.

Despite the high specialisation (specifically clothing and bar-restaurants in RF2.a and RF2.b), stores classified in RF2 present high values of store variety: minimum value of 18/22 retail categories is found within the first 300 metres from each store.



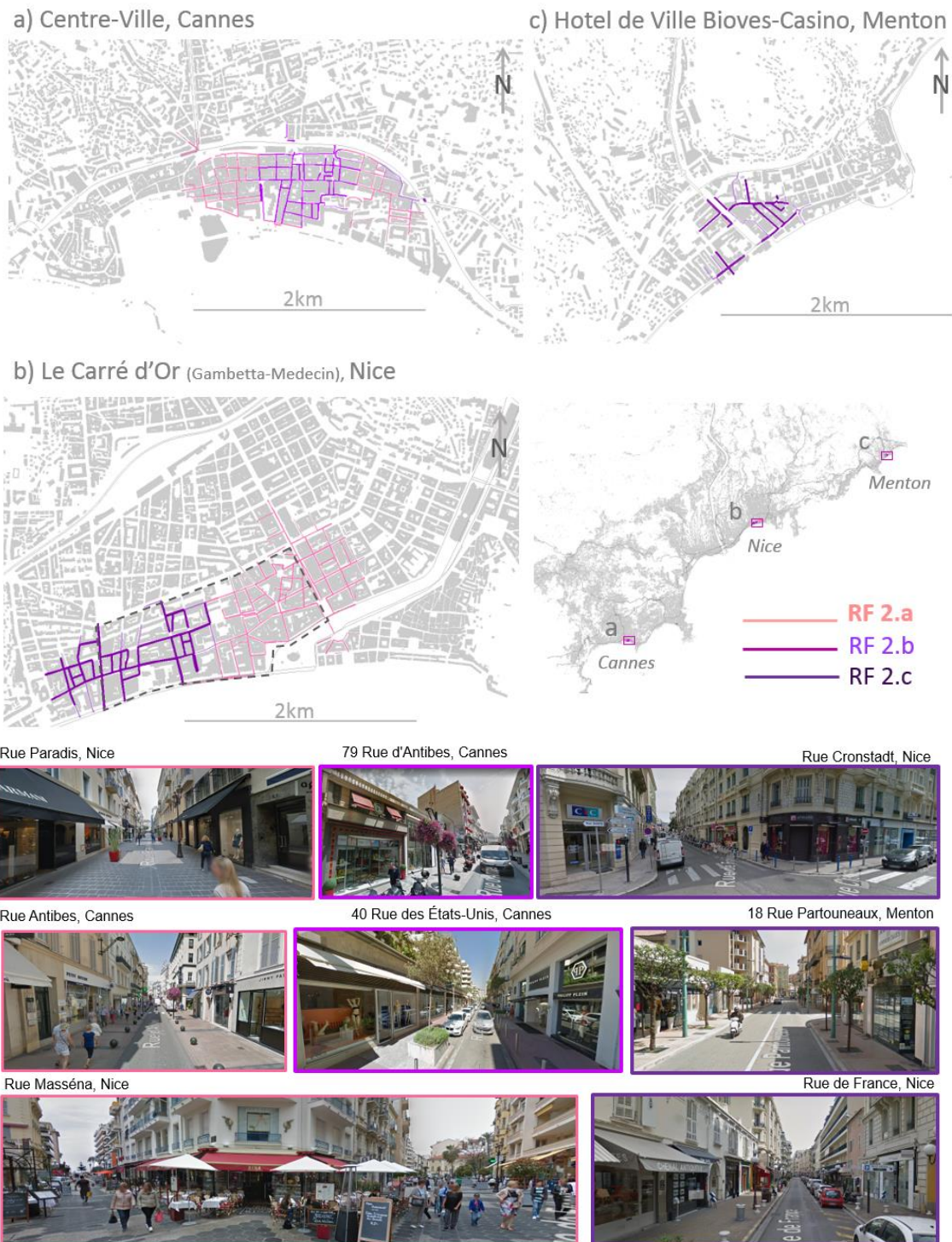


Figure 36 RF2 Clothing and Leisure-oriented agglomerations in Cannes, Menton and Nice (above); street view of RF2 (Source: Google Street View 2019)

Two Retail Fabrics **RF3.a** and **b** gather 6.75% of the retail activities (Fig.38-39). In these specific areas, we do not observe a specific functional orientation as for RF1 (touristic) and RF2 (clothing): the highest values of retail category variety (Rich300) are observed with an average of 20.8/21 categories represented in the 300 surrounding metres. These high values might be explained because of the relative position of these urban spaces, located between touristic/leisure retail fabric (RF1 and RF2) and more residential areas (RF4). The functional

mix of these urban spaces is also enriched by those stores attracted by the proximity to train stations: in Nice and Antibes a higher number of hotels, bars, small restaurants and fast food stores might be detected. In Cannes, there are only a few stores located along those streets close to the train station and between the historic centre and the clothing district.

The absence of a large district of category RF3 in Cannes might find, again, explanation in a specific urban morphological characteristic: at the north of the clothing district, the city centre is delimited by the train line and an urban roadway. There exists a strong urban fracture which defines a net limit between leisure-clothing-touristic oriented spaces at the south and residential areas at the north without a gradual transformation of the retailscape as observed in Nice or Antibes.

Some functional over/under representations might be still recognised: RF3.a is characterised by a high presence of leisure, clothing, and home-related large stores. Bar and restaurants are overrepresented only in specific subspaces (e.g. the train station, but also where larger pedestrian sidewalks allow bars and restaurants to extend their surface). When compared with RF2, what most distinguishes these areas is the higher mix of the retail offer with a more important presence of large leisure surfaces (cinemas) and furniture stores; a higher presence of small food stores is also detected. RF3.b represents a specific variation of RF3 which is found in relatively peripheral areas of the centre; compared to RF.a, lower values of large stores and overrepresentation of financial services and real estate agencies are detected in these subspaces.

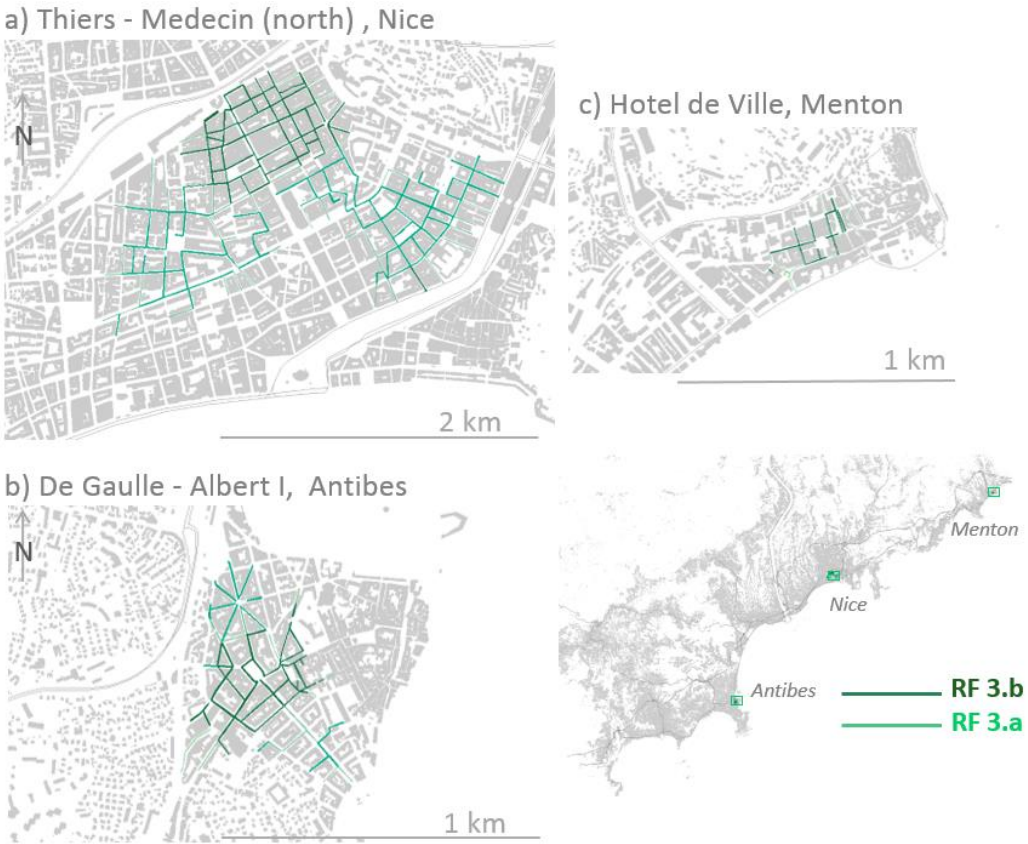


Figure 37 RF3 Mixed and Service-oriented agglomerations in Nice, Menton and Antibes

a) 29 Rue d'Angleterre, Nice



b) 4 Rue de la République, Antibes



c) 28, Rue de la République, Menton



b) 21 Avenue Robert Soleau, Antibes



a) 37 Rue Gioffredo, Nice



Figure 38 Street view of RF3 (Source: Google Street View 2019)

As previously mentioned, from RF1 to RF3 despite similar form characteristics of the retail distribution we might observe a gradual reduction of the functional specialisation toward more mixed and residential oriented areas. Following this trend, **RF4** is found in the most marginal spaces of the core-centres.

From the functional point of view, we observe in RF4 underrepresentation of real-estate agencies, hotels and large surfaces of leisure and clothing stores. The residential oriented nature of these retailscapes are perceivable by the overrepresentation of food, furniture, and large do-it-yourself stores; similarly, a higher presence of small food stores and beauty/health services (such as pharmacies and hairdressers) might be observed

Despite the high level of pervasiveness at both the district and neighbourhood scales, not every street of these neighbourhoods is described by a continuous distribution at the street level: some streets might present specific morphofunctional conditions such as the presence of churches, schools, or local administrative services which might cause a higher local fragmentation. RF4.b identifies those spaces where a lower number of stores is detected. The indicator which most differentiates from the RF4.a retail fabric is the Concentration (Ratio150-300) showing a higher presence of stores at further distances. The same information is also confirmed by the geometry variable at street level describing a less continuous retail presence (10fold analysis show us as the R150 and the C150300 are the most important indicators to distinguish from the other classes and in particular from the blue one).

These retail fabrics are distributed in what we might designate as buffer zones: the transition between the core and semi-peripheral retail agglomerations. A less specialised retail agglomeration with a complete and pervasive retail offer for residents (both in large and small retail surfaces).

Three areas are identified (Fig.30): Vernier-Liberation and Carabaçel-Port-Riquier in Nice while in Cannes the Prado neighbourhood.

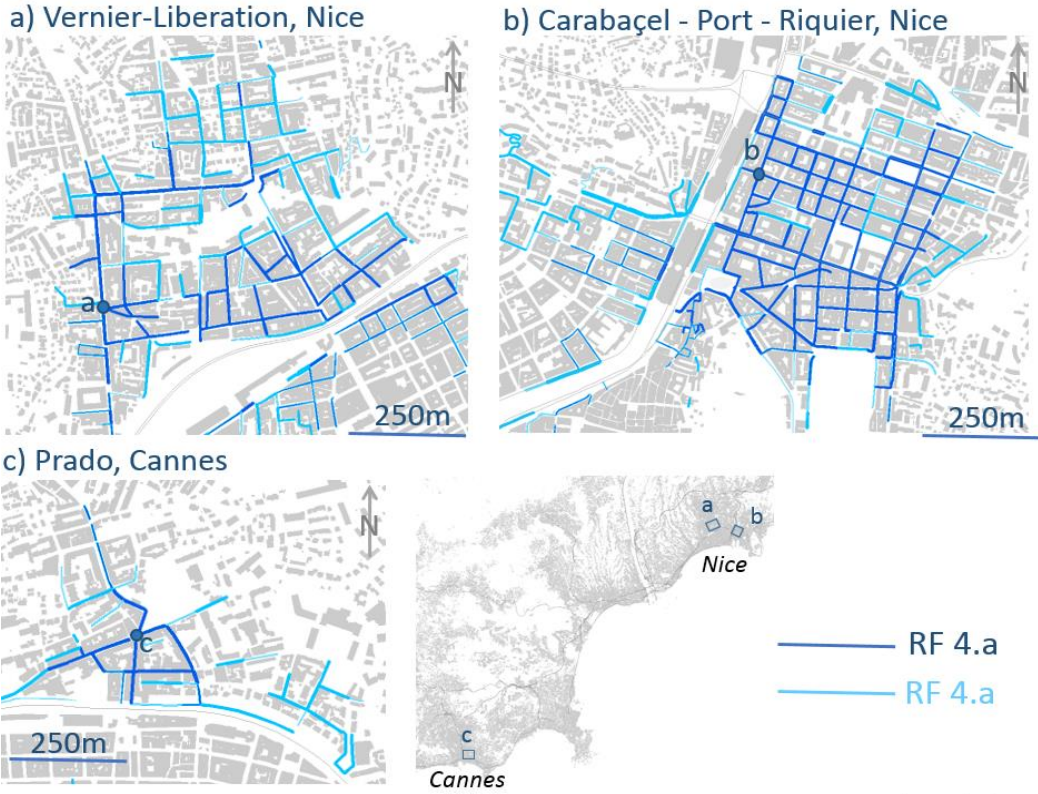


Figure 39 RF4 Residential pervasive agglomeration in Nice and Cannes



a) 111 Bd.  
Gambetta, Nice

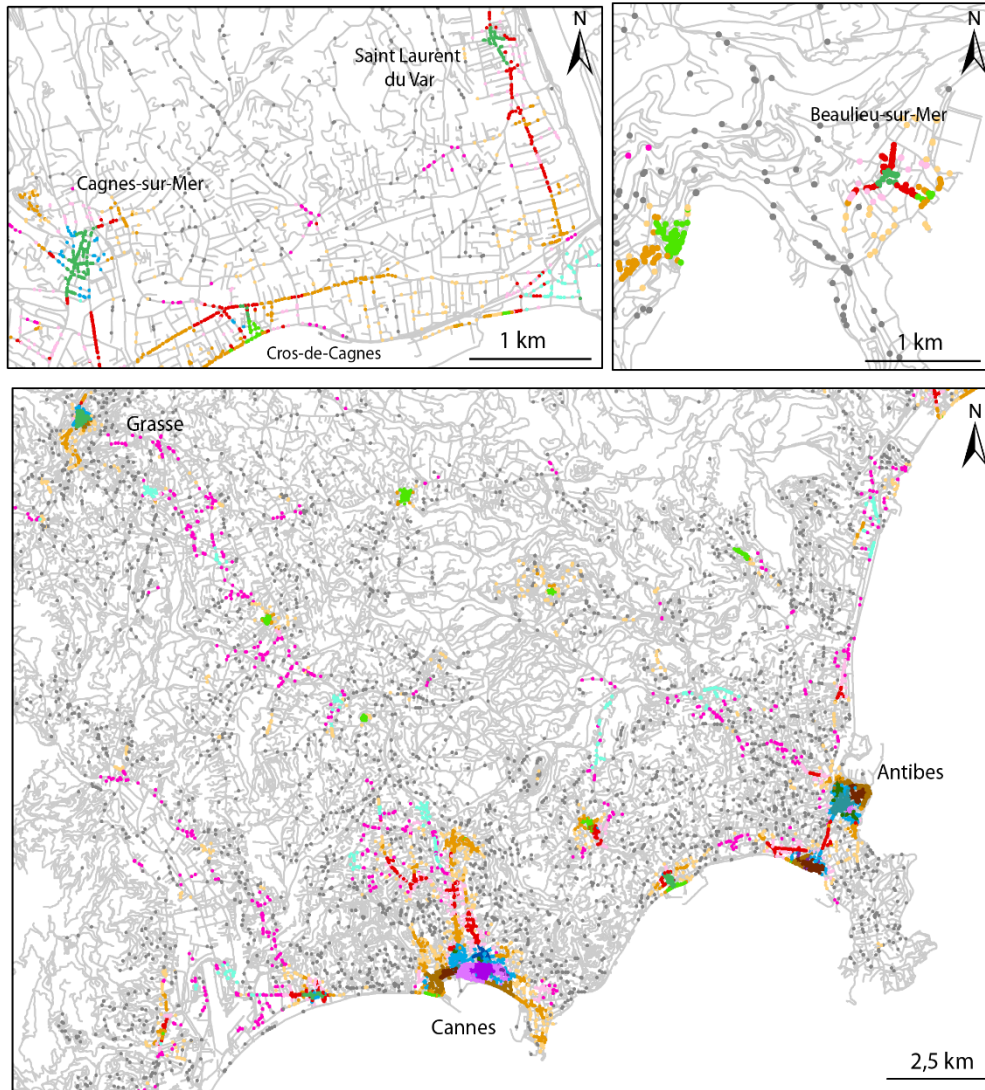
b) 30 Av. de la  
République, Nice

c) 28 Rue Jean  
Haddad Simon,  
Cannes

Figure 40 Almost-360° street view of RF4 (Source: Google Street View 2019)

## Metropolitan Peripheral Areas

Outside of the core areas of these larger cities, RFA protocol identifies 8 Retail Fabric gathering 42.1% of the total retail volume. While Retail Fabrics previously described have been mostly recognised through those variables delimitating functional clusters within a continuous and pervasive distribution of stores, different sized and shaped agglomerations are found in semi-central and peripheral areas.



Data source: BD TOPO IGN 2017, CCINCA 2017, Areldi A., Fusco G. UMR ESPACE Laboratory

### Retail Fabrics of the French Riviera

- |  |   |  |                         |
|--|---|--|-------------------------|
| <ul style="list-style-type: none"> <li>● <b>RF1.a</b> : Pervasive touristic retail fabric (culture and leisure oriented)</li> <li>● <b>RF1.b</b> : Pervasive touristic retail fabric (small restaurants and clothing stores)</li> <li>● <b>RF2.a</b> : Pervasive continuous retail fabric (clothing district)</li> <li>● <b>RF2.b</b> : Pervasive continuous retail fabric (clothes, real estate, dep. stores)-Cannes</li> <li>● <b>RF2.c</b> : Pervasive continuous retail fabric (clothes and real estate agencies)-Nice</li> <li>● <b>RF3.b</b> : Pervasive continuous service-oriented retail fabric with anchor stores</li> <li>● <b>RF3.a</b> : Pervasive continuous mixed retail fabric with anchor stores</li> </ul> | <ul style="list-style-type: none"> <li>● <b>RF4.a</b> : Pervasive continuous residential retail fabric</li> <li>● <b>RF4.b</b> : Pervasive fragmented residential-ancillary retail fabric</li> <li>● <b>RF5.a</b> : Residential ribbons with large store presence</li> <li>● <b>RF5.b</b> : Scattered out-of-RF5a-ribbon stores</li> <li>● <b>RF6.b</b> : Specialized Nucleated retails with food anchor stores</li> <li>● <b>RF6a</b> : Specialized nucleated or vehicle-oriented ribbon , (furniture-clothing-leisure)</li> <li>● <b>RF7</b> : Nucleated, mixed retail fabric (small, medium-sized cities)</li> </ul> | <ul style="list-style-type: none"> <li>● <b>RF8</b> : Nucleated isolated touristic oriented retail fabric</li> <li>● <b>RF9.a</b> : Residential fragmented ribbons, small stores /services</li> <li>● <b>RF9.b</b> : Scattered out-of-RF-9.a ribbon stores</li> <li>● <b>RF10</b> : Isolated stores</li> </ul> | <p>— Street Network</p> |
|--|---|--|-------------------------|

RF in Core-centers

RF in semi-peripheral and suburban areas

Figure 41 RFA outcomes in peripheral areas

The first retail agglomeration we find surrounding the centres of larger cities is **RF5.a**. As we might observe in Table 7, the most characterising aspect of this agglomeration typology is the linear shape at the district and neighbourhood level. The geometry indicator at the street level shows an average value of 0.6 describing frequent interruptions of the continuity of store alignment; several factors might explain the local fragmentation such as vehicle entrances, buildings with larger hall entrances, larger store surfaces with a longer frontage, etc. (Fig.42).

From a functional point of view, together with a large variety of small stores, we also observe an overrepresentation of large food stores (mainly national and international retail chains such as Casino, Carrefour, Lidl etc.); other large surface categories are underrepresented. These characteristics are also confirmed by the 10-fold analysis specifically carried out for this class showing the high importance of variety (Rich300), size, and geometry indicators at the neighbourhood level, while redundancy is detected for anchor stores and hotels functional variables.

These agglomerations can be explained when considering the relative location within the surrounding urban context: RF5.a are frequently found along main boulevards and avenues rooted in large and average-sized city centres and spatially arranged towards/penetrating residential neighbourhoods with compact built up. These streets have the twofold advantage to be located at the core of a catchment area rich in customers, but also it represents the main access/passageway collecting daily journeys of the surrounding dwellings toward city centres or more peripheral areas. The high proximity both between stops and to residential places allows both vehicle and pedestrian multi-purpose trips.

Based on the specific characteristics here described, we labelled these agglomerations as 'pedestrian ribbons' which might find a certain degree of correspondence to the 'Urban arterial ribbons' of Berry (Fig.3, Section1.3) or, similarly, to the 'Traditional Arterial Routes' of Brown (Fig.5, Section1.3). The specific pedestrian adjective is here used to outline how these axes are found in more compact built-up spaces typical of the Mediterranean-European context with a less vehicle dependent mobility; there is an explicit intention to highlight the difference from the original Berry's notion of vehicle-based ribbons, traditionally associated to routes of more peripheral/suburban streets in a North-American context, and which we further describe with a specific retail fabric typology.

RF5.b identifies those stores located in the proximity of pedestrian ribbons; this specific class might be distinguished by the previous mainly for the different values assumed by local geometrical and concentration indicators. A fragmented and scattered distribution of stores located at a certain distance from higher concentrations (confirmed by low values of C150300 and C300600). From the functional point of view, the high values of diversity found along the main pedestrian ribbons are reduced but still detected in its proximity. There is a significant underrepresentation of those activities which need the visibility such as banks, real estate, and beauty/health stores; on the contrary, there is a higher probability of finding furniture and garages and repair shops requiring larger surfaces. In Figure 43, some street view photos of the lateral spaces of the Boulevard Saint Roch in Nice show the residential nature of the lateral areas of pedestrian ribbons; on the contrary, the central boulevard collects through-movement and, consequently, the higher retail concentrations.

City	Street Name	Length [m]
Nice	Bd. Saint Roch (a)	1000
	1-60 Bd. Gorbella (b)	900
	60 Av. Alfred Borriglione -17 Av. du Ray (c)	700
	60-154 Bd. Celssole(d)	300
	1-280 Bd. Madeleine (e)	1800
	1-40 Bd. Carlone (f)	400
	103 Av. de la Californie -121 Rue de France (g)	1300
	212 Av. de la Californie- 30 Av. Saint-Agustin (h)	600
Beaulieu-sur-Mer	Bd. Marinoni	200
	Bd. du Maréchal Joffre	600
Cagnes-sur-Mer	Av. François de May	500
	60-95 Av. de Nice (l)	600
Roquebrune-Cap-Martin	22-75 Bd Maréchal Juin (m)	650
Saint Laurent-du-Var	26-46 Av. Auguste Renoir	250
Menton	Av. Aristide Briand	500
Saint Laurent-du-Var	Av. du General Leclerc - Av. du General De Gaulle (i)	900
	1-30, Avenue Carnot	300
Vence	Av. Henri Isnard,	300
	Av. de la Resistence	350
Antibes	40-100 Bd. du Président Wilson	700
	1-40 Avenue Philippe Rochat	500
Antibes (Juan les Pins)	Bb. Raymond Poincaré	700
Cannes	70-120 Bd. De la Republique	800
	109-97 Bd. Carnot	600
Le Cannet	Bd Paul Doumer	800
	69 Av. Franklin Roosevelt- 10Route Napoléon	400
	Av. Francis Tonner	400

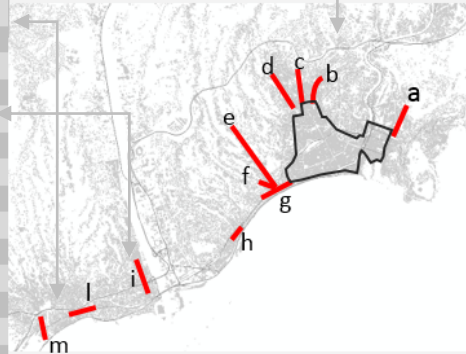


Figure 42 RF5.a Pedestrian Ribbons in Nice metropolitan area(above); street view of RF5.a (Source: Google Street View 2019)



Bd. Cessole, Av. Borriglione, Bd. Gorbella, Bd. Madeleine - Nice



Figure 43 Pedestrian Ribbon (RF5.a) and scattered out-of-ribbon (RF5.b) in Nice (above). Some street view pictures of RF5.b (below) (Source: Google Street View 2019)

While RF5a-b are found in urban semi-central areas, the next two Retail Fabrics are identified in suburban and peripheral areas and characterised by small concentrations of both large and small stores.

**RF6a** identifies that 6.2 % of stores are located in small and dispersed concentrations: in the surrounding 150 metres from each retail area we count an average of 8 stores (standard deviation of 5) which is the lowest after an isolated store; nonetheless, compared with isolated stores, RF6a is described by a higher number of stores at further distances. Consequently, other form indicators such as concentration (C150300 and C30600) and geometry describe, again, a sparse distribution of punctual concentrations. It is only when projected in the

geographical space that we might perceive the overall linear organisation along the main departmental routes connecting the main centres and strongly influence by the relief (Vallée de la Siagne, Vallée du Var, Valle du Paillon, Vallée de Carei, etc.) or a more nucleated disposition in the peripheral areas of villages and cities.

As we might see from the aerial pictures in Fig.44, these corridors might intersect areas with different uses, from the urban sprawl with a residential prevalence to more mixed land uses with a higher presence of productive, industry and logistic functions.

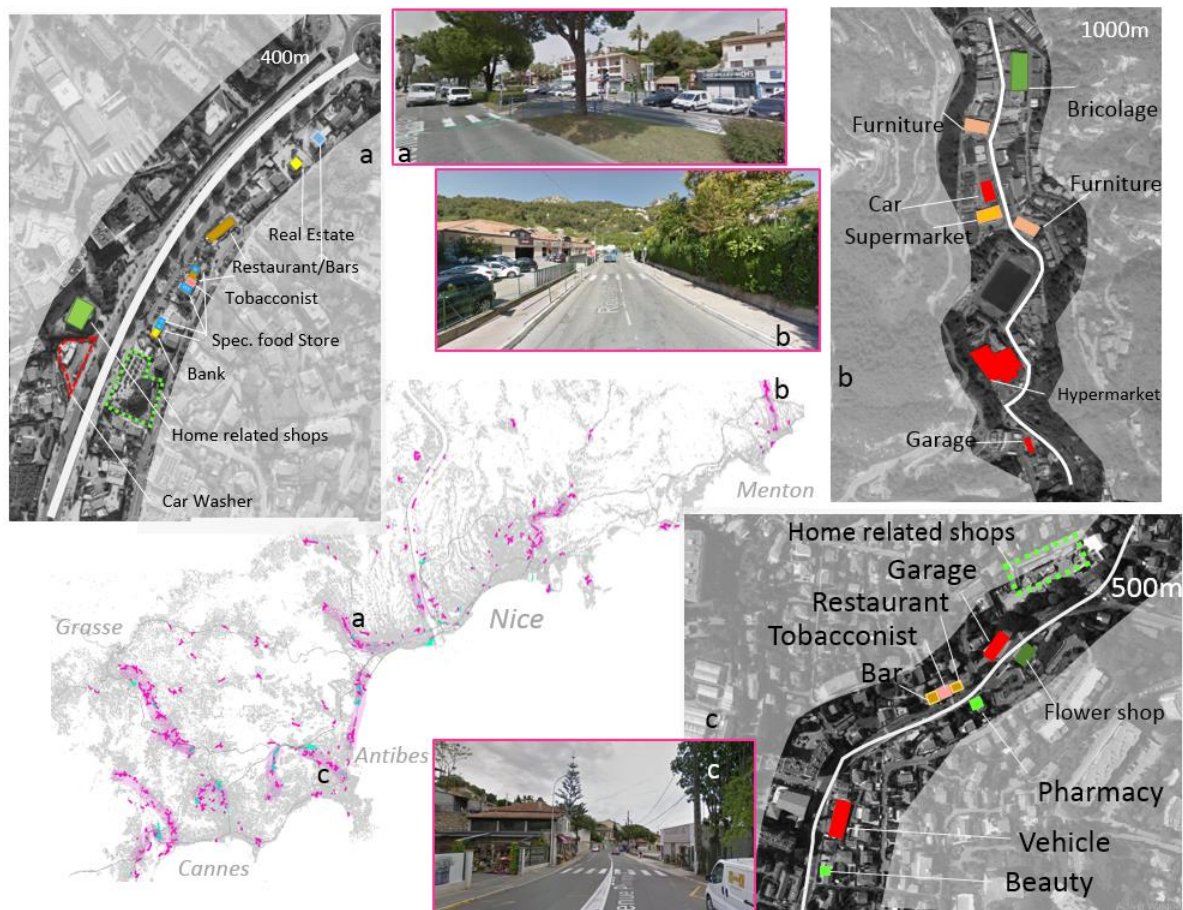


Figure 44 RF6.a: Scattered Vehicle Ribbons in the French Riviera and some local examples in Cagnes-sur-Mer(a), Menton(b) and Antibes(c). (Source: Google Map-Google Street View 2019)

**RF6.b** gathers 2.1% of the overall number of stores of our space study, representing those specialised in nucleated or linear agglomerations in urban peripheral areas (Fig.45). There exists a strong functional prevalence of vehicle and home-related large surfaces (furniture, electronic apparels, bricolage, garden-related stores, etc.); in some of these nucleated agglomerations, clothing anchor stores (outlet format) complete the retail offer. On the contrary, large food stores are underrepresented, probably not because of their absence, but because of the relatively lower proportion compared to the overall number of non-food stores.

Hotels, leisure, beauty, small clothing, and almost every proximity service are absent in these areas with a consequent low value of richness (average of 15/21 retail typologies). The co-occurrence of furniture and vehicle stores is not a recurrent aspect of every agglomeration:

mono-functional subspaces are also detected (such as the concentration of vehicle-related stores in Fig.46). When compared to RF6.a, these specialised store concentrations show a higher density at the district level with an overall nucleated or linear geometry at the neighbourhood level.

One additional peculiarity of these concentrations is their relative isolation from the rest of the retail distribution: the average value C300600 is the second highest (after RF9.a, see further in this section). These clustered concentrations, differently from the other Retail Fabrics, do not represent the result of a spontaneous self-organised phenomenon, rather the result of strategic choices of local administrative authorities. Indeed, these agglomerations are frequently located within the perimeter of specific zoning areas <sup>59</sup>. The urban landscape in these areas is almost entirely occupied by large surface stores surrounded by car parking areas and logistics and transportation infrastructure.



Figure 45 Specialised Nucleated Peripheral agglomerations (RF6.b)

<sup>59</sup> 17 agglomerations of category RF8 are located within special productive and commercial zones defined by local authorities as activity centres (Zone d'Activité-ZA) or regulated activity centres (Zones d'activités réglementaires-ZAE): ZA de Campon, ZA Les Tourrades, ZAE Saint-Isidore Sud, ZAE Cap3000, ZAE de l'Ariane, ZAE Pôle Marina 7 / Les Maurettes, ZAE Les Terriers Nord, ZAE Saint Bernard, ZAE Saint Martin La Plaine, ZAE Sainte Marguerite, ZAE Minelle, ZAE de Saint-Laurent-du-Var, ZAE Sudalparc, ZAE Grand Arenas, ZAE Fuon Santa, ZAE La Paoute.

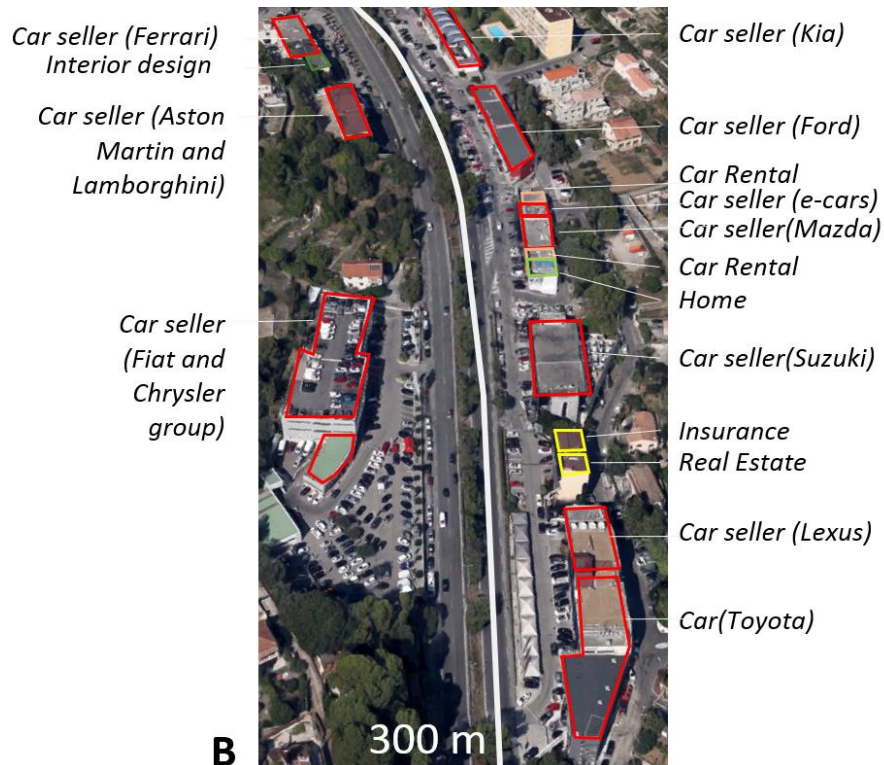


Figure 46 The concentration of Car Seller along the Avenue du Campon, Le Cannet

So far we have provided a qualitative description of RFA outcomes starting from core areas of larger cities (from RF1 to RF4) and, in the second part, we focused on suburban/peripheral agglomerations (RF5 and 6). Nevertheless, we have not discussed yet the distribution of retail activities within smaller towns and villages.

With **RF7**, as we previously introduced, we are able to identify the 3.2% of stores composing the retailscape of the core of eight average-sized cities and four secondary centres within the municipality of Nice. As we previously observed, this retail fabric presents intermediate characteristics with respect to every retail fabric described so far.

From a functional point of view, we might observe: i) a general underrepresentation of clothing, vehicle, home-related large retail surfaces, and hotels; ii) overrepresentation of food and leisure large surface stores; ii) a higher over-representation of small food, clothing, beauty/health stores, bars and restaurants.

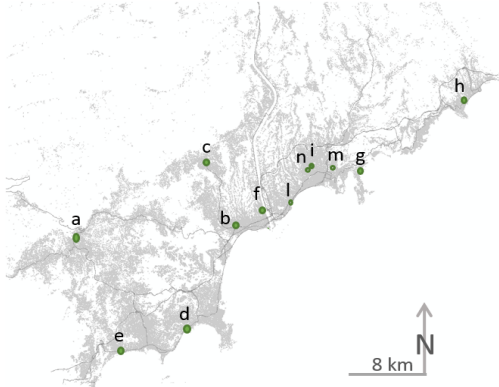
The size indicators associated with stores of type **RF7** are associated with average values: smaller than what is observed in core areas of larger cities; higher than semi-central and peripheral agglomerations. A strong concentration characterises these distributions: of the overall number of retailers in a 600-meter radius, almost 60% is found in the first 300 metres. We also observe a high number of stores at the street scale: this variable assumes greater values than those observed in historic centres of core areas (RF1) and pedestrian ribbons (RF5.a) and find its explanation in the hyper-concentrated distribution in few streets and squares representing the main urban spaces gathering the socio-economic life.

When focusing on geometrical indicators, we observe a strong linear continuity at the street level along the main town axes (see streets views in Fig.47); the linear shape is also detected

at both the district and neighbourhood scale. Similarly, in relation to what has been previously explained as core areas, even average-sized city centres identified by RF7 represent the central points from which pedestrian ribbons are originated. The same model is replicated, in smaller proportions, in four sub-centres within the municipality of Nice. When considering Massena Square as the main centre of the urban settlement of Nice, these four neighbourhood-centralities are all located at a similar distance, between 3 and 4 km.

If we compare the profiles of average-sized city centre agglomerations (RF7) and pedestrian ribbons (RF5.a), besides the evident differences of form characteristics, the former shows a higher overrepresentation of bar and restaurants, clothing, beauty/earth stores and a fewer presence of IT and vehicle-related stores.

Municipality	Size	
a. Grasse	268	
b. Cagnes-sur-Mer	204	
c. Vence	160	
d. Golfe Juin, Vallauris	117	
e. Mandelieu la Napoule	105	
f. Saint Laurent-du-Var	74	
g. Beaulieu	46	
h. Roquebrune-Cap-Martin	7	
Nice	i. Quartier Liberation Nord, Borriglic	70
	l. Quartier Carras	15
	m. Quartier Saint Roch	11
	n. Coeur Quartier Cessole	6



42 Avenue Marcellin Maurel, Vence



15 Boulevard du Jeu de Ballon Grasse



Quartier Liberation Nord, Borriglione, Nice



Figure 47 RF7 The identification of the 12 average-sized cities and neighborhood centralities (above); some Street View of RF7 (Source: Google Street View 2019)

### **Isolated centralities of small stores**

The seven retail Fabrics described so far are located in large and average-sized cities and peripheries with a significant presence of large anchor stores; three additional classes of retail fabric (gathering the 13.3% of the overall retail number) identify those isolated agglomerations composed only by small stores and services.

**RF8** (3.1%) is mainly defined by its small size and its high compactness and isolation. The concentration indicator (the ration C300600) is the highest observed among the eighteen classes: when considering the number of commercial activities found in a radius of 600 m from each store, on average, the 71% of them is disposed in the first 300 metres and 38% in the first 150 metres. Functional indicators of small stores describe a low richness, an over-representation of bar and restaurants, and an underrepresentation of vehicle, real-estate and beauty/health stores.

When projecting these classes in the geographical space, we identify nine historical inland villages (such as Saint Paul-de-Vence or Èze), four coastal centres which correspond to small settlements where traditionally fishermen villages were located (such as the Cros-de-Cagnes and Saint-Jean-Cap-Ferrat), and four small centralities located along the Quai of port areas of villages and towns (Villefranche) (Fig.48).

What these centralities have in common is the touristic and leisure-oriented function, similarly to RF1. Nevertheless, they might be easily differentiated from RF1 by the absence of anchor stores and the extremely distinctive form characteristics.

Values of geometry indicators describe another specificity of this fabric is identified: store distribution at the street level is almost linear but with high fragmentation while at the district and neighbourhood scale a linear layout is described. When observing these thirteen nuclei more in detail, three configurations might be identified. Two of them are shown by the relative location of the historical settlement, its main square, and the route connecting the village to the departmental transport network: the main route could pass through the village (Fig.48.a) or in its proximity (Fig.48.b-c). In the former case, the retail agglomeration is linearly disposed along with this main axis; in the latter, the agglomeration is centred on the main square of the village and stretched toward the departmental route. Finally, the fourth configuration considers coastal agglomerations in correspondence to port activities where retails are distributed along the coastline (Fig.48.d).

**RF9.a** (7%), similarly to the previous class, is defined by an isolated agglomeration of few small stores; nonetheless, we might still recognise some main features which differentiate this specific retail fabric. From the functional point of view, RF9 has a lower presence of bars and restaurants compared to RF8, but a higher overall variety. These agglomerations lose their touristic orientation and provide a more residential environment. What the two classes have in common is the lower presence of hotels, IT and finance, and a higher prevalence of beauty and health (pharmacies and hairdressers).

	Municipality	Commercial core
Inland Villages	Mougins	2 Rue du Maréchal Foch
	Valbonne	Place des Arcades
	Mouans-Sartoux	Garbejaire
	Vallauris	Eplanade Route Napoléon
	Tourrettes-sur-Loup	2 Rue François Donnet
	Èze	Place de la Libération
	St Paul de Vence	Rue du Châteaux
	Sospel	Place du Tilleul
Coastal Villages	Villefranche-sur-Mer	Village and Port- Quai de l'Amiral Courbet
	Saint-Jean-Cap-Ferrat	Port
	Cagnes-sur-Mer	Village- Cros de Cagnes
	Golfe Juin, Vallauris	Port- Avenue des Frères Roustan
Within Small Towns	Mandelieu-La Napoule	Port
	Beaulieu-sur-Mer	Port- Quai É Whitechurch
Within Large Towns	Nice	Port(East side)-Quai des Deux Emmanuels
	Cannes	Beaches- Boulevard du Midi Jean Hibert

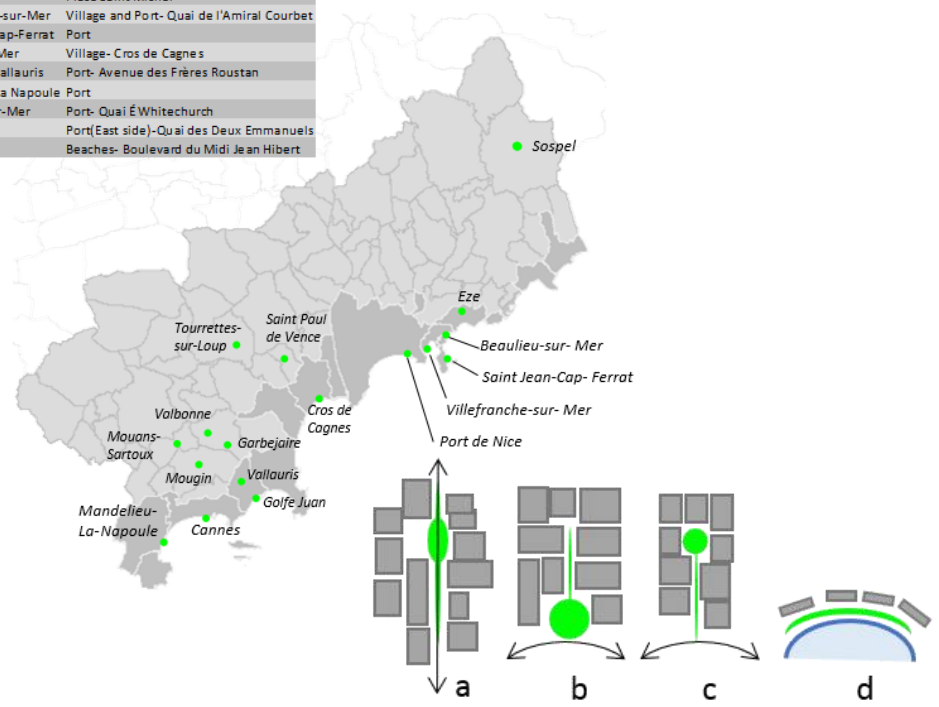


Figure 48 The identification of the isolated touristic/leisure oriented retail agglomerations, associated with small historical centres (above); Street View of RF8 (Source: Google Street View 2019)

When analysing form variables, the fragmented linear layout at the street and district level become linear only at the neighbourhood scale. The geometry of these agglomerations is fragmented at the street and district level but linear at the neighbourhood level; lower compactness is detected between the district and neighbourhood scale. When projected in the geographical space we might recognise fragmented linear distributions which are disposed along main axes in peripheral or semi-central areas.

In order to better understand the nature of this class, we might observe the following three examples:

- In Fig.49.a we highlighted the central area of the municipality of Le Cannet (local centre incorporated within the northern part of Cannes conurbation). RF9.a is detected along Bd. Carnot and Rue St. Sauveur. As we might perceive from the aerial image, the distribution of the buildings along Bd. Carnot is more fragmented, with a higher heterogeneity of building types which alternate their presence with empty spaces such as private gardens or small parking lots. On the contrary, Rue St. Sauveur corresponds to the historical core of Le Cannet with the presence of a more continuous distribution of small buildings on its sides. These two main axes represent the core of the surrounding residential neighbourhoods and give to this area an overall stretched and fragmented linear shape, following the morphology of the site and with walkable distances between these stores.

This specific example gives us also the opportunity to show another peculiarity of BN classification: when plotting the second most probable classes associated to the same area (lower part of Fig.49.a), we might observe how different subsections of this agglomeration show specific local characteristics revealing more information about their nature and possible similarity to other RF. For instance, the higher store continuity at the street level in the southern section of Bd. Carnot (towards Cannes) shows higher similarity to the continuous linear ribbons (RF5.a). Rue St. Sauveur presents characteristics closer to those found in the historic villages of RF.8: a linear distribution with a higher presence of bars and restaurants, which finds its centre in a small square (La Piazzetta). The analysis of the probability outcomes associated with each store might help in a finer understanding of the nature of each retail agglomeration and its possible hybrid and composite nature.

- Fig.49.b provides a second example of agglomeration type RF9.a: here again, we observe a fragmented linear distribution of stores along Av. Maréchal Foch, Villefranche. The streets highlighted in this specific space are disposed along the main departmental route connecting Nice and Monaco (*Baisse Corniche*) in correspondence with the historic centre of Villefranche (identified as RF.8). Similar to the previous example, this agglomeration provides the local services for the surrounding community, with a more residential orientation.
- Finally, a third example is provided in Fig.49.c where show how we might find these areas not only scattered around cities and villages but also close to core-centres agglomerations of larger cities where a more urban compact characteristics are



observed. The example shows the northern section of Bd. Grosso, Nice, where the higher fragmentation of the buildings along the main streets or the residential use of the ground-floors causes frequent interruptions in the store linear distribution at the street level. Nonetheless, the overall linear direction of this boulevard is prevalent. The fragmented pedestrian ribbons become more continuous in the southern section where higher probability of RF8 is observed.

Considering the descriptions and examples provided, we might label these agglomerations as ‘fragmented pedestrian ribbons for a residential local-neighbourhood community’. Like in pedestrian ribbons (RF5.a), we have walkable distances between stores distributed along the main vehicle/pedestrian flow direction; nonetheless, we observe a relative higher fragmentation and spacing of shops at the local level and the absence of large surfaces. When the spacing becomes more important, the linearity at the neighbourhood level is lost, leaving these concentrations of a few small stores without a coherent integration with the surrounding areas; in this case, RF9.b is observed. This retail type identifies small scattered agglomerations with street-to-district importance, last gradation of store agglomeration before the identification of isolated stores (RF.10).



Figure 49 RF9.a Fragmented pedestrian ribbons in a residential local-neighbourhood community. Some examples in Le Cannet (a), Villefranche (b) and Nice (c); Street View of RF9 (Source: Google Street View 2019).

## Isolated stores

Finally, we count 7830 isolated stores in RF10, 23.9% of the overall retailer volume. The main indicators defining this family is the small size and the low variety observed in the proximity of these stores (with an average of 3.6 typologies and a standard deviation of 3 stores in 300meter radius). While for closer radii we might observe complete isolation (27% of the stores have no other retailers within 150 m radius, 18% at 300 metres), in the 600 metres around the stores belonging to this class we still find few small retailers.

The direct consequence of the small number of stores observed in the three scales (where we have an average of, respectively, 3.4-6.1-21.8) is the punctual layout detected by geometry indicators (average geometry indicator of, respectively, 0.15-0.09-0.08); moreover, the low number of stores does not allow the identification of statistically significant patterns of functional under or over-representation. For this reason, the high redundancy of functional indicators is confirmed by the 10-fold analysis which identifies, again, in the low variety (Rich300) as the most important indicator

The quality of RFA outcomes could be evaluated when compared to expert-based assessments carried out by local agencies together with the local chamber of commerce (Agence de Déplacements et d'Aménagement des Alpes-Maritimes, 2012). Within this document, the procedure implemented for the identification of commercial districts consists of a visual delimitation of retail agglomerations and the relative expert-based interpretation. Five classes of retail districts are detected by local authorities (Fig.50): touristic-oriented (pink) and proximity retail activities (green) find a straightforward correspondence to some retail fabrics identified through RFA (RF1.a and RF1.b for the former, RF5.a and RF5.b for the latter). Within this document, medium-sized cities and small villages are not described, while specialised agglomerations are represented by their geographical locations and NAF classification group. The application of RFA allows a finer geographical characterisation of all stores in the study areas considering cities of all sizes and specialised suburban clusters within a common framework.

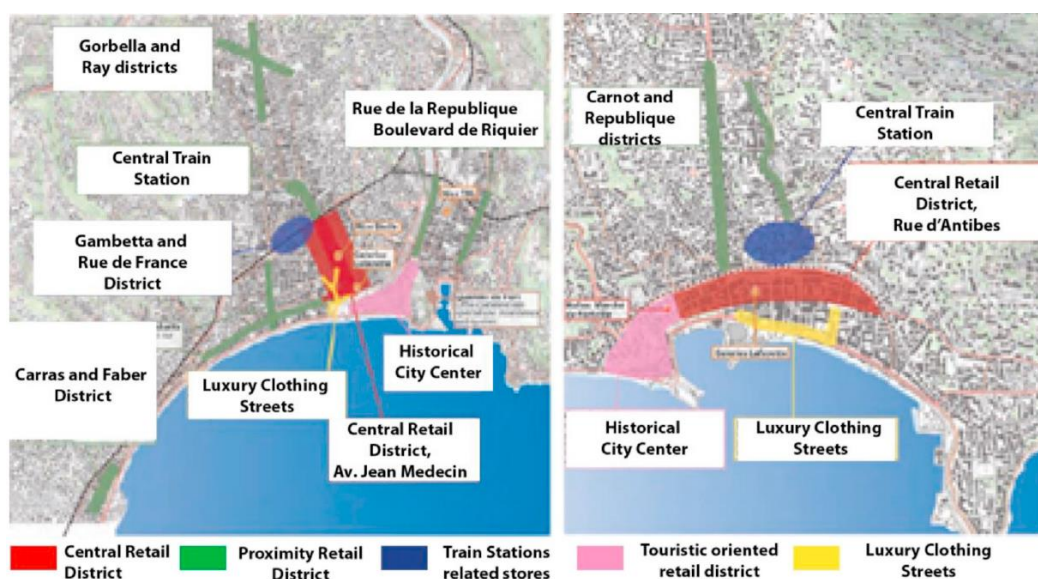


Figure 50 Retail districts identified by local authorities. Source: Interscot Commerce (2012).

## The spatial organisation of retail fabrics

The projection in the geographical space of all the 18 retail fabrics has been separately illustrated for core centres in Figure 34 and for peripheral areas in Figure 41. Despite the point-based representation, the reader might already be able to recognise clusters and their spatial organisation.

We might summaries the spatial retail organisation as follow (Fig.51):

- i) Along the coastline, larger cities with extended core areas allow the coalescence of functional clustering. Touristic-leisure oriented concentrations are always found in the correspondence of more historical urban areas; clothing and service-oriented districts surround this nucleated centrality. At higher distances from the city centre, the functional specialisation becomes weaker, and a residential-oriented retail offer composes the more or less fragmented store frontage. When the core centre is not sufficiently extended, the functional specialisation collapses in the same urban city centre. Mixed-functional city centres and small historic villages are scattered in the immediate hinterland.
- ii) Rooted in core areas and city centres, pedestrian ribbons are radially disposed, pointing to other cities/villages directions; these linear distributions are traditionally the prolongation of main axes already perceptible in more central areas. Fragmented pedestrian ribbons are found in more suburban areas and less compact urban fabric.
- iii) This centre-ribbons scheme is a module repeated several times with few variations: 4 large cores along the coastline surrounded by 12 smaller city centres and 13 villages. Surrounding these pedestrian agglomerations, suburban car-oriented ribbons are found along the main axes connecting city centres to the departmental or national network. Along these axes are nucleated planned specialised concentrations.

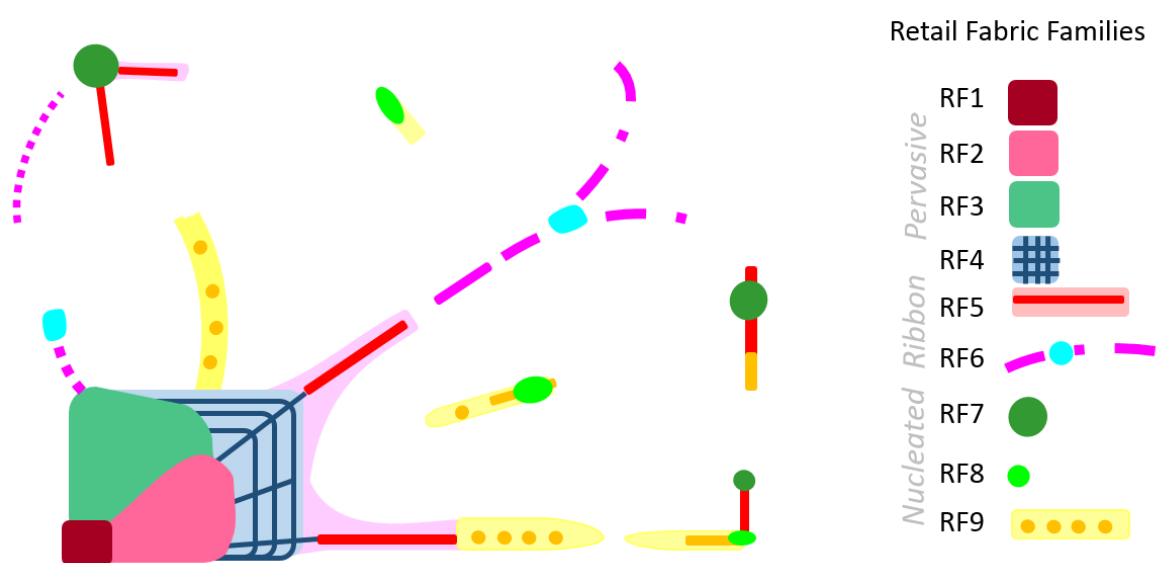


Figure 51 The spatial organisation of retail fabrics: a model for the French Riviera Conurbation

In this second Part of our dissertation, we proposed a multi-step geoprocessing approach for the identification of retail agglomeration typologies in a vast metropolitan area. It integrates a theory-based matrix of morpho-functional indicators taking into account micro-retail locations, network-constrained geo-statistics and Bayesian clustering. The method proved successful in identifying retail district typologies and their spatial arrangements in the real-world case of the French Riviera metropolitan area. Through the probabilistic content of the Bayesian clustering results, each retail fabric has been described and illustrated. The heterogeneity and polycentrism of the coastal region analysed proved the capacity of RFA method to detect a wide range of commercial districts, providing a complete analysis of the retail fabric for both central and peripheral areas. Their spatial organisation has been finally summarized in a simplified diagram describing the archetypical organisation of retail agglomerations in the French Riviera.

### **Analysing the spatial organisation of retail: recent methodological developments**

In order to better understand strengths, limitations, and perspectives of this specific approach, this conclusive section is developed around three main points: in the first part (as announced in the methodological literature review of Section 8, we provide a comparative analysis between our work and the one proposed in Dolega et al. (2019). Based on this comparative analysis in the second part, it would be possible to outline and discuss future developments and research perspectives. Finally, we discuss how RFA results could be used in urban morphology studies and, more specifically, how we integrate this work within the more global framework of this dissertation.

In the recent work of Dolega et al. (2019) entitled “Beyond retail: New ways of classifying UK shopping and consumption spaces”, a specific procedure is proposed for a multidimensional taxonomy of consumption spaces in Great Britain. From the title, we might already understand how the goal of the authors is to describe the retail agglomerations not only with retail data but with the integration of other information. Independently from the ultimate purpose of the study, when observing the overall methodological architecture of this study, we can still recognise the traditional three-step procedure previously highlighted in empirical retail geography studies (Fig.24).

In the first stage of the analysis, the identification of retail clusters is obtained through density-based spatial clustering of application with noise, DBSCAN (discussed in an earlier work of the same research group: Pavlis, Dolega and Singleton 2018). This analysis allows the authors to delineate the boundaries of 3110 retail centres in Great Britain, representing the spatial unit on which the following steps are based on.

In the second step, these agglomerations are described by both retail and non-commercial socioeconomic variables. The authors implement a matrix of 52 indicators, grouped into four categories as follow: “*Composition classifies shopping spaces by the type of store and shopping trip purpose, measured by the proportional presence of retail and service categories; Diversity focuses on the variety of goods sold and services offered and included store ownership (i.e. multiple, small multiple, independent); Size & Function identifies the various roles shopping*

*spaces have and the ways in which they interact with catchment demographics; Economic health explores both the cause and effect of a retail centre's economic performance by measuring the most popular drivers of its vitality and viability and links these to the information on each centre's hinterland."* (p.6).

Concerning the specific Size and Function group of variables, the authors implement measures of size, composition diversity and retail functional proportions.

These indicators are obtained from several data sources at different spatial resolutions: they consider variables such as economic health, retail vacancy, etc. from specific surveys about retail agglomerations (from the Consumer Data Research Centre-CDRC), socioeconomic variables such as unemployment rates, income variable, etc. at the Lower Layer Super Output Areas LSOA level (from the UK Office of National Statistic), and crime data at the postcode level from the national police website. The integration and combination of different sources of data in retail classificatory approaches represent one of the main peculiarities of this specific work.

Finally, after the exclusion of 8 variables for multicollinearity problems, the classification analysis considers the remaining 46 indicators. The authors implement a specific variation of centre-based algorithms, the k-medoids, in order to overcome possible problems generated outliers. A specific graphical representation of the classification results is visually analysed (a specific tree graph is known as clustergram, Schonlau 2002) and "used to determine the optimal number of clusters for the data at hand".

The authors select five coarse classes: i) Local retail and service centres, ii) Retail shopping and leisure parks, iii) Leading comparison and leisure destination, iv) Primary food and secondary comparison destination; v) Traditional high streets and market towns. From this first classification, 15 nested sub-clusters are further described.

The underlying goal of both Dolega and our works is the elaboration of a horizontal classification allowing a comparative assessment of sub-regions, opposed to the more traditional vertical approach seeking for a hierarchical ranking/classification of centres. Comparative studies are only allowed by systematic and robust procedures enabling the reproducibility in different spatial and temporal contexts (Reynolds & Schiller 1992, Guy 1998).

Nonetheless, the two works show substantial differences both on a theoretical and, consequently, methodological point of view. Dolega provides a taxonomy of urban centres, following the traditional approach of the Anglo-Saxon School of Potter and Davies, where the main source of innovation consists of the agglomeration detection and the integration of a large number of non-commercial sources of data. In our case study, the classification procedure is applied to the notion of retailscape: instead of grouping retail point patterns, we consider the retail context surrounding each store, individually considered without a preconceived cluster detection. Moreover, we use a relatively lower number of variables focused only on the morpho-functional characteristics of the retail distribution; our main

innovation resides in the pedestrian-behaviour approach (and scale) in the description of the spatial units with the implementation of network-based point pattern detection methods.

Both works integrate a number of dimensions and scales that are essential to capture the complexity of retail centre ecologies (Batty 2008, Coca-Stefaniak 2013, etc.): in Dolega's work, the notion of multiscale is considered from a functional point of view with a differentiation between local and national diversity through the elaboration of separate indicators for independent storekeepers, chain stores, and national firms. In our study, we consider a spatial multiscale approach considering the retailscape at the street-district-neighbourhood levels; despite the great interest of functional scale of each store, this information is not provided in our original database.

In conclusion, we might observe how the comparative goal seems to be achieved in both cases. While Dolega's procedure has been applied at the (almost) national scale, our case study has been applied to the (almost) departmental scale, allowing in both case studies a comparative analysis of retail agglomerations at two different scales. Nonetheless, the use of specific indicator based on different sources of data is both an innovative point and a limitation: Dolega's method is replicable only to the British context. On the contrary, the RFA approach considers only store location and category (marked point pattern) which is not only available at the French national scale, but also in other international contexts and could be applied both to official and non-official datasets such as Google or OSM POI.

*In this Second Part of the dissertation, we focused on the analysis of the retail spatial distribution on the French Riviera.*

*In Section 4.1, we reviewed computer-aided procedures for the analysis of the spatial organisation of micro-retail developed within the retail geography domain of studies. In Section 4.2, more recent procedures for the point pattern analysis on network-constrained spaces were also presented. Considering the specific goal of this dissertation, we discussed in 4.3 how traditional approaches present several conceptual and methodological limitations.*

*In order to overcome these limitations, a new protocol is proposed and developed in Chapter 5. The Multiple Retail Assessment is a new procedure for the street-based, bottom-up identification of retail fabrics.*

*After the identification of a location-based spatial unit in Section 5.1, a specific set of morpho-functional indicators are elaborated in Section 5.2. The specific need for an indicator describing the geometry of the spatial layout of points stimulated the development of a new indicator based on fractal theories as described in Section 5.3. Once the morphofunctional indicators are selected and implemented, a Bayesian Network classification approach is finally implemented in Section 5.4.*

*Finally, Section 5.5 describes the results of the RFA procedure. The method proved successful in identifying retail district typologies and their spatial arrangements in the real-world case of the French Riviera metropolitan area; a spatial model for the retail on the French Riviera is finally proposed.*

*The Second Part of this dissertation has been entirely focused on the description and analysis of the retail spatial organisation. Together with the simple count representation of stores describing both presence/absence and magnitude of the retail system, we are now also able to associate to each segment a categorical variable describing the morpho-functional characteristics of the retail distribution.*

*In the next part, the focus of the analysis will be given to the second component of this dissertation: the description of the physical properties of the urban form.*

*Dans cette seconde partie de la thèse, nous nous sommes concentrés sur l'analyse de la distribution spatiale des commerces de détail de la région de la Côte d'Azur.*

*Dans la Section 4.1, nous avons examiné les procédures assistées par ordinateur pour l'analyse de l'organisation spatiale de la micro-distribution développée dans le domaine d'études de la géographie du commerce. Dans la Section 4.2, des procédures plus récentes d'analyse de modèles de points sur des espaces restreints au réseau viaire ont également été présentées. Considérant l'objectif spécifique de cette thèse, nous avons expliqué dans la Section 4.3 comment les approches traditionnelles présentaient plusieurs limitations conceptuelles et méthodologiques.*

*Afin de surmonter ces limitations, un nouveau protocole est ainsi proposé et développé dans le Chapitre 5. La Retail Fabric Assessment, une nouvelle approche ascendante (bottom-up) d'identification des tissus commerciaux.*

*Après avoir déterminé une unité spatiale basée sur la localisation dans la Section 5.1, un ensemble spécifique d'indicateurs morpho-fonctionnels est élaboré dans la Section 5.2. Le besoin d'un indicateur spécifique décrivant les caractéristiques géométriques de la distribution des points, a conduit à la mise au point d'un nouvel indicateur basé sur les théories fractales, qui est décrit dans la Section 5.3. Une fois les indicateurs morpho fonctionnels sélectionnés et mis en œuvre, une approche de classification par réseau bayésien est finalement employée à la section 5.4.*

*Enfin, la section 5.5 décrit les résultats de la procédure de RFA. Cette méthode a permis d'identifier les typologies de quartiers de vente au détail et leurs arrangements spatiaux dans notre zone d'étude de la région métropolitaine de la Côte d'Azur. un modèle spatial pour le commerce de détail sur la Côte d'Azur est finalement proposé.*

*La deuxième partie de cette thèse a été entièrement centrée sur la description et l'analyse de l'organisation spatiale du commerce. Parallèlement à la simple représentation du nombre de magasins décrivant à la fois la présence / absence et la quantité d'activités commerciales pour chaque segment de rue, nous pouvons également associer une variable catégorique décrivant les caractéristiques morpho-fonctionnelles du tissu commercial.*

*Dans la partie suivante, l'analyse portera sur la deuxième composante de cette thèse: la description des propriétés physiques de la forme urbaine.*



**Part 3**  
*Measuring the Form  
of the Physical City*

*Similarly to what has previously been observed in retail geography, and also in urban morphology, we might identify several approaches for the analysis of the physical form of cities.*

*In the second half of the XX<sup>th</sup> century, new concepts have been developed in the discipline of urban morphology with different approaches inside the British, Italian, and French traditional schools. At the same time, in the North American context, the notion of urban design is developed and rapidly spread in the north European countries. While the urban morphology focuses on the study of the urban form system and spatial organisation, the urban design focuses on aesthetical aspects of the urban form at a smaller scale, which covers the design of individual façades, to a street environment, all the way up to a small urban fragment.*

*In these two disciplines, quantitative approaches developed between the '60s and '80s relied heavily on manual measurement and the interpretation of calculus. Starting from the '80s, computer-aided mathematical models were developed for the analysis of the morphological characteristics of the urban form.*

*During the same period, the implementation of graph theories to the specific case of the street network took place, and a new approach was proposed for the analysis of the urban form.*

*For several years, these three approaches (urban morphological, design, and street-network configurational studies) have mostly been separately developed and implemented.*

*It is only in more recent years that several academics highlighted the need to bridge these different methodological approaches within a common theoretical and methodological framework for the study of the urban form (Kropf 2009, 2017). The idea underlying this concept is that the combination of different procedures might be able to highlight the relative importance of various facets of the urban physical form.*

*As discussed in the first part of this dissertation, we are specifically interested in understanding how the combination of these facets might explain the retail distribution. To achieve this goal, we proposed to implement the analysis of urban form with street-based approaches, sharing the same definition of spatial unit and point of view.*

*While configurational, and, more recently, urban design both propose the integration of the notion of pedestrian perspective in their quantitative analysis of the urban form, the procedures proposed in urban morphological schools traditionally consider the aerial point of view.*

*The implementation of a coherent set of measures of urban form sharing the same street-based perspective becomes the main challenge of the third part of the dissertation.*

*Based on these considerations, Part 3 is dedicated to the description and implementation of urban form measures with both traditional and new procedures developed, ultimately to respect the thematic and methodological framework of this dissertation.*

*De manière analogue à ce que nous avons pu observer précédemment pour la géographie du commerce, nous pouvons également identifier plusieurs approches dans le domaine de la morphologie urbaine pour l'analyse de la forme physique des villes.*

*Dans la seconde moitié du XXe siècle, de nouveaux concepts ont été développés dans la discipline de la morphologie urbaine avec une approche différente selon les écoles traditionnelles britannique, italienne et française. Parallèlement, dans le contexte nord-américain, la notion de design urbain s'est développée et rapidement étendue dans les pays d'Europe du Nord. Alors que la morphologie urbaine se concentre sur l'étude du système de formes urbaines et de leur organisation spatiale, le domaine du design urbain met l'accent sur les aspects esthétiques de la forme urbaine à une échelle plus fine, prenant notamment en considération les façades, l'environnement d'une rue, voir un petit fragment urbain.*

*Dans ces deux disciplines, les approches quantitatives développées entre les années 60 et 80 reposent largement sur la mesure manuelle et l'interprétation du calcul. À partir des années 80 cependant, des modèles mathématiques assistés par ordinateur ont été développés pour l'analyse des caractéristiques morphologiques de la forme urbaine.*

*Au cours de la même période, lors de l'application des théories des graphes au cas particulier du réseau routier, une nouvelle approche est avancée pour l'analyse de la forme urbaine.*

*Depuis plusieurs années, ces trois approches (études de morphologie urbaine, design urbain et configuration de réseau urbain) ont été principalement développées et mises en œuvre séparément.*

*Ce n'est que plus récemment que plusieurs universitaires ont souligné la nécessité de relier ces différentes approches méthodologiques dans un cadre théorique et méthodologique commun pour l'étude de la forme urbaine (Kropf 2009, 2017). L'idée sous-jacente à ce concept est que la combinaison de différentes procédures pourrait permettre la mise en évidence de l'importance relative des différentes facettes de la forme physique urbaine.*

*Comme discuté dans la première partie de cette dissertation, nous souhaitons comprendre la façon dont la combinaison de ces facettes pourrait expliquer la distribution du commerce. Pour atteindre cet objectif, nous avons proposé de mettre en œuvre l'analyse de la forme urbaine avec des approches orientées rue partageant la même définition d'unité spatiale et de point de vue.*

*Alors que l'approche configurationnelle et, plus récemment, le design urbain proposent tous deux l'intégration de la notion de perspective du piéton dans leur analyse quantitative de la forme urbaine, les procédures proposées dans les écoles de morphologie urbaine considèrent traditionnellement le point de vue aérien.*

*L'utilisation d'un ensemble cohérent de mesures de la forme urbaine partageant la même perspective de rue devient le principal défi dans la troisième partie de la thèse.*

*Sur la base de ces considérations, notre troisième partie est consacrée à la description et à la mise en œuvre de mesures de la forme urbaine en ayant recours aux méthodes traditionnelles mais aussi grâce à de nouveaux procédés développés dans le but final de respecter le cadre thématique et méthodologique de cette thèse.*

## 6. Computer-Aided Approaches for the Analysis of the Urban Form

As stated by Erin et al. (2017, a collective paper in which I personally took part), urban morphology is defined as the study of the urban form; it deals with the analysis of the physical layout and design of a city, a town or an urban fragment within these. Urban morphology involves the study of patterns of its components and the processes of their formation and transformation (Vernez-Moudon 1997).

As for economic geography (Section 1.2), the literature on the analysis of urban form also has its origins in the seminal works of urban geographers of the early XX<sup>th</sup> century in German-speaking countries. The works by Fritz, Schlüter, Ratzel, and later Hassinger, Geisler, Bobek and Louis, laid the foundations of an accurate historical and geographical analysis of urban forms of the European cities, taking into consideration street layouts, building styles and arrangements, function locations and centralities (Hofmeister 2004). The classical urban morphology schools after WW2 in Britain, Italy, and France developed new original concepts in the study of urban form.

The British school found its origins in the micro-scale studies of settlements with both a historical and geographical approach and with subsequent development of 'town-plan analysis' by M.R.G. Conzen (Whitehand et al. 2001). Similarly, the Italian school is more attentive to the historical description of building typologies, their variations and transformations, and their role in the genesis of the urban form, from the seminal work of Muratori (1959) in Venice, to the analysis by Caniggia and Maffei (1979). Finally, the French school is based on the analysis of the relationship between parcel, streets, and buildings, and the relation to space and society in the works of Depaule, Castex, Panerai, Borie, Denieul (Samuels et al. 2004).

The works of the Conzenian and Muratorian schools jointly developed the concept of the urban fabric (or 'urban tissue' as in Larkham and Jones 1991). This is defined as a typical pattern of buildings, parcels and streets, characterising the form of the physical city in a given historical and cultural context, at the scale of an urban fragment (Muratori 1959, Conzen 1960). The concept of the *morphological region* is also proposed to describe urban sectors characterised by consistent extents of a single urban fabric and to produce more global schemes of the morphological organisation of the city.

Although urban morphology in its early stage did not feel the need to engage in the algorithms of spatial analysis, the endeavour of the French school of urban morphology gives rise to the systematisation of quantitative measurements of urban forms. Yet, quantitative urban geographers (like Haggett et al. 1977) were more interested in questions of land use and location of urban functions than in the analysis of urban form at the fine grain of urban fabrics. It is no surprise that detailed measurements were often carried out on limited case studies (paradigmatic examples of urban fabrics, small towns, city-centres).

The Versailles school (i.e. Castex et al. 1980), as well as the Parisian school (Borie and Denieul 1984), proposed a systematic way to qualify and quantify spatial relations among form

elements. Buildings, plots, and streets are thus characterised by dimensional, geometrical (i.e. angular) and topological relations that define an urban fabric much more than building styles and functional occupation. Site characteristics (topography, hydrography) and their relations to urban fabrics can also be analysed, as already shown by the German forerunners of urban morphology. Such analytical approaches can be applied via the manuals developed by Borie and Denieul (1984) and Allain (2004).

Quantitative approaches developed between the '60s and '80s relied heavily on manual measurement and interpretation of calculus. The only remarkable exception in this overall picture was the research carried out at the Centre for the Land Use and Build Form Studies (LUBFS) at the University of Cambridge. Martin et al. (1972) first introduced the use of computer-aided mathematical models to analyse the morphological characteristics of the built environment. Kruger (1977) and Steadman (1973, 1983) also introduced the use of graph theory to analyse and characterise both urban fabrics and three-dimensional buildings. For the first time, quantitative approaches in urban morphology became geoprocessing methods that were being developed by spatial analysis.

From this moment on, we can distinguish two broad directions of research within quantitative urban morphology: i) configurational analysis and ii) spatial morphological analysis.

## 6.1 Configurational approaches

Inspired by the seminal works of the LUBFS, Hillier and Hanson (1984) and Hillier (1998) developed a coherent theoretical framework, as well as methodological procedures for the implementation of graph theories to the physical organisation of spaces, named Space Syntax. The object of the analysis is the configuration of spaces both at the building level and the urban grid aiming at measuring the capacity of a spatial configuration to produce encounter patterns. Space is simplified as the juxtaposition of voids separated by visual physical obstructions: while, at the building level, configurational approaches focus on the organisation of rooms enclosed by internal walls, at the urban scale, they analyse the spatial structure of public spaces.

From these theoretical premises, two main procedures have been developed. Visual Graph Analysis (VGA, Turner et al. 2001b) performs SSx analyses at the human-scale of a building or a small urban fragment: the original notion of isovist (Benedikt 1979) is implemented to points, paths or perimeters within planar spaces. The final goal is to obtain an undirected graph connecting all the inter-visible points of planar surfaces (Fig.52-left). A second procedure, the Segment Analysis, is implemented considering the urban grid at a wider scale: adjacent convex spaces are in this case simplified as a network of interlinked visual axes (Fig.52-right) (street-like elements, Hillier 1993). Independently of the modelling approach to space, VGA and Segment Analysis implement the same graph-theoretical algorithms to quantify and describe configurational properties.

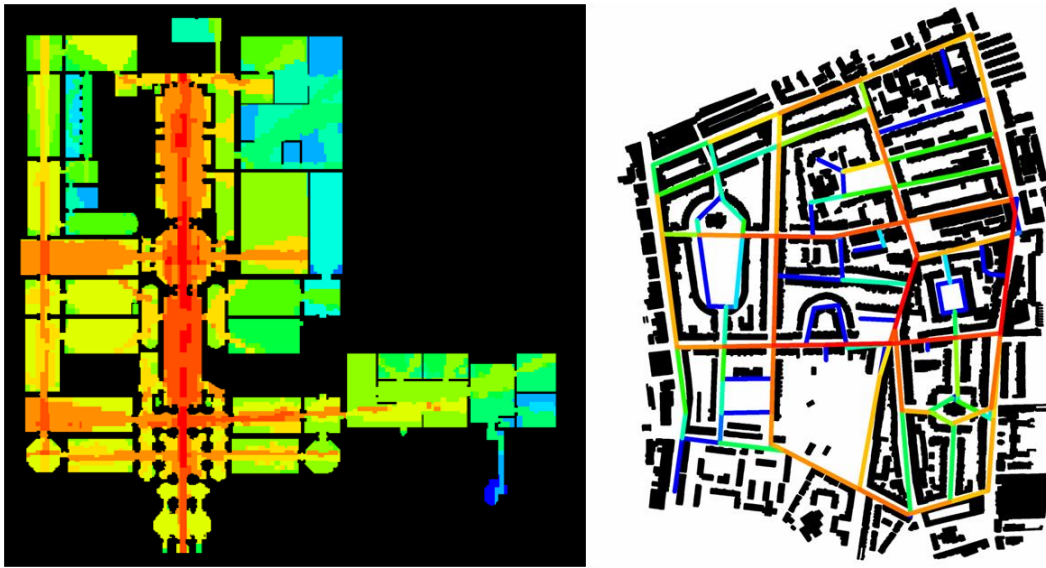


Figure 52 Example of VGA analysis (on the left) and segment analysis (on the right) (source: UCL Space Syntax: <http://otp.spacesyntax.net/>).

While SSx segment analysis investigates the dual<sup>60</sup> graph representation of the axial network, where axes are represented by nodes and intersections by arcs, a second group of configurational analysis have been more recently proposed considering a primal graph representation of the street network. Multiple Centrality Assessment of Porta et al. (2006a) and the metrical/directional centrality indicators of Peponis (Peponis et al. 2007, 2008) independently developed graph theories from structural sociology (Freeman 1977, 1979) on the traditional representation of street networks proposed by transportation models. Street networks in the primal approach are modelled by graphs where junctions and streets are respectively represented by nodes and arcs. These alternative approaches overcame the main critics about the need of a metrical measure of centrality (Ratti 2004).

Further developments and proposals have been considering alternative modelling possibilities: street names (Jiang and Claramunt 2004a, 2004b), angular analysis (Turner, 2001b), continuity analysis (Figueiredo and Amorim, 2005), Intersection Continuity Negotiation (Porta et al., 2006b, Lagesse 2015), Mark Point Parameter Analysis (MAPPA, Cutini et al., 2004), natural roads (Jiang et al. 2008) etc. Each of these approaches proposes alternative ways to interpret the human movement behaviour in urban spaces by simplifying two elements: the street representation and the modelling of shortest paths.

Place Syntax Theory (PST, Ståhle et al. 2007) and Urban Network Analysis (Sevtsuk and Mekonnen 2012) extended the configurational approach introducing buildings and parcels served by the street network. While in urban studies these approaches allow measuring configurational properties at the buildings/blocks level, in transport studies this same

<sup>60</sup> Recent work of Marshall et al. 2018 detect some inconsistencies in the scientific literature in the use of the notions of *primal* and *dual* representation of the network graph (for further details see in Marshall et al. 2018 section 3.2); nevertheless, “a definitive conclusion on this issue awaits a dedicated review of all relevant kinds of graph and their relations”. For this reason, in this work we still use the conventional primal/dual graph distinction as originally proposed by Porta et al. (2006)

procedure helps to estimate the movement/accessibility given by the overall characteristics of the urban form.

Graph analysis approaches applied to these different modelling procedures might highlight different network properties; specific advantages and limitations are associated with each modelling procedure and the choice of the specific method is strictly dependent on the purposes of the research.

A detailed description and comparison of every configurational analysis approach developed in the last twenty years goes beyond the goal of this dissertation. Nonetheless, understanding the main differences among these analytical procedures is of paramount importance to understand how we chose the most suitable approach to our research.

In order to better understand the overall framework of configurational studies, the recent work of Marshall et al. (2018) provides an important overview of different modelling approaches going beyond the traditional definitions from the urban morphological domain of studies (including analytical procedures independently elaborated in the research fields of transportation and physics). The goal of this work is the reorganisation of the analytical approaches and vocabulary independently developed in different research domains for the analysis of street network procedures. On the contrary, Fusco and Tirico (2016), provide a systematic description of configurational approaches within the urban morphology domain of study.

This second work reorganises configurational studies in a simple scheme (reported in Figure 54) which might help us in a synthetic description of methodological differences among configurational approaches. The authors propose a categorisation through the combination of three fundamental aspects: (i) graph representation (ii) street network modelling, and (iii) network metrics.

Graph Model		Primal			Dual			
Network Distance		Topological	Angular	Dimensional	Topological		Angular	Dimensional
Modelling of Network Elements	Angular	Ma.P.P.A. (Cutini, Petri, Santucci 2004)			Axial Analysis (Hillier and Hanson 1984)	ICN (Porta, Latora, Cardillo 2006 and Lagesse 2015)	Angular Analysis (Turner 2000)	
	Topologic			MCA (Porta, Latora, Cardillo 2006)	Road centrelines (Turner 2007)		Continuity Analysis (Figueiredo and Amorim 2005)	
	Dimensional			MCA (Sevtzuk and Meckonnen 2012)				
	Socially Defined			Metrical/Directional Analysis (Peponis et al. 2008)		Street Names (Jang and Claramunt 2004a,b)		

Figure 53 Overview scheme of configurational approaches in urban morphology studies. (Readaptation from: Fusco and Tirico 2016)

After a first broad differentiation between primal and dual models of the street network, two further aspects should be considered. “The first one concerns the modelling of the network, i.e. the way we identify the physical elements of the street network. There are at least four

*different options: topological elements (defined by connections which are consistent through deformation and magnification/ reduction of the urban space), angular elements (defined by directions which are consistent through magnification/reduction but won't resist deformation), dimensional elements (segments of a given length, which won't resist neither deformation nor magnification/reduction), socially defined elements (like street names or any other cognitive and social recognition of form elements, which depend on the social representation of urban space). A second additional aspect to be considered is the way distance between form elements is defined within the network: it can be topological (number of connections, whatever the nature of the connections is), angular (modelling the psychological impedance to change of direction in movement) or dimensional (metric or temporal distance on the network, modelling the physical impedance to movement)."* (Fusco and Tirico 2016, p.2).

Based on these large methodological choices, configurational approaches developed specific vocabularies to describe network configurational properties. On one side, Peponis, Porta and Sevtzuk adopt the traditional parameters and names originally proposed by Freeman (1977, 1979). On the other, Space Syntax and derivatives use the vocabulary specifically developed by Hillier and Hanson (1984). Nonetheless, configurational indicators are comparable when differences in vocabulary are overlooked.

- i) The simplest indicator is the **Reach / Node Count** (NC), quantifying the number of form elements which can be reached from every node within a given radius of analysis. As discussed in 5.3, the count of the number of neighbouring elements is a traditional metric which configurational analysis shares with other quantitative spatial approaches such as statistics and fractal studies; street network constrained counting is a specific (and more recent) variation adopted by social science disciplines in order to integrate a more accurate measuring of spatiotemporal distances.
- ii) The **Farness / Total depth** (TD) indicator is a measure of how far a form element is from all other elements in a given radius; differently from the previous indicator, this measure considers the metrical/angular/topological distances of every destination. By normalising it through the number of form elements (Reach/NC) within the same radius, we obtain the **Normalized Farness / Mean Depth**.<sup>61</sup>
- iii) The **Integration** index of Axial SSx is a further normalisation of mean depth, considering as values of reference theoretic street networks (Hillier 1996). Integration is traditionally considered as a predictor of the potentials of each element within a given radius to be a highly desired destination. Hillier's integration measure is implemented as  $NC^2/TD$ . The equivalent measure of the Integration in the primal graph might be obtained as  $Reach^2/Farness$ ; likewise, considering the *Closeness* as the inverse of the *Farness* we might measure the **Normalised Closeness** as  $Reach^2 * Closeness$ .

---

<sup>61</sup> On a small radius (<500 metres), Integration will simply approximate Node Count.



- iv) The geometrical shape of individual streets and their layout represent two aspects that SSx integrates with the modelling of the street network elements as longest visual axes; the angular version of SSx further explores how the geometrical relative disposition of axes considering the shortest angular path. On the contrary, in the primal representation, the sinuosity of both street segments and of shortest paths are specifically considered by the **Straightness**. This indicator is a centrality measure similar to the Reach, where each reachable destination is weighted by the ratio between Euclidean and street-network distances.
- v) The **Betweenness / Choice** is a measure of how many shortest (metrical, angular, topological) paths pass through a form element when considering all shortest paths within a given radius. While the three groups of indicators presented so far describe, with different metrics, the proximity of each street network element to the other elements (to-movement centrality potential, Hillier 1996, 2012), this measure considers the intermediary position of a form element considering the movement generated by the network configuration (through-movement potential according to Hillier).

Further normalisations of Choice and Integration indicators through the Total Depth are necessary to enable the comparison between systems of different sizes (Hillier 2012, Al-Sayed 2018, etc.).

As previously indicated, the goal underlying configurational studies consists in finding the best generalisation of human individuals' behaviour, which allows for describing the overall movement patterns within urban spaces; the only way to prove the superiority of a specific methodology requires its validation through empirical data collected of pedestrian counts.

The debate comparing and evaluating the goodness of a specific configurational procedure is far from being closed, and we expect in the next few years a growing number of studies supported by the diffusion of advanced tracking technologies, providing new sources of information (i.e. cameras, Wi-Fi, GPS tracking systems) and substituting traditional mechanic or electronic counting methods. Nonetheless, the regulatory framework might still impose significant constraints on their application, delaying the diffusion of these advanced counting procedures.

While in the scientific literature we might find a large number of works implementing and comparing different configurational approaches for the estimation of movement flows, when studying retail distribution, traditional Axial SSx and MCA represent the two main choices applied by academics (see Tab. 2, column "Configurational Approach"). A significant growth of the latter approach might be observed in more recent years. These works illustrate how configurational properties represent an essential locational factor for retail distribution beyond the specific network modelling approach.

Comparative studies between configurational approaches usually consider pedestrian counts to measure and validate both network modelling procedures and the relative importance of each indicator. When the time comes to explain the retail distribution, we have not found any work demonstrating the superiority of a specific configurational approach: of all

configurational approaches developed in the last twenty years, traditional axial SSx and MCA are the most implemented, independently considered, in urban studies.

Future endeavours should consider this specific goal (as we further discuss in the conclusive section): only the implementation of different configurational approaches within the same study area might provide further insights about the most adapted procedure to use when describing the localisation of retail distribution. However, this specific methodological comparison goes beyond the goals of this dissertation.

As discussed in the first part of this dissertation, network configurational properties are here studied in combination with other urban morphological characteristics to evaluate their collective importance as *retail locational factors*. The final goal of this research is to better understand the role of the urban form characteristics (the combination of the *fabric* and *network effect*) within an 'extended' Movement Economy theoretical framework.

Therefore, in this dissertation, we limit our choice between those configurational approaches which have already been studied and which have demonstrated their efficacy in the study of retail distribution (with the awareness that other configurational approaches might show similar or even higher correlations).

Since one of the four main goals of this dissertation is also the introduction of detailed urban morphometric descriptors (Section 2.3), we should also consider how VGA or other derivatives of visual analysis could be used as possible solutions. VGA has been implemented for the study of retail centre organisation in Penn and Turner (2002), Fong (2003,2005), Koch (2007), Min et al. (2012) etc.; Mohamed and Van Nes (2017), Vialard (2017), etc. compared network configurational properties of planned and unplanned shopping centres. The focus of these studies is the description of both the internal and surrounding configurational properties of these retail complex formats; their final goal is to provide a systematic analysis of the architectural properties of shopping centres and their impact on human behaviour (similarly to those works in behavioural retail geography described in Section 1.3).

Nonetheless, when studying large urban spaces within an entire region (extended for 1500 square kilometres and almost 13 thousand kilometres of streets), ranging from compact centres to natural areas, it is evident how VGA approaches would require extreme computational power and memory capacities.

Beyond the technical challenges imposed by VGA, the difference between VGA and Segment Analysis is mainly related to the spatial scale of the problem under analysis. Moreover, when studying planned shopping centres, what we should consider is that locational factors of these specific retail formats operate under specific criteria which go beyond the urban morphological and configurational properties of space. Rather than being influenced, they directly modify the morphology and accessibility of the surrounding urban spaces. Moreover, retail typology, mix and spatial organisation within these formats are also directly planned with the goal of maximisation of the overall incomes and attractiveness of the shopping centre. In this dissertation, on the contrary, we focus on the spatial organisation of individual

stores seen as independent agents within a vast geographic region. Our goal is the identification of those urban physical conditions where micro-retail distribution and agglomerations occur, without the direct influence of a controlling economic agent.

Before discussing the specific configurational approach and the set of indicators implemented in this dissertation, it is therefore essential to investigate the second group of urban analysis approach developed in the last few decades. Specific analytical procedures have been developed, implementing geoprocessing and spatial analysis for the typo-morphology description of urban fabrics.

## 6.2 Quantitative Typo-morphological approaches

A second direction of the research in quantitative analysis of urban morphology considered the advances of GIS platforms and spatial analysis algorithms; from the late '70s, new computer-aided approaches have been proposed for the spatial analysis of the urban form.

While simplicity and robustness represent the two main factors at the origin of the success of configurational analysis, on the contrary, a spatial analysis approach gathers a broader range of more or less sophisticated procedures considering specific goals, spatial scales and data sources. The combination of these aspects gives to the urban spatial analysis literature a wider richness of methodological procedures with higher flexibility to new data sources and data analysis techniques, such as artificial intelligence and data mining techniques.

Here again, proposing a complete literature review of every methodological approach for the urban morphology analysis would result in a complex challenge, going beyond the core goals of this dissertation. The endeavour to summarise and describe with a systematic approach the quantitative analysis of urban form has been recently undertaken by Fleischmann (2017). The author considers five main criteria for the classification of urban form studies: Scale (from the neighbourhood to metropolitan level), Goal (comparison/classification, evolutionary patterns and performance measuring), Metric (single or multiple), Time (synchronic or diachronic), Mode (quantitative, qualitative or mixed approach). Following the Fleischmann classificatory approach, we can better specify the characteristics of the spatial analysis works here reviewed: in this section, we focus on studies investigating urban form with a classificatory, synchronic, quantitative approach at the intraurban scale.

The implementation of classificatory approaches in urban morphology studies finds its theoretical background in the aforementioned European schools, where the notion of typo-morphology has been developed. Typology refers to the systematic study of architectural forms, urban design, and landscape architecture, while the notion of Morphology denotes the study of larger urban structures and patterns. The final goal of typo-morphology studies is to recognise the spatial structure of an urban fragment, village, town, city or metropolitan area by examining the patterns of its component parts (Moudon 1994).

*“Typomorphology is the study of urban form derived from typical spaces and structures (Moudon 1994). It presumes the existence of certain morpho-structural similarities among the variety of existing built forms, which allow for their grouping into intrinsically distinct categories, called types. The type is an elemental object that embodies a morphological idea, and has little or nothing to do with function, even if it is sometimes wrongly equated with it. Indeed, buildings and cities are largely independent of their use at one certain moment in time and can shelter many different functions through history, while keeping their form generally intact (Rossi 1966). Urban typo-morphological studies are important for two main reasons. Firstly, they constitute a way of understanding the shape of cities and their evolution [sic]. They try to unveil the fundamental elements that construct the city at each moment in history, their permanence, disuse or mutation through time, and the fundamental physical and spatial relations that they establish between themselves. Secondly, typo-morphological studies can also contribute to inform urban planning and design, whether as preliminary analytical procedures of the design and planning processes, or as prescription instruments for the implementation of urban plans and policies concerning the physical and spatial structure of the city (Moudon 1994).” (Serra et al. 2012).*

Three main differences should be highlighted between configurational and typo-morphological approaches. Firstly, the spatial approach: the former investigates the relative/systemic dimension of form elements (street-like segments) and how they aggregate into neighbourhoods and cities. The latter focuses on the geometrical/compositional description of individual components of urban form (Marshall 2005) and the relationship between contiguous, directly connected, closely related elements. Secondly, the urban form element under investigation: while configurational approaches focus on street (street-like) network analysis, typo-morphology analyse different components of the urban fabric among which include buildings, streets, parcels, and the site morphology. Thirdly, the goal of the analysis: configurational studies propose measures able to estimate the potential movement, while typo-morphology focuses on the identification of urban fabrics and their geographical and historical evolution.

Classical typo-morphological schools privileged qualitative approaches; quantitative analyses were limited by both the computational power and data availability to small urban fragments. Only in the last few decades have quantitative computer-aided procedures seen an impressive development providing more robust and scientific support to this relatively young discipline.

Quantitative methodological procedures developed in urban typo-morphology studies share a similar three-step *modus operandi* (i.e. Gil 2012, Song et al. 2013 etc.). **Step A)** A specific spatial unit is chosen, considering the urban elements and the scale of the analysis. **Step B)** A set of size and shape descriptors quantify the characteristics of the urban form (morphometric descriptors). **Step C)** The implementation of multivariate analysis and classification procedures are supported by graphical representations or more sophisticated classificatory approaches.

The development of these quantitative multivariate protocols offer two main advantages: *“firstly, they allow for the definition of types based on multiple variables in a precise and repeatable manner, enabling the study of large samples and the comparison between both*

*cities and regions; secondly, they frame design choices in terms of types without being fixed and so open up for design explorations where the relation between the variables can be challenged to propose new types.” (Berghauser Pont et al. 2017)*

As the reader might notice, this three-step procedure is similar to the traditional approach of retail geography previously discussed (Chapter 4) for the identification of morpho-functional retail agglomerations and their spatial organisation.

An alternative research design might also be implemented in comparative/classification studies: archetypical typologies are firstly defined and spatially delineated through the expert-based knowledge of the analyst and subsequently, the internal morphometric properties are measured and compared (as in Hermosilla 2014, Dibble et al. 2016, Venerandi et al. 2015, Fusco and Tirico 2017, etc.). This second procedure is traditionally implemented with the final goal of studying the relative importance of a specific morphological descriptor of phenomena (or set of variables) within emblematic typo-morphological regions.

### **Streets, Blocks and Building**

After this broad description of quantitative typo-morphological approaches, we can now review more in detail some of the most relevant procedures representing the main methodological references at the origin of our work.

Considering the specific focus on street-based approaches for the identification of urban morphological context, the first reference which we should mention is the work of **Marshall (2005)**. The author proposes an innovative quantitative approach for the description, classification, and identification of street layouts considering geometrical descriptors such as topology, length, area, and orientation measures. Firstly, he proposes a taxonomic organisation of street patterns based on geometrical characteristics. Every street pattern might be directly fitted in Marshall’s taxonomy or recognised as hybrid patterns considering specific spatial combinations and the scale under analysis. Secondly, the author proposes a simple methodological procedure for the recognition of these patterns in empirical studies based on few quantifiable form descriptors: the number of street-cells (surface partition delimited by streets), of dead-ends (cul-de-sac), of T and X junctions. These four simple measures are combined in two couples of complementary parameters: the proportion of T or X-junctions on the overall number of intersections and, similarly, the ratio of street-cell and cul-de-sac. These two parameters plotted on a two-dimensional diagram (ABCD diagram, Fig.54), allow an easy visual comparison of their location in the space of all possible street layout form combinations.

Marshall’s ABCD procedure is one of the first quantitative analytical methods allowing the typo-morphological description of urban spaces with an objective and reproducible approach. Typologies of patterns are defined at a given scale and for a given urban fragment. Nonetheless, similarly to configurational analysis, the only planar street pattern is considered; buildings, plots and their spatial arrangement, which represent essential elements for the urban fabric definition, are here overlooked.

Following the approach of Marshall, more recent works such as Louf and Barthelemy (2014), Boeing (2017), etc. propose innovative classificatory approaches for the identification and quantification of street patterns, considering several street network indicators. These studies are traditionally applied to metropolitan areas, cities and large urban sectors (i.e. boroughs of New York).

The application of street pattern analysis as in the aforementioned references might be employed here at the neighbourhood scale. Nonetheless, as discussed in Section 2.3, we think that the physical context of the urban fabric can be an important factor in explaining retail distribution in the city, beyond the mere description of the street-network. For this reason, in this dissertation, we are interested in the identification and description of urban fabrics as defined in the traditional typo-morphological theories considering the layout made by the combination of streets, buildings, and blocks in a given urban fragment.

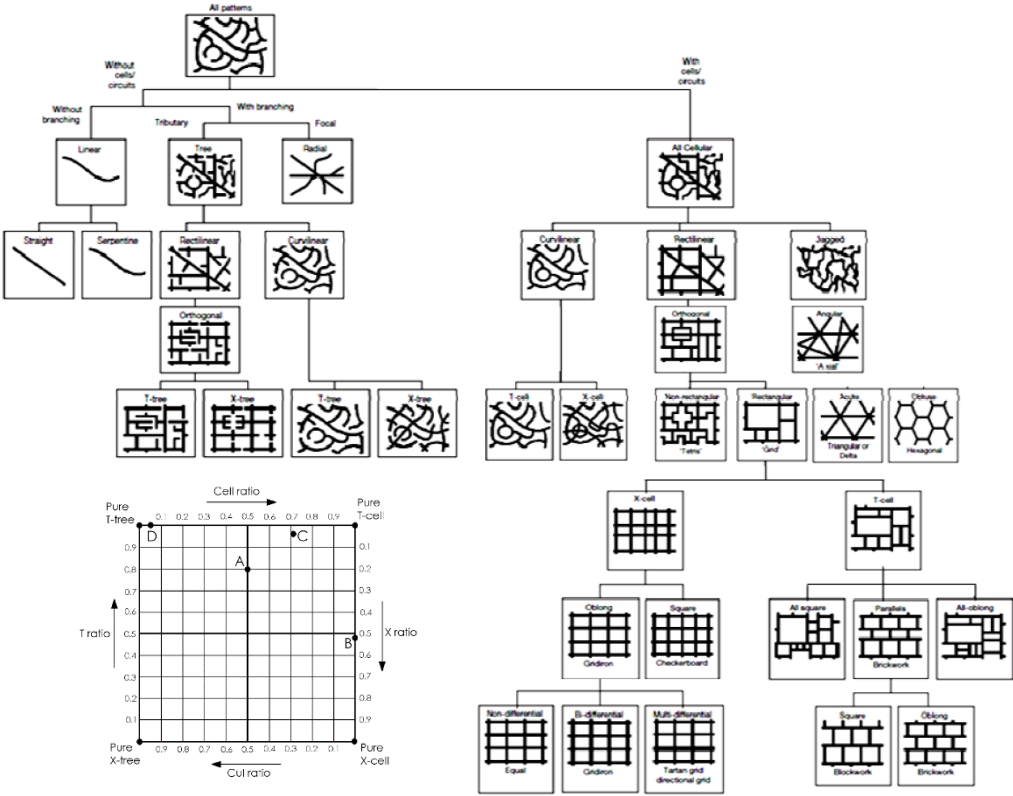


Figure 54 Marshall's Street pattern taxonomy and ABCD diagram (Source: Marshall 2005)

The idea of describing different urban contexts through the implementation of a limited number of simple geometrical parameters is also shared by the procedure developed in **Berghauer Pont and Haupt (2004, 2007, 2010)**. The traditional notion of building density is here developed as a combination of four indicators, implemented at the block level: i) Floor Space Index (FSI) describing the built-up intensity, ii) Ground Space Index (GSI) measuring the compactness, Open Space Ratio (OSR) as a measure of pressure on non-built space and, finally iv) the Building Height (L). Each block is numerically distinguished and described by these four variables. As for the graphical representation of the ABCD method of Marshall, these values

are visually represented on a specific graph, the Spacemate diagram (Fig.55). This graphical procedure allows, here again, the visual comparison and classification of urban block typologies. In the last ten years, a large number of studies have been implementing and developing Spacemate approaches both from both empirical and theoretical point of view (Steadman et al. 2013).

Spacematrix gives the possibility to describe urban morphological patterns considering both buildings (in 2.5 dimensions) and street-blocks. Street patterns are only implicitly considered. What Marshall calls the street-cell is here used as the spatial unit of the analysis; three of the four indicators are therefore strictly dependent on the underlying street network spatial partition. Nonetheless, information about the connectivity of the street network and the relationship between built form and streets are overlooked by Spacematrix.

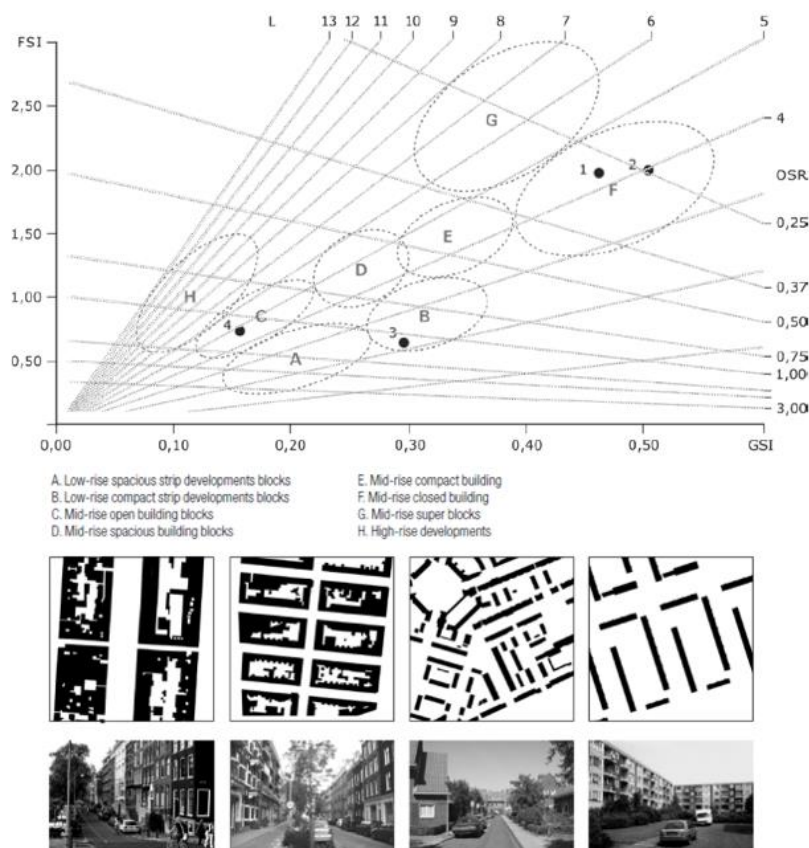


Figure 55 Spacematrix (source: Berghauser Pont and Haupt 2010)

Before proceeding with the description of other typo-morphological approaches, a specific observation about the spatial unit choice should be analysed.

Street-block partitions represent a common choice among urban morphologists; as in Spacematrix, in the scientific literature, we might find a large number of works implementing morphometric measures and classification using plot/street-block as spatial unit (i.e. Bernabé et al. 2013, Yoshida and Omae 2005, etc.). An interesting variation might be found in the work of Hermosilla et al. (2014), where surfaces of streets surrounding the block-faces are included in the analysis, enabling the consideration of the vegetation index in the statistical analysis of urban fabrics.

Urban morphologists and planners associate with street-block partition a specific epistemological meaning. As described by Pinon (1991), plots and streets represent those critical elements which define the organisation of the ground plan making up what Sheer defines as the boundary matrix (Sheer 2001, 2015). On the contrary, buildings, infrastructure, and every built-up human-made object represent a second category of urban form elements composing the built form.<sup>62</sup>

*“In almost all urbanized places, paths and plots continuously underlie the entire built form, providing a slow-changing game board upon which built form plays. As complicated and varied as built form is, the boundary matrix is far simpler, and that simplicity helps us see the structure and the containers in which all built form largely rests. By isolating the elements that constitute the boundary matrix, it is possible to isolate and name some very clear patterns that appear in cross-cultural comparisons”* (Scheer, 2001, p.10).

Two different temporal evolutionary processes might be associated with the built form and the boundary matrix. The use of spatial partition made by traces and plots would allow not only a cross-cultural comparison (i.e. Siskna 1997), but also the study of the temporal transformation of urban fabrics.

Moreover, Sheer also highlights how *“boundary matrix is also the preferred many other kinds of data, including land use, taxation, ownership, land value, construction data, soil conditions, and demographics. Already, many studies and plans use the containers/cells/spaces of the boundary matrix in GIS to record many kinds of data, which assists urban morphologists in their quest to connect physical form with other conditions of urban life, economy, and history.”* (Moudon et al. 2000).

Nonetheless, when the historical evolution of the urban fabric is not the main goal of the analysis and other phenomena are under investigation (social, environmental, economic, etc.), the meaning of this boundary matrix might vary (i.e. data resolution availability), becoming of secondary importance, or even a possible source of statistical biases.

This specific importance attributed to street-block partition is less considered, for instance, in urban geography studies. Despite studying the *same object* (the spatial distribution of urban form elements), urban geographers might be less familiar with notions developed in urban morphological and planning studies (Sheers 2015); moreover, *different research goals* encourage the adoption of alternative spatial unit partitions more adapted with the specific requirements of each work.

Other spatial units adopted in urban studies are, for instance, grids (i.e. Long and Kergomard 2005), administrative boundaries (i.e. Puissant et al. 2010, Tsai 2005), or specifically conceived spatial units as generalised Thiessen tessellation (Thiessen 1911)<sup>63</sup> applied on built-up units as in Voiron (1995), Hamaina et al. (2013), Sevenet (2013).<sup>64</sup>

---

<sup>62</sup> In the French school of urban morphology, the difference between ground plan and built form is named by Pinon as morphological sub-structure and super-structure (Pinon 1991).

<sup>63</sup> Also called Dirichlet or Voronoi Tessellation from their author's name (Dirichlet 1850, Voronoi 1908)

<sup>64</sup> While Hamaina et al. (2013) and Shirmer and Axhausen (2015) implement generalized Thiessen polygon starting from vector definition of built-up footprint, Voiron (1995) and Sevenet (2013) implement skeletisation procedures from



Methods based on grids, administrative boundaries, or generalised Thiessen polygons applied on buildings, completely overlook the importance of the *ground plan*, and more particularly of street patterns, focusing on the spatial patterns of building volume distribution.

As for Retail Geography in Section 8, we might categorise spatial units used in urban morphology studies in three typologies. *Endogenous* partitions, when the spatial tessellation is defined or based on urban morphological components, and further divided in *ground-plan-based partitions*, when derived by plot or street elements or *built-form-based partitions*, considering the only distribution of building footprints. *Exogenous* partitions, such as administrative boundaries, are traditionally used when dealing with socioeconomic data from governmental statistical institutions. *Neutral* partitions, such as regular grids, concentric circles, etc. can be implemented in environmental or fractal approaches.

The implementation of the three-step classificatory procedure to a fine-grained partition (such as streets, blocks, grid cells, etc.) traditionally provides a taxonomy of the partition regions. The identification of urban fabrics or morphological regions is not a direct result of these procedures; they are rather visually seized by the analyst when plotting the classification outcomes in the geographical space.

This task is made possible by the presence of spatial autocorrelation in the original morphometric variables. Similar typo-morphologies of blocks, streets, cells, etc. are therefore found in spatially more or less adjacent subspaces, allowing an easy recognition of continuous sub-regions. When the local heterogeneity becomes higher (as in more peripheral areas) the classificatory approaches result in a patchwork of different classes (of streets, blocks, cells) leaving to the analysts a much more complicated task and limiting interpretations to the scale of the individual unit (Fig.56).

As discussed in Section 4, when the three-step classification procedure is applied on the intrinsic (compositional) descriptors of each spatial unit, the consideration of the spatial organisation of the urban form is only limited at the first step of spatial partitioning. Consequently, i) the classification is strongly determined (and potentially biased) by the initial choice of the spatial partition definition; ii) the classification is based on the intrinsic (compositional) morphometric descriptors overlooking their spatial distribution. Using the same archipelago example previously described, when shuffling the spatial distribution of units (blocks or buildings), the classification would not change the outcomes while its spatial pattern is completely modified.

---

morphological mathematics approaches based on a raster representation of data. The two procedures reach the exact same spatial partition: each point of the planar space is associated with the perimeter of the closest built-up footprint.

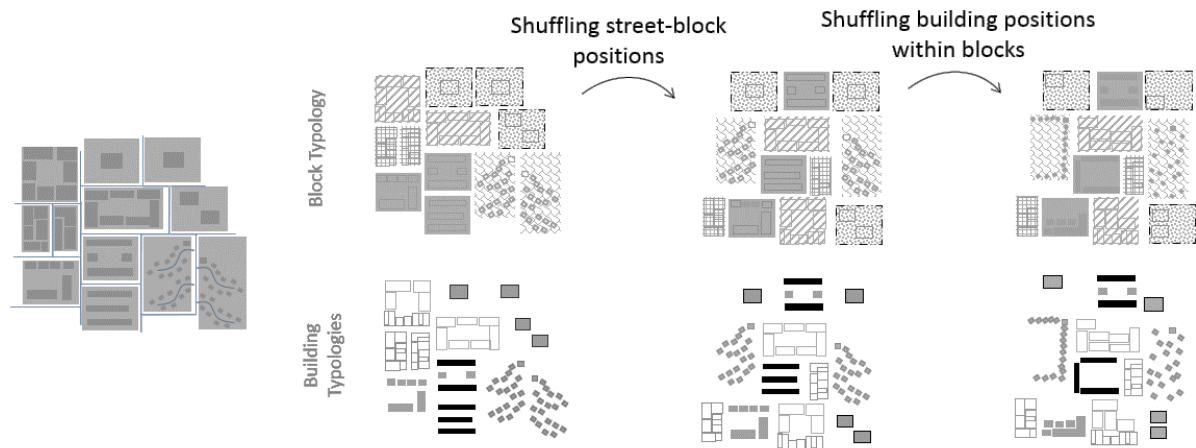


Figure 56 Traditional classification of blocks and buildings do not consider their relative spatial distribution. Shuffling street-block position or building position within blocks, the outcomes of the traditional three-step classification would result in the same typologies of building/block morphological type.

### **The location-based approach to typo-morphological analysis**

When the goal is the identification of typologies of urban fabrics or urban morphological regions, two approaches should be considered: on the one hand we might implement the same three-step procedure at wider scales (such as a neighbourhood or district level partition). On the other, we might preserve a fine-grained spatial partition and morphometric descriptor, but include an alternative solution able to assess their spatial distribution.

Traditionally, the first approach is the most adopted one: once defined as the new coarser spatial partition, a multivariate analysis of the internal morphometric characteristics allows the definition of typo-morphological neighbourhoods. Here again, the same problems about spatial unit definition, statistical biases (MAUP), and spatial distribution of morphometric variables might arise.

*“All studies that focus on neighbourhoods must confront the difficult problem of defining neighbourhoods. Past studies used census tracts, traffic analysis zones (TAZs), zip codes or other pre-defined neighbourhood boundaries to compute measures of neighbourhood types. These measures might be misleading since the unit of analysis may have been too large, neglecting the fact that characteristics of development patterns vary across different parts of the predetermined neighbourhood.”* (Song and Knaap 2007, p.7).

To partially overcome the issues related to exogenous partitions, Song and Knaap (2007) propose a *neutral location-based* definition of spatial units: for each of the almost 7 thousand locations (around newly developed single-family parcels), a ¼ mile radius buffer is traced and 21 indicators (about physical form and accessibility) are collected. Factor analysis and classification allows the identification of six typologies of location-based neighbourhoods in Portland’s (Oregon, US) metropolitan region. In a second work, Song et al. (2013) implement the same procedure with some refinements: 27 indicators measured at 1, 3, 5 and 8-mile location-based buffers from more than 20 thousand locations spread around the United States and allow the description of seven location-based neighbourhoods.

As observed by the same authors, the buffer definition does not avoid the MAUP, but it allows to focus on a more consistent definition of neighbourhoods, enabling the analysis and classification of the relative *immediate* urban context Song et al. (2013).

Similarly, a recent work of Serra et al. (2018) proposes a comparable location-based approach for the classification of the immediate environment, surrounding 77 schools in Liverpool. Here again, a Euclidean buffer allows to delimit and collect local and social and morphological descriptors, subsequently combined in contextual typologies.

Throughout the research of Song and Serra, the context surrounding specific locations is evaluated with Euclidean buffers; Berghauer Pont and Marcus (2014, 2017) propose a similar location-based approach to measure and classify built form context introducing network-constrained distances. From each location is a set of morphometric descriptors implemented on a network-based floating catchment area.

Location-based approaches might be considered as smoothing approaches, where instead of measuring indicators at the small scale definition of each spatial unit (i.e. the individual block) a wider subspace at a given (Euclidean or network) distance is considered. When implementing these more recent procedures, the identification of the spatial radius becomes a key factor. The radius might vary hypothetically between zero, describing the individual properties of the individual spatial unit (building/block/plot/street) and the maximal distance (diameter of the area/network under analysis) for which we should obtain the morphological values corresponding to the overall study area. The values between these two extremities, describe, for each location, the evolution of the morphometric indicator at different spatial scales; consequently, curves associated to different locations would converge to the same point representing the indicator measured at its largest scale<sup>65</sup> (Fig.57). In this same area we find the works of Badariotti et al. (2007, 2009): the authors studied the urban fractures also called catastrophes (in their mathematical definition) and studied the integration of buildings at different network radii. Street network, functional network and neighbourhood networks are investigated.

These approaches are well known in quantitative geography analysis: geostatistics, fractal analysis (as described in Section 5.3), but also methods such as variography (Dauphiné and Voiron 1998) and mathematic morphology (Voiron 1995) have been developed since the '80s and applied on network constrained environments in the last two decades. Nonetheless, all these approaches focus on the description of the overall urban settlement with an overall function which does not allow the identification and delimitation of morphological intraurban regions (fabrics).

---

<sup>65</sup> The precise convergence on the same point might depend on the border effect/conditions imposed by the analyst.

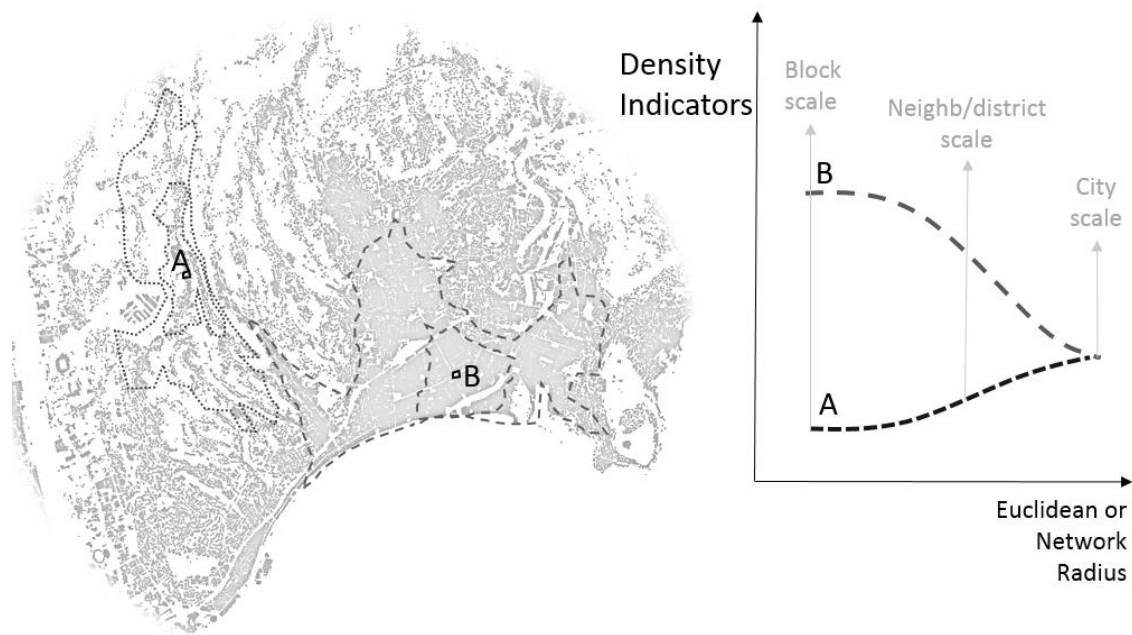


Figure 57 Location-based measures: while Song and Knaap (2007) and Berghauer Pont and Marcus (2014, 2017) focus on specific radii for the identification of local morphological regions, Fractal and geostatistical studies show the evolution of density at different scales and specific behaviours (such as, respectively, fractal transition or catastrophes in Badariotti).

The innovative aspects of the works of Song and Berghauer Pont approach consist, on the contrary, on the specific local implementation at the intraurban scale and integrate notions of urban form/design developed in the urban morphology schools (while geography and statistics focus more on the study of the land use occupation at the inter-urban scale).

The methodologies developed by Song and Berghauer Pont, instead of considering a spatial tessellation of the space, propose a location-based density measure, which, similar to a moving window, adapts to the specific surrounding conditions of each location. This smoothing operation allows for the removal of hyperlocal specificities and to account for the spatial autocorrelation of morphological features.

These approaches implement a methodological innovation of the first phase of the traditional three-step procedure, leaving the following two unchanged.

What is more relevant to the focus of this section is that the implementation of this procedure to all the units of a fine-grained spatial tessellation (such as blocks or plots), results in a bottom-up identification of coherent urban morphological regions without a preconceived delineation of neighbourhoods.

The same goal might also be reached with other modifications of the traditional three-step procedure as, for instance, in the work of Hamaina et al. (2013). Firstly, the authors define a generalised Thiessen tessellation based on built-up footprints; secondly, 16 indicators, including both traditional volume measures (such as GSI and FSI) and innovative visual measures generated by empty spaces (such as isovists and sky view factors). Thirdly, instead of traditional classificatory analysis, the authors implement spatial clustering approaches

(geographical self-organised maps Geo-SOM, Bação et al. 2004) enabling the identification of twelve categories of urban fabrics coherently distributed in the urban space.<sup>66</sup>

In the last works reviewed, we discussed how recent developments might allow for the identification of urban morphological contexts/fabrics with bottom-up procedures; instead of an a priori definition of urban morphological regions, these procedures enable the assignment of each location to a typology of urban morphological context.

These approaches might be implemented in order to answer to the specific goals of this dissertation. Nonetheless, as discussed in Chapter 2.5, one of our specific goals is to study both retail and urban morphological variables from the street point of view, developing an entire analytical framework able to capture the urban spaces as perceived by the street perspective.

The procedures mentioned above only partially integrate the notion of street perspective.

The set of morphometric measures in Hamaina et al. (2013) includes with the traditional density measures (i.e. FSI) several innovative indicators of visibility (such as isovists for ground openness and sky view factors for the spatial openness). The indicator of High-Width ratio is implemented, considering the skeleton defined by the generalised Thiessen partition as an artificial street network. An ill-defined notion of the street network, which does not correspond with any modelling definition traditionally described in urban, transport or configurational analysis (Marshall et al. 2018), limits any possibility to bridge and cross analyse the results with outcomes from other approaches (as, in this dissertation, with street network configurational properties). Moreover, the geo-SOM clustering considers the Euclidean distances between regions of a partition overlooking the importance of network-based distances.

On the contrary, the location-based approach of Berghauer Pont and Marcus (2014) integrates street networks' distances, allowing the implementation of several street network/metric modelling approaches (angular, topological, and metrical as in configurational analysis). The integration of network-constrained distances leads to the description of accessible density. Nonetheless, density measures are implemented as in traditional urban planning approaches, in order to consider the aerial point of view and overlook the street perspective; for these reasons, this approach should not be understood as a measure of street-based perceived density.

The two main natural actions which every pedestrian does in urban spaces are *observing* and *moving*, independently of (and simultaneously with) more specific behaviours/activities. In order to assess the urban form as perceived from the pedestrian point of view, we should consider both the geometrical distribution of urban morphological elements around each public open space, but also how they might be perceived when pedestrians explore different spaces along their paths. Our goal is not the estimation of the pedestrian flow and ease of

---

<sup>66</sup> The only limitation of this approach is associated with the application of measures such as isovists and sky view factors, which (as previously discussed) might limit the implementation to a large metropolitan area. The example of Hamaina is indeed implemented on a 5 square kilometre area with 2.8 thousand buildings (in our case we have more than 314 thousand buildings on 1500 square kilometres)

movement as in SSx, but rather the identification of urban fabric typologies from the street point of view, as perceived from the street.

In order to integrate these two aspects, we should combine the advantages of both Hamaina et al. (2013), where the urban fabric is considered as visually seized by a pedestrian among buildings, and the street network distances as in Bergauer Pont and Marcus (2014) where real streets define the accessible spaces among urban fabrics.

The pedestrian perception of the urban fabric should consider how built-up elements are organised around the public space instead of the traditional aerial perspective of urban morphologists, geographers, and planners. More attention should be given to the interface between streets and buildings/blocks.

This perspective in the analysis of the urban space is traditionally investigated in the urban design domain of study where urban morphological characteristics of street environments are studied in relation to human perception and behaviour.

### **Street-Based measures of Urban Form: the urban design approach**

Urban design focuses on aesthetical aspects at a smaller scale (compared to urban morphology), which covers the design of individual façades, to street environment up to a small urban fragment; it is an interdisciplinary domain in the middle ground between architecture and urban morphology/planning. Urban design *“is concerned with the extension of architecture beyond the design and construction of individual buildings and with the attainment of environmental quality, broadly defined”* (Rowley 1994).

While the European schools developed the concept related to the urban morphology, urban design finds its origins in the North American context of the late '50s; in 1957, the American Institute of Architecture established a Committee on Urban Design and the first course of urban design at Harvard University in 1960. The seminal works of Lynch (1960) and J. Jacobs (1961), Alexander (1977), are joined by north European authors such as Cullen (1961), Ghel (1971), etc.

As discussed before surrounding urban form (Section 2.1), the notion of urban design is also a multifaceted concept at the crossway of several domains of study (Rowley 1994); in particular, urban design studies are strictly associated with notions of environmental psychology, human perception, cognition, and behaviour. For example, studies in urban design investigate the relationship of urban design descriptors with walkability, liveability, physical activity, appeal, perceived safety, etc.

Within urban design literature, numerous perceptual qualities have been described and their importance has been widely debated in the last decades; nonetheless, it is only in more recent years that urban design literature has attempted to objectively measure perceptual qualities, rather than simply assert their importance (Ewing et al. 2005).

As for the other disciplines discussed in this dissertation, in the last twenty years, urban design has also seen a growing number of studies implementing quantitative approaches. The necessity of a systematic approach for the operationalisation from notions to quantitative

measures has been undertaken in the work of Ewing et al. (2005); the qualitative discussion of seminal urban design studies is reviewed and summarised in 51 variables.

Not all urban design qualities might be reinterpreted in quantitative terms; some are more amenable to measurement than others. Based on the importance assigned in the literature, Ewing et al. (2009) select and successfully operationalised five main variables defined as follows:

- a) **Imageability**, after the seminal work of Lynch (1960), “*is the quality of a place that makes it distinct, recognizable and memorable. A place has high imageability when specific physical elements and their arrangement capture attention, evoke feelings and create a lasting impression*”.
- b) **Enclosure** “*refers to the degree to which streets and other public spaces are visually defined by buildings, walls, trees and other vertical elements. Spaces where the height of vertical elements is proportionally related to the width of the space between them have a room-like quality*”.
- c) **Human scale** “*refers to a size, texture, and articulation of physical elements that match the size and proportions of humans and, equally important, correspond to the speed at which humans walk*”.
- d) **Transparency** “*refers to the degree to which people can see or perceive what lies beyond the edge of a street and, more specifically, the degree to which people can see or perceive human activity beyond the edge of a street*”.
- e) **Complexity** “*refers to the visual richness of a place. The complexity of a place depends on the variety of the physical environment, specifically the numbers and types of buildings, architectural diversity and ornamentation, landscape elements, street furniture, signage and human activity*”.

These five main aspects of urban design are quantified by 28 indicators, such as counts and proportions of specific micro-grained features which might be observed in urban spaces; walkability measures are then individually correlated to each variable.

Purciel et al. (2009) operationalise streetscape measures using a variety of GIS data sources; in this study a set of 22 “*objective and theoretically-grounded indicators*” quantify physical and functional variables of the urban spaces of 588 block-faces in New York (one or two-sided measures of the streets depending on the indicator). Similarly to what observed for configurational studies, individual characteristics of urban design at the street level are correlated to pedestrian counts.

GIS algorithms are also developed in the works of Harvey et al. (2015a, 2015b) and Harvey et al. (2015, 2017) focusing on a key set of physical descriptors of “*skeletal streetscape*”.

*“A chief limitation of GIS methods is exclusion of microscale design characteristics (e.g. materials, architectural styling, ornamentation, fixtures, cleanliness), activity, and nonvisual sensations that contribute in important and nuanced ways to user experience. Such microscale elements may be considered the skin of a streetscape. Field audits are advantaged in capturing these characteristics (Clifton et al., 2007). Streetscape skeletons, based on more macroscale*

characteristics such as the dimensions and arrangement of buildings, are more measurable with a GIS because they are based on common data inputs and evaluation of geometric relationships for which GIS tools are well suited. The relationship between skeleton and skin is analogous to that between a wireframe drawing and an architectural rendering. The former supplies the spatial structure for a scene, defining the size and shape of space; the latter embellishes with visual texture. The boundary between skeleton and skin is undoubtedly nebulous. While characteristics at scalar extremes, such as building height and siding material, clearly contribute to the definition of streetscape space, characteristics between these scales, such as lampposts, street furniture, awnings, or vegetation, also contribute. On one hand, these objects may define important subspaces within streetscapes. Alternatively, they simply embellish the broader streetscape already defined by buildings that dwarf them in size. [] skeletons of streetscapes, delineated by the massing of surrounding buildings and trees, provide spatial proportions that may be elemental to perceptions of them as comfortable urban spaces”.

Based on these definitions of the skeletal streetscape, seven variables are proposed and implemented: width, length, building height, cross-sectional proportion, street wall continuity, buildings per length, and tree canopy coverage. These variables propose the analysis of the urban volumes as perceived from the street point of view and are individually correlated to different phenomena, such as crash severity (Harvey et Aultman-Hall 2015b), and perceived safety (Harvey et al. 2015).<sup>67</sup>

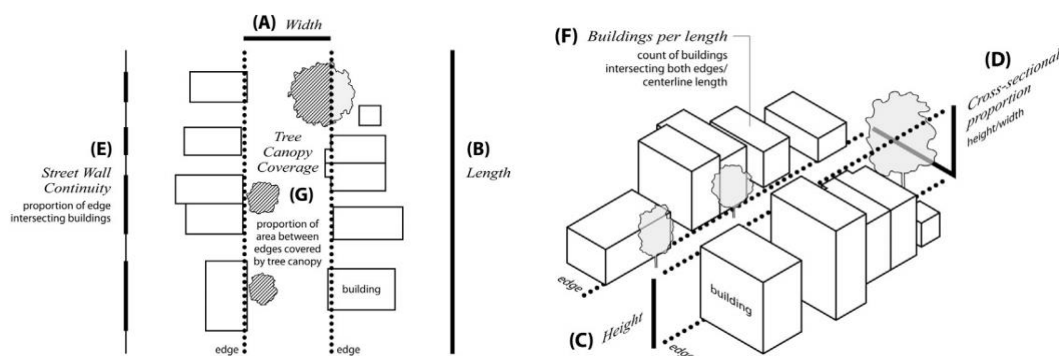


Figure 58 Skeletal streetscape (Source: Harvey 2014)

In a more recent work by Harvey et al. (2017), 12 streetscape skeleton variables were measured on more than 120 thousand block-length streetscapes in three northeastern U.S. cities: Boston, New York, and Baltimore. Cluster analyses were then used to identify four streetscape skeleton classes that were also consistent between cities: **upright**, **compact**, **porous**, and **open**, which is proposed as an alternative to the functional classification traditionally proposed by transport planners (arterial, collector, and local). As we might observe, the multivariate three-step procedure (that we described in both retail geography and typo-morphological studies) is here again applied to the intrinsic

<sup>67</sup> In this second work correlation with Perceived safety scores of streetscape images, collected by researchers at the MIT Media Lab using an internet-based survey called Place Pulse (2014), were available for more than six hundred sites throughout the city (Salesses et al., 2013).



geometrical/compositional description of the physical form as perceived from the street point of view.

These last works focus on GIS data and on the physical layout of the built form along the streets. We should also mention that, in the last few years, data science is contributing to urban design with innovative algorithms for the operationalisation of qualitative measures from street view images (street frontage in Law et al. 2018 and Liu et al. 2017; greenery in Dubey et al. 2016; *scenicness* in Seresinhe et al. 2017, etc.). These highly sophisticated approaches based on imagery analysis will provide in the next years interesting insights assessing and developing the traditional notions proposed by the seminal works of urban designers.

An interesting observation about these two approaches should be highlighted: while algorithms for street view image analysis propose a pedestrian perspective in the analysis of urban spaces as an alternative to traditional satellite/aerial images, analogously GIS skeletal streetscapes represent the street-based alternative for traditional urban morphology and planning.

The skeletal streetscape is not only the interface between private and public or street and block spaces; it might be also considered as the interface between the urban design and urban morphology (traditionally focusing on two different scales of the urban form). For these reasons, the skeletal streetscape assumes in this dissertation a particular role which allows us to integrate the urban morphological characteristics of a street as it could be perceived from the street point of view in order to answer to the specific questions discussed in Section 2.3.

Moreover, skeleton streetscape measures represent an interesting compromise from the computational point of view when compared to visual analysis such as isovists and derivatives: instead of using any position of the planar space, streets centrelines might be used as a reference for the elaboration of geometrical measures. In this way, geoprocessing algorithms are more economic to implement, and, therefore, wide geographical regions as our study region might be easily analysed.

When considering urban morphology studies, the role of the streetscape has received little attention; the only exception might be found in the work of Vialard (2013) on Atlanta. From the original definition of block-face in Purciel (2009), Vialard specifically focuses on the interface between streets and blocks. In her work, *“The block-face emerge as the morphological element that naturally associates the properties of the street network and the properties of the buildings”* (Vialard p.36).

The notion of block-face is differently conceived and operationalised compared to the notion of skeletal streetscape of urban design. The author, indeed, considers the block face as the line where street edges and the block perimeter overlap; *internal and external forces* are successively defined and measured. The former corresponds to configurational properties (using metrical and directional Reach indicators as proposed by Peponis 2007, 2008). The internal *load* is implemented with compositional/geometrical measures based on a radial-buffer method (Boyce and Clark 1964) by means of radial lines from the barycentre of the

block (further description in Chapter 8). Several classifications are finally proposed, considering internal and external forces, using both blocks and block-faces as spatial units.

Despite the great interest and innovations proposed in the work of Vialard, her analysis of the *block-face* is still developed from a planar approach to urban space: i) the absence of the third dimension (traditionally associated to human scale variables) ii) the implementation of built form indicators using the barycentre of the block as reference instead of the centreline; and iii) the implementation of independent block-face measures which do not allow the contemporary evaluation of the two sides of a street co-participating to the definition of a street environment, or streetscape.

Two final observations should be made on streets-based approaches. Firstly, we might notice how the works of urban designers as in Purciel (2009) and those of Harvey and Aultman-Hall (2005a, 2005b) correlate detailed local urban morphometric variables with pedestrian counts or human behaviour/perception similarly to what is proposed in street network configurational studies. In both cases, these works overlook the importance of the urban morphological context, which might allow outlining the role of a specific street design indicator in different urban contexts: variables as height, length, or continuity of facades along streets might play different roles in historical centres, planned or peripheral neighbourhoods, etc. Considering these variables in large and heterogeneous regions, it might result in a non-significant statistical relationship.

We believe the urban morphological context is, therefore, an indispensable variable for investigating how both the *network effect* (Omer and Goldblatt 2015) and *street design effect* might vary in different urban regions when considering their influence on the study pedestrian presence, and, as specifically investigated in this dissertation, on retail distribution.

Secondly, almost every quantitative approach reviewed in this section overlooks the importance of public open spaces, such as squares (even more evident in network configurational approaches where open spaces are simplified as street segments). Squares represent an important aspect in the Mediterranean/south European context (Cutini 2003, 2014) both for the urban morphological organisation of space, and, more importantly for this research, for the socioeconomic distribution of activities. Recent works by Lopes et al. (2015) and Cutini (2017), etc. have been focusing on these spaces, implementing different morphometric perimetral and visual analysis (VGA sky view factors etc.) on 2D and 3D GIS/CAD representations of squares. In these works, a classification of squares is provided, which implements, here again, the traditional three-step classification procedure. Similarly, the works of Beirão et al. (2014, 2015b) and his research group at the University of Lisbon, focus on the parametric description and architectural analysis of public open spaces (convex and solid void models).

Although this work also focuses on street-based quantitative analysis of the urban form, specific descriptors of square presence in the city should be integrated, which we will further discuss in the next sections.

In the next table, the most relevant works reviewed in this section are summarised.

Authors	Year	City	Method	Spatial Unit	Innovative Aspect for this work
<b>Typo-morphological approach</b>					
Marshall	2005	-	ABCD Diagram	Small to large urban fragment	Diagram approach (4 measures) for the street pattern recognition
Song and Knaap	2007	Portland	Factor Analysis + K-means	Buffers of ¼ mile around 6788 sites	Location-Based context analysis
Berghauser-Pont and Haupt	2004-2007-2010	-	Spacemate Diagram	Blocks	Spacemate diagram
Gil et al.	2012	2 Neighb. of	K-means	Streets and Blocks	Combination of conf. and combina
Sevenet	2013	Nice, France	CHA	Building-based Voronoi	Thiessen tessellation, (Nice case study)
Vialard	2013	Atlanta, US	K-means	Blocks and Block-face	Operationalisation of block-face measures
Hamaina et al.	2013	District of Nantes, France	Geo-SOM	Building-based Voronoi	Spatial tessellation + GeoSOM
Song et al.	2013	20,467 Residential homes in US	Factor Analysis K means	Euclidean buffers of 1,3,5,8 km around 20,467 sites	Location-Based context analysis + Multiscale classification.
Hermosilla et. al	2014	Valencia, Spain	ANOVA, Decision Tree	Blocks and surrounding streets	Introduction of street and vegetation index
Berghauser Pont et L Marcus	2014	Stockholm, Sweden	Spacemate Diagram	Network-Constrained Location-Based Convex hull	Network-Constrained Location-based density and Spacemate
Berghauser Pont et al.	2017	Amsterdam London and Gothenburg with Stockholm			
<b>Urban Design approach</b>					
Purcel	2009	New York, US	-	Streetscape	GIS Operationalisation of Streetscape
Harvey et Aultman Hall	2015a - 2015b-2017	Baltimore, Boston, NewYork, US	-	Skeletal streetscape	Operationalisation of Skeletal streetscape measures
Harvey et al.	2017	Baltimore, Boston, NewYork, US		Skeletal streetscape	Classification of Skeletal Streetscapes

Figure 59 Methodological literature review of typo-morphological and urban design studies developing quantitative approaches for the analysis of the urban form.

## 6.3 Finding a common ground for configurational, typomorphological and streetscape analyses

In the two previous sections, we reviewed quantitative approaches for the analysis of the physical form of cities with a specific attention to street-based methods.

In the first section, we focused on street network configurational analysis approaches, which, by definition, focus on streets and their spatial organisation. Throughout the past two decades, several procedures resulting from the combination of different street and graph modelling approaches have been proposed (Fig.53), discussed, and implemented in a large number of studies. An important number of tools have been specifically developed by several research groups (see Section 2.2) for the implementation of street network configurational analysis. Each approach is able to capture different nuances of the network effect; the debate about the most adapted approach, if *one* exists, goes beyond this work. The implementation of a specific configurational approach should be justified when considering the whole research framework and the specific goals of each research.

In the second section, we reviewed typomorphological approaches, providing quantitative procedures for the identification of urban morphological contexts (within which the network effect might play different roles, Section 2.2). Multivariate classificatory analysis applied to the neighbourhood scale presents several inconveniences related to the definition of the spatial unit. As a response, new location-based approaches have been discussed in more recent studies, allowing a location-based identification of morphological regions. Nonetheless, the street-based perspective is traditionally overlooked in quantitative typomorphological studies.

On the contrary, urban design studies focus on the specific analysis of the streetscape at the hyper-local level of a single street. The operationalisation of the notion of skeletal streetscapes have been recently proposed; this specific approach allows a compositional analysis of built forms from the street point of view. As for configurational studies, urban design approaches traditionally overlook the importance of the urban fabric (contextual description) within which they are located.

Three main quantitative approaches for the study of the urban physical form have been outlined: i) *configurational*, from street network studies, ii) *compositional*, from both urban morphology and urban design, and iii) *contextual*, from typomorphological studies.

The first two approaches put the street at the centre of the analysis. The configurational approach describes the relative spatial position of each street element within the surrounding network while compositional analysis describes the built form, including the two sides of each street, integrating the street point of view. On the contrary, contextual approaches have been traditionally conceived from the aerial point of view and streets are only considered in location-based methods in order to implement *accessible density measures*. While the first two approaches might be implemented with few adaptations (Section 7 and 8), the third approach needs to be profoundly revisited in order to better integrate the street point of view (Section 9).

In the scientific literature, a large number of studies have implemented the first two approaches, individually, in order to study the relationship between configurational and design properties of streets with pedestrian presence/movement and human behaviours.

Few studies have tried to analyse their combined effect on both pedestrian presence and (more importantly for this work) retail distribution. Representative of this method, we might find the work of Remali et al. (2014): in this work, both configuration and street quality indicators are implemented on three neighbourhoods of Tripoli (Libya), specifically selected through the expert knowledge of the authors.

Similarly, few studies have tried to assess the importance of the urban morphological context when evaluating configurational or street design properties; recent works have been tackling the problem of the *network effect* in different morphological regions, as in the works of Omer and Berghauser Pont research groups. The high costs of data collection (surveys and pedestrian counts) limit these applications on small selected areas chosen by the expert knowledge of the authors.

While configurational and compositional approaches are individually studied, on the contrary, in typo-morphological studies, the implementation of multivariate analysis and classification procedures frequently consider both morphometrics and configurational descriptors.

This widespread practice might be seen as the general response of researchers in urban morphology with the need to integrate different analytical approaches, with the final goal being to obtain a more complete set of urban environmental attributes. The underlying assumption of doing research in this way is that the combination of different procedures might be able to highlight the relative importance of different facets of the urban physical form in highlighting different aspects of the city and its functioning (Kropf 2009, 2017; Conzen 2010; Wineman et al. 2009, etc.).

Several works have proposed different techniques for the integration of configurational and typo-morphological approaches. Three main approaches might be described:

- **juxtaposition** of outcomes from specific approaches, as in the works of Gil et al. (2012), Berghauser Pont et al. (2017), etc. where specific approaches are implemented for different urban components (street configuration and block typo-morphology), and the combined cartography of the outcomes allows the analyst to visually compare and discuss the results and their spatial co-occurrences.

-**incorporation** of configurational variables into typo-morphological multivariate analysis as in Marshall (2005), Vialard (2013), etc. or, reversely, the introduction of built form measures in configurational analysis with weighted centrality assessment, as in MCA of Sevtzuk (2012, 2014) or in Place Syntax (Stale 2005), etc.

-**mix**, as in Oliveira (2013), Oliveira and Medeiros (2016), or in Ye et Van Nes (2014), where individual configurational and morphometric indicators are linked and combined in the same spatial unit, providing elaborated measures. From this last approach, important questionings about statistical and methodological accuracy, due to its high manipulation of data, might arise.

Independently of the specific combinatory approach, this procedure should not be reduced to a simple procedural issue focusing only on the most adapted geoprocessing protocol to join data of different nature. On the contrary, the analyst should give specific attention to the notions underlying each approach, their relative spatial scale, and the specific phenomena under study.

As highlighted by Kropf (2017), when the goal is to combine different approaches, urban form analysis should be reconceptualised. The scientific analytical procedure should start from defining the specific phenomena under study, identify a common theoretical and methodological framework, and finally adapt traditional approaches in order to allow a more coherent combination.

The goal of this research is not to combine, but rather to assess the different aspects of the urban form: configurational, compositional, and contextual procedures are here separately investigated and only in the last part of this dissertation will their relative importance be assessed in relation to retail distribution.

While on several previous occasions we insisted on the importance of the street point of view, it is now necessary to define a common spatial unit on which the different modelling approaches might join or converge with the different aspects of morphology and retail distribution.

In our case study, we privilege the primary approach where the street segment is defined as the centreline segment between two intersections. Several reasons motivate this choice:

- As observed by Marshall et al. (2018) *“from Network Analysis in Geography (Haggett and Chorley 1969) to a recent review on spatial networks (Barthelemy 2011), the dominant network model is one that represents the street junctions as vertices in the graph and the linear street segments as its edges”*.
- Using street centrelines and the primal approach would allow the identification of configurational properties independently of the physical shapes or sizes of the objects constituting the streetscape (in SSx, the definition of axial lines is influenced by the distribution of the built form). From this observation arise two advantages: firstly, the primal approach is recognised as a standard in the representation of the urban street network which is available in several cities without any specific elaboration (enabling the reproducibility of the method for future studies). Secondly, because the separation of network configurational properties from other urban landscape elements are further considered for the definition of urban typo-morphologies (and reducing the inevitable correlation between these aspects).
- The primary graph allows one to implement centrality indicators based on a combination of specific impedances and subnetworks; these two characteristics allow an easier interpretation within retail geography context where customer movement measures are traditionally expressed in terms of distance-based accessibility between stores and store agglomerations. Additionally, the use of the same urban network for both RFA and MCA has the advantage of reducing any bias.

- The use of centreline assumes the role of geometrical reference when studying the streetscape from the street point of view: as we will see in the following sections, measures of setback, parallelism of facades, etc. use street edges and street centrelines as reference points. The use of visual axes as in SSx or alternative street-like representations of the street network would provide a *distorted* reference system. The adoption of different street representation is still possible, but it would require the use of different definitions of streets with possible theoretical and statistical inconsistencies and biases when combined.

From this definition of the street network, we present in the next section how configurational, compositional and contextual measures are implemented and associated with each street element.

## 7. Measuring the street configuration

### **Selecting Indicators, radii and metrics**

Reach, Closeness, Straightness and Betweenness represent the four main indicators traditionally studied in those works assessing the relationship between retail distribution and network configuration with a primal approach (i.e. Porta 2006, Porta 2012, Sevtsuk 2014, Scoppa and Peponis 2015, Wang et al. 2018, Lin et al. 2018, Fusco et al. 2016; Section 2.2, Table 2). Despite sharing the same primal approach, differences among these works might still be detected when considering parameter combinations and modelling choices. In this chapter, we describe how configurational measures have been implemented for our specific purposes.

The elementary unit of the analysis, the street, is defined as the centreline segment between two intersections independently identified from the surrounding built-form characteristics (which is further studied in Section 8 and 9). This choice differs from the previous works of Peponis et al. (2008) or Sevtsuk and Mekonnen (2012) where the authors implement metrical configurational indicators considering, respectively, the street network length and the building surface accessible in the proximity of each location at a given distance. Our definition of configurational properties follows the formulation of Porta et al. (2006) where topological measures of the street network are implemented. Differently from Porta (2006) who implements configurational properties on street junctions, we implement configurational descriptors on street midpoints (similarly to the *betweenness centrality* implemented by Scoppa and Peponis 2015).

This specific approach allows us to focus on the configurational structure of the street network independently of its dimensional characteristics, further measured by compositional descriptors (Section 8). Since our final goal is the study of the relationships among different aspects of the urban form, the distinction between configurational and compositional

descriptors is a fundamental methodological choice for the reduction of redundancy and collinearity between indicators (See Chapter 11.2).

Our four centrality measures are defined as follows:

- The **Reach**  $R_r(i)$  of a street midpoint  $i$  lying on the network  $G$  informs us about the number of *streets*  $j$  reachable through the shortest path, having a maximum distance given by the search radius  $r$ . The five metrical and temporal distances define the local sub-networks within which the number of destinations  $j$  are counted. For each street midpoint  $i$ , this indicator might be implemented as follow:

$$Reach_r(i) = \sum_{j \in G - \{i\}; d[i,j] \leq r} j \quad (7.1)$$

Where  $d[i, j]$  represents the distance of the shortest path between the reference midpoint  $i$  and each destination midpoint  $j$ . This centrality measure might be easily interpreted as the number of accessible streets within a given radius  $r$  (instead of the number of street junctions as in Porta 2006, the street length as in Peponis 2008; or the building surface as in Sevtsuk 2012).

- The **Closeness** represents the inverse of the Farness  $F^r(i, r)$  (or Remoteness in Sevtsuk 2010) which is the sum of the distances  $d[i, j]$  determined by the shortest paths between the midpoint-origin  $i$  and the midpoint-destination  $j$  within the sub-network identified by the radius  $r$ ; the indicator  $C_r(i)$  for each point  $i$ , is therefore given by the expression:

$$Closeness_r(i) = \frac{1}{F^r(i,r)} = \frac{1}{\sum_{j \in G - \{i\}; d[i,j] \leq r} d[i,j]} \quad (7.2)$$

This measure might be interpreted as a frequency and it represents a concept of proximity between streets, which, differently from the Reach centrality, is expressed in metric terms, capturing the notion of accessibility (Porta 2009).

- The straightness  $s_r(i)$  of a node  $i$ , is obtained by calculating how much the shortest path connecting a pair of points  $d[i, j]$ , deviates from the straight Euclidean distance  $\delta[i, j]$ . If the two paths have an equivalent value,  $s(i,r)$  is equal to 1; for increasing tortuosity, the difference between  $d[i, j]$  and  $\delta[i, j]$  increases and consequently  $s(i, r)$  will tend to 0. The **Straightness** indicator  $S_r(i)$ , is obtained as the sum of all the ratios  $s_r(i)$  calculated for each target point identified within the lattice radius  $r$  (or  $t$ ); this centrality value must be read as a *Reach* weighted based on the tortuosity of the paths designed to reach the destination points.

$$Straightness_r(i) = \sum_{j \in G - \{i\}; d[i,j] \leq r} \frac{\delta[i,j]}{d[i,j]} \quad (7.3)$$

- **Betweenness** centrality provides the information about the number of shortest paths lying on the point  $i$  and implemented considering every pair of points having a relative distance less than or equal to the selected radius  $r$  (or  $t$ ). This index will be calculated according to (12.4)

$$Betweenness_r(i) = \sum_{j,k \in G - \{i\}; d[j,k] \leq r} n_{jk}[i] \quad (7.4)$$



Where  $n_{jk}[i]$  is the number of minimum paths from node  $j$  to node  $k$  on network  $G$  passing through point  $i$ , with  $j$  and  $k$  at a distance less than or equal to  $r$  (or  $t$ ). The peculiarity of this index, unlike the three centrality measures previously described, is that it provides the information about the intermediary centrality (Freeman 1977) or through-movement potential (Hillier 1996).

As we might see from their definitions, one of the main advantages of the primary approach is that configurational measures might be implemented considering specific subnetworks at a given radius allowing the investigation of configurational properties at different scales as suggested by Porta et al. (2012), Sevtsuk and Meckonnen (2012). Moreover, while the former focuses on topological measures, the latter also considers the possibility of implementing different impedance attributes in order to consider different transportation modes beyond pedestrian movement.

Despite the great interest in exploring different modelling choices for the best representation, we should remind the reader how this work is not focused on network configuration or accessibility properties for the assessment/description of the overall study area. On the contrary, we are interested in its relationship with stores distribution; therefore, the choice of **scale** and **impedance** are here selected with the ultimate aim being to acquire a better comprehension of the different configurational properties in relation to the retail system.

For the definition of the first aspect, we should consider that, as discussed in Part 2, retail geography traditionally considers local scales when describing store agglomerations (Section 9). RFA procedure developed in this dissertation is focused on three spatial scales: street (150metres), neighbourhood (300 metres) and district scale (600 metres).

At the first scale, as we further develop in the next section, specific urban design properties describe the streetscape with a compositional rather than a configurational approach. Neighbourhood and district scales, as previously discussed in retail geography studies, are here implemented for the investigation of local configurational properties. These two spatial scales also close to the traditional thresholds investigated in urban and transport studies focusing on the relation between network configuration, retail presence (250 and 800 metres in Scoppa and Peponis 2015; 600 metres in Sevtsuk 2010, 2014; 800 metres in Porta et al. 2012, etc.) and pedestrian counts (250 metres in Lerman et al. 2014, 600 metres in Bielik et al. 2018, etc.). In this work, the choice of these two thresholds should guarantee the coherence of the spatial scales used in both retail (RFA) and configurational analyses, reducing possible biases related to scale discrepancies. A third scale is further considered: 1200m is traditionally considered as the maximal distance which pedestrians are willing to walk for shopping purposes as discussed in behavioural retail geography; similarly, Genre-Grandpierre (2008) has discussed how beyond this radius, vehicle-based movement prevails on pedestrian-based<sup>68</sup>. We can also add empirical evidence from the study area. The 2008 Household Mobility Survey on the French Riviera shows that 66.3 % of pedestrian trips are less than 1200

---

<sup>68</sup> The highly sloped morphological characteristics of the French Riviera might also be considered as a possible limit in the pedestrian movement at higher distances. Further studies might consider the role of street acclivity in configurational properties; here we focus on the traditional planar representation of the network.

m long and that 79.1% of trips shorter than 1200 m are done on foot, while more than 80% of trips longer than 1200 m are done with motorised modes (mainly car).

The choice between two different mobility alternatives might explain why in RFA outcomes we identified vehicle-oriented retail agglomerations found in peripheral spaces of our study regions (RF6 described in Subsection 5.5). Moreover, as previously discussed, in the same suburban spaces, small retail concentrations are distributed along the main suburban routes, connecting cities and villages to the highway and to the departmental street network. Small retail agglomerations spread in the peripheral areas might not be explained by the traditional MET, which has been developed for central areas of large and compact cities (Van Nes and Yamu 2017). Outside of the centres, where the urban spaces are predominantly car-oriented, the network configuration/accessibility should not ignore the mobility characteristics of inhabitants.

We therefore apply the same four configurational indicators using a vehicle-oriented impedance: to each street segment, we assign a travel time value derived by the congestion-free speed allowed on each street<sup>69</sup>, considered as a proxy of real average speed time. We then define two thresholds at the meso (5 minutes) and macro (20 minutes) scales. The macro-scale level has been here again chosen as the maximal distance citizens are willing-to- drive to reach retail locations; 20-25 minutes are traditionally considered in the empirical analysis of the local Chamber of Commerce of Nice for the study of anchor stores catchment areas. For the implementation of these indicators we reintegrate tunnels, motorways, urban highways and every street segment not accessible by pedestrians which have been previously excluded in pedestrian-based configurational analysis.

The investigation of different scales (network-constrained radius) and street-network mobility modelling has been the focus of the DAMA project (Fusco et al. 2016) before the commencement of this dissertation. In this work, where the focus of the analysis was only on the city of Nice and its surrounding municipalities of Saint-Laurent-du-Var and Cagnes-sur-Mer, we outlined the importance of the combined effect of vehicle-oriented macro-scale and meso-scale pedestrian-oriented centralities. A similar approach has also been investigated in Cutini et al. (2018) where the SSX approach considers a combination of pedestrian, bicycle and vehicle-oriented configurational analysis.

When implementing vehicle-oriented configurational indicators, Reach (12.1), Closeness (12.2) and Betweenness (12.4) do not change their signification. Reach is still a measure of street count within a given temporal distance, Closeness describes the proximity (in terms of temporal frequency) of streets, and Betweenness is the number of shortest paths defined by the street vehicle impedance. However, the Straightness indicator on a vehicle-based network with temporal impedances has a totally different meaning: the straightness  $sr(i)$  of a node  $i$  represents the ratio between the straight Euclidean distance  $\delta[i, j]$  and the time needed to travel between  $i$  and  $j$ . The resulting spatial unit is [meters/time] (speed) describing the network vehicle efficiency and it might vary not only considering the tortuosity but also the different speed limitations of each street. Since vehicle-based measures are implemented on

---

<sup>69</sup> The speed for each street segment is defined by the “importance” functional classification provided by the BD Topo.

a complete street network with tunnels, motorways, urban highways, and access ramps (not accessible by pedestrians), this *Efficiency* measure is profoundly different from traditional Straightness. Road-network Efficiency might highlight different configurational properties that would be hidden from the traditional pedestrian-based approach and which might explain specific retail distribution characteristics in peripheral sectors.

The combination of different scales and impedances should better represent how configurational properties influence pedestrian and car flows within the large conurbation of the French Riviera. While pedestrian-based approaches at smaller scales might be associated with compact continuous centres, vehicle-oriented at local mesoscales describe the network characteristics of suburban areas. These methodological choices allow the study of configurational properties applied to wide heterogeneous urban spaces with different mobility logics within a common and coherent framework.

While in this work we focus on the two main mobility choices—pedestrian vs vehicle oriented—future studies might consider a different and more detailed representation of mobility impedances for the identification of specific configurational/accessibility indicators (i.e. bicycle, public transport, etc.)

In summary, four configurational indicators implemented at five different local scales are considered, describing flows and configurational properties of the street network both from a pedestrian and vehicle-oriented perspective.

In Figure 61, we observe the outcomes of Reach and Betweenness indicators at the five local scales: while the three local pedestrian scales define different properties within the core areas of cities and villages of the study area, by only focusing on the wider vehicle-based scales we can highlight the peripheral structure of the French Riviera. At a five-minute distance, the central area of Nice extends its influence over the western municipalities of Saint-Laurent-du-Var and Cagnes-sur-Mer. A similar situation is observed in Antibes towards its west side over Juan-les-Pins and Cannes towards the north including the municipality of Le-Cannet.

When considering a 20-minute radius, high centrality values highlight two large central spaces. On the eastern side, a triangle can be traced along the cities of Cannes, Grasse and Antibes, while, on the central section of the Riviera, we might identify a “U” shaped region defined by the coastline between Cagnes and Eastern Nice and the two linear strips towards the inland along the Var Valley and the Cagnes Valley. It is in these areas that major retail projects of regional importance are under construction. On the contrary, the core of Nice becomes relatively less central. Monaco and Menton are described by nucleated and isolated centralities which, at wider scales, seems to strongly reduce their centrality values.

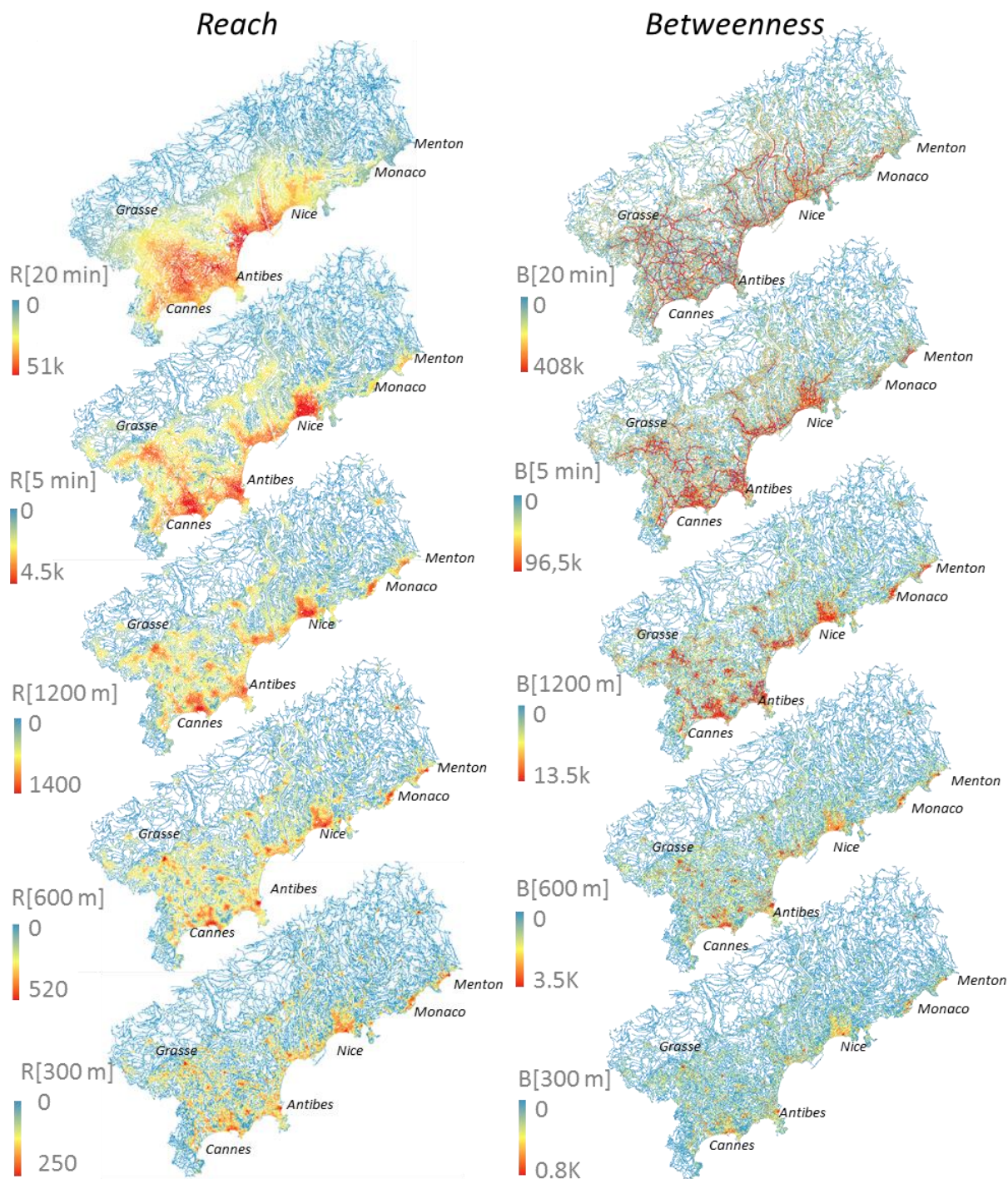


Figure 60 Reach and Betweenness configurational centralities implemented at different scales (pedestrian-oriented on the left and vehicle oriented on the right).

### **A new proposal for the normalisation of MCA street network centralities**

The four centrality measures, in originally conceived, are strongly influenced by the local densities of street segments. As previously introduced in Subsection 6.1, several normalisation approaches have been proposed with the intention to remove the effect of the number of streets and unveil further connectivity properties of the street network. In SSx, normalisation

methods are implemented and survey the total number of streets (more precisely visual axes) of the network; the total depth or comparing local centrality values with their average on the overall space (Hillier 2012). Normalisation procedures in the primal approach of MCA have received minor attention.

We propose here a relatively recent and simple approach for the normalisation of the four traditional centrality indicators of the Multiple Centrality Assessment. The implemented procedure finds its origins in the two-step floating catchment area (2SFCA) developed by Luo and Wang (2003) and traditionally implemented in studies of health/facilities spatial accessibility.

2SFCA analysis finds its origins in a gravity-based accessibility model where the spatial interactions between supply and demand locations play important and distinct roles. The general formula provided by Hansen describes the accessibility of a location as follows:

$$A_r = \sum_j \frac{W_j}{f(c_{ij}; \beta)} \quad (7.5)$$

Where the total accessibility  $A_r$  at a given radius  $r$  is given by the sum of the destinations  $j$  within a distance  $r$  (similarly to the Reach in 6.1); each destination  $j$  is weighted with a proportional measure  $W$  describing the mass of a location  $j$  and inversely proportional to the distance.  $f(c_{ij}; \beta)$  is a function which considers the cost needed to overcome the distance between  $i$  and  $j$  while  $\beta$  is a friction distance decay parameter.

The first main difference which we might notice between accessibility and configurational studies is that in the former, origin  $i$  and destination  $j$  are associated with a specific function which needs to be defined by the analyst in order to assess the relative importance of each point and to estimate the interaction between origins and destinations. In the latter, origins and destinations are equally considered; their spatial organisation on the network at the origin of the hypothesis of natural movement is used to estimate flows, and, consequently, the location of urban functions (as described in Section 2.2).

From the measure of accessibility defined in the seminal work of Hansen (1959), Sevtsuk and Mekonnen (2012) derive a network centrality measure called Gravity centrality and measured as:

$$G_r(i) = \sum_{j \in G - \{i\}, \{d_{ij} \leq r\}} \left( \frac{W[j]}{e^{\beta \cdot d[i,j]}} \right) \quad (7.6)$$

As in the 7.5,  $W[j]$  is a weight for each destination  $j$  and  $\beta$  is the exponent controlling the distance decay function for each shortest path between two destinations. When the origin and destinations  $i$  and  $j$  correspond to every network junction (or, as in our case, to street midpoints) and the coefficient  $\beta$  is set to zero, the accessibility value  $G_r(i)$  is equal to the street network configurational Reach  $R_r(i)$ .

In health care accessibility studies, the simple accessibility implemented as in 7.5 only assesses the destination (services) distribution in the geographic space without considering the relative demand distribution associated with each service.

The application of the traditional accessibility measure to the accessibility to, for instance, health care services, would describe high values in densely populated areas, with lower values in peripheral areas. This estimation does not consider that the heterogeneous distribution of the population (demand side) might overcrowd the services in central areas, reversing the real accessibility to the service. Measures of spatial accessibility should consider what is called the *supply-demand match ratio* of a region, as firstly proposed by Joseph and Bantock (1982). From the 7.6, the accessibility at location  $i$  is then implemented as

$$A_r(i) = \sum_j \left[ \frac{S_j d_{ij}^{-\beta}}{\left( \sum_k P_k d_{kj}^{-\beta} \right)} \right] \quad (7.7)$$

Where  $S_j$  is the capacity of the health facility  $j$ ,  $P_k$  is the population surrounding the location  $k$ ,  $d$  is the distance to travel and  $\beta$  is the exponent controlling the distance decay function. The numerator is the traditional Hansen accessibility of a health Service  $S$  (Demand) while the denominator considers the Population  $P$  (Supply) surrounding each facility.

*“Although conceptually advanced, the model is not intuitive or transparent to public health professionals. Moreover, its distance friction parameter  $\beta$  requires additional data and work to define and might be region specific (Huff 2000).”* From this observation, Luo and Wang (2003) proposed a simplified version of the *supply-demand match ratio* and an easy procedure for its implementation in GIS environment called the 2SFCA.

As originally explained by the authors: in the **first step**, for each physician location  $S_j$  we implement the sum of the population  $P_k$  for every location  $k$  within a radius  $r$  which informs us about the *physician-to-population* ratio. In the **second step**, for each population location  $i$ , the sum of the *physician-to-population* ratios are found within the same radius  $r$ .

$$A^{2SFCA}_r(i) = \sum_{j \in \{d_{ij} \leq r\}} \left( \frac{S_j}{\sum_{k \in \{d_{kj} \leq r\}} P_k} \right) \quad (7.8)$$

2SFCA might be also considered as a *normalised accessibility*, where each origin  $i$  accessibility to a given facility/service is assessed not only evaluating a function of the distance and of the number of direct accessibility, but it also considers the size of population/customers which each facility has to serve.

When implementing 2SFCA considering  $\beta=0$  and origins-destinations representing street network nodes (as previously discussed for  $G_r(i) \equiv R_r(i)$ ), the *supply-demand match ratio* might be extended to the configurational studies considering the *node-to-node match ratio* (or, *street-to-street match ratio*).

Similar applications have been recently provided in a sDNA tool from the Cardiff University (Chiaradia et al. 2012) where these measures are named as Two Phase Destination (TFD) and Two Phase Betweenness (TFB), defined as the probability of people-flows passing-through the street link. Only more recently TFD and TFB have been studied in pedestrian flow studies in Cooper et al. (2018) and in retail-related studies in Grimaldi et al. (2019).

The rationale of the 2SFCA in configurational studies might be the following:

Every origin (node or street)  $j$  within a subnetwork centred in a given destination (node or street)  $i$ , participate in an equal way to the determination of the centrality measures in  $i$ ; every  $j$  participate in equal way to the determination of  $R_r(i), C_r(i), S_r(i), B_r(i)$ . Nonetheless, when considering every origin-point  $j$  the possibility to choose  $i$  as destination-point from  $j$  is given by the inverse of the amount of possible destination  $k$  surrounding  $j$  in the same radius  $r$ . The possibility of  $i$  to be chosen as a destination from  $j$  is given by the inverse of the number of available destination  $k$  surrounding  $j$ .

Using this information to weight the traditional MCA configurational measures would reveal

$$Reach_r^N(i) = \sum_{j \in G - \{i\}; d[i,j] \leq r} \left( \frac{j}{\sum_{k \in G - \{j\}; d[j,k] \leq r} k} \right) \quad (7.9)$$

Which is equal to:

$$Reach_r^N(i) = \sum_{j \in G - \{i\}; d[i,j] \leq r} \left( \frac{j}{R_r(j)} \right) \quad (7.10)$$

And which correspond to the TPD of Cooper and Chiaradia.

This indicator would reveal the possibility of a street segment to be chosen as a destination from its neighbouring street element and which is strictly dependent on the connectivity properties of the network independently by the density (each street is “weighted” on the number of surrounding street elements).

Similarly, the other centrality indicators might be normalised as follow

$$Closeness_r^N(i) = \frac{1}{\sum_{j \in G - \{i\}; d[i,j] \leq r} d[i,j] \cdot \frac{1}{R_r(j)}} \quad (7.11)$$

$$Straightness_r^N(i) = \sum_{j \in G - \{i\}; d[i,j] \leq r} \frac{\delta[i,j]}{d[i,j]} \cdot \frac{1}{R_r(j)} \quad (7.12)$$

$$Betweenness_r^N(i) = \sum_{j,k \in G - \{i\}; d[j,k] \leq r} n_{jk}[i] \cdot \frac{1}{R_r(j)} \quad (7.13)$$

This approach allows us to apply a local normalisation which considers the network configurational properties.

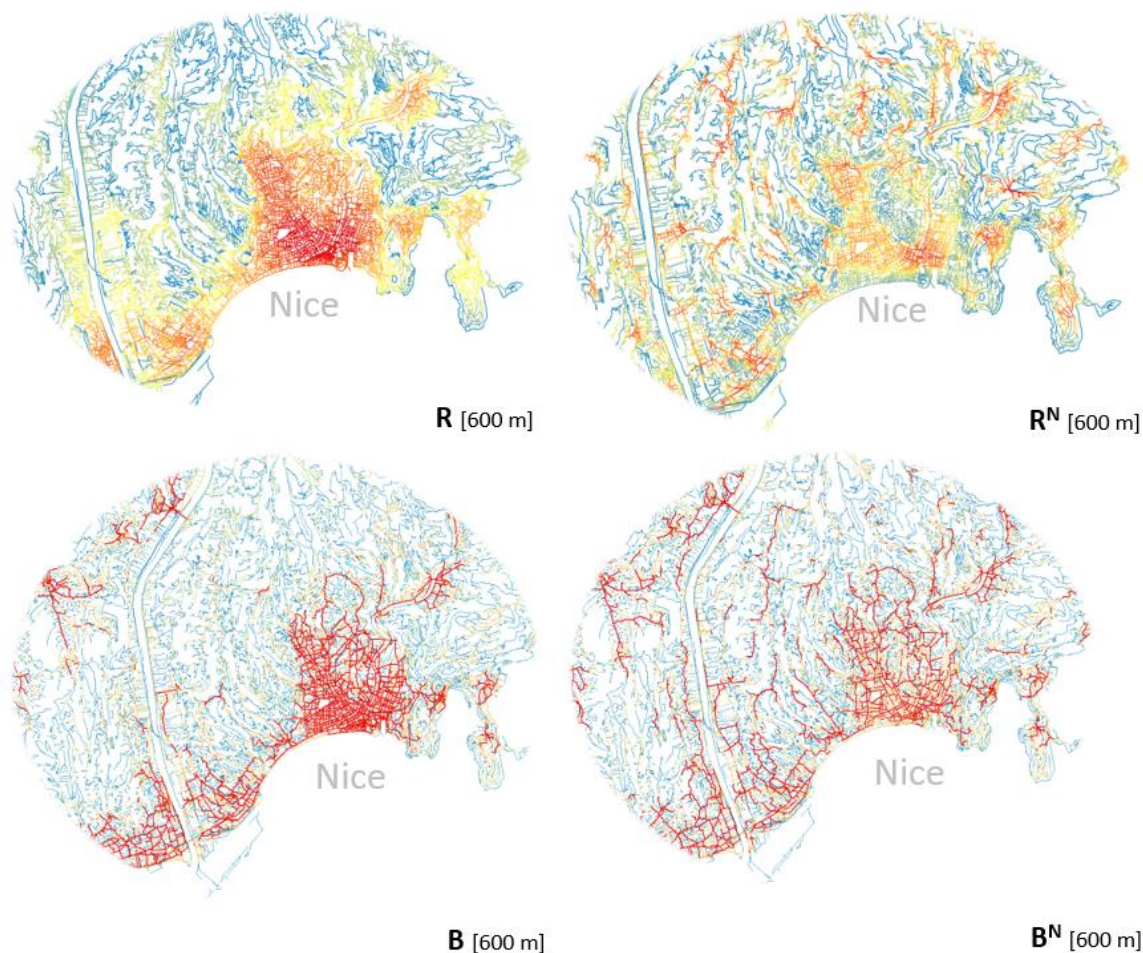


Figure 61 Projection in the geographic space (Nice) of Reach-R and Betweenness-B indicators (left) and their normalized version through the 2SFCA method (right)

While the simple Reach informs us about the number of streets in a given network radius, the normalised Reach allows one to identify the relative importance of each street within their surrounding subnetwork: as we might see from Figure 61 when normalising the Reach, bottleneck structures within the city and not just with bottleneck axes (that are already highlighted by the normalized Betweenness). After all, comparing the two maps on the left of Figure 61, we see that bottleneck structures normally have one or two bottleneck axis, but concern wider areas around these axes.

In the core area of Nice, the Reach provides a sort of density measure with a negative gradient from the centre towards the peripheral areas, following the more or less well-meshed street grid. Normalised Reach highlights two main bottlenecks. The first (butterfly-shaped) structure is found at the core of Nice city: its shape is probably determined by the presence of the two hills at its north and south sides (respectively Cimiez and Castle Hills), which leaves only a small corridor bridging the Eastern and Western parts of the city. The second structure is found at the north of Nice where the tree-shaped street networks climbing up the surrounding hills are grafted to the meshed grid of the low valley. While these two bottlenecks are influenced by the morphology of the site, the third area with higher Normalised Reach values might be observed around the Central Station of Nice: only six streets allow for the connection/accessibility between the central area and the northern districts. The train line



and the superposed urban route represent the major urban fracture between northern and southern Nice. Finally, in the peripheral hill spaces, several local bottlenecks are also highlighted along the main axes of the tree-shape urban network.

The second images (Fig.62-right) we highlighted are the fourth quantile of Betweenness and Normalised Betweenness: here again, we might see how the former focus on the core areas and almost every street is described by high values. On the contrary, the Normalised Betweenness selects only the most intermediary axes within central areas, and, contemporarily, allows the identification of the streets with the highest Betweenness in peripheral areas. This last aspect seems to us a promising indicator when investigating the distribution of socioeconomic activities and phenomena in semi-peripheral areas (traditionally overlooked in those studies of Table 1-2). In the following image, the 8 values assumed by the four indicators of centrality (at the same metrical radius) are projected on the space of Cannes and Antibes (Fig.62).

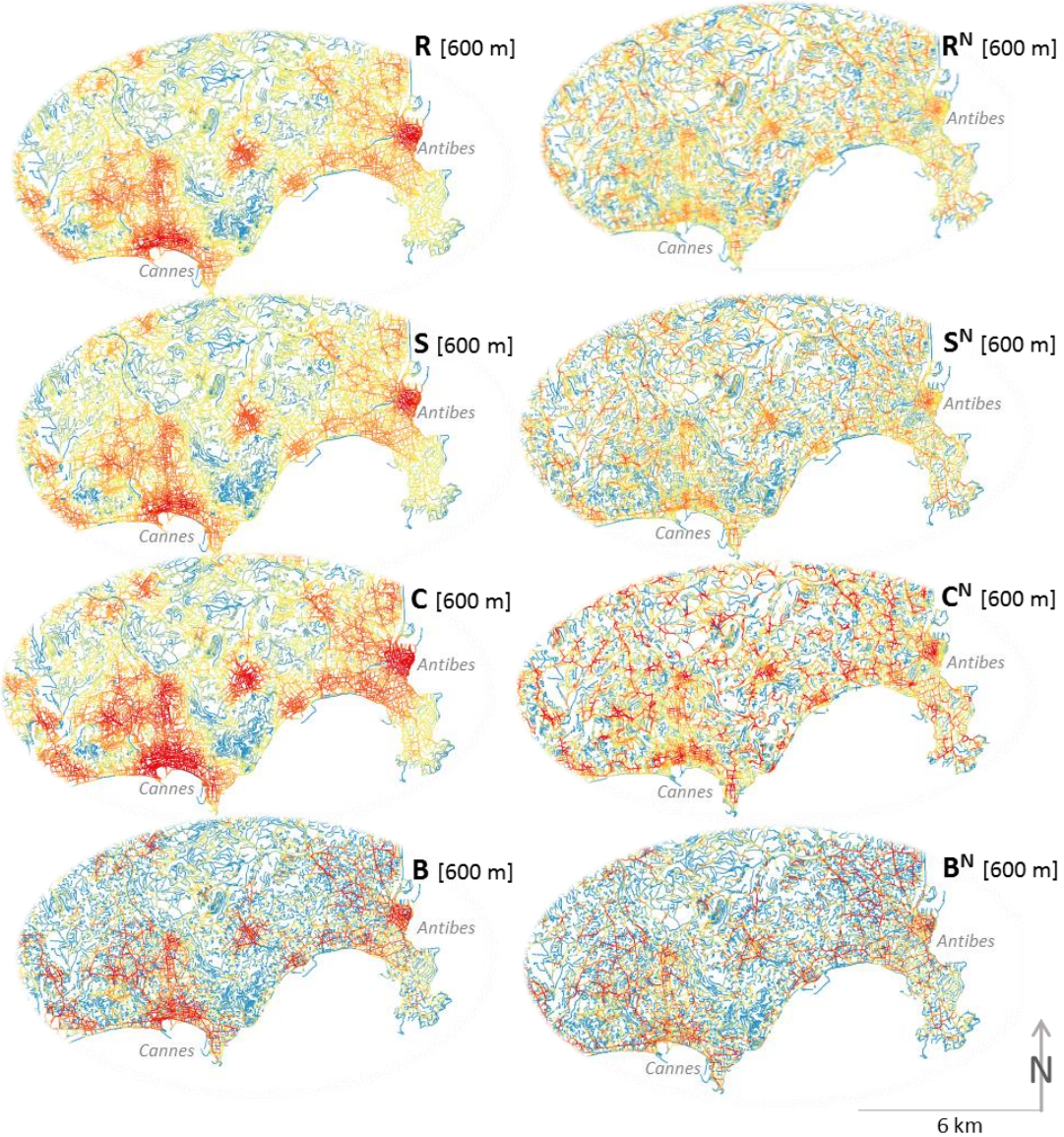


Figure 62 Projection in the geographic space (Cannes and Antibes) of network configurational indicators (left) and their normalised version (right).

The following table proposes a synthetic view of the 40 indicators which have been implemented for the description of the street network configurational properties, which implements five spatial scales and two impedance metrics, pedestrian and vehicle oriented (Tab.8). Each configurational indicator provides a rich description of the study areas which would need an attentive description and interpretation. Nonetheless, this work goes beyond this dissertation when our main goal is the identification of the urban morphological characteristics correlated with retail distribution and organisation.

	Pedestrian r [meters]			Vehicle [minutes]	
	300	600	1200	5	20
<i>Reach<sub>r</sub></i>	$R_{300}$	$R_{600}$	$R_{1200}$	$R_5$	$R_{20}$
<i>Reach<sub>r</sub><sup>N</sup></i>	$R_{300}^N$	$R_{600}^N$	$R_{1200}^N$	$R_5^N$	$R_{20}^N$
<i>Closeness<sub>r</sub></i>	$C_{300}$	$C_{600}$	$C_{1200}$	$C_5$	$C_{20}$
<i>Closeness<sub>r</sub><sup>N</sup></i>	$C_{300}^N$	$C_{600}^N$	$C_{1200}^N$	$C_5^N$	$C_{20}^N$
<i>Straightness<sub>r</sub></i>	$S_{300}$	$S_{600}$	$S_{1200}$	$S_5$	$S_{20}$
<i>Straightness<sub>r</sub><sup>N</sup></i>	$S_{300}^N$	$S_{600}^N$	$S_{1200}^N$	$S_5^N$	$S_{20}^N$
<i>Betweenness<sub>r</sub></i>	$B_{300}$	$B_{600}$	$B_{1200}$	$B_5$	$B_{20}$
<i>Betweenness<sub>r</sub><sup>N</sup></i>	$B_{300}^N$	$B_{600}^N$	$B_{1200}^N$	$B_5^N$	$B_{20}^N$

Table 8 Street Network Configurational Indicators

### **Directional configurational properties**

When implementing configurational analysis, the importance of each urban space is only defined by its relative position within the surrounding urban network system. Nonetheless, several works in urban studies focused on the influential power that specific urban features might have on their close surrounding spaces, modifying their attractiveness, economical value, accessibility, etc. In retail geography, the presence of urban magnets is also seen as an important locational factor (i.e. ancillary and parasitism locational behaviour described by Lebrun 2002, Section 2.4).

The presence of these specific morpho-functional features might be at the origin of local discrepancies between the observed value of flow and retail presence and the estimated values from the traditional configurational approaches.

In this work, where specific attention is given to urban morphological elements, three main components are studied with more attention: squares, the coastline, and large anchor stores.

As previously discussed, urban squares represent an essential morphological component, especially in Mediterranean cities (Cutini 2003, 2014; Fusco 2012<sup>70</sup>, etc.). Specific commercial activities might be found in these spaces and within their proximities. Festivals, temporary fairs, open-air markets, flea markets, Christmas villages, and several entertainment events take place in squares attracting local residents and tourists all throughout the year. Reducing the whole urban space to the street network would overlook the importance of these peculiar spaces. Using SSx vocabulary, we might assume that the encountering and retail potential associated with these spaces is higher than what the network configuration might determine. Squares might represent one of those factors, triggering the multiplicative effect described by Hillier as his MET.

Similarly, in a coastal urban region known internationally for the beauty of its seaside, we should not ignore the importance of the coastline attractiveness. The seaside is not just a natural amenity. It has been integrated into the urban design of the cities of the French Riviera in the form of boardwalks (*promenades* in French). The *Promenade des Anglais* in Nice, the *Promenade du Soleil* in Menton and the *Croisette* in Cannes are emblematic urban features attracting movement flows beyond the municipal and even national level. Tourism, as previously described, has been the economic engine and the first economic resource of this department for the last 150 years. The hospitality business industry, blue economy, and secondary home real estate, described in Section 3.4, find their economic power in the coastline and specific retail agglomerations types have been detected along these spaces.

These two components, coastline and squares, represent two important morphological characteristics of the urban spaces under investigation: the former is a linear feature which is traditionally identified as a limit, doubled by a path in the presence of a boardwalk; the latter is traditionally identified as a node and, sometimes, as a “core”, i.e. a micro-district within a district. These two morphological features have influenced the urban form of the urban cities and which still stimulate the movement, the economy, and the retail distribution of the French Riviera. Nonetheless, configurational analyses are blind to these two aspects (as the further analysis of Chapter 7 and 8).

The third component which deserves specific attention is the presence of large anchor stores. When considering large planned centres (over 2000 square metres of retail surface), it is clear how their location is not anymore influenced by the properties of the urban morphology, but other locational factors going beyond the departmental level should be considered (such as vehicle accessibility at the regional scale, land availability and price, presence of competitors, special agreements with local authorities, etc., as for example shown by Bauer 2016 for the case of IKEA). Moreover, these large projects might take several years (even one or two decades) between the moment of its first conception and the final realisation with the important participation of local stakeholders and of local and national authorities. These projects modify their surrounding urban morphology and independent retailers are also attracted by these locomotives. The different localisation factors between large and small stores and the consequent importance of large stores have also been discussed in Ortiz-Chao (2008) where differences in spatial organisation of these two retail formats and the network

---

<sup>70</sup> The role of squares in Nice is described in: <http://unt.unice.fr/uoh/espaces-publics-places/>

organization are investigated. Cutini (2001) differentiates between stores operating in the free market and “monopolistic” stores, recognising that the presence of monopolistic activities are non-attracted attractors distorts the close correspondence between configuration and land use.

To consider the importance of the influence of large retail centres on small/average-sized stores, we decided to introduce a specific set of indicators describing the directional configurational properties. Instead of weighing centrality indicators with buildings (as proposed in Sevtsuk and Mekonnen 2012), we use a specific weighing matrix integrating (separately) the three aforementioned features.

Firstly, we select those street segments associated with squares and similar open spaces (street elements named as *Cours*, *Esplanade*, *Place*, *Placette*, and *Square* in the IGN BDTopo database). Secondly, the length of the selected street segments is implemented and used as a proxy of the correspondent place size; other street segments are associated with a null value. Finally, this new weight system is implemented at three local radii (600, 1200 and 2400 meters).

This same procedure has been further applied by considering a second weight matrix based on the length of coastal street segments (with the exclusion of commercial port and airport platforms and other logistic/industrial sites), and a third time by considering the surface of the 135 largest anchor stores.

Reach value in these three cases informs us, for example, on the amount of squares/coastal streets/large anchor stores within a given radius  $r$ ; Betweenness informs us of the number of shortest paths of maximal length  $r$ , having one of the two extremities lying on the square/coast/large anchor store streets. Closeness informs us of the proximity to and Straightness of the linearity of the shortest paths towards these three destinations.

Table 9 provides an overview of the weighted configurational indicators.

	Towards Squares $r$ [metri]			Towards Coastline $r$ [meters]			Towards Anchor Stores $r$ [meters]		
	300	600	1200	300	600	1200	300	600	1200
<i>Reach<sub>r</sub></i>	$R_{300}^S$	$R_{600}^S$	$R_{1200}^S$	$R_{300}^C$	$R_{600}^C$	$R_{1200}^C$	$R_{300}^{AS}$	$R_{600}^{AS}$	$R_{1200}^{AS}$
<i>Closeness<sub>r</sub></i>	$C_{300}^S$	$C_{600}^S$	$C_{1200}^S$	$C_{300}^C$	$C_{600}^C$	$C_{1200}^C$	$C_{300}^{AS}$	$C_{600}^{AS}$	$C_{1200}^{AS}$
<i>Straightness<sub>r</sub></i>	$S_{300}^S$	$S_{600}^S$	$S_{1200}^S$	$S_{300}^C$	$S_{600}^C$	$S_{1200}^C$	$S_{300}^{AS}$	$S_{600}^{AS}$	$S_{1200}^{AS}$
<i>Betweenness<sub>r</sub></i>	$B_{300}^S$	$B_{600}^S$	$B_{1200}^S$	$B_{300}^C$	$B_{600}^C$	$B_{1200}^C$	$B_{300}^{AS}$	$B_{600}^{AS}$	$B_{1200}^{AS}$

Table 9 Directional Centrality Indicators (Squares, Coastline and Anchor Stores)

## 8. Measuring the streetscape

### Definition of street-based spatial unit

While the previous chapter described the methodological choices for the implementation of configurational indicators, this chapter will focus on the compositional descriptors of the physical streetscape.

A wide number of GIS procedures allow for the measurement of different aspects of the streetscape and the same descriptor might be obtained in several ways with more or less efficient algorithms. The previously reviewed works of Harvey et Aultman Hall (2016a, 2016b, 2017) and Vialard (2013) describe two different GIS procedures for the extraction of these indicators. There is not a right or wrong set of attributes and beyond the technical efficiency of each procedure (which is mainly determined by hardware/software characteristics and the accuracy of the available dataset) what is more interesting for the purpose of this dissertation, is the ability of each measure to account for the street-based perspective.

Harvey et Aultman Hall (2016a, 2016b, 2017) and Vialard (2013) propose two different reference systems for the measurement of compositional/geometrical indicators; the following image is a synthetic representation of these two approaches.

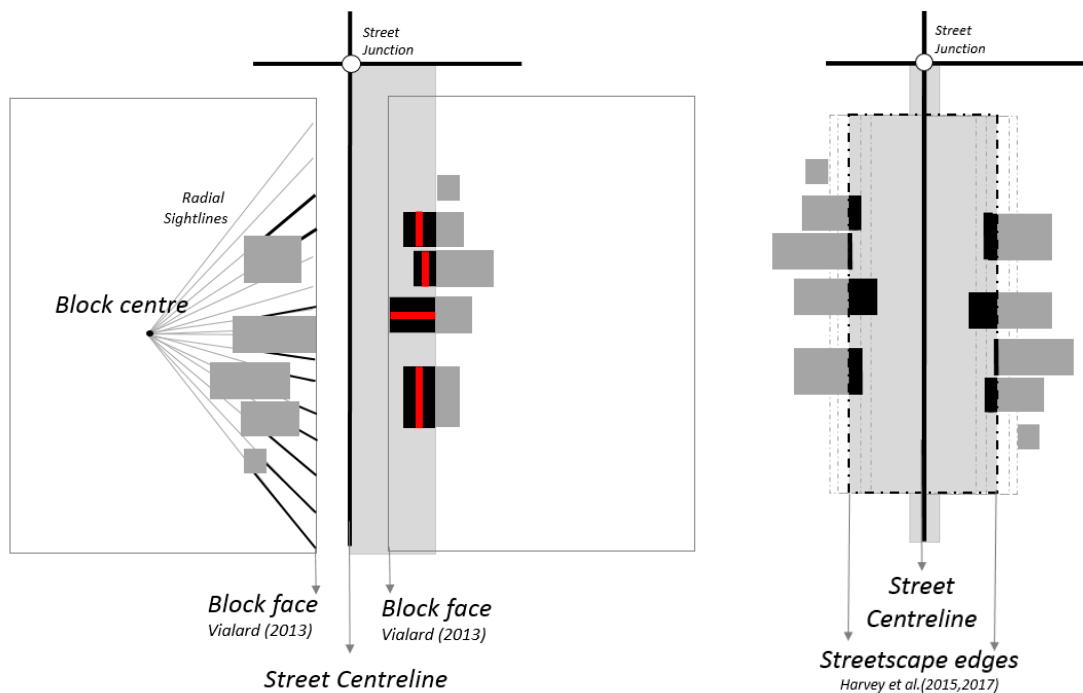


Figure 63 Two different approaches for the compositional analysis of the interface between streets and blocks. On the left, Vialard (2013) block-face approaches (radial sightlines and 20-metre buffer); on the right, the skeletal streetscape of Harvey et Aultman Hall (2015).

Vialard (2013) implements two procedures (Fig.63-left). The first is based on the *radial line method* from Boyce and Clark (1964): from the barycentre of the street-block, radial sightlines are regularly drawn up to the block-faces. The subpart of each sightline between the block-faces and the building footprints are successively used to implement measures of setback

(dark grey segments in Fig.63-left). The second indicator of *façade length* is obtained considering the sum of the largest width of each building footprint within a 20-meter buffer from the street centreline (red segments in Fig.63-left).

As previously discussed, Vialard (2013) focuses on the characteristics of block-faces following the morphological tradition based on the description of street-blocks from the aerial point of view.

On the contrary, from the urban design tradition, Harvey and Aultman Hall (2015a, 2015b) develop specific indicators describing the skeletal streetscape as visually observed considering its three dimensions.

Harvey and Aultman Hall (2015a, 2015b) propose a street centreline-based approach (Fig.63-right): after erasing street extremities closest to the street network nodes (with a buffer of 0.5 metres around intersections), only street segments within 20 and 500 metres are considered for the analysis. A specific algorithm is implemented for the identification of skeletal *streetscape edges*: a set of buffers is created around each street segment with one-metre intervals and within the first 40 metres from the street centreline. Each buffer is evaluated with a geometrical rule considering its internal and external building footprint; the algorithm allows detection of the buffer most consistently intersected by buildings, and the street edges are likely to be found. From these edges, several indicators are implemented such as: street width (distance between streetscape edges), average building height, Cross-sectional proportions (ratio between average height and average width), street wall continuity, variability of building height (standard deviation of the building height intersecting street edges), variability of width (proportion of area between edges which overlap with buildings), street length and sinuosity (ratio between street length and the straight line distance between the ends of segment).

The edge detection protocol proposed by the authors is based on the premises that, from a street-level perspective, pedestrians recognise an edge where facades align at a predominant setback distance. In our case study, these premises become weaker due to several urban peculiar characteristics: compared to the three North American cities studied by Harvey and Aultman Hall (New York, Boston and Baltimore; similar to Vialard's studies on the city of Atlanta), streetscapes in the French Riviera conurbation are characterised by higher heterogeneity in slope morphology, higher sinuosity, and, in some areas, a higher influence of the CIAM precepts (especially in semi-peripheral developments of the '70s, see Box..). Consequently, the premises of Harvey and Aultman Hall might be recognised only in planned central areas of the late 20th century, representing a small subpart of the region under investigation.

In the next images, the reader might perceive the low regularity in the spatial disposition of streets and buildings in the French Riviera semi-peripheral areas (Fig.64 left) compared to North American cities (Fig.64 right), which does not support the notion of predominant setback distance developed in Harvey and Aultman Hall.



*Figure 64 Some example of semi-peripheral suburbs in Nice, France (left) and Baltimore, USA (right).*

The procedures developed here use the street segment centreline, like in Harvey and Aultman-Hall (2016, 2017), as a geometrical reference system. This approach is better suited when the urban streetscape design should be analysed from the street point of view. As previously discussed, the use of the street segment is also a necessary condition to enable the cross-analysis with outcomes from configurational analysis implemented in the previous section. Moreover, preserving the street network connectivity (differently from Harvey and Aultman-Hall's works) would allow the implementation of network-constrained and location-based studies of the street facades distribution in the following Chapter (Chapter 9).

Inspired by the procedures here reviewed and considering the specificities of our case study, we elaborated on two different protocols, enabling us to describe different aspects of streetscapes.

The first procedure is based on sightlines perpendicularly and homogeneously distributed along the street centreline. A point every 3 metres is disposed along the street centreline; from each point  $j$  two sightlines have been drawn on the right and left side (respectively  $S_r(j)$  and  $S_l(j)$ ) up to the building facades or to a maximal length of 50 metres (Fig.65). In our case study we distributed almost 4.2M points on our street network and, consequently, 8.4M sightlines.

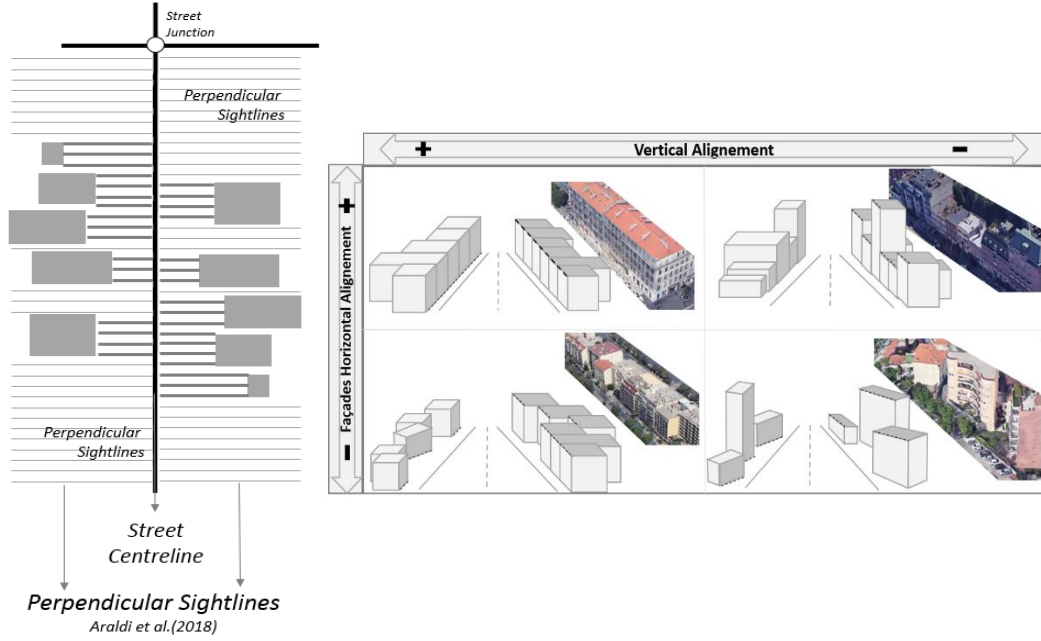


Figure 65 Streetscape measures implemented with sightlines. Graphical representation of the procedure (left), and combination given by Façades alignment and Vertical Alignment.

The length of all these sightlines is used for the implementation of two indicators: Openness informs us about the average open space width along the street, while Openness Roughness describes its variability. For each street  $i$  these measures are implemented as follows:

$$Openness(i) = \frac{1}{N} \sum_{j=1}^N S_r(j) + S_l(j) \quad (8.1)$$

$$Openness\ Roughness(i) = \sqrt{\frac{(\sum_{j=1}^N (S_r(j) - \overline{S_r(j)}) + \sum_{j=1}^N (S_l(j) - \overline{S_l(j)}))}{N-1}} \quad (8.2)$$

Where the 8.1 is the average of the sum  $S_r(j) + S_l(j)$  for each of the  $N$  points along the street  $i$ , and it could vary between 0 and 100 meters. The 8.2 is the standard deviation of the sightline lengths; right and left errors are separately evaluated in order to avoid the possible compensation effect given by the right/left alternation of buildings (resulting in a null standard deviation in case of perfect alignment and alternation of built forms). The same procedure might be further applied considering only the subset of right and left sightlines touching building façades, respectively  $W_r(j)$  and  $W_l(j)$  (dark grey in Fig.65). Of the original 4.2M sightlines, those touching buildings at least on one side represent 2.2M (which represent more than 52%<sup>71</sup>). Measures of **Building Setback** and **Façade Misalignment** are thus implemented as follows:

$$Building\ Average\ Setback(i) = \frac{1}{n} \sum_{j=1}^n W_r(j) + W_l(j) \quad (8.3)$$

$$Façade\ Misalignment(i) = \sqrt{\frac{(\sum_{j=1}^n (W_r(j) - \overline{W_r(j)})}{n_r - 1} + \frac{(\sum_{j=1}^n (W_l(j) - \overline{W_l(j)})}{n_l - 1}}{2}} \quad (8.4)$$

<sup>71</sup> A smaller value compared to streets units, where 71% of streets segments have at least one building in its 50-meter proximity.



Here again, the 8.3 represents the simple average of the building setback evaluated on the two-sided buildings every three metres; the 8.4 requires to be separately evaluated considering the sided-average  $\overline{F_r(j)}$  and  $\overline{F_l(j)}$  and the relative number  $n_r$  and  $n_l$  of sightlines touching buildings.

When associating height information of the buildings found at the extremity of each sightline, respectively  $H_r(j)$  and  $H_l(j)$  (only on those touching the façades previously selected), we might implement measures of Average Façade Height and Façade Height Misalignment. From the 8.3 and 8.4 we obtain:

$$\text{Building Average Height}(i) = \frac{1}{n} \sum_{j=1}^n H_r(j) + H_l(j) \quad (8.5)$$

$$\text{Building Height Misalignment}(i) = \sqrt{\frac{(\sum_{j=1}^n (H_r(j) - \overline{H_r(j)}))}{n_r - 1} + \frac{(\sum_{j=1}^n (H_l(j) - \overline{H_l(j)}))}{n_l - 1}} \quad (8.6)$$

Finally, for each sightline touching a building, we might apply the ratio  $HW_r(j) = H_r(j)/W_r(j)$  and likewise for the left side. This measure informs us about the semi cross-sectional ratio of the building height on its distance from the street centreline for each point j. Evaluating average and standard deviation along the street segment, we might obtain the street **Cross-sectional ratio** and its **Variability**, implemented as follows:

$$\text{Cross-sectional proportion}(i) = \frac{1}{n} \sum_{j=1}^n HW_r(j) + HW_l(j) \quad (8.7)$$

$$\text{Cross-sectional Variability}(i) = \sqrt{\frac{(\sum_{j=1}^n (HW_r(j) - \overline{HW_r(j)}))}{n_r - 1} + \frac{(\sum_{j=1}^n (HW_l(j) - \overline{HW_l(j)}))}{n_l - 1}} \quad (8.8)$$

Each of these eight descriptors might be implemented from the simple street centreline representation of the urban spaces. Nonetheless, the 8.3 and 8.7 might be biased by the different width of each street lane; Building Setback and the (semi) Cross-sectional Ratio should be measured from the street edges and not from the centreline. This information might be provided in the street dataset, using other sources of information (such as cadastre) or approximated by specific procedures like the one presented in Harvey and Aultman-Hall (2016, 2017). In our case study, we estimate this information, using the number of lanes provided by the BDTopo, multiplied by three metres; this measure might be considered as a proxy of the street edges separating the car and pedestrian pavement area. Before implementing the 8.3 and the 8.7, we subtracted from each sightline the semi lane width. Through this correction *Building Setback* and *Cross-Sectional Ratio* are thus measured from lanes edge and not from the centreline.

Through the sightlines approach, a large number of geometrical measures and combinations might still be possible; other indicators which we have explored (but omitted in this dissertation) are, for example, measures of symmetry and asymmetry<sup>72</sup> on the two sides of the street and measures of *streetscape entropy* combining the previous indicators (8.1-8) the

<sup>72</sup> As presented in Fusco G, Araldi A, Borgo S, Stufano R (2018) "Asymmetry along the way. Applications and conceptualization for morphology of urban fabric", First European Asymmetry Symposium -FEAS 2018, Nice, France

street and measures of *streetscape entropy* combining the previous indicators (8.1-8) aiming to measure the streetscape global regularity/legibility. Future studies will provide further details about this specific approach. In this dissertation, we focus on the eight main measures previously described and summarised in Table 10; Figure 66 illustrates the geographical projection of 6 indicators in some subspaces within our study area.

Streetscape indicator from Street Sightlines		
Urban Streetscape Component	Indicator	Implementation Formulae
Open Space	Openness	$\frac{1}{N} \sum_{j=1}^N S_r(j) + S_l(j)$
	Openness Roughness	$\sqrt{\frac{(\sum_{j=1}^N (S_r(j) - \overline{S_r(j)}) + \sum_{j=1}^N (S_l(j) - \overline{S_l(j)}))}{N-1}}$
Facades-Street Network-Parcels Relationship	Building Setback *	$\frac{1}{n} \sum_{j=1}^n W_r(j) + W_l(j)$
	Facades Misalignment	$\sqrt{\frac{(\sum_{j=1}^n (W_r(j) - \overline{W_r(j)}))}{n_r - 1} + \frac{(\sum_{j=1}^n (W_l(j) - \overline{W_l(j)}))}{n_l - 1}}$
	Average Building Height	$\frac{1}{n} \sum_{j=1}^n H_r(j) + H_l(j)$
Facades Cross-sectional Ratio	Building Height Misalignment	$\sqrt{\frac{(\sum_{j=1}^n (H_r(j) - \overline{H_r(j)}))}{n_r - 1} + \frac{(\sum_{j=1}^n (H_l(j) - \overline{H_l(j)}))}{n_l - 1}}$
	Cross-sectional proportion *	$\frac{1}{n} \sum_{j=1}^n HW_r(j) + HW_l(j)$
	Variability of Cross-sectional proportion	$\sqrt{\frac{(\sum_{j=1}^n (HW_r(j) - \overline{HW_r(j)}))}{n_r - 1} + \frac{(\sum_{j=1}^n (HW_l(j) - \overline{HW_l(j)}))}{n_l - 1}}$

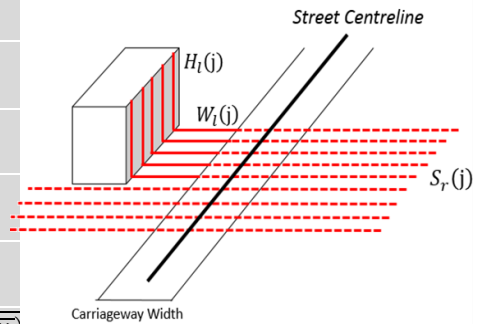


Table 10 Streetscape indicators from Street Centerline Sightlines; (\* corrected measures with lane width)

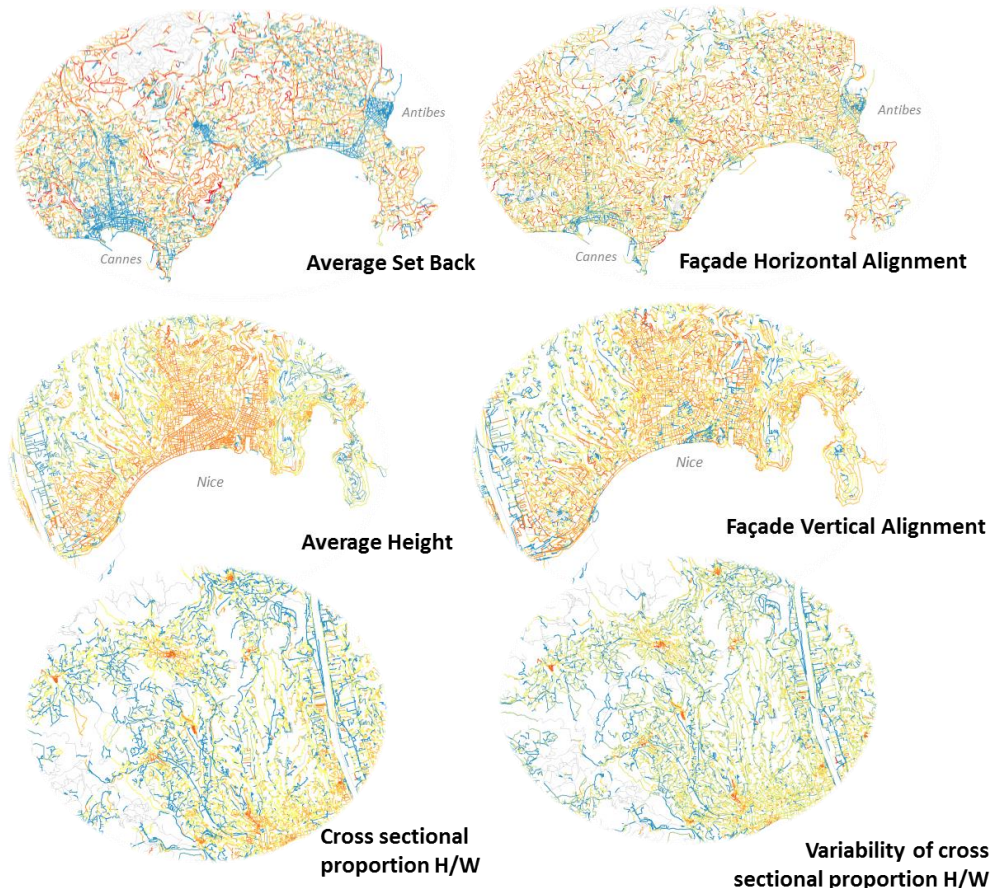


Figure 66 Projection in the geographical space of streetscape indicators.

All these indicators, similarly to those previously described by Vialard, and found in Harvey's works, consider façade distribution (along street centreline, in our approach, or along the block-face and streetscape edge in Vialard and Harvey).

Other measures of the urban form traditionally investigated in typo-morphology studies, such as surfaces and volume occupation of the streetscape (Subsection 6.2) are here overlooked. The streetscape should not be considered only for the disposition of the block front faces distribution (as in a theatre cyclorama background); the volumes of buildings should be considered as well in their three (or 2.5) dimensions.

A second procedural approach is therefore proposed, providing a set of indicators describing the perceived volumes and the traditional urban form measure considered in urban typo-morphology studies for the description of the urban fabrics.

This second procedure is derived from a street-based tessellation of space. Studying different issues from urban form/design, Okabe (2012) proposed a partition of urban space, namely network-based Thiessen tessellation, with the purpose of redistributing sociodemographic data on the street network (Okabe 2012, p.36). As discussed in Section 6.2, Thiessen tessellation applied to building footprints has been implemented by Voiron (1995), Sevenet (2013) and Hamaina (2013) for the study of the built form distribution. Inspired by these studies, we propose a new subdivision of urban space. A generalisation of Thiessen polygons is firstly created around each street segment, to identify the portion of planar space conventionally served by each segment.

The intersection with a two-sided buffer around each street allows for the identification of the portion of the Thiessen regions which is closest to street-edges, approximating the visible space from the street point of view (Figure 67). This resulting partition, named Proximity Bands (PBs) can consider different visual depths, depending on the specific aspect of the urban fabric under analysis.

In a North-American context, Vialard (2013) uses thresholds distances between 20 and 30 metres from the street centreline for frontage and setback measures; Talen (2005) suggests 9-15 metres from the block perimeter when analysing the sense of enclosure; Harvey implements buffers within the first 40 metres from the street centreline.

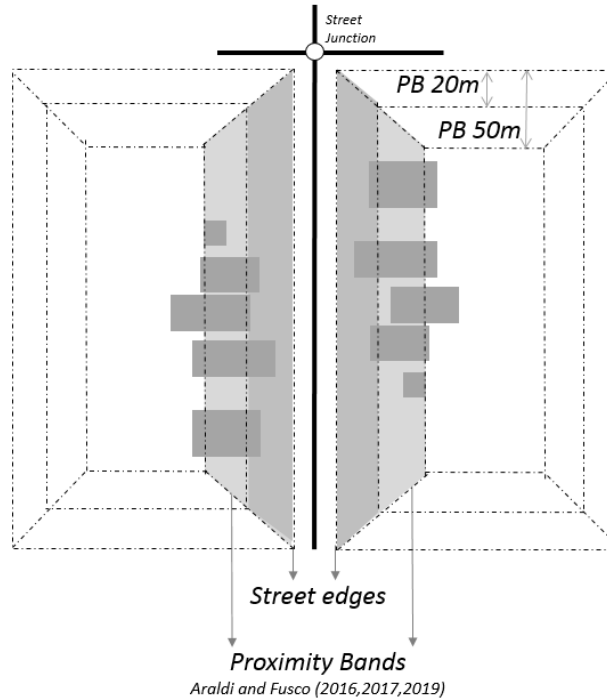


Figure 67 Streetscape measures implemented on the Proximity Band definition.

As in the previous procedure, the connectivity information of the underlying street network is transferred to PBs, allowing the implementation of network-constrained spatial statistics.

Finally, the small size variability of the PB surfaces reduces the potential bias induced by MAUP in statistical analysis. In our study area, 99,562 PBs were identified; 90% of the street segments measured between 4 and 300 m and an average surface of 4,680 m<sup>2</sup> when a 20-m band of visible space is considered (Table 11).

Spatial Unit Partition	Count spatial units	Mean spatial units size	Median unit size	Standard Deviation of size distribution	Skewness of size distribution
Administrative units (IRIS)	443	4,903,194 m <sup>2</sup>	599,539 m <sup>2</sup>	11,542,676	4.01
Urban Blocks	20,149	63,831m <sup>2</sup>	5,185 m <sup>2</sup>	286,931	13.52
Plots	488,858	2,976 m <sup>2</sup>	701 m <sup>2</sup>	20,813	57.25
Streets	99,562	129 m	77.65 m	181	4.69
Proximity Bands 20m	99,562	4,680 m <sup>2</sup>	2,608 m <sup>2</sup>	7,248	4.67

Table 11 The statistical dispersion of different spatial unit dimensions at the origin of the MAUP

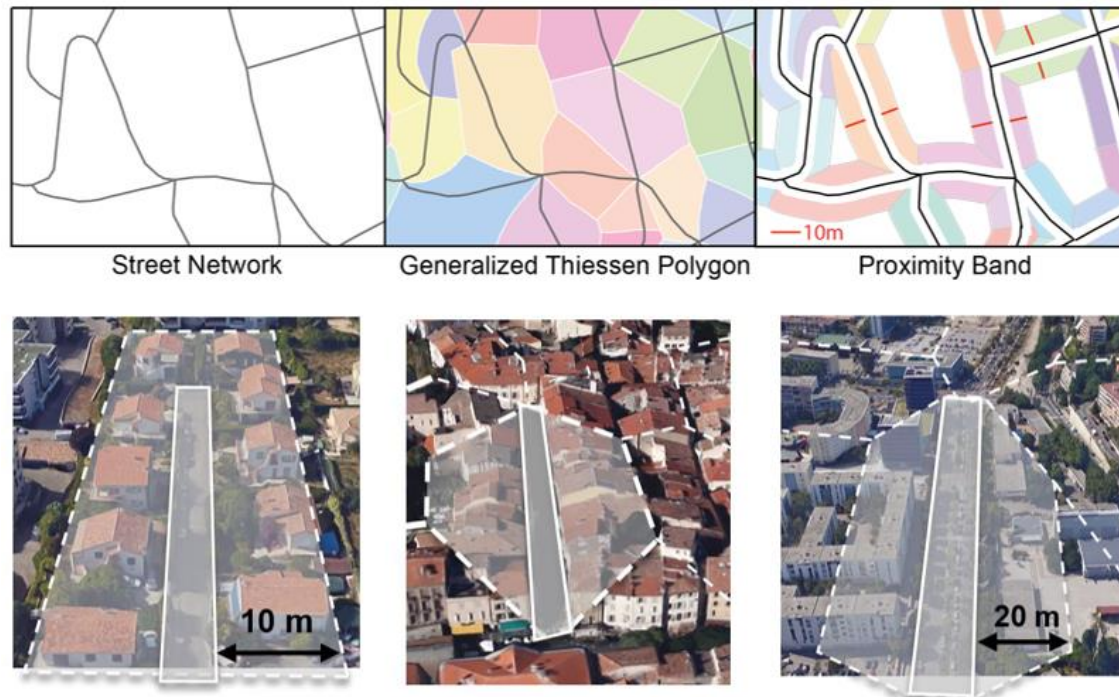


Figure 68 Proximity Band implementation from street segments (above) and some projection on real case studies (below).  
(Source: Araldi and Fusco 2016)

Twenty-one indicators obtained through geoprocessing in GIS were calculated for each PB. Selection criteria of indicators are the relevance and complementarity of the information carried by each of them when describing the urban fabric following the theory of urban morphology analysis (Borie and Denieul 1984) reinterpreted from the pedestrian point of view.

**Network morphology** is analysed through the street segment **Length** ( $L_{street}$ ) and **Linearity** (or inversely, its Windingness) (Harvey 2016) measured as the ratio between the street length and the Euclidean distance between its junctions.

Three indicators measure the **Local Connectivity** of the street network: each street junction is described by Boolean values that indicate the presence/absence of nodes of degree one (cul-de-sac), four (crossings) and other (mainly bifurcations), reflecting the main distinctions already made by Marshall 2005). Their averages are attributed on each street segment. Three indicators (for each node class) assuming three possible value: zero if both the nodes are different from the node class under analysis; 0,5 if one of the two extremities belongs to the node class under analysis; 1 if both are of the same node degree group (except by cul-de-sac which assumes only 0 or 0.5 values). As we describe later, when studying location-based patterns, these indicators would allow for the identification of areas characterised by a prevalence of tree-like network structure or meshed street grids (as in Marshall 2012).

**Built-up morphology** is studied within a 50 m wide PB and is first characterised by the traditional building Coverage Ratio index (as in Berghauser-Pont and Haupt 2010) as the ratio of the building footprint surface included within the PBs and the total surface of the PBs.

$$PB \text{ coverage ratio}(i) = \sum S_{tot} / \sum S_{PB50} \quad (8.9)$$

This indicator is then enriched by information about Built-up Type prevalence. The building footprint area is used as a proxy for categorisation of buildings (Sevtsuk 2014). For the French Riviera case study, the footprint segmentation is defined as follows: 0-125 m<sup>2</sup> (independent small houses), 125-250 m<sup>2</sup> (large villas and small traditional multi-family buildings which are particularly frequent in our study area), 250-1000 m<sup>2</sup> (townhouses and medium-sized blocks of flats), 1000-4000 m<sup>2</sup> (modern large blocks of flats, commercial, industrial or service buildings), greater than 4000 m<sup>2</sup> (mainly functionally specialized large buildings). For each building class  $j$ , we implement the following ratio:

$$Prevalence\ of\ Building\ Types(i) = \sum Si / \sum S_{tot} \quad (8.10)$$

Going beyond the expert-based (and surface-based) classification of building typologies, specific procedures might be considered as for instance the procedure we developed in Perez et al. (2018, 2019) where we propose the identification of building types through Bayesian clustering of few building morphological descriptors. Independently of the building classification, the 8.10 enables the attribution of the prevalence of building classes to each street segment.

We calculate the Specialization of Building Types by considering the presence/absence of ordinary dwellings.

$$Specialisation\ of\ Building\ Types\ (i) = \sum S_{spec} / S_{PB50} \quad (8.11)$$

Building Contiguity informs us whether the urban fabric is characterised by adjoining buildings linked to the compactness of the urban built-up environment at a very fine grain.

$$Building\ Contiguity(i) = \frac{\sum S_{b-u(k)} \left( \frac{1}{N_{build\ in\ b-u(k)}} \right)}{\sum S_{b-u(k)}} \quad (8.12)$$

Where  $S_{b-u(k)}$  is the surface of the built-up units  $k$  (defined by adjoining buildings) and  $N_{build\ in\ b-u(k)}$  is the number of individual buildings contained in each built-up unit.

The role of these two indicators in the characterisation of the urban fabric is recognised by classical urban morphologists (Caniggia and Maffei 1979, Pinon 1991).

**Site Morphology** measures the presence of steeply sloped surfaces in the 50-m proximity band as:

$$PB\ Surface\ slope(i) = \sum Sloped\ Surf_i / S_{PB50} \quad (8.13)$$

Where  $Sloped\ Surf_i$  is the surface of the PB where a slope higher than 30° is observed and  $S_{PB}$  is, again, the total surface of the PB at 50 metres depth. The angle of 30° (known as degree slope of soft natural sand) is chosen as the theoretical limit from which retaining walls are most probably necessary to allow the construction of building structure.

The relationship between street Network and Site Morphology is analysed through Street Acclivity as the average of the arctangent of the slope measured on a set of  $n$  points  $j$  (in radiant) homogenously distributed along the street (every 3 metres as previously defined).

$$Street\ Acclivity(i) = E [\arct(slope)_j] \quad (8.14)$$

The combination of *PB Surface slope* and Street Acclivity allows us to distinguish between streets following contour lines of site morphology (described by low acclivity and high PB Surface Slope) and streets perpendicular to the contour lines (where acclivity is high both along the centreline and on its sides).

**Network-Building Relationship** indicators describe the building geometry analysed in relation to the street segment. For this reason, they are computed on narrower PBs of 20 m. They include measures of Building Height, Open Space Width and Cross-sectional proportion obtained from the volumetric distribution of built form on the Proximity bands:

$$PB\ Building\ Height(i) = \sum V_{tot20} / \sum S_{PB20} \quad (8.15)$$

$$PB\ Open\ Space\ Width(i) = \sum S_{tot20} / \sum S_{PB20} \quad (8.16)$$

$$PB\ Height\ Width\ Ratio(i) = \frac{PB\ Building\ Height(i)}{Open\ Space\ Width(i)} \quad (8.17)$$

As the reader might notice, these three measures are similar to the (8.1,8.5,8.7) but while in the previous approach they are measured by only considering the building facades oriented to the street centreline, here (8.15-17) they are measured using proportions of volumes ( $\sum V_{tot20}$ ) and surfaces ( $\sum S_{tot20}$ ) in the 20-meters PB ( $\sum S_{PB20}$ ). Moreover, these three volumetric-based measures are closer to the urban typo-morphology approaches and they are easier to implement.

The Building frequency along the street network informs us about the serial vision of different buildings along a street (Hallowel and Baran 2013, Harvey et al. 2017).

$$Building\ Frequency(i) = \frac{N_{build}}{L_{street}} \quad (8.18)$$

Following the same approach, Land ownership fragmentation is used as an indicator of Network-Plots Relationship.

$$Land\ Ownership\ Fragmentation = \frac{N_{plot}}{L_{street}} \quad (8.19)$$

This indicator is still important from the visual point of view, especially when considering peripheral regions where buildings are not visible but the fragmentation of the landscape is still perceivable from the differences in the fence, hedges and dividing walls between different properties and private gardens.

Plot morphology in itself, although central to the study of morphogenetic processes (Conzen 1960, Caniggia and Maffei 1979) is considered less relevant for pedestrian perception.

Finally, **Street Corridor Effect** (Araldi and Fusco 2016) is implemented as a ratio between the total length of parallel façades and the street segment. This indicator is similar to the building wall indicator of Harvey 2016, and, similarly, Vialard 2013 considers the total length of building facades as previously described. Nonetheless, in our study the indicator of Street Corridor Effect focuses only on parallel facades ( $L_{parallel\ facades}$ ).

$$\text{Street Corridor Effect } (i) = \frac{L_{\text{parallel facades}}}{L_{\text{street}}} \quad (8.20)$$

The selection of facades parallel to the street centreline is obtained through a specific procedure implementing two sets of sightlines parallel to the street centreline (with spacing between lines as follows: 2.5 meters; and spacing between a set of parallel lines being: 1.25 meters). Through a combination of using perpendicular and parallel sightlines, only building faces parallel to the street are identified. Talen (2005) suggests focussing on the first 9-15 metres from the block perimeter when analysing the sense of enclosure; following this notion, Street Corridor Effect is implemented on a 10-metre PBs.

The synthetic definition of this second group of indicators is provided in Table 11<sup>73</sup>.

Streetscape indicator from Proximity Bands				
Urban Fabric Component	Indicator	Definition and implementation formulae		Proximity Band width
Network Morphology	Street Length	Street segments length between two intersections	$L_{\text{street}}$	/
	Windingness	$1 - (\text{Euclidean distance} / \text{Network distance})$ between two intersection	$1 - L_{\text{eucl.}} / L_{\text{street}}$	/
	Local connectivity	Average of the presence nodes of degree 1 (ND1)	$\sum ND_i [0,1] / 2$	/
		Average presence nodes of degree 4 (ND4)		/
Average presence nodes of degree 3, 5+ (ND35+)		/		
Built-up Morphology	Prevalence of Building types	(0:125] m2 building surf. / total built-up surf.	$\frac{\sum S_j}{S_{\text{builted}}}$	50
		(125:250] m2 building surf. / total built-up surf.		
		(250:1000] m2 building surf. / total built-up surf.		
		(1000:4000] m2 building surf. / total built-up surf.		
PB coverage ratio	Built-up Surface / PB Surf.	$\sum S_{\text{tot}} / \sum S_{\text{PB}}$	50	
	Building Contiguity	Weighted average of buildings frequency on built-up units		$\frac{\sum S_{b-u(i)} (1/N_{\text{build in b-u}(i)})}{\sum S_{b-u(i)}}$
	Specialisation of Building Types	Specialized Building surf. / PB surf.		$\frac{\sum S_{\text{Spec}}}{\sum S_{\text{PB}}}$
Network-Building Relationship	Street corridor effect	Parallel façades length / street length	$L_{\text{par.fac}} / L_{\text{street}}$	10
	PB building height H	Building volume / PB surface	$\sum V_{\text{builted}} / \sum S_{\text{builted}}$	20
	Open Space Width W	(PB surf. - built surf.) / street length	$(S_{\text{PB}} - S_{\text{built}}) / L_{\text{street}}$	
	Height/Width Ratio	PB Building Height / Open Space Width	$H / W$	
	Building frequency along SN	Number of Buildings / Street length	$N_{\text{build}} / L_{\text{street}}$	
Network-Plot Relationship	Land ownership fragmentation along the street network	Number of Plots / Street length	$N_{\text{plot}} / L_{\text{street}}$	50
Site Morphology	Surface slope	High sloped surf. ( S>30%) / PB Surface	$\sum \text{Sloped Surf}_i / S_{\text{PB}}$	50
Network-Site Relationship	Street acclivity	Avg. arct(slope) along the street centerline	$E [\text{arct}(\text{slope})]$	/

Table 11 Streetscape indicators implemented through Proximity Band procedure.

<sup>73</sup> These indicators are the most important and easy to implement, which we developed at the beginning of the research activities (between 2016 and 2017, before the publication of Harvey et al. 2017 published in the same period). Moreover, during the last two years we implemented similar methodologies to other case studies such as Brussels, Marseille and Osaka in collaboration with colleagues of the University of Louvain-la-Neuve and The University of Tokyo. These collaborations gave us the opportunity to test and improve these indicators (at the origin of the MFA procedure, see next chapter) as discussed in Araldi et al. (2018); future publications will disclose the results and new outcomes on these specific case studies.



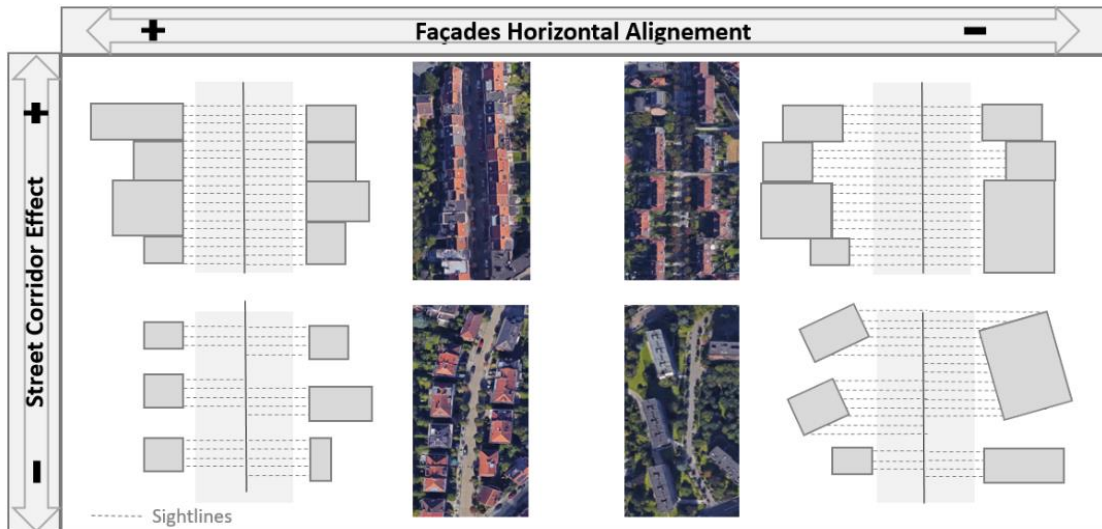


Figure 70 Example of combinations of Street Corridor Effect and Façade Misalignment.

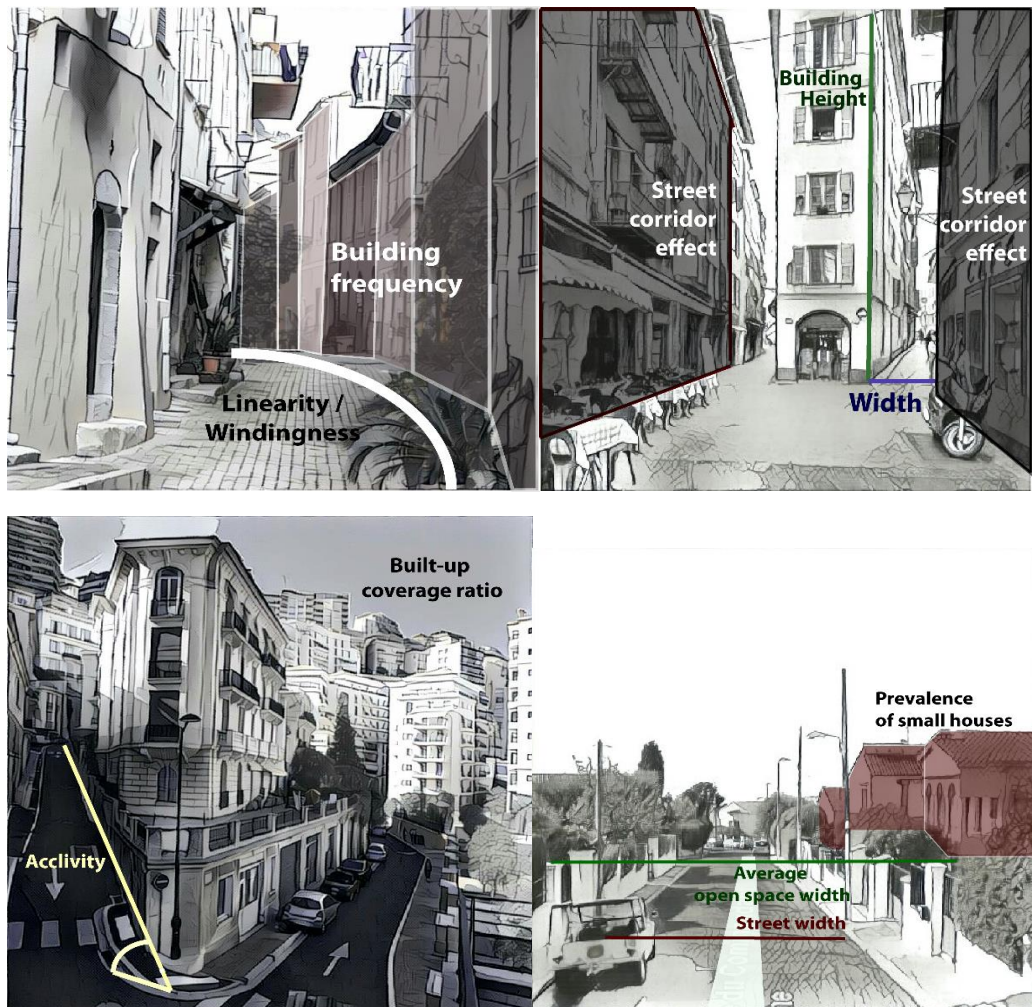


Figure 69 Skeletal streetscape morphological indicators.

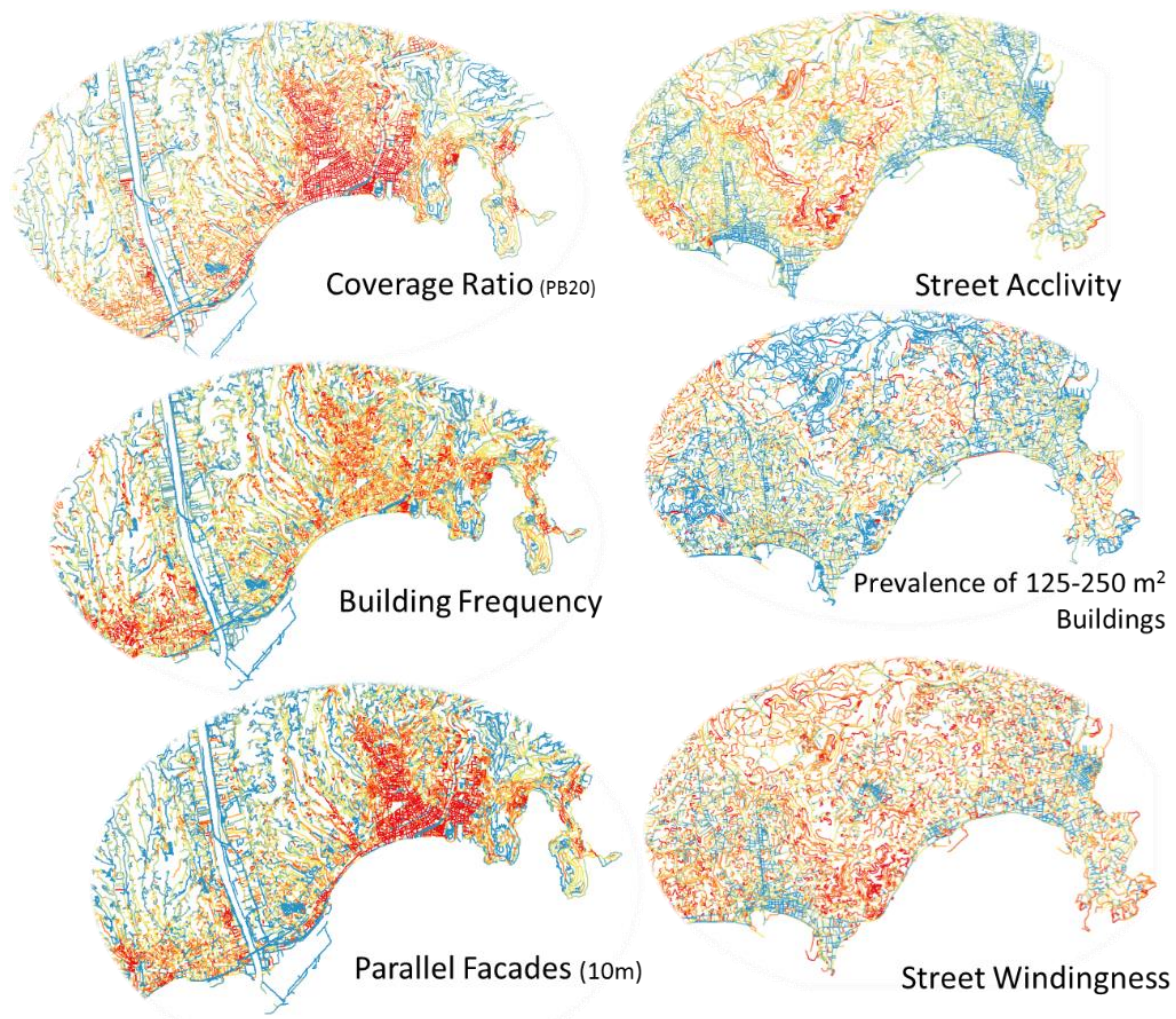


Figure 71 Projection in the geographical space of some skeletal streetscape morphological indicators.

## 9. Urban fabrics from the pedestrian perspective

In the two previous chapters, we identified street network configurational properties (Chapter 7) and streetscape compositional descriptors (Chapter 8). In urban studies, the relationship between these two sets of indicators and socioeconomic phenomena are traditionally investigated (along with pedestrian movement and human behaviour). Nonetheless, as previously discussed (Sections 2.3 and 6.2), when implementing these approaches to large urban settlements encompassing a wide heterogeneity of urban spaces, a contextual variable is required: the effect/influence of a given (street network or streetscape) indicator might be differently related to a phenomena depending on the surrounding urban fabric context.

As previously discussed in Chapter 6, we are looking for a new methodology producing an automated delimitation of urban fabric typologies:

- I. without any preconceived or exogenous partition of space or expert-based recognition of neighbourhoods;
- II. integrating the analysis of built forms as perceivable by pedestrians from the street;
- III. able to associate different compositional descriptors of the urban form elements (streets, blocks, parcels, sites) within an acknowledged theoretical framework;

IV. able to consider the spatial relative disposition of urban form elements.

Together with these objectives, some methodological constraints are included:

- V. network constrained distances to consider how built forms are experienced by a pedestrian exploring urban space;
- VI. minimal information focused on urban compositional indicators (excluding any additional information such as age of buildings, functional or qualitative characteristics) required to guarantee the reproducibility and adaptability to other case studies.

In this Chapter, we discuss a new approach for the identification of urban fabrics which might be able to respond to the different theoretical and technical requirements abovementioned.

**Multiple Fabric Assessment (MFA)** is a new methodological procedure for the identification of urban fabric typologies and the analysis of their spatial organisation. The three-step protocol of the traditional typo-morphological analysis is here renewed integrating geoprocessing approach and the combination of geostatistical procedures.

MFA can be summed up as follows:

- A. Definition of a street network-constrained partition of urban space and identification of street-based compositional morphometric descriptors (streetscape indicators);
- B. identification of spatial associations of indicator values on a network constrained space;
- C. Clustering of morphometric spatial patterns through Bayesian methods in order to identify and characterise typologies of urban fabrics.

## 9.1 Network Constrained Streetscape Patterns

The first step (**Step A**) of MFA procedure has been already developed in the previous chapter. Among the 29 streetscape descriptors, we focus here on the 21 obtained with the Proximity Bands approach. These are the three reasons behind this choice: firstly, this set of indicators proposes a synthetic description of the compositional characteristics of different elementary components of the urban form (streets, buildings, parcels, site morphology and their relative relationships). Secondly, these indicators focus on the volumetric characteristics of the streetscape (and not only on facade distribution) allowing one to grasp the notion of *perceived volumes* as discussed in Alexander (1993). Thirdly, their definitions are derived from the traditional urban morphometric descriptors of Marshall (2005), Bergauer Pont and Haupt (2010), etc. and applied to an urban street-base fragment (PBs), following the urban typomorphology tradition.

The implementation of classification procedures to this set of intrinsic indicators would describe, as in Harvey (2017), morphological typologies of streetscapes. In our view, the local value of an indicator for each spatial unit is not adapted to describe the wider spatial patterns that are typical of the urban fabric, the latter being characterised by the arrangement in space of urban components with consistent morphological values. Spatial continuity and

aggregation are key-factors for their significance, especially if we adopt the pedestrian point of view. In the traditional three-step classification procedure, spatial relations between adjoining spatial units are then overlooked and clustering of spatial units does not necessarily lead to the identification of urban fabrics.

In MFA, we propose the implementation of geostatistical procedures for the detection of patterns for each streetscape morphological descriptor, considering the spatial distribution of its values as observed by a pedestrian exploring the city. We are interested in assessing the spatial consistency of morphometric characteristics, or, in statistical terms, assess the intensity of the spatial autocorrelation of morphometric values.

Among the different statistical approaches for the investigation of the spatial autocorrelation, we focus on a subgroup of spatial statistic procedures derived from the Moran Indicator (Moran 1950).

Given a spatial partition P where a specific observed measure  $p$  (i.e. building coverage ratio) is associated to each spatial unit, Moran's I could be used to estimate the level of clustering of the spatial pattern given by the values  $p$  in the overall partition P. Moran's I coefficient is defined as:

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=i}^N w_{i,j}} \frac{\sum_{i=1}^N \sum_{j=i}^N w_{i,j} (p_i - \bar{p})(p_j - \bar{p})}{\sum_{i=1}^N (p_i - \bar{p})^2} \quad (9.1)$$

Where N is the number of regions of the spatial partition P,  $p_i$  and  $p_j$  are the observed measures in the spatial units  $i$  and  $j$ ,  $\bar{p}$  is the average of the values observed in all units (i.e. the average of the coverage ratios in every  $p$ ) and  $w_{i,j}$  is the weighting matrix between subareas  $i$  and  $j$  (traditionally, the rook/queen contiguity matrix of the partition P).

The resulting global Moran's I ranges between -1 and 1, corresponding to strong positive and negative spatial autocorrelation, while values close to zero indicate no spatial autocorrelation (Goodchild 1986). In other terms, -1 corresponds to a "chessboard" regular pattern, 0 to random distribution and 1 to a strong clustering of high and/or low values in specific sub-areas.

When assessing the spatial autocorrelation of a given variable, Moran's I represents a better solution compared to other approaches such as Geary's C or semi-variograms for one important reason: Moran's I is based on the mean value of a distribution (measuring differences from the average and not between neighbours) with a consequent lower sensitivity to outliers.

The original Moran's I was conceived at the global scale and where contiguities between units of the partition are observed on a Euclidean space. In the '90s, two major developments of the Moran's Index were proposed. On one side, Black (1991, 1992) focused on its extension to a network-constrained space; on the other, Anselin (1995) developed a local version of the Moran's I, and of other spatial autocorrelation indicators, named LISA-Local Indicators of Spatial Association.

In Anselin's local version of Moran's I, each partition  $p_i$  might be described by a local spatial autocorrelation value  $I_i$  as follows:

$$Local\ I_i = \frac{n(p_i - \bar{p}) \sum_j w_{ij} (p_j - \bar{p})}{\sum_i^n (p_i - \bar{p})^2} \quad (9.2)$$

Where, similarly to the 9.1,  $p_i$  is the observed measure in the unit  $i$ ,  $\bar{p}$  is the average of the values observed in the overall space study, and  $n$  is the number of the local surrounding units of  $i$  as defined by the contiguity matrix  $w_{ij}$ . Local Moran's I helps us in the identification of the positive and negative spatial autocorrelation, respectively localised around each unit.

As we might see in the 9.2, the *Local  $I_i$*  decomposes the Moran's index by isolating the importance of each individual unit in the overall definition of the measure of spatial autocorrelation. The main interest in this separation is to propose a proportionality measure between global index and local index; the average of the local indices is, therefore, equal to the overall Moran's I.

*Local  $I_i$*  might be differently interpreted in relation to the corresponding *Global I*: when no global spatial autocorrelation exists, a significant local I value would suggest a local spatial cluster around unit  $i$ . When global spatial autocorrelation is significant, the local I statistic can be also interpreted as a local instability (Yamada and Thill 2010).

A second innovative aspect of the LISA approach consists of the evaluation of the statistical significance of the *Local  $I_i$* : since its exact distribution is unknown, Anselin (1995) proposes a conditional randomisation approach where the target variable  $p_i$  is fixed, while the surrounding values are randomly permuted over the other regions, to draw a reliable inference about the local I statistic.

The projection of the *Local  $I_i$* s in the traditional Moran's Scatterplot (Fig.72) might help to visualize and understand the result of the LISA analysis: the slope of the regression line corresponds to the value of the Moran's I, describing the type and strength of the spatial autocorrelation in the overall study region. On the x-axis of the scatter plot is displayed the standardised value of the variable  $p_i$  while on the y-axis, the standardised spatially lag of the same variable is shown.

Each unit  $i$  of the space is plotted in the scatterplot. **Quadrant I** contains those units where high values of  $p_i$  are surrounded by high values (HH, in red). **Quadrant II**: low values of  $p_i$  surrounded by high values (HL, in light blue). **Quadrant III**: low values of  $p_i$  surrounded by low values (LL, in blue). **Quadrant IV**: high values of  $p_i$  surrounded by low values (HL, in pink). Finally, grey colours identify those regions with non-significant patterns where the values of  $p_i$  and the surrounding regions are close to the average of the overall study area. Two statistical conditions might be at the origin of the absence of significant local patterns: i)  $p_i$  are all close to the average  $\bar{p}$  (homogenous pattern of average values); ii) mixed values of  $p_i$  with a local average close to  $\bar{p}$  (heterogeneous pattern with *normal* distribution centred in  $\bar{p}$ ).

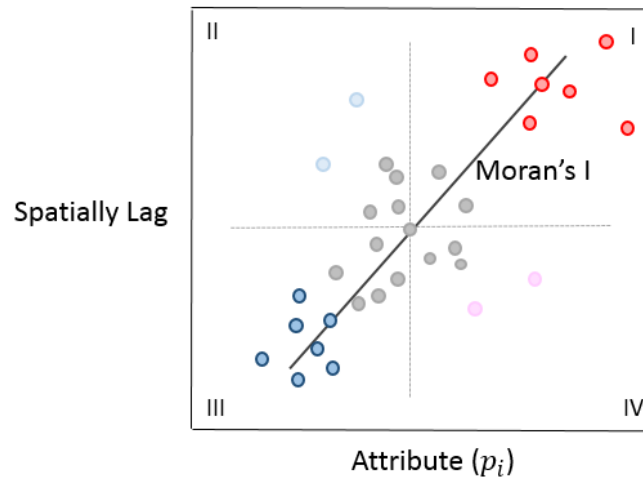


Figure 72 Moran's I Scatterplot.

More recently, Yamada and Thill (2007, 2010) extended the local version of the Moran's I to a network-constrained version (ILINCS I statistics in Local Indicator of Network-Constrained Clusters) where the matrix  $w_{ij}$  is defined by the network contiguity between street network segments and each  $p_i$  corresponds to the values observed along the links of a street network.

Despite the large application of both Global and Local Moran's I in several research fields, this statistical approach has so far been used relatively little in urban morphometric studies. Tsai (2005) implemented the Global Moran's I with the goal of quantitatively characterising the overall urban form of metropolitan cities, while the local version is implemented by Torrens (2008) and Serra (2018). These three studies implement Moran's I statistics on population distributions; to our knowledge, local Moran's I (both LISA and ILINCS) have not yet been implemented in urban morphometric or typo-morphology studies.

We first proposed the application of ILINCS in Araldi and Fusco (2016) to the study of streetscape morphometric descriptors. In this work we compared the application of LISA and ILINCS in the specific case of morphometric streetscape descriptors: firstly, we implemented the local Moran's I, considering the Euclidean contiguities given by the generalised Thiessen partition (Fig. 73.b) and, following that, considering the street network contiguities between Proximity Bands at two different radii (i.e. five steps in 73.c, three in 73.d). ILINCS are best fitted when analysing patterns from the pedestrian point of view: since the contiguity matrix is defined by the street network connections (as for the configurational analysis previously implemented), morphological values are therefore evaluated compared to the values of the surrounding connected street segments. This analysis highlights patterns of morphological features closer to what a pedestrian might observe when exploring the urban space.

When applying the ILINCS procedure to a morphological indicator such as, for example, the indicator of street corridor effect, four resulting categories HH, HL, LH, LL classify each street considering the neighbouring context. The overall local coverage ratio might be High (HH) or Low (LL); HL identify those streets with a small agglomeration of buildings along their edges within a sub-regions characterised by low values and, on the contrary, LH identifies those small open spaces immersed in high built up sub-regions (such as squares, urban gardens and parks

etc.). These two last examples might be observed in Fig.74c where LH (light blue) street segments are found at the interface between compact and semi-peripheral regions or in the proximity of the motorway line while HL (pink) is found in correspondence with a small concentration of buildings within a low-built neighbourhood. The not significant class might be associated both to those spaces where the coverage ratio is close to the mean value and to those heterogeneous areas characterised by a mix of streets with both high and low values. Image 75 shows some applications to the streetscape morphometric indicators previously measured.

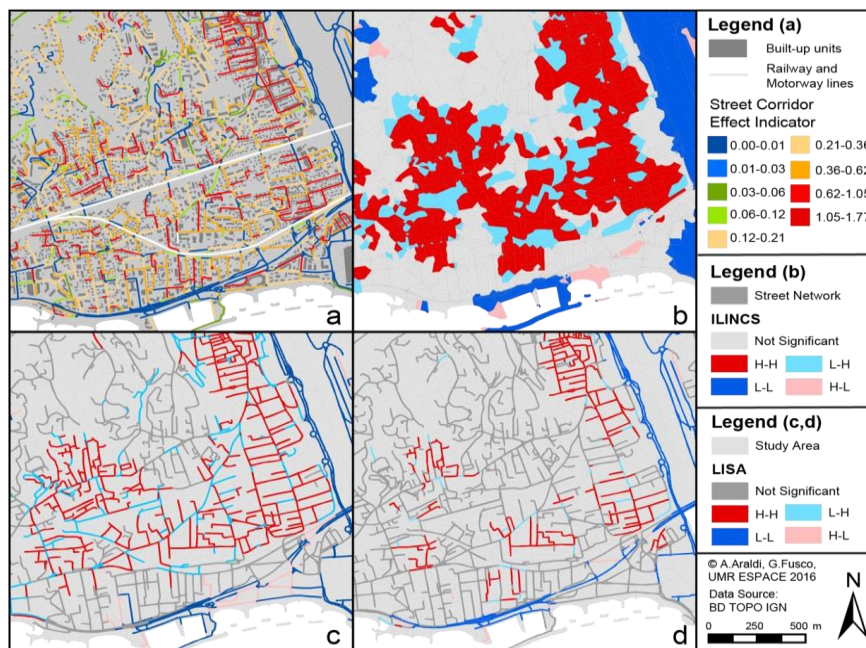


Figure 73 Outcomes from the implementation of LISA (on Generalized Thiessen Partition) and ILINCS (on Proximity Bands partitions) to morphological streetscape indicators (Corridor Effect). Zoom on the urban area between Cagnes-sur-Mer and Saint Laurent-du-Var. (Source: Araldi and Fusco 2016).

PBs represent a fine-grained partition of the space that it might be considered as a possible solution for the reduction of statistical biases (MAUP), as traditionally suggested in the scientific literature<sup>74</sup>. Table 11 shows how PBs are characterised by a lower size variability compared to other endogenous/exogenous partitions.

Nonetheless, statistical biases are still present when implementing ILINCS. As observed by the same Yamada and Thill (2010, p. 273): “Under the null hypothesis that there exists no spatial pattern in the configuration of  $x_i$  [...], the Moran  $I$  statistic asymptotically follows a normal distribution when either normality assumption or randomization assumption holds. The former means that the regional  $x_i$  values arise from a normal distribution with mean and variance that are constant over the study region. The latter means that all possible permutations of the regional  $x_i$  values are equally likely. Unfortunately, when the data are in the form of counts or ratios to the base distribution, either of the assumptions is hardly met because of anticipated spatial variations in the base distribution over the study area. Although several methods have been proposed to take into account such variations in the Moran  $I$  statistic [...], none is fully

<sup>74</sup> Another suggestion for the reduction of the MAUP would be the subdivision of longer street segments in subparts (i.e. defining a length max and splitting those segments at point=round()).

*satisfactory and Monte Carlo simulation is often carried out to estimate the null distribution of the I statistic.”*

In another study (Fusco and Araldi 2017), we focused on the analysis of the statistical distribution of morphological indicators addressing the problem of the appropriate corrections for the implementation of local Moran's I in the specific case of urban morphometric variables.

Streetscape morphological indicators developed in the previous section and used for the geoprocessing of urban fabrics are often rates, where the base population is the street length or the PBs partition (for instance, building coverage rates on PBs etc.)<sup>75</sup>. The spatial analysis of these indicators poses the question of the influence of the base population size on rate variability among the spatial units. If this variability is overlooked, LISA/LINCS analysis can produce spurious spatial patterns. Epidemiology has been one of the first and is still one of the most important fields of application of spatial clustering techniques (Jacquez et al., 1996a; 1996b).

Tackling the case of spatial clustering of rate values, Marshall (1991) and Assunção and Reis (1999) highlight that the size of the base population over which the rate is calculated has an impact on the variability of the rate. When dealing with spatial units having strong differences in population size, the variability of the prevalence rate of a given disease in large population units is lower than the variability measured in small population units. This is due to the limits of statistical estimators on small populations. When this dependence of variability on base population size is overlooked, spatial clustering techniques could detect spurious hot/cold spots, linked to the variability of estimations on small populations and not to spatial structure.

The same authors thus propose an empirical Bayesian correction taking into account the knowledge of the base population size to correct the a priori assumptions on the variability of the rates. This approach, which is now commonly accepted in epidemiology, considers that spatial unit definition (normally administrative units like municipalities in a given region) is independent of the phenomena under investigation (here disease prevalence). Epidemics do not define spatial units and the epidemiological phenomenon (the prevalence rate) can be modelled as a Poisson stochastic process whose outcome (the number of infected people) is directly proportional to the population exposed in the spatial unit. Under this assumption, Empirical Bayesian correction of Moran local I (EBI) is proposed by Assunção and Reis (1999).

LISA/LINCS clusters are aimed at verifying whether, on the one hand, the observed rates are the likely result of random outcomes of a spatially homogenous Poisson process with a unique rate or with different rates without spatial structure (null hypothesis) or, on the other hand, if they are the random outcome of a spatial structure of different underlying rates (alternative hypothesis). Bayesian corrections have a direct impact on the calculus of LISA/LINCS clusters: deviations from the study area average values tend to be more significant when they are calculated for units with large populations and less significant when they are calculated for units with small populations. Accordingly, the usual categories of High-High, Low-Low,

---

<sup>75</sup> Only exception: street connectivity and the street length measure itself.



High-Low and Low-High spatial patterns tend to be less significant when calculated on units having mainly smaller populations.

To consider the heteroscedasticity of morphological rates, we applied Empirical Bayesian Corrections (EBI, Assunção and Reis 1999) considering both the asymmetry and the base value of the ratios (length and PB). Each pattern has been evaluated with a Monte Carlo 99999 permutations and a p-value of 0.05<sup>76</sup>

In conclusion, we might observe how the implementation of ILINCS on PBs spatial partition presents several advantages: firstly, the combination of a fine-grained tessellation and statistical corrections allows to assess spatial patterns with lower statistical biases compared to other traditional approaches (using endogenous/exogenous partitions). Secondly, PBs partition also follows the suggestions of a behavioural-oriented partition of Zhang and Kukadia (2005), which is better suited for socioeconomic phenomena (such as the retail distribution in urban space). Thirdly, PBs assumes a deeper meaning in the urban design/morphology context. Size variability is not randomly distributed. Different subspaces might be described by a higher or lower street length: shorter segments in historic centres and longer in peripheral areas. When using the same PBs contiguity depth (i.e. three steps radius as in our case study), the resulting sub-region on which the Local's I is assessed will adapt to different morphological contexts depending on smaller or wider sub-spaces as a function of the intensity of the urban form characteristics.

For these reasons PBs might be thus considered as a good compromise, incorporating statistical robustness, a behavioural based definition of spatial units, and a spatial distribution of its size variability which finds a specific urban morphological meaning.

Through these network-constrained geostatistical analyses, we pass from a collection of morphological values for each spatial unit to statistically significant spatial patterns that correspond to hot- and cold-spots of given morphological characteristics, using the usual categories of LISA analysis (High-High, Low-Low, Low-High, High-Low and Not-Significant). These geostatistical classifications are the result of the analysis of the values of each PB, considering the local neighbourhood connected through the street network, with a topological depth of 3 units.

---

<sup>76</sup> Further studies might explore different significance values.

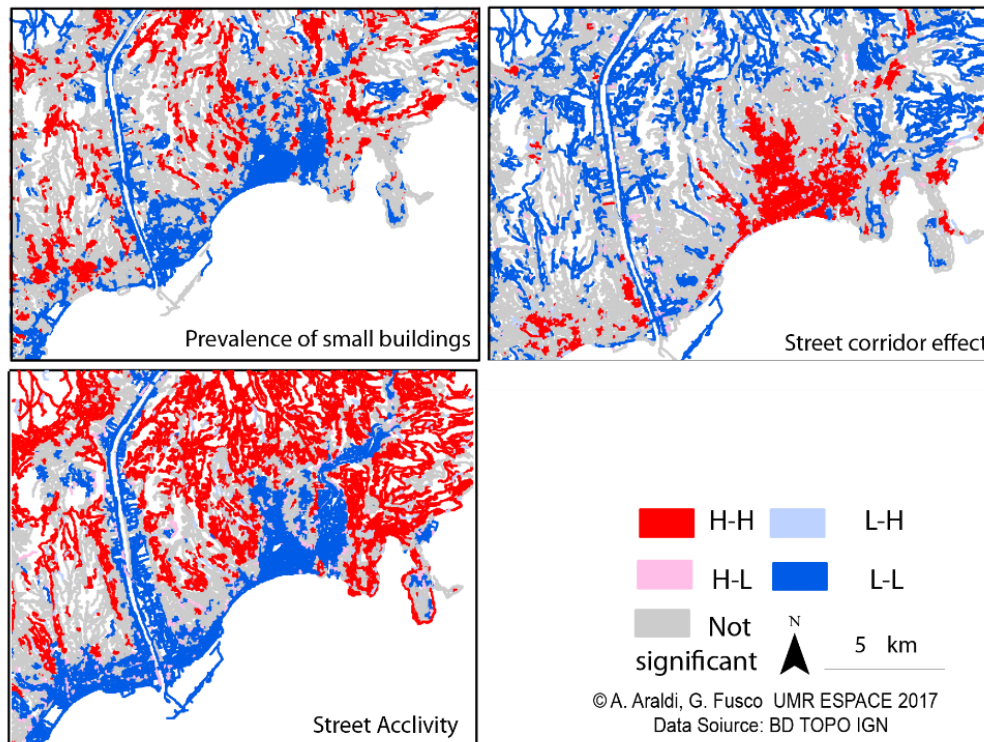


Figure 74 Patterns of morphological streetscape indicators, resulting from the application of network constrained local Moran's I statistics (ILINCs). (Source: Araldi and Fusco 2017).

## 9.2 Clustering morphological patterns

The geostatistical analysis produced 21 sets of significant spatial patterns. The next step consists of studying how they overlap to yield specific combinations. The urban fabric from the pedestrian point of view is precisely the perception of persistent co-occurrences of given morphological characteristics upon exploring urban space. At the same time, a few key characteristics could identify a given urban fabric even in the presence of a certain variability of other morphological aspects. For example, a specific compact and dense traditional urban fabric could be characterised by a consistently high building coverage ratio, street corridor effect, building contiguity and height/width ratio, while some heterogeneity of building types and network connectivity is seen. As previously discussed in Part 2 (for RFA procedure), clustering approaches aiming to achieve high intra-cluster homogeneity based on variance minimisation of the values of all the indicators (like k-means, SOM and geo-SOM) might not be well fitted to this task.

Here again, we propose a Naïf Bayesian approach for the clustering of streetscape morphological patterns (instead of row values), using Proximity Bands as spatial units (while for RFA we used points).

Bayesian clustering associates to each PB a set of probabilities of belonging to each Urban Fabric class  $UF_j$ , considering its description given by the set of attributes  $A_{1i}, A_{2i}, \dots, A_{ni}$  (morphological patterns).

$$p_{PBi}(UF_j | A_{1i}, A_{2i}, \dots, A_{ni}) = \frac{\prod_{k=1}^n p(A_{ki} | UF_j) p(UF_j)}{p(A_{1i}, A_{2i}, \dots, A_{ni})} \quad (9.3)$$

with the constraint

$$\sum_j p_{PBi}(UF_j) = 1 \quad (9.4)$$

As previously discussed, two main reasons led to the decision to use a Bayesian clustering approach: first, it allows the identification of groups of individuals sharing a few key characteristics among the set of attributes  $A_{1i}, A_{2i}, \dots, A_{ni}$  which could vary from one group to another. Secondly, Bayesian probabilities can address the uncertain assignment of observations: hybrid street segments could have non-null probabilities of belonging to different urban fabrics.

The combination of ILINCS and Bayesian clustering is a key feature of MFA. Instead of grouping streets based upon the similarity of their internal geometrical profile (row indicators on spatial units as in Harvey 2017), we will instead perform Bayesian clustering of categorical values (HH, LL, HL, LH, NS) which, in turn, summarize the statistical significance of morphological spatial patterns defined on a street network.

As previously discussed, HL and LH patterns identify those urban exceptions where Low and High values are located within an overall surrounding space of, relatively, Low and High values. Since our goal is to find the predominant urban pattern in a given region, before the classification procedure we regrouped HL with LL and LH with HH values. This further reduction of the variability allows us to simplify the classification which is finally based on 21 indicators, each one having three levels.<sup>77</sup>

Using 8 different random seeds, 1000-step random walks explored the solution space to obtain optimal Bayesian clustering of spatial units. The search constraints were a minimum cluster content of 1% of the spatial units, a minimum average probability of 0.9 for the assignment of units to each cluster and a maximum of 20 clusters. The optimal solution is obtained by minimising Akaike's Information Criterion (AIC, Akaike 1974), which combines the log-likelihood of the clustering solution given the data with penalisation for the increasing number of clusters.

Within our 8 random-walk searches, the 9-cluster solution was always found to be the optimum. A successive refining phase of the 9-cluster solution was carried out by imposing this fixed number of clusters to 8 new random walk searches. The results obtained were almost identical, with the best one showing a contingency table fit score of 59.4%.

---

<sup>77</sup> Future studies might implement different classification procedures applied to the original 5 categorical levels provided by ILINCS statistics, or with a combination of the continuous values of I'Moran. Other approaches might consist of the study of only LH and HL values in order to study the distribution of urban fractures and exceptions (for example, counting for each street the proportion of HL and LH on the total number of indicators).

## 9.3 Results: the nine urban fabrics of the French Riviera

### Overall classification outcomes

The projection of the 21 variables (resulting from geostatistical analyses) in mutual information space (Fig.75) shows little redundancy among morphological indicators. Variables are well-separated, with the exception of *Building Coverage Ratio*, *Building Contiguity*, *Height/Width Ratio* and *Building Frequency*. The latter also has the highest mutual information with the cluster variable (Table 13-left).

However, their redundancy does not play an essential role in the clustering results. When asked to select the ten collectively (rather than just individually) most informative variables for each spatial unit (as explained in Section 5.5), the Bayesian classifier identifies the indicators listed in Table 13-right, in descending order of frequency, and obtains a clustering accuracy of 93.14%. *Height/Width Ratio* and *Building Coverage Ratio* are thus normally coupled with *Street Acclivity*, *Building Frequency*, *Street Length* and the presence of *Small Buildings*, and identify the same urban fabric cluster obtained with all 21 indicators. Table 13-right also tells us that the last five morphological indicators are seldom needed to correctly predict the cluster variable; street-building relationship indicators are less frequently observed compared to building typologies and street indicators (windingness, linearity and slope).

Here again, we might observe the quantitative description of the role of each variable in the overall classification that might also be obtained for each class: this information is collected in Table 14 where each cell informs us about the targeted mutual information  $I$  (Indicator, Class). These values are disposed in the same descending order defined by the indicators' mutual information for the overall classification; this specific disposition allows us to easily identify which subgroup of indicator is the most correlated to each class.

The clustering results are relatively robust in terms of sensitivity to input data. The use of the 10 most informative variables on 10 test sets obtained in a 10-fold cross-validation schema produces a mean clustering accuracy of 92.62%, just half-a-point less than the one for the model learned on the whole dataset.

MFA on these spatial units identifies nine different families of urban fabrics on the French Riviera: when we observe the projection of the optimal classification in the mutual information space (Fig.75-right), these families are particularly well-defined and well separated. They also have a different weight in the study area varying between a minimum of 6% and a maximum of 20% of the total number of streets (values given in Fig.75 and Tab.13 are calculated on number of street segments and not on street segment length). The number of clusters was identified as optimal by Bayesian algorithms of cluster optimisation, with the only constraint being that each cluster should represent at least 1% of the spatial units in the study area, in order to avoid data overfitting. As they are arranged in Fig.75-right, the clusters identified make up a continuum gradient of decreasing urban character. In interpreting the cluster content, we will use the concept of "families" of urban fabrics. Clusters are often made

of slightly different urban fabrics, representing synchronic or diachronic variations of a basic type (Caniggia and Maffei 1979). Sometimes similar urban fabrics can also be produced in phases of urban history, which are well separated in time. They will nevertheless be included in the same family if they share the same key morphological characteristics defining the family.

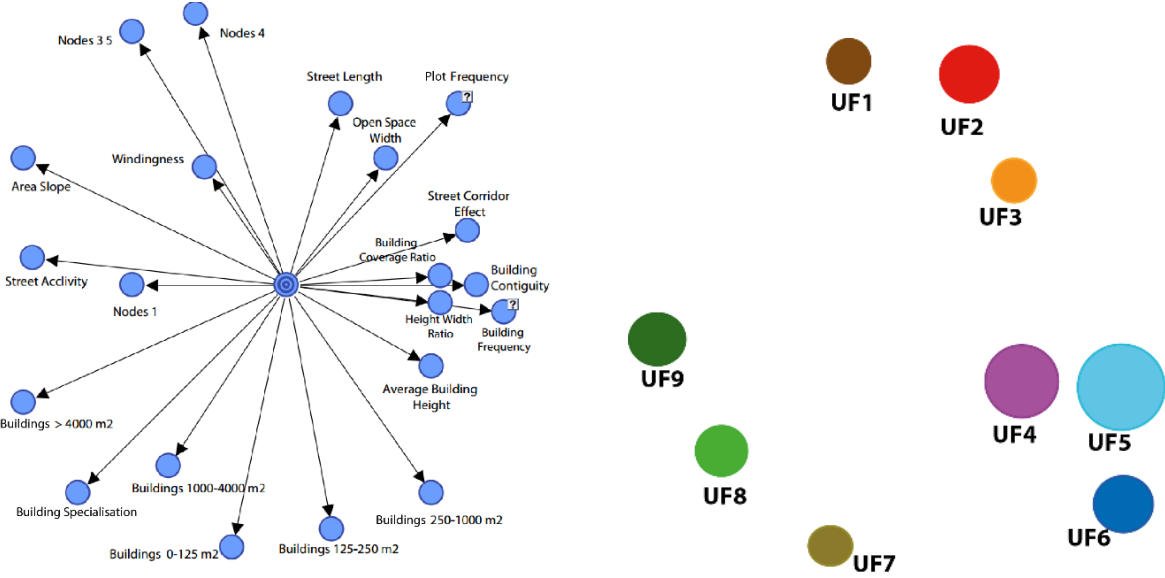


Figure 75 Projection in mutual information space of the MFA variables (left) and outcomes (right).

Indicator name	Mutual Information (%)	Frequency of Observation (%)	Indicator name
Height Width Ratio	37.68	100,00	Height Width Ratio
Building Coverage Ratio	37.46	97.13	Building Coverage Ratio
Building Frequency	33.88	79.63	Street Acclivity
Building Contiguity	29.27	75.61	Building Frequency
Street Corridor Effect	28.22	72.01	Street Length
Average Building Height	26.87	64.18	Buildings (0-125 m <sup>2</sup> )
Street Length	21.80	61.70	Street Windingness
Open Space Width	21.27	59.10	Buildings (1000-4000m <sup>2</sup> )
Plot Frequency	16.98	51.56	Average Building Height
Street Windingness	16.34	49.57	Buildings (125-250 m <sup>2</sup> )
Street Acclivity	15.83	47.91	Street Corridor Effect
Buildings (0-125 m <sup>2</sup> )	14.05	43.15	PB Slope
Buildings (125-250 m <sup>2</sup> )	12.83	41.13	Plot Frequency
Buildings (250-1000 m <sup>2</sup> )	12.46	38.94	Building Contiguity
Buildings (1000-4000m <sup>2</sup> )	11.31	36.62	Open Space Width
PB Slope	8.90	35.57	Buildings (250-1000 m <sup>2</sup> )
Nodes 1	7.11	11.77	Nodes 1
Building Specialisation	4.60	9.59	Building Specialisation
Nodes 4	3.90	4.31	Buildings (> 4000m <sup>2</sup> )
Buildings (> 4000m <sup>2</sup> )	2.38	3.19	Nodes 4
Nodes 3, 5+	1.82	0.00	Nodes 3, 5+

Table 13 MFA Indicator mutual information with cluster variable (left) and frequency of observation among the 10 collectively most informative variables (right.)

Indicator name	Global	Mutual Information [%]								
		UF1	UF2	UF3	UF4	UF5	UF6	UF7	UF8	UF9
Height Width Ratio	37.7	45.6	52.1	11.2	14.4	27.2	16.6	28.8	40.0	41.3
Building Coverage Ratio	37.5	42.0	49.4	22.1	20.0	32.1	9.6	12.9	39.8	41.1
Building Frequency	33.9	38.1	25.1	32.6	19.9	18.9	15.1	25.6	32.5	41.0
Building Contiguity	29.3	40.7	34.3	15.1	16.4	22.0	16.7	16.6	24.3	29.8
Street Corridor Effect	28.2	44.3	39.8	16.7	16.3	15.3	6.3	14.6	28.2	30.0
Average Building Height	26.9	22.3	46.5	9.2	2.9	19.2	17.3	25.1	32.2	23.1
Street Length	21.8	32.6	15.2	8.9	8.5	16.3	12.3	15.0	12.8	38.2
Open Space Width	21.2	48.0	33.5	9.1	7.8	7.5	5.4	8.4	19.7	26.8
Plot Frequency	17.0	44.2	6.0	9.2	2.5	12.3	5.4	0.4	15.2	35.0
Street Windingness	16.3	5.3	24.7	3.0	4.6	10.1	18.3	5.2	6.2	38.0
Street Acclivity	15.8	2.3	13.8	1.1	8.2	8.0	26.0	11.1	5.0	33.3
Buildings (0-125 m2)	14.1	30.7	10.7	9.8	8.1	11.1	12.1	17.5	5.8	0.3
Buildings (125-250 m2)	12.8	2.3	3.0	4.8	6.6	28.3	7.6	15.8	5.9	10.6
Buildings (250-1000 m2)	12.5	2.1	36.5	2.4	11.5	4.7	2.2	3.7	7.1	18.3
Buildings (1000-4000m2)	11.3	2.4	24.6	2.0	16.8	8.8	3.7	2.4	4.9	10.4
PB Slope	8.9	0.0	1.3	0.2	0.6	8.8	13.4	3.0	3.9	27.5
Nodes 1	7.1	7.2	12.1	1.7	0.6	11.4	6.8	9.4	0.7	0.2
Building Specialisation	4.6	2.6	4.4	0.5	7.9	5.3	1.4	3.4	2.5	2.7
Nodes 4	3.9	1.5	14.8	0.7	0.1	4.9	1.6	3.4	0.2	0.7
Buildings (> 4000m2)	2.4	1.6	1.4	0.9	3.0	1.8	1.6	4.5	0.8	1.6
Nodes 3, 5+	1.8	0.8	6.6	0.2	0.1	1.6	0.8	1.2	0.9	1.1

Table 14 MFA Mutual Information of each Urban Fabric Class.

### **Urban Fabric Classes description**

The first family **UF1** gathers 6.3% of the streets (2% of the total street length) and it represents the smallest class identified by MFA protocol. UF1 corresponds to the old constrained urban fabrics of town-houses making up the medieval and late-medieval villages of the French Riviera. They are dense (in the sense of high coverage ratio), made up of adjoining townhouses and small buildings with a high height/width ratio on short streets that can be more or less steep but always present a high street-corridor effect, and channels the pedestrian visual field while moving. Street networks are highly connective and irregular. Street acclivity and windingness are higher in the perched villages and in the oldest city-centres. The “bastide” towns of Valbonne, Mouans-Sartoux and Vallauris in the western Riviera are exceptions of street-network regularity but possess all the other defining characteristics of the family. Exceptionally, a few residential developments were produced in the last decades to replicate the urban fabrics of the old villages (Fig. 76-left).

The second family **UF2** accounts for 9.3% spatial units (4.8% of the total street length). It corresponds to the traditional urban fabrics with adjoining buildings having characterised the strong urban expansion of French Riviera cities on the flatlands from the early XVIII<sup>th</sup> century until the Second World War. Buildings and parcels are bigger than in the previous family and, above all, street networks are much more regular. The highest regularity characterises the

city-centre of Nice (but not its old town, mixing medieval urban fabric with early XVIII century expansions), particularly the neighbourhoods planned by the Consiglio d'Ornato (the local planning body) in the XIX<sup>th</sup> century (Graff 2000). Street corridor effect remains particularly important, as well as building frequency, because mega-buildings are only to be found as exceptional additions of the last decades. Sometimes, social housing projects of the last decades or new developments of high-rise buildings are also included in this family. The high-rise developments in Monaco are also exceptional within this family: their street network, although well-connected, is particularly irregular and the height/width ratio would make it a class on its own, had the 1% constraint not been applied in the cluster search (Fig. 76-centre).

The family of discontinuous and irregular urban fabrics mixing individual houses and bigger buildings **UF3** accounts for an additional 8.1% of the study area (5.3% of the total street length). Many of these urban fabrics started being developed in the late XIX<sup>th</sup> or early XX<sup>th</sup> century, in the form of small subdivisions or even individual self-construction. They have later undergone building infilling, additions and replacements, resulting sometimes in important transformative processes. They are thus characterised by less pronounced specificities than the two previous families, as they often mix characteristics of denser traditional urban fabrics, modern discontinuous urban fabrics, and even suburban fabrics (Fig. 76-right).

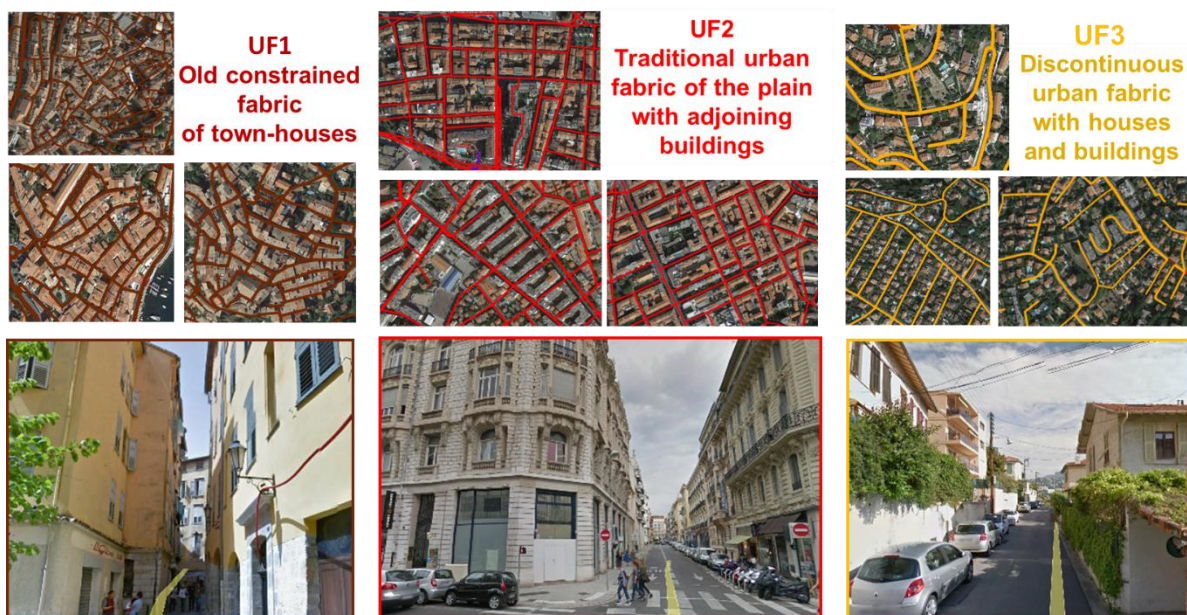


Figure 76 Aerial and Street View images of streets belonging to UF1, UF2 and UF3.  
(Source: Google Map and Google Street view 2017).

**UF4** represents the family of modern discontinuous urban fabrics with big and medium-sized buildings is particularly important on the French Riviera, gathering 15.5% of the study area (10.1% of the total street length). These urban fabrics are particularly the product of post-war modern developments, following the blueprint of CIAM precepts. The traditional relationship between parcel, streets and buildings is often lost (Panerai et al. 1997, Graff 2013). Street networks are also relatively irregular as meshed arterial roads are superposed to more local tree-like fragments. It is true that some developments like the late XIX<sup>th</sup> century garden city of belle-époque big buildings on the hill of Cimiez (Nice) are part of this family of urban fabrics, despite their pre-modern appearance. In these neighbourhoods, the discontinuity between

big buildings is indeed a main departure from more traditional urban fabrics and is correlative to the project of inserting buildings in large, wooded parcels (Fig. 77-left).

The next two families, **UF5** and **UF6** are made up of two different strains of suburban fabrics. On the French Riviera, suburbs lack the regularity and monotonous patterns typical of North-America, but also of British and other French residential urban peripheries. Land ownership fragmentation, topographic constraints and the importance of self-construction produced landscapes of particularly irregular and heterogeneous suburban fabrics over the last 50 years. Irregular, tree-like, winding street networks with numerous cul-de-sacs and prevalence of individual houses with low coverage ratios are the main characteristics of all suburban fabrics in the study area. A difference can nevertheless be made: on the one hand, UF5 the suburban residential fabrics on the hills or in the plain are particularly important in the western section of the French Riviera, where street acclivity and area slopes are low and houses tend to be bigger villas. On the other hand, the constrained suburban fabrics of more spontaneous areas UF6 occur especially in the eastern section of the French Riviera, where steep winding roads are the norm and houses tend to be smaller. The prestigious capes (Cape d’Antibes, Cap Ferrat, Cap Martin) are strongly characterised by the first form of suburban fabrics, with exceptionally big villas and swimming-pools. Old villas and mansions from the late XIX and early XX<sup>th</sup> century are here mingled with more modern villas, sharing few architectural and stylistic features, but contributing both to the same spatial logics of the suburban fabric. These two families of suburban fabrics make up almost a third of the street-segments in the study area: UF5 and UF6 gather, respectively, the 16% and the 15% of the street segments (20% and 12% of the overall street network length) (Fig. 77).



Figure 77 Aerial and Street View images of streets belonging to UF4, UF5 and UF6. (Source: Google Map and Google Street view 2017).

**UF7** represents the family of connective artificial fabrics with sparse specialised big buildings; it is perhaps the most peculiar urban form of the last 40 years of metropolitan development. Its quantitative importance is not huge, occupying only 7.7% of street segments (4.8% of the overall street network length), but it is present in almost all peripheral areas in the interstices between the other families of urban fabrics. Artificial connective fabrics are usually to be found around main arterial roads, highway interchanges and roundabouts. The biggest buildings of the metropolitan area (shopping malls, factories, hospitals, air terminals,



warehouses, hangars) are built along these street segments, but, at a distance, that is not at the scale of pedestrian movement and perception. Most of the technological park of Sophia-Antipolis, north of Antibes, is characterised by this emerging metropolitan fabric. Functional developments around ports and marinas also belong to this family. Of course, street-corridor effect is inexistent in these areas, and coverage ratio is low, but, unlike suburban fabric, houses and other small buildings are almost non-existent and cul-de-sacs are rare: the street network is highly connective in order to ease vehicular traffic in pedestrian unfriendly environments (Fig. 78).

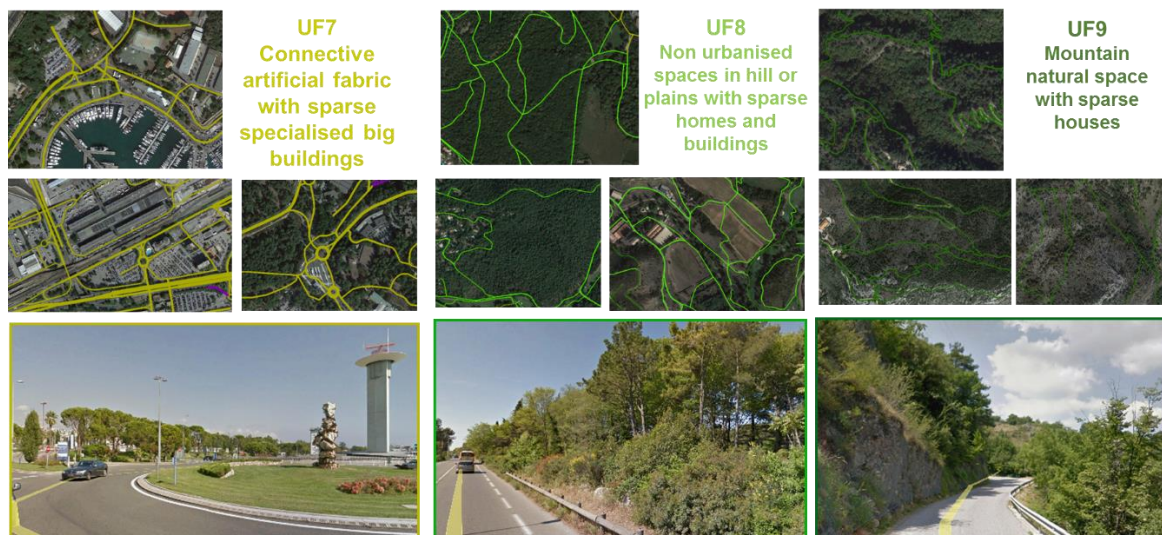


Figure 78 Aerial and Street View images of streets belonging to UF7, UF8 and UF9.  
(Source: Google Map and Google Street view 2017).

The last two families, **UF7** and **UF8**, characterise the less developed areas of the metropolitan space. Most of them are made of natural land and, more seldom, agricultural land. Buildings are nevertheless always present here and there, and the roads and trails (more than the streets) serving these areas are still a channel for pedestrian movement and visual perception. Like for suburban fabrics, a difference can be made between non-urbanized space in the hills and in the plains UF7, where landscapes are more open and smoother, and both buildings and individual houses can punctuate the landscape, and natural space in the mountains UF8, with long, winding and steep street segments, high terrain slopes, rare small houses, and sometimes panoramic visual fields. Once again, the former is more present in the western section of the Riviera, and the latter are more frequent in the eastern section, where the Pre-alps almost tumble into the sea. Respectively 11.3% and 11.2% street segments in the metropolitan area are characterized by UF7 and UF8; two families of essentially non-urban fabrics. If calculated in street length, their share would be significantly higher, as the longest street segments are to be found here (14.5 and 26.6% of the overall street network length). We will also remark that metropolitan parks with organised networks of long pathways are to be found in these two families, whether in the plain or in the mountains. On the contrary, most urban public gardens, because of their higher connectivity, shorter segments, and presence of technical facilities, are much more often classified in the family of the connective, artificial fabrics, even if, for once, they are much more pedestrian-friendly than a car-friendly environment (Fig. 78).

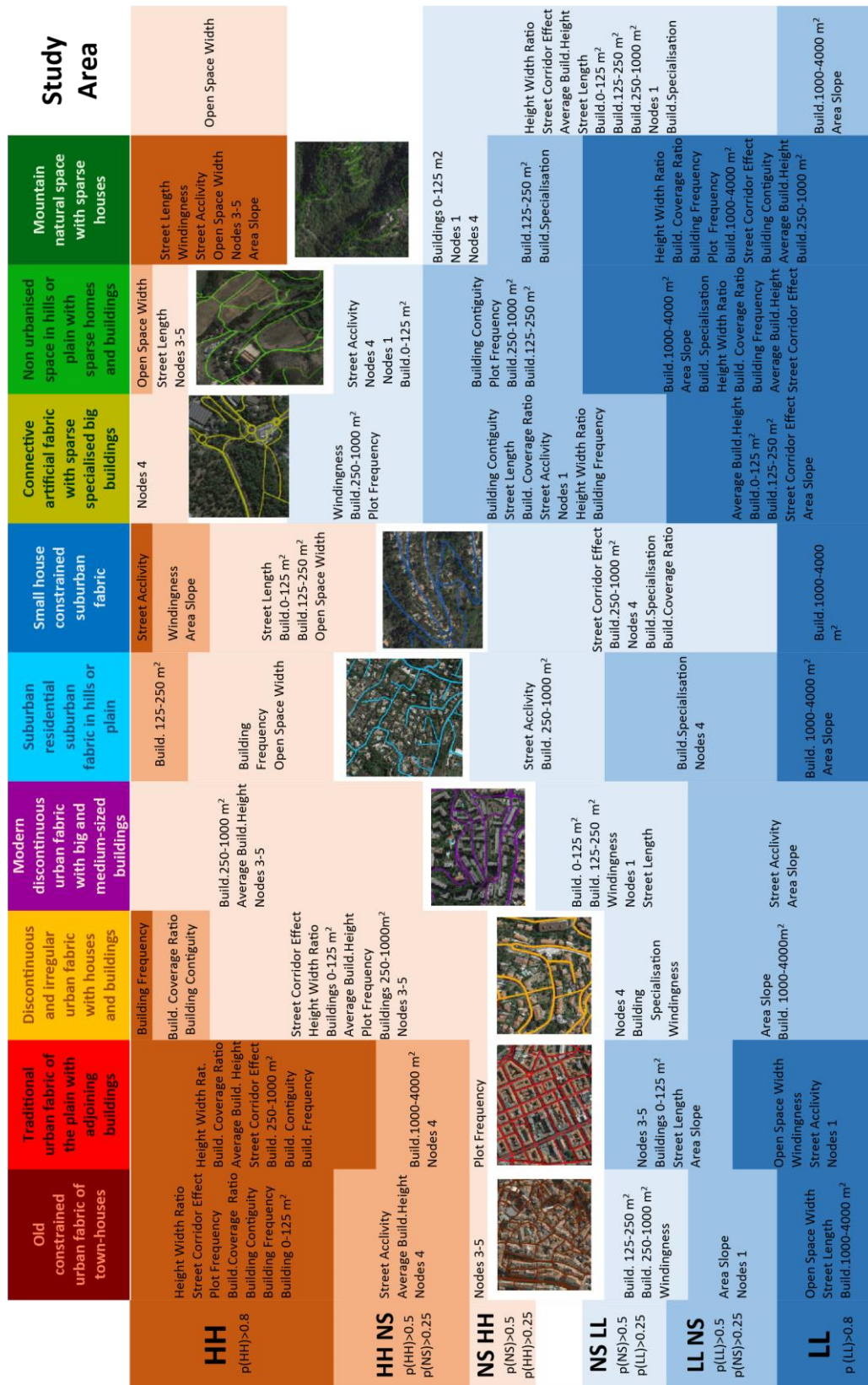


Figure 79 Characterising urban fabrics in the French Riviera through probabilities of morphological patterns; synthetic descriptions of MFA outcomes (Source: Araldi and Fusco 2019).

The quantitative probabilistic profiles provided by the Bayesian Classification for each class might be found in Annex.2.

In Figure 79, we gather the 9 Urban Fabric families, allowing for the comparison of the spatial patterns underlying the definition of each class. In Figure 80, we projected in the geographic space the probabilities  $p_{PBi}(UF)$  of the street segments  $i$  to be associated with each Urban Fabric class. Darker/lighter colours correspond to higher/lower probabilities, respectively; grey segments correspond to a null probability.

As we might observe from these nine projections, the probabilistic content of each class describes a different spatial organisation. UF1 is characterised by small nucleated centres, scattered almost equally in the urban area under study. UF2 delimitate well defined central areas, prevalently located along the coastline; their differences in size reveal the importance level of each major city within the metropolitan area. From the largest to the smallest, we find, along the coast, the conurbation of Nice in the central part, Cannes and Antibes on the western side, while Monaco and Menton rest on the Eastern side. In the inland section, UF2 identify small-sized cities such as Grasse and Vence.

While UF1 and UF2 are spatially well defined (probability values close to 0 or 1), the other UF classes show a higher variability in the probability of association resulting from a higher heterogeneity and mix of urban morphological characteristics. UF3 and UF4 are found in the close peripheries surrounding the centres defined by the previous classes; UF5 and 6 describe the suburbs at higher distances from the core centres. UF5 and UF6 show a peculiar spatial organisation: while the former is found in small or large pockets on hilly or planar areas of the western inland section, the latter shows a more linear-shaped distribution, following level lines in the eastern area. The same difference between eastern and eastern sectors is found again when observing the different distribution of UF8 and UF9.

The separation between these two sectors is made even more evident when observing UF7: the Var Valley is mainly associated to a connective artificial fabric; similarly, the technopark of Sophia Antipolis is highlighted at the Eastern inland sector. These two spaces, as further detailed in Text-Box 3 in this section, represent the two main spaces which have known a strong process of urban transformation in the last decades, which is still undergoing.

So far we have described each class identified by the MFA, individually through the probabilistic parameters provided by the Bayesian Classification of the MFA procedure. In what follows, we will focus on how these urban fabrics are spatially organised in the French Riviera.

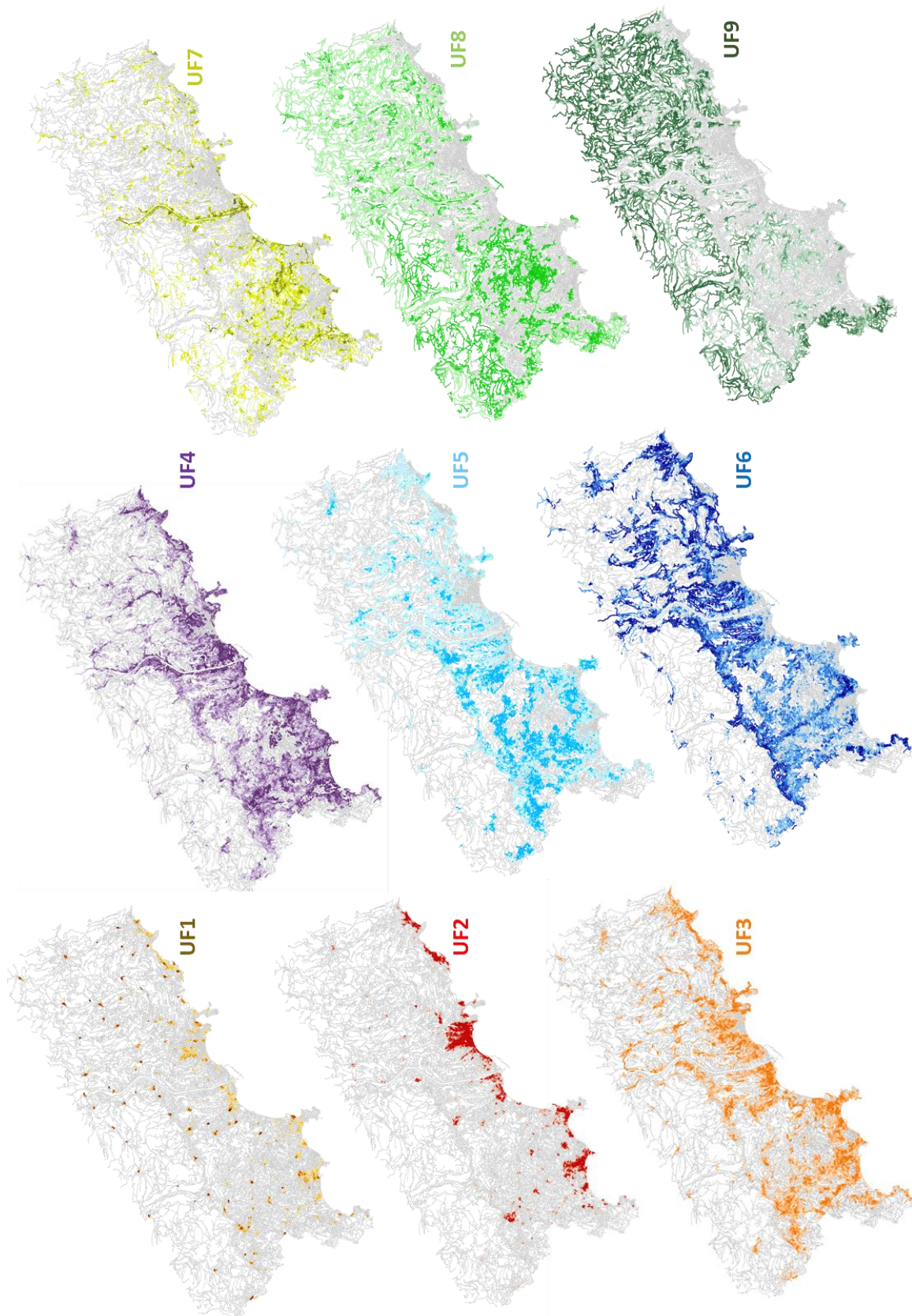


Figure 80 Projection in the geographical space of the probabilistic values associated with each Urban Fabric class. Darker-lighter colours correspond to higher-lower probabilities; grey segments correspond to a null probability.

## **From the urban fabric to morphological regions: the spatial organization of the French Riviera**

When the most probable class associated to each street segment is projected in geographic space, we are able to observe well defined, large spatial patterns of connected spatial units making up peculiar morphological regions in the metropolitan area. A few sectors mix two families of similar urban fabrics (UF 1 and 2, UF 3 and 4, UF 5 and 6, UF 8 and 9). Areas of a fine-grained patchwork of several urban fabrics are particularly rare.

Figure 82-83 maps the western section of the French Riviera, around the cities of Antibes, Cannes, and Grasse and Figure 84-85 the central section, around the city of Nice. The old city centres of Antibes, Nice, and Cannes (including Grasse in the hinterland) are clearly distinguishable in brown colour (UF1), surrounded by the traditional urban fabric of the late XIX<sup>th</sup> and early XX<sup>th</sup> (UF2). More peripheral urban sectors belong to families UF3 and UF4, often in opposite sectors of urban development, whereas vast suburban areas (mainly family UF5 around Antibes and Cannes, family UF5 west of Nice, but UF6 north and east of Nice) border the outskirts of the main cities.

The traditional fabrics of smaller villages in the close hinterland are identifiable as islands within suburbia. We can also observe the importance of suburban fabrics of family UF6 in the capes. Finally, connective artificial fabrics correspond to the urban fringe belts west of Nice and north of Antibes, but also to the technical space around ports and airports, and to the core of the technological park of Sophia Antipolis (where a few modern developments have also recreated the morphological characteristics of traditional urban fabrics).

The results found for the city of Nice, showing the substantial continuity of traditional urban fabrics of the plain from the most central area towards the North and the East, as well as the relative importance of modern discontinuous urban fabrics in the western section of the city, in the north and on the hill of Cimiez, are in agreement with more qualitative expert-based analysis carried out by Graff (2013).

The striking similarities in the spatial organization of morphological regions around the coastal cities of Cannes, Antibes, Cagnes-sur-Mer, Nice, Monaco, and Menton brought us to the conclusion that, despite the local specificities (like the planned urban grid of Nice or the constrained high-rise developments of Monaco), the cities of the French Riviera show a common genotype of spatial organization. This is represented through the simplified diagram of Figure 81.

The archetypical Riviera city has a concentric double core of inner old city and traditional pre-modern urban fabric (Levy 1999). Discontinuous irregular fabrics and modern discontinuous fabrics develop on well-defined and opposed sectors around them. Further out, different forms of suburbs were added in the last fifty years, with specific forms on the prestigious capes. A further belt of more or less topographically constrained natural space lies around the suburbs. However, the spatial organisation of the metropolitan area cannot be understood without taking into consideration the three following aspects:

- Old villages are scattered in the immediate hinterland. When they are small cities (like Grasse or Vence), they show the same juxtaposition of urban forms already found for the coastal cities.
- Connective artificial fabrics are a very peculiar new urban form of the metropolitan development of the last fifty years. They are to be found in the interstices of the urban periphery and suburbs, along main transportation axis and sometimes on the very coastal line, connecting ports, airports, shopping malls, and other large urban facilities.
- This scheme is a module repeated several times within the metropolitan area around each major coastal city. The spatial organisation of the French Riviera metropolitan area is thus a doubly self-similar structure: there is repetition of the same module, on a similar scale, along the coast, and repetition of miniature version of it in the close hinterland.

A few exceptional areas depart from this scheme. This is namely the case for the Sophia-Antipolis technological park (see Box), north of Antibes. Here, a large matrix of artificial, connective urban fabrics is surrounded by natural areas. Within it, we find enclaves of modern discontinuous urban fabrics and of more traditional urban fabrics with adjoining buildings. The morphological region of the technological park is thus more a mosaic of these three families than a homogeneous area characterised by the presence of a single type of urban fabric. Finally, apart from a few noticeable exceptions, the metropolitan area is structured by an urban scheme mixing concentric belts and sectors, interwoven with new connective fabrics, that repeats itself several times from east to west, with some self-similarity at different scales.



Figure 81 Archetypal spatial organisation of morphological regions in French Riviera cities. (Source: Fusco and Araldi 2017)

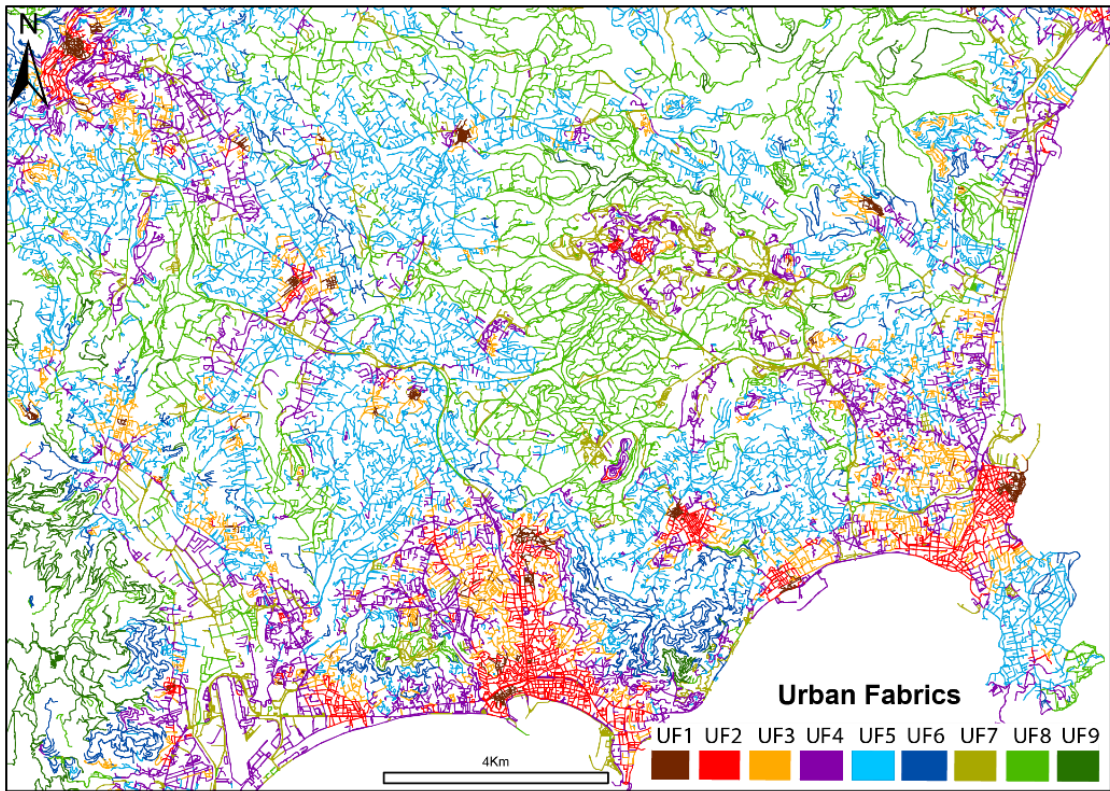


Figure 82 Projection in the geographical space of the nine Urban Fabric: Grasse, Cannes and Antibes (Source: Araldi and Fusco 2019).

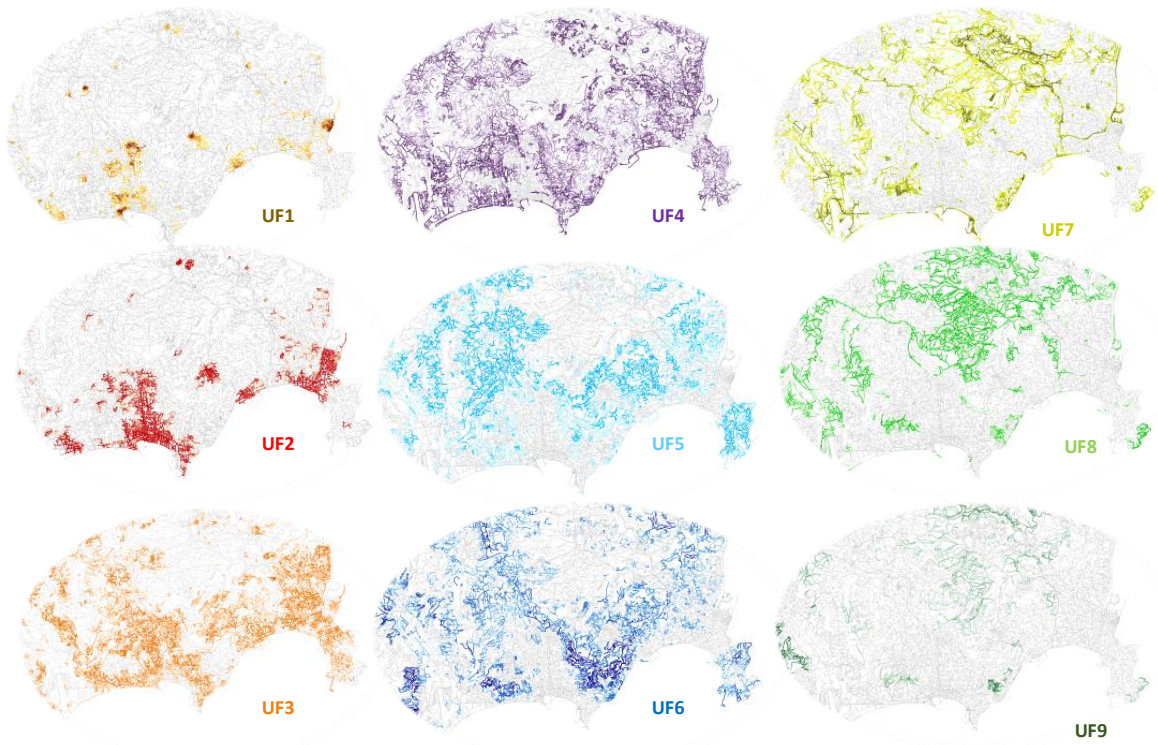


Figure 83 Projection in the geographical space of the probabilistic MFA outcomes: Cannes and Antibes.

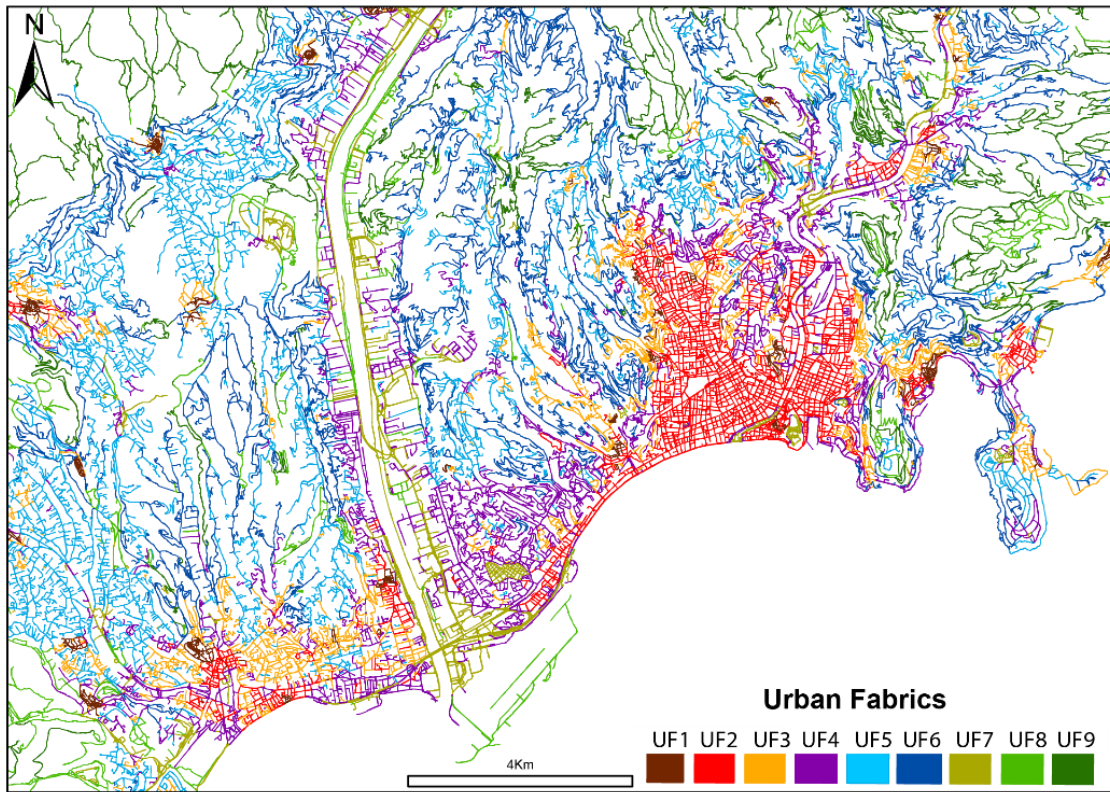


Figure 84 Projection in the geographical space of the nine Urban Fabric: Nice metropolitan area (Source: Araldi and Fusco 2019).

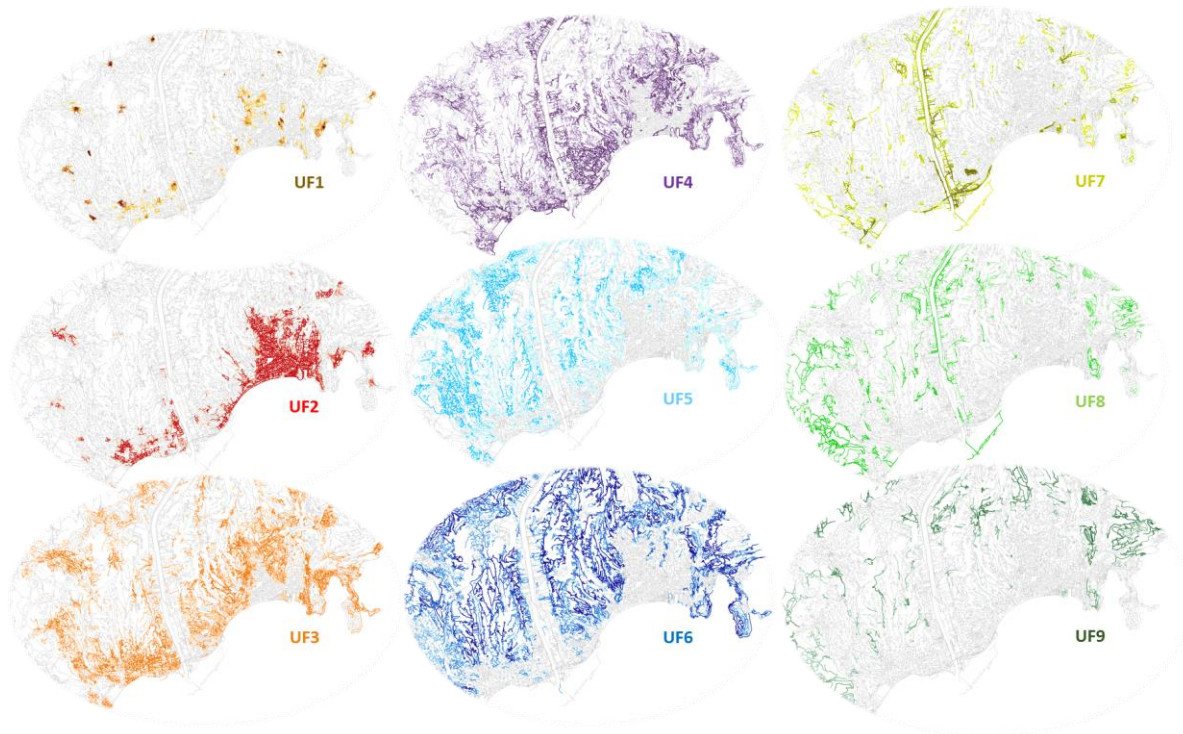


Figure 85 Projection in the geographical space of the probabilistic MFA outcomes: Nice.



### **TEXT-BOX. 3: Focus on some urban spaces of the French Riviera to better understand MFA results**

In this box of text, we focus on the description of some subspaces of our study area: Nice, Antibes, and their surrounding peripheral spaces provide good examples for understanding the urban form characteristics of French Riviera cities. The reason for looking at these spaces is due to their peculiar urban morphological and historical evolution. In this box, we try to provide the reader with the essential notions about the main historical phases of the urban expansion. Comparing this historical information, as well as old maps to MFA outcomes, the reader can appreciate the strengths and limitations of the analytical procedures developed in this dissertation.

#### **Nice**

*“for a geographer, it is a real privilege to study Nice, prosperous and charming city, whose name alone is a caress”* (Blanchard 1960, p.151).

The first evidence of a human presence in the French Riviera dates back to 400 thousand years ago when the sea level was higher and the actual Castle Hill which we now find in the city centre, was a small island. The first edifications close to the actual old centre date to 350 a.c. by the Greeks (The Phocian population came from a close region to the actual Izmir, Turkey; they founded the City of Marseille, also known as *La Phocéenne*, and successively the city of Nikaia), who named this first settlement Nikaia (Nissa, Nice, Nizza). Three centuries later, the Romans came to the region and established a second centre *Cemeneleum* (the actual hill of Cimiez) with its baths and theatre. This city was established on the *Via Julia* connecting Genoa with Aix-en-Provence.

Between the years 400 and 1081, Nice was at the edges of the Holy Roman Empire; throughout all the medieval age, continuous invasions caused by its marginal position pushed the local population to the feet of the Castle Hill, abandoning the historical centre of Cemeneleum.

Starting from the 1388, the County of Nice is annexed to the Duchy of Savoy. The latter<sup>78</sup> played an important role in the urban history of Nice. The Baroque period, starting from the late 16<sup>th</sup> century in Italy, blossomed throughout all of Europe. In 1563, the dukes of Savoy moved the Capital of the Duchy to Turin, transforming the latter from a small city to a modern capital recognised at an international level. The new plan of the city became one of the early urban models in history (previous to the modern movement of Cerda, etc.); the long period without wars between 1562 and 1675 allowed the dukes of Savoy to focus their attention to urbanism, a symbol of the power of a new-born regional power in the European political panorama. In 1773, an architectural congress was charged with the responsibility to plan an expansion of the city of Turin, with new rules about proportions of streets, facades, heights, building continuity, and even aesthetical elements such as cornices. The city became an object of art, which has to contribute to the image of the city, influencing its international

---

<sup>78</sup> Originally extended from actual region of Savoy in France over to the Aosta Valley and Piedmont regions in Italy.

importance. Turin became a laboratory where for the first time planning, architecture and urban design converged, and building extensions were the perfect occasion to test new urban forms (Pollak 1988, *“From Castrum to Capital”*). The model of the capital was successively exported to the other cities of the Duchy of Savoy, among which was the city of Nice.

Nice had a strategic role, being the only access to the Mediterranean Sea for the Duchy, who had become a Kingdom at the beginning of the XVIII<sup>th</sup> century with the annexation of the island of Sardinia.

A first extension of the southern side of the historical centre was planned between the end of the XVI<sup>th</sup> century and XVIII<sup>th</sup>, and after the demolition of the defensive walls in 1706. A second extension came about in 1718 with the construction of the Opera and the pre-modern urban forms (wider streets with aligned buildings). Within this period, the historical centre of Nice was enriched by Palaces and Churches<sup>79</sup>. For three centuries the architectural and urban activities were concentrated in the small triangle between the Paillon, the hill and the seaside, and they gave origin to those urban forms which remained unchanged until today (Fig.86).



Figure 86 Plan de Nice, anonyme, 1790. Archives municipales de Nice.

Between 1830 and 1860, following the model of Turin, the king Charles-Albert of Sardinia established in Nice the “Consiglio d’Ornato” (from the Latin *ornare*, to decorate or adorn with, to embellish with); this public institution, was charged of the organization and the extension of the city on the right river of the Paillon. Compared to the other French cities, the urban planning and the abundance of urban projects give this city its accolade of “urban exception” (Graff 2000). Inspired by the model of the capital, the plans of the architects Scoffier and Vernier (between 1825 and 1843) set up the guidelines for a first extension between the seafront and the train station (in 1860), and successively replicated to the eastern side of the city, behind the Port Lympia.

<sup>79</sup> Today, only in the 24 ha of the old town of Nice, we count one cathedral, seven chapels and five churches.

The Consiglio d'Ornato aimed for the urban project to help with the organization of the city around public spaces, recognising that places, gardens and squares had an important role for the life of the community, to its commerce, and to everyday exchanges. The *piazza* represents a pivotal element that aided the articulation between different parts of the city; from this urban space, a main street is traced pointing to natural, architectural or symbolic features and represents the main axe coordinating the street grid of the surrounding spaces. Two examples might be found in the main squares named (today) Place Massena and Place Garibaldi, representing the pivotal element between the old town and the new extensions: from the former, the Avenue de la Gare (the present-day Avenue Jean Médecin) is traced and it represents the reference for the extension of the city on the right side of the Paillon. From the latter, the Royal Avenue (present-day Avenue de la Republique) organises the east side of the city (Riquier Quarter). Similarly, the port Lympia (1748) represents an “*aquatic square*” geometrical reference for the surrounding grid.

Differently from other rigid grids of the Roman and Spanish foundation cities, the grid is not rigid, but rather accommodated and adapted to the natural morphology and local landscape. The regular grids follow the coastline and the Paillon river. Grids organised from different squares are perfectly joined (Fig.87).



Figure 87 The city of Nice (grey) and the new plan of the Consiglio d'Ornato (orange) using squares with a key role for the street grid organisation: A) Place Massena; B) Place Garibaldi; C) Port Lympia. (Source Graff 2013)

With the annexation of Nice by France in 1860, the Consiglio d'Ornato is dismissed; while the aesthetical notions and criteria are corrected into a more functional version, the outlines of the original plan have been maintained. The street grid is replicated with less regularity in future extensions until the second world war in the three directions allowed by the planar morphological conditions (dashed polygons in Fig.88).

The period between the annexation to France and the first World War correspond to *La Belle Époque*, where technological innovations in transport and construction<sup>80</sup> helped the city of Nice become one of the largest cities of France. In fifty years, the French Riviera becomes an international touristic destination of the aristocracy of all Europe headed by the royal families of Europe (Queen Victoria and the Romanov family in the first place), but also intellectuals and artists<sup>81</sup>.

While the city-centre is filled with high standard hotels, the surrounding hills see the development of residential areas: the slopes of the Mont Boron, Cimiez hill, and the western hills (Fabron and Saint-Helene) become the preferred site of mansions (*Maison de Maître*) and villas<sup>82</sup> surrounded by large gardens and parks reflecting the same eclecticism of the interior design of these structures (Durandy 1960, p.24). As previously introduced (Section 3.1) from this moment, the economy of the region is transformed: from a place of passage between northern Europe and Italy (with agricultural production of olive oil since the Greek period), tourism becomes the main economical source. And even the architectural and urbanism is a function of this economic activity: the paradigmatic example of the Excelsior Hotel Regina built for Queen Victoria on the top of Cimiez Hill has stimulated the alignment of the road network serving the hill along the Boulevard de Cimiez. A sort of competition between the cities of the French Riviera, particularly between Cannes and Nice, encouraged the architecture of the city to become more attractive for aristocracy and wealthy clients.

The First World War (1914-1918) and the Russian Revolution (1917) ended the extravagance of the Belle Époque, as well as most of the European aristocracy. Nonetheless, between '20s and the '30s, the socioeconomic changes in the market economy and the prosperous evolution ("*Les années folles*") allowed for a more democratic tourism, shown by the construction of new buildings in the Art Deco style, and here again Nice became a laboratory for the modern art competing at an international level with architects of Miami and Casablanca. From the urban point of view, the large properties and mansions are sold (in particular those of the bankrupted Russian aristocrats who came back to the French Riviera, this time escaping from the Russian Revolution, Ellis 1988). The large parks of the villas were divided into smaller lots and filled with family houses and buildings for the growing middle class.

The second after-war period was characterised by a strong economic expansion of several countries such as Germany, Italy, Sweden, and Japan. In France the period between the 1945 and 1973 is known under the name of "*The Glorious Thirties*" (*Les Trente Glorieuses*). The economic development of this period was accompanied by increasing leisure time, higher incomes, and more efficient transport means which transformed every citizen to a potential customer of the tourism industry. Nice, as more generally the French Riviera, became the

---

<sup>80</sup> Train line Marseille Saint-Charles and Ventimiglia have been realized between 1858-1872. This enabled the connection both to the French national rail system and with the Italian network, important for the Russian tourists.

<sup>81</sup> In 1869, the historian Brunel wrote, "The city of Nice became, with Naples, Rome and Florence one of the *salon d'hiver* of the European high society. The Russian colonies is full of life and entertainment: [] they organise balls, soirées, garden-parties with open air dances, in Nice as well as in the villas bought or rented in the surrounding areas".

<sup>82</sup> The Excelsior Regina Palace 1879, Le palais Pauline 1904, Le Chateaux de l'Anglais 1856, Le Chateaux de Valrose 1867 (now head office of the University Cote d'Azur), are just some of the most magnificent and extravagant examples of the "*follies de la belle époque*" in Nice of the European aristocracy.

attraction of mass tourism. From the landscape point of view, the traditional effect of urban sprawling, which characterises every large city in North America and in Europe in the same period, was heightened in the French Riviera. Real estate agencies, as well as private households, materialised the desire for freedom with a wide development of secondary homes. Costumers with a lower interest in the creativity and exuberance of the architectural characteristics (and quality) of the previous decades originated a strong development of residences easily recognisable by the prominent presence of terraces, balconies and loggias as an answer to the collective imaginary of the holiday home. High standing blocks of flats on the hills further from the centres or private homes sprawling from of the centre were built. As we might observe from Fig.88, at the end of the '50s, almost one hundred years after the annexation of Nice to France, the city has known a large expansion filling the planar spaces towards the north with a dense and ordered urban fabric (light orange). On the surrounding hills we find a patchwork of areas with dense and disordered fabric (red), low density ordered fabric (blue, such as real estate developments), and low-density areas with disordered buildings (yellow). This map represents an essential process the city underwent in its urban evolution, which shows the urban morphological regions previously to the arrival of the big transport infrastructure (Fig.88).

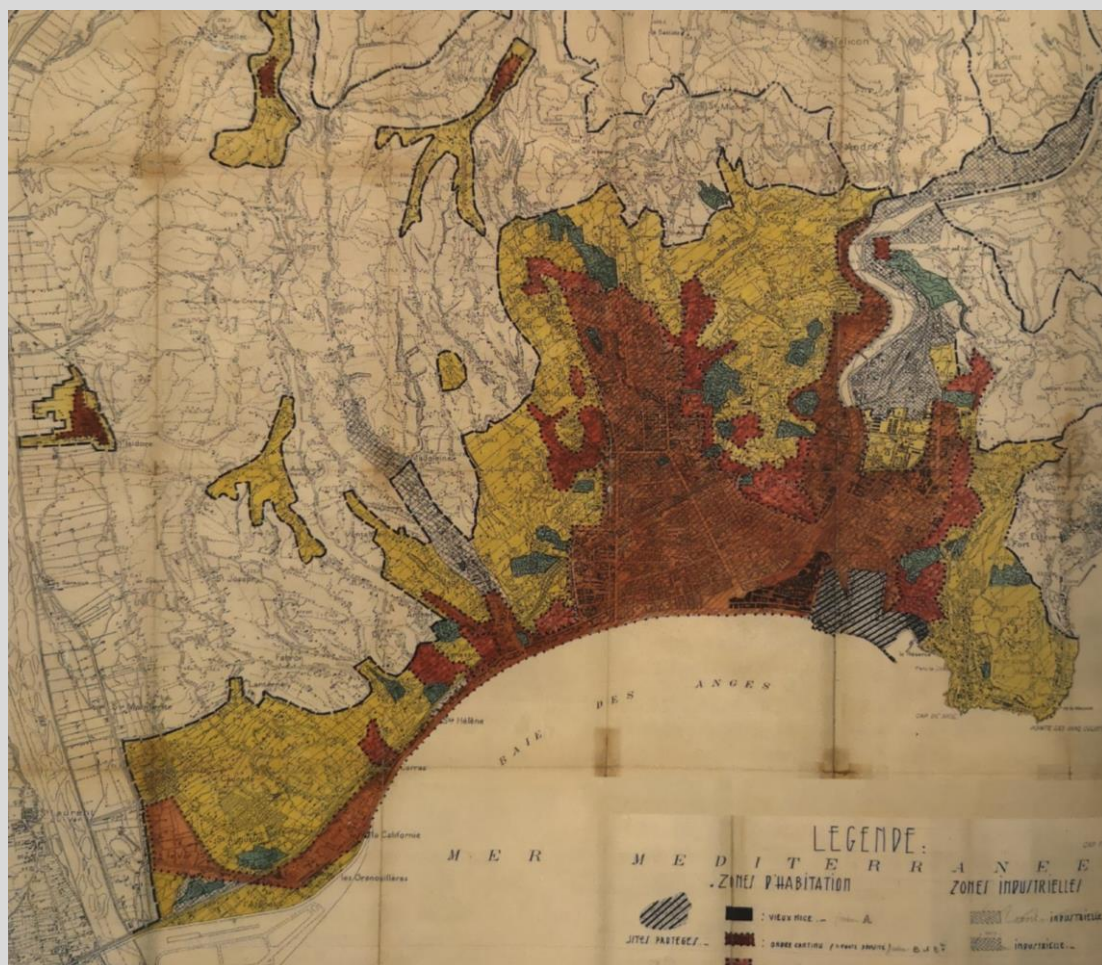


Figure 88 "Aublet" Zoning Plan of Nice 1950. Archives municipales de Nice.

During the '60s, the A8 motorway connected the coastal cities between the Italian border and Aix-en-Provence; in Nice, the A8 passes through the hills and valleys of the northern sector of the municipality and once reaching the Var, it descends the valley towards the coastline. Within the city, the train line is doubled by an urban highway, which allows one to drive across the city from its eastern quarter of Riquier, while, on the western side, the urban highway joins the motorway behind the Airport.

In the following decades after the opening of these transport infrastructures, while the compact ordered urban fabric was little modified (with a few extensions in the northeastern sector towards the motorway, and in the south-western sector towards the airport), conversely, the inland boundary has been continuously rising towards the north with an urban fabric made by small individual homes (especially during the '90s, Graff 2013).

In the meantime, between the '50s and the '60s, another phenomenon affected the city of Nice: the independence of the north African countries engendered the exodus of the Pied-Noirs<sup>83</sup> (French settlers). Between 1954 and 1962, 100 thousand people moved to the department of the Maritime Alps, of which 48 thousand came to Nice. A demographic explosion (+20%) to which the city of Nice was not prepared, despite the program of rent-controlled housing (HLM *Habitation à Loyer Modéré*), which, since the '50s, was providing housing solutions for the baby boomers of the second after war. The exceeding demand of housing produced squatters/shantytowns in two peripheral areas today known as *Les Moulins* (behind the Airport) and *L'Ariane* (at the northeastern extremity of the city). The intervention of the public administration created two social housing programs isolated from the rest of the city, both in terms of morphological continuity and in terms of transport accessibility, with, consequently, several social issues related to the phenomena of ghettoisation.

Throughout the last two decades, several projects have been mainly focusing on the activity of infilling and the requalification of the compact urban fabric, especially in the more semi-peripheral compact areas. Nice is a continuously *changing city* ("*la ville en mutation*", Graff 2013) and a large number of projects have been modifying the quality of its urban public spaces: from the three tramway lines (the first one in 2003-2007, the second and third in 2015-2019), the linear park (*Coulée Verte*) on the Paillon River cover, as well as the areas behind two train stations (the Liberation Quartier and the Saint-Jean d'Angély university campus are rehabilitating an industrial area). The suppression of the bus lines, which will be substituted in the future by tramway lines, are also going to be transformed into bicycle lanes and green corridors. While all these projects have been requalifying neighbourhoods of the compact central fabric, a new urban extension has also been planned in the last two decades along the Var valley and is today taking shape.

---

<sup>83</sup> Algerians, Moroccans and Tunisians with European (mostly French) origins.

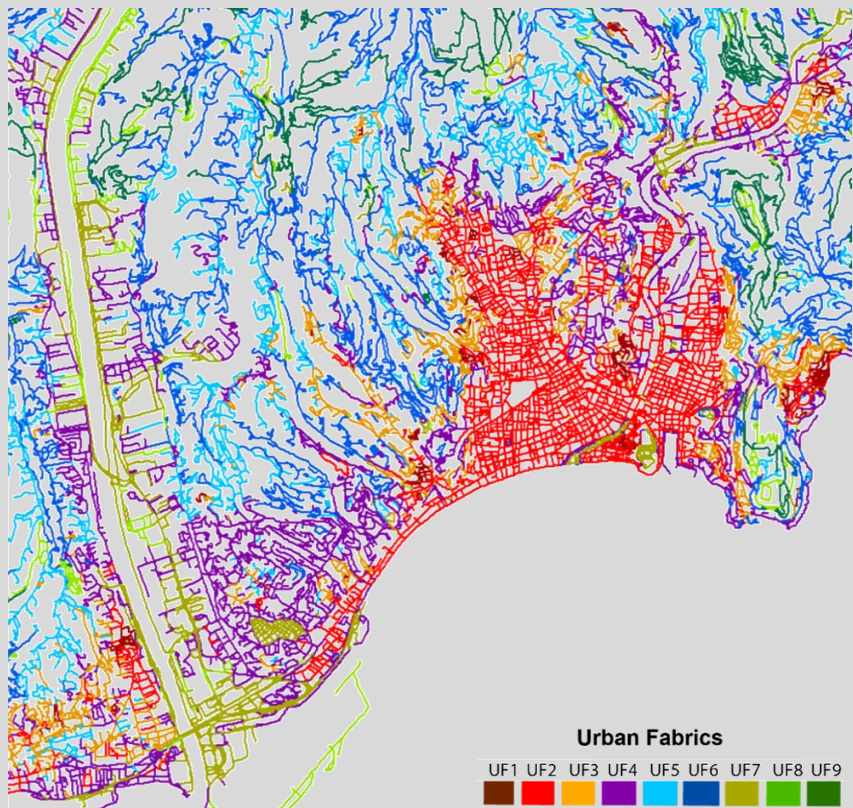
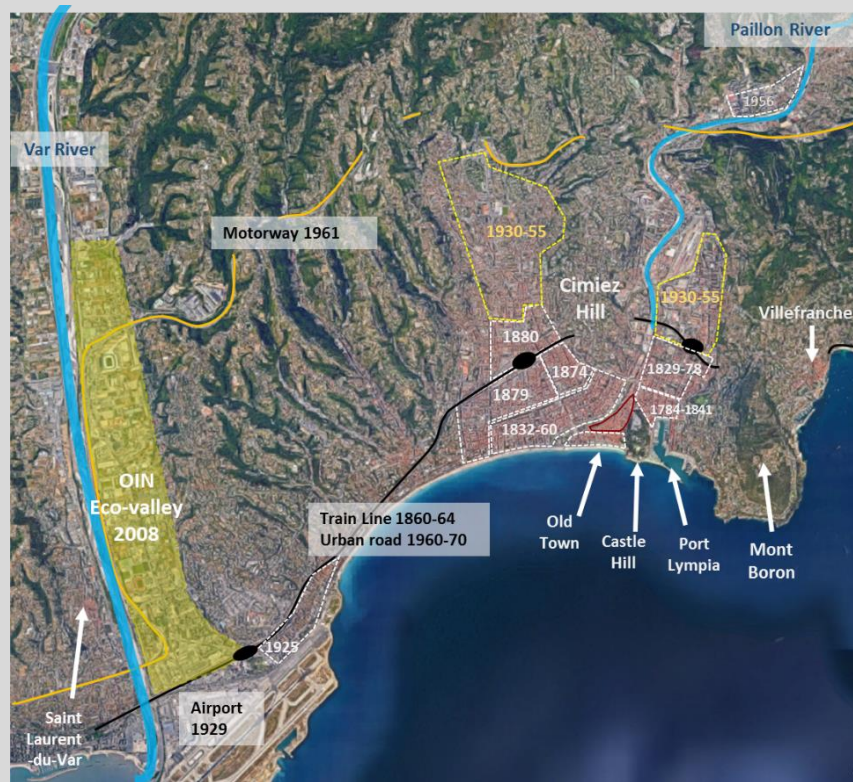


Figure 89 The city of Nice. On the top: the historic development of the urban system and its main transport infrastructures. On the bottom: MFA outcomes

## The Var Valley

At the west side of the municipality of Nice, the Var valley separates the Nice metropolitan area from its eastern part (Saint Laurent du Var and Cagnes-sur-Mer). As already presented in the study area description (Chapter 3), the Var river represented the natural border dividing kingdoms and states of the Italian peninsula from the French kingdom and nation for centuries (Fig.90). The Var river appears in several manuscripts throughout history for its unpredictable nature, and for a long period it was seen as a dangerous natural obstacle to overcome<sup>84</sup>; crossing the Var symbolised, for the religious and intellectual, the entrance into the Italian peninsula. Indeed, the French Riviera, before becoming an inevitable intermediary destination of the aristocracy for the Grand Tour, the same route was travelled by religious pilgrims heading to the capital of the Roman Church. For this reason, along the French Riviera, we might find several monasteries and villages founded by religious orders. The city of Saint-Laurent-du-Var is one of these centres: an original hospital/rest home to assist pilgrims crossing the Var. With time, the village has been developed around the original religious site, attracting the attention of business families from the close Ligurian region for its strategic position.



Figure 90 The Var Valley in 1780. Archives départementales-Conseil général des Alpes-Maritimes

In 1860, Saint-Laurent-du-Var lost its status of frontier-city, the wooden bridge was removed, and the two banks of the river were finally connected with a stone-and-metal bridge. “This village passed from a border guard-city to a vegetating and sleeping city immersed into silence” (Durandy 1912, p 18). The Var Valley was no longer a battlefield and the area became an agricultural land for Mediterranean cultures (to grow Mediterranean *essences* such as lemons and olives).

---

<sup>84</sup> As reported since the roman period from Julio Cesar, defined “fou et gueux” (mad and beggar) from the French general Vauban. Durandy (1960) describes the Var river as « capricieux, sournois, de colère prompt et meurtrière, s’enflant démesurément avec les pluies d’orages et cachant des crevasses pleines de trahison ».



Along the 20th century, the right bank of the river, under the administrative control of the municipality of Saint-Laurent-du-Var, transformed its landscape with a slow and confused growth following the example of the larger cities of the French Riviera. The outcomes of the MFA might help us in identifying the historical centre of Saint-Laurent-du-Var (UF1), its urban planned development towards the south (UF2) and the pavillonnaire fabric (UF3); modern urban fabric of large buildings (UF4) has conquered the more hilly spaces at the north of the municipality and along the waterfront.

When considering the left bank of the river, a different urban development has been defined by its marginal position within the municipality of Nice. While the top of the surrounding hills has been pervaded by modern urban fabric and small houses of the more recent sprawl, in the lower part of the valley are functional large infrastructure, such as the the airport at the beginning of the century, one of the first shopping malls of France (Cap3000), and the national roadway A8 in the '60s. This transport infrastructure gave the Var Valley important advantages for vehicle-oriented accessibility. With the increasing use of the car in the second half of the 20th-century, this space became the receptacle of other functional built forms, such as regional administrative offices and store agglomerations of large surfaces, such as car sellers, garden appliances, furniture and home appliances. More recently, other large buildings have been taking place, such as the concert hall Nikaia (2001), the Allianz Riviera Stadium (2016), etc. The transformations of the second half of the century make of the Var Valley a sort of fringe belt with a gradual process of infilling, transforming marginal spaces from agricultural to functional, and, more recently, to a mixed-use urban fabric. This process has been recently, and, impressively, accelerated by the constitution in 2008 of a National Interest Project (OIN-“Operation d’Interet National”), the Eco-Valley project.

The aim of the OIN Eco-Valley is to transform the Var Valley into a new urban model: from a divisional element to a space of reunification of the Nice metropolitan area with its eastern municipalities; from a mono-functional and crossing space to a new multifunctional centrality; from an environmental threat to an environmental resource and opportunity for sustainable development.

The analysis of the project and its future urban development goes beyond the goal of this dissertation<sup>85</sup>; moreover, the ESPACE laboratory has been recently awarded a research grant from IDEX UCA JEDI to study the first ten years of evolution of the southern part of the Var Valley (the TransMetroMed project is particularly focussed on the analysis of the physical form of the Meridia neighbourhood, Fig.91).

What is more interesting within the context of this chapter is the possibility to use the MFA procedure to assess and compare the urban morphological characteristics of the Var Valley previously to the profound transformations under development, providing analytical observations for future diachronic comparative studies.

In Figure 91, we might see the urban landscape of a fragment of the Var Valley in 2016; this space was characterised by a connective urban fabric (UF7) with an overall up-side-down “T”

---

<sup>85</sup> If the reader would like to have a more detailed description of the project and the master plan of the Eco-Valley, all official information is available at the following link: <http://www.ecovallee-cotedazur.com/>

shape constituted by essential transport infrastructures such as the train line, the airport, and the roadway. Modern urban fabric of large buildings (UF4) is found in small fragments grafted on the connective backbone; each small concentration gathers administrative, commercial, or private offices functions. The residential function is found in the site of Les Moulins, one the City of Nice realised in 1975 to alleviate the housing crisis of the '70s.

In Figure 91(bottom-right) is the new masterplan for the central section of the Var valley: as we might see, at the end of the implementation, a finer meshed grid will connect these spaces with a higher built density and more organised urban spaces. The urban fabric of this fragment will assume morphological characteristics more similar to the planned areas from the XIX century of central Nice. Several questions arise about the future of this area both from the morphological and functioning point of view.

For the goal of this chapter, we demonstrated how MFA might help in the identification and description of the urban morphological characteristics, even when considering a more complex and composite urban space such as the Var Valley.

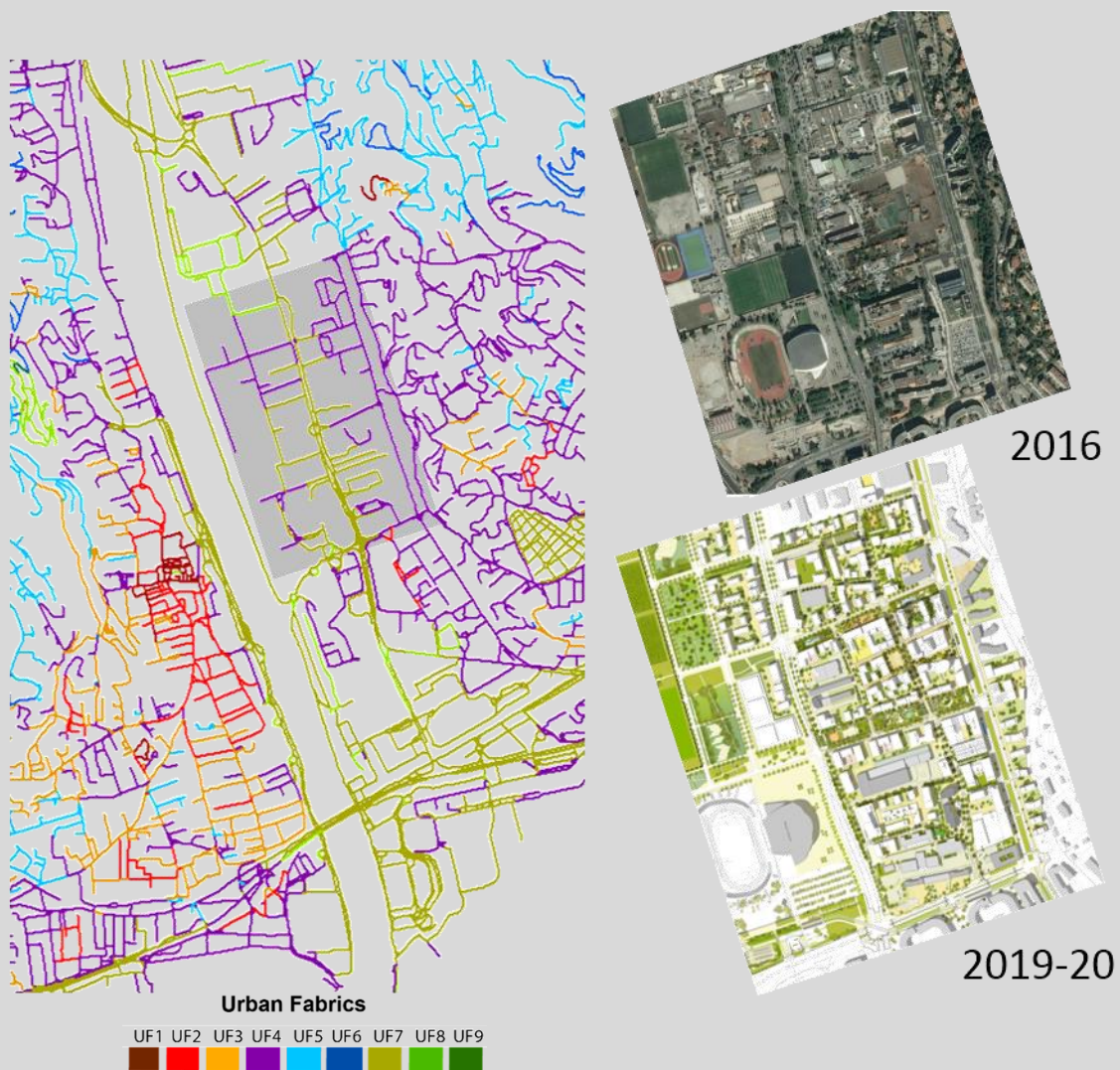


Figure 91 On the left: MFA outcomes in the southern section of the Var Valley; on the right: the current(2016) landscape of the neighbourhood of Meridia and its master plan within the OIN Eco-Valley program.

## **Antibes**

In the IV century B.C. the city of Antibes was founded by Phocaeen colonisers (coming from the nearby city of Marseille founded two centuries earlier). Originally named Antipolis (meaning the city at the opposite side [of Nikaïa]), the location of this first settlement was founded for its strategic morphological condition. It was a natural cape (Cap d'Antibes) surrounding and protecting a natural bay from winds and sea currents, as well as a natural port site, providing favourable natural conditions for the development of maritime trade and a strategic intermediary position on the route between Greece and Marseille. In the following centuries, a process of Romanisation of the city reinforced its commercial importance, developing the original settlement. Romans (coming from the inland, See Chapter 3) expanded the first urban nucleus to its western side with a second line of walls. Fig 93.a shows the original extent of the Greek and the following Roman street grids based on two different orientations.

The fall of the Roman Empire marked the beginning of a long period of instability, which characterised the whole region (Chapter 7); from this moment, Antibes saw a gradual transformation from commercial city to military fortress. In order to defend from the continuous Saracen attacks, the Roman walls were reinforced around 1300 for the first time.

From the Middle Ages until the 18th century, the city became a military base; the political tensions between the French Kingdom and the Duchy of Savoy (allied with the Spanish and the Habsburg Family) brought about the fragility of this city during the Spanish sacking of Antibes in 1524-36. Under the kingdom of Louis IX and Henri II, Antibes was the first south-western city of France, the gateway to the Italic peninsula. Between 1506 and 1588, the walls were reinforced for a second time; their width become more significant and extended around the port with the construction of the Fort Carré (Fig 93.b). These military infrastructures restricted every possible expansion of the city for more than three centuries. Until 1860, Antibes and its military construction represented a defence point ("Ville de Garnison") against the neighbouring country. In 1860, with the annexation of Nice to the French Empire and the changes in military technologies, Antibes lost its military function.

After only three years, in 1863, a train station was inaugurated, bringing the first tourists to Antibes. The wide military walls surrounding the city were limiting the expansion of the city and did not serve the new touristic economy. In 1886, Juan Les Pins was created on the western side of the cape as a solution: a private domain was sold to a real estate agency ("Société Foncière de Cannes et du Littoral") and with a plan from the architect Ernest Macé, new villas, hotels and buildings in the modern style Art Deco were built. From this moment onwards, two urban centralities were founded on the two sides of the Antibes Capes; a smaller but modern touristic centre on the west side and an ex-military fortress on its east side.

In the meantime, the city of Antibes entered into the process of decommissioning its military role, obtained in 1895 with the permission to remove the military walls which had been defending the city since the Roman age. Consequently, Antibes became a large building site of 24 thousand hectares; it took five years to remove the military walls and smooth out the site. The same architect of Juan les Pins, Ernest Macé, proposed a new plan for the urban

expansion, organised around the central place and with the typical aspects of the Modern-Hausmann traits (Fig93.c).

While Antibes and Juan-les-Pins were developing urban centres with the compact and regular precepts following the Cerda and Hausmann example, the southern extremity of the Antibes Cape was occupied by the large villas and mansions of the European aristocracy and wealthy international traders (i.e. Villa Eilenroc in 1860). Towards the north, beyond the railway, the landscape was dominated by greenhouses for the floriculture. Between 1900 and 1960 the city was known as “Antibes-les-fleurs” (Antibes the flowers).

In 1960, with the construction of the A8 motorway and a second demographic explosion, a new unplanned and disordered urban growth came about in the agricultural spaces at the north of the city. The urban development began with more intensity in spaces surrounding the motorway exit where new commercial activities and large retail surfaces are settled (section 3.2); the axes connecting this new commercial centre to the traditional city centre became the modern backbone along which spaces were gradually filled with single-family houses. A disordered urban development of low-quality characteristics threatened the equilibrium between natural and built spaces.

A second demographic growth started in 1972 with the new plan of the Sophia-Antipolis techno-park in the northern part of Antibes, beyond the motorway, which attracted more traffic on the local street network, which was not prepared for high traffic volumes. In about 60 years, the city passed from 6.8 thousand inhabitants to more than 60 thousand, becoming the 2nd city of the French Riviera.

This brief description of the urban history of Antibes might explain the urban and economic characteristics of this city which is observed today: there are three distinct commercial subspaces in the centres of Antibes, Juan-les-Pins, and around the roadway exit.

Observing the MFA outcomes Fig.93.f, the different phases of the urban evolution of Antibes become easily recognisable. The Greek-Roman centre is identified by UF1. The original plans of Ernest Macé for Juan-les-Pins and the extension of Antibes are found in UF2; the less regular space in between these two centres is also defined by UF2. The train line defines the northern border of UF2. From the train station towards the islands is the urban development of '70s houses (UF3) and buildings (UF4). UF5 identify both the villa southern region of the Antibes Cape, and the inland region impacted by a more recent urban sprawl. Finally, UF7 identify the connective fabric around the roadway exit and around the Port Vauban.

The expert-based cartography produced by local authorities (Fig.93e) shows a striking similarity with the patterns described by the MFA bottom-up procedure.

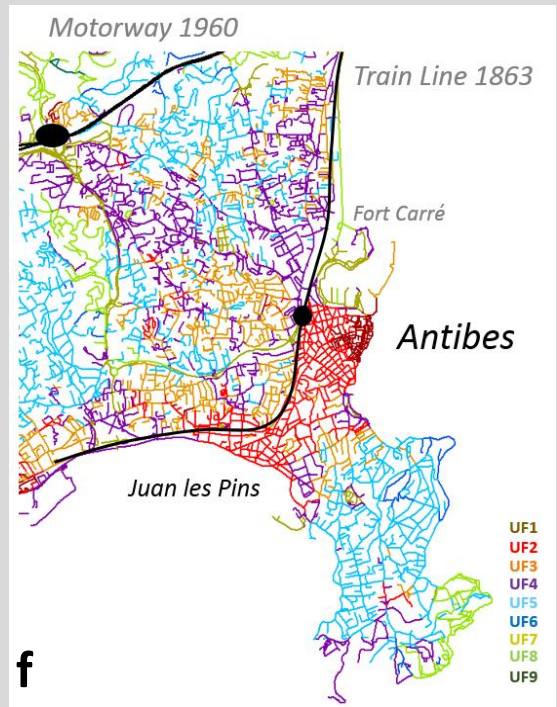
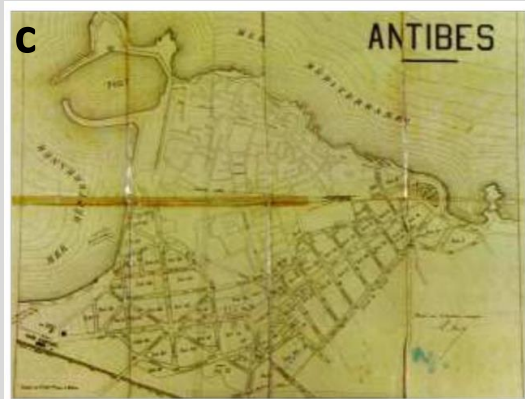
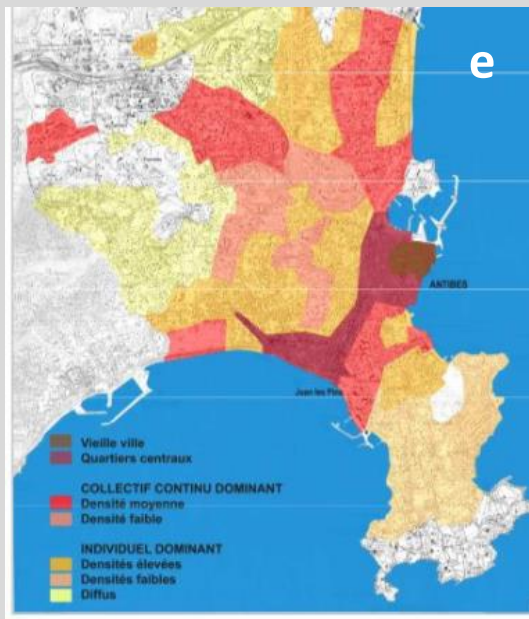
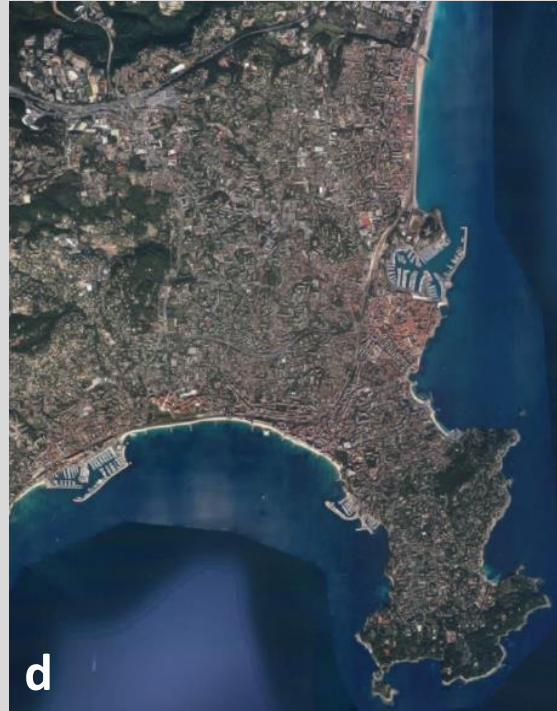
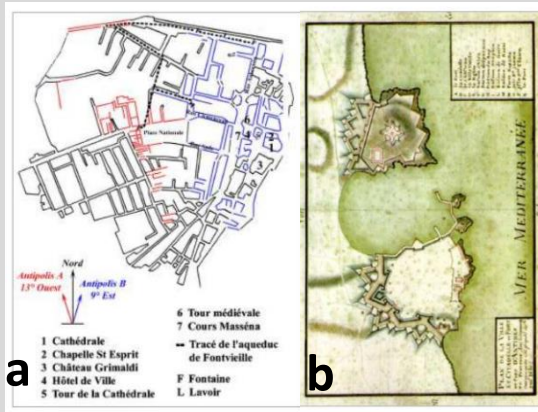


Figure 93 Antibes and its urban evolution: a) the Greek (IV b.C.) and Roman (I a.C) grids of Antipolis; b) the plan for the fortification of Antibes and the construction of the Fort Carré (XVI century); c) Ernest Mace urban plan for the extension of Antibes in the region previously occupied by the military walls (1895); Today's Antibes as perceived from aerial image from Google Earth (d) and the urban fabric classification of the local Authorities (e); in (f) MFA outcomes. (Sources: Google Maps 2019, Archives Municipales de Antibes)

## **Sophia-Antipolis**

As introduced above, 1972 marks a significant turning point the urbanisation process of Antibes with the new urban project of the techno-park of Sophia-Antipolis.

Following the Californian example (Rosemberg 2001), in 1969, the French government acknowledged the techno-park project of Sophia-Antipolis as a National Interest Project (OIN-“Opération d’Intérêt National”), a specific labelling of the French government recognising this project and its strategic importance at the national level. The 2400 hectares of the natural plateau of Valbonne became a laboratory of planning policies for technological innovation. An inland region extended over the different municipalities: Antibes, Biot, Mougins, Valbonne and Vallauris.

While the plans of Ernest Macé for Antibes and Juan-les-Pins at the turning of the century had been influenced by the fore-runners of the modernist movement (Cerda, Haussmann), Sophia-Antipolis is a functionalist car-oriented techno-park following modernist approach to planning and denying street-based forms: a tree-like system of roads connects a variety of buildings freely located in the natural space.

The original vision of the politician Pierre Laffitte was to create a technological centre able to stimulate the encounter between researchers, entrepreneurs, professors and scholars. The three main ideas underpinning this project were: i) “cross-fertilisation” between universities and enterprises; ii) natural and built-up space harmony; iii) reconstruction of a “Latin Quarter among fields”. From these three notions arose the name of the technological park: Sophia-Antipolis. Sophia, the Greek term for knowledge, and Antipolis, the original name of the Greco-Roman settlement of Antibes.

The first sketch of Sophia-Antipolis drawn by Yves Bayard (Fig.94a) proposed the innovative concept of a diffused socio-cultural corridor; the strong functionalist vision of the territory (inherited by the CIAM precepts) is still significant, with a separation of housing, offices, and recreation areas (Grondeau 2018). Environmental matters are also imposed: in 1976, an environmental Charter is signed by the French government, defining the irreversible proportions of one-third of the buildings for two-thirds of green spaces. The same Chart imposed the creation of two departmental natural parks (Valmasque and Brague Departmental Parks for a total of one thousand five hundred hectares) and the protection of hilltops of Sophia-Antipolis.

Nonetheless, as we discussed in the previous section, starting from the ‘70s, the modernist notions of the CIAM enters into a profound crisis. The unstable economic context of the last decades of the 21st century entreated a continuous adaptation of the original plan to the evolving technological and economic global context. From the notion of a natural corridor of Bayard’s plan, the project grew into a more traditional de-structured and less coherent urban approach.

Sophia-Antipolis became a real laboratory of urban planning and different phases might be identified along the past 50 years (Fig.94.b): i) from the utopian period of the Latin Quarter in the countryside, ii) to a second phase where the corridor idea left the place to the creation of

a mono-centrality. iii) A third phase saw the multiplication of micro-centralities and the notion of urbanism of dispersion; iv) a more recent urban planning approach gave more importance to the road infrastructure. This continuous change of urban paradigm is symptomatic of the absence of a strong urban model and the continuous research of an alternative to the Second-Age cities (Portzamparc 1995, Levy 1999). Only more recently have we observed the research of a multi-use of space as a late solution to the absence of urban dynamism.

Two aspects persisted during the evolving project of Sophia Antipolis: the first being the two village centres, Haut-Sartoux et Garbejaire, constructed from scratch and reproducing the continuous and almost compact urban fabric of the traditional city. The second aspect being the irreversible protection of the natural landscape. This latter represented an important constraint which contributed to the process of sanctuarisation of Sophia-Antipolis and protecting these spaces from the explosion of the massive urbanisation of the last fifty years (Grondeau 2018).

From a strictly morphological and organisational point of view, the techno-park might be summarised as a car-oriented urban space, with reduced functional diversity and limited accessibility for pedestrians. A small number of scattered public spaces result from a low urban density as imposed by rigid urban planning rules.

MFA (Fig. 94d) here again is able to well identify these morphological spaces: the two recent centres of Haut-Sartoux et Garbejaire are identified by UF2; the sparse distribution of large buildings, typical from the '70s and identified by UF4. The vehicle-oriented connective fabric (UF7) joining together the previous concentrations of buildings, and, with a bottleneck to the motorway (A8) where the new projects are under construction/planning. UF8 identify those natural hilly areas protected by the departmental parks and which has limited the expansion of the sprawl: UF3, UF5 and UF6 are found only in contiguous municipalities and their expansion has been limited by the borders of the park.

Today, Sophia-Antipolis counts 2230 enterprises gathering both start-ups and large international firms, and more than 36 thousand people from 63 different nationalities (among which 4 thousand researchers and 5 thousand students). The exceptional "cadre de vie" (quality of life) represents one of the main attractive aspects of these spaces (Grondeau 2018). Only nine thousand people (about 3.5 thousand families) live in different small concentrations; residents of Sophia-Antipolis are not the same people working in the techno-park in order to avoid the social segregation of the area. The vehicle-based transport system is overloaded, and daily congestion problems affect the bottleneck that links between Sophia-Antipolis with the roadway (A8).

The recent work of Grondeau (2018) is a socioeconomic analysis of the evolution of these spaces and its new plan for Sophia-Antipolis 2030. In this study, a survey measuring the city-user perception reveals how social activities are limited during the lunch break; 86% of the surveyed desire a reinforcement of the sociability and dynamicity of urban spaces beyond the

working hours. The research of urbanity and sociability of Sophia Antipolis spaces are displayed as the central notion of the new plan for Sophia 2030. Nonetheless, instead of working on the evolution, re-adaptation and integration of the existing spaces in a more socially coherent urban form, stakeholders of the Sophia-Antipolis 2030 plan are focusing on four new large projects imposing new socio-economic centralities. As highlighted by Grondeau (2018), less than 30% of those surveyed support this new plan, and, on the contrary, half of those surveyed would prefer the reactivation of the centralities already in place (Garbejaire and Haut Sartoux).

When observing RFA outcomes in this area, we might find only a few scattered vehicle-oriented retail configurations (RF) and one small nucleated agglomeration (RF8) centred in Place Mejane of Garbejaire. The small concentrated overrepresentation of restaurants and bars in Garbejaire finds a specific explanation: this centrality represents the main space of social encounters for residents within the entire park. As discussed by Grondeau, the creation of the shopping centre of a new generation in proximity to the motorway access (Sophia Antipolis 2030) will compromise the already fragile survival of these few independent commercial activities, threatening the social life of these urban spaces.

This area represents an interesting region and a specific case study for this dissertation as opposed to the traditional compact centres. Its interest is not limited to the urban morphological point of view but also from the retail spatial organisation.



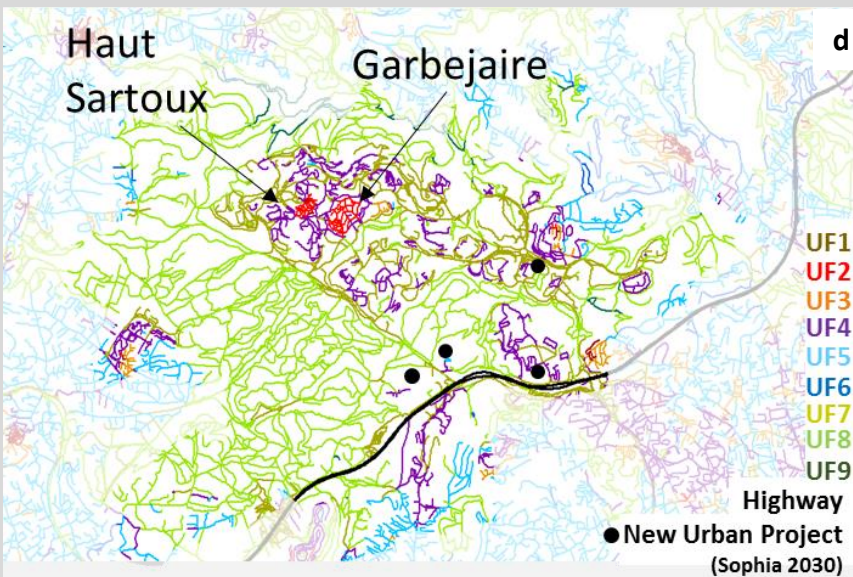
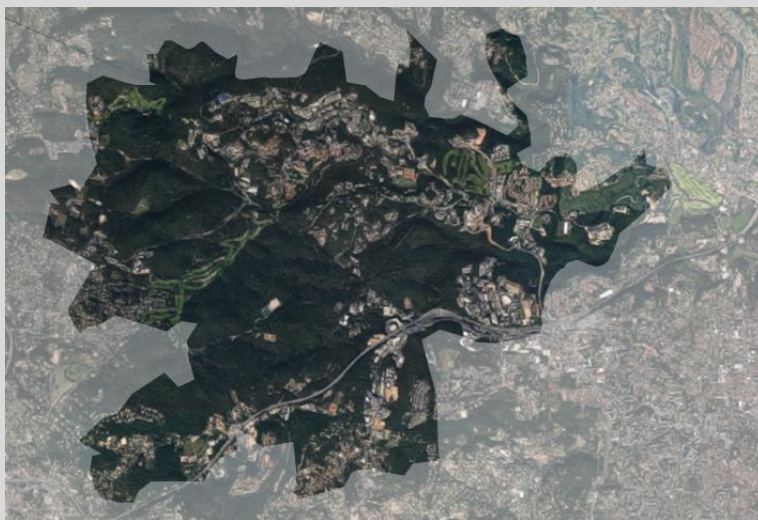
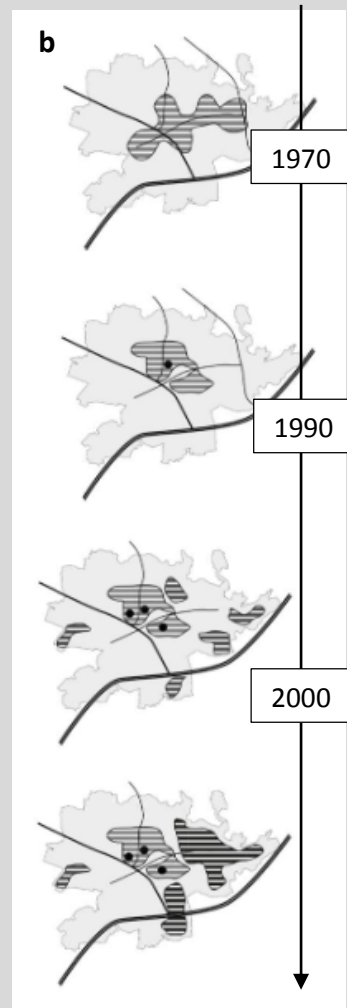


Figure 94 a) Yves Bayard original plan for Sophia Antipolis; b) the four phases of the planning evolution of Sophia Antipolis in the last 5 decades; c) Aerial image of the Sophia Antipolis region from Google Earth 2019; d) MFA outcomes.

### **Urban fabric and Urban morphological regions**

MFA procedure is able to identify nine typologies of urban fabrics in the French Riviera. Nonetheless, as previously described (and demonstrated in Text-Box 3), urban morphological regions might be defined by the combination of two or more urban fabrics.

Urban fabrics with compact and continuous morphological characteristics (UF1, 2, and 3) are mainly found in large coastal urban centres of Nice, Cannes, Antibes, Monaco, and Menton (with the exception of Grasse in the hinterland); few streets and fragments of connective and natural areas correspond to urban parks and gardens. On the contrary, peripheral regions and inland municipalities form a macro-region where historical villages and small centres are spread within large sectors of peripheral urban fabrics (UF4, 5, and 6), mainly defined by *open block* and *urban sprawl* characteristics of more recent urban development.

This dual nature and representation of the urban system has been referred to by French architects and researchers as “*la ville age I et la ville age II*”, referred here as the First and Second-Age cities (Mongin 1995, Portzamparc 1995, Levy 1999, etc.). In particular, Potzamparc (1997) proposes a description of the First and Second-Age city based only on their morphological characteristics and making abstractions of the related functions.

*The First-Age City* is the city organised following the millennial model based on its streets and its dual, the blocks as closed and continuous entities; this simple model is comparable to the traditional city. In the First-Age City “*the making of the city is dominated in its form by the principles of classical urban composition. The prevailing city idea is that of a dense, continuous, compact closed space organised around the street and the island, ordered and regularised, or reign symmetries and perspective for the development of monument. It is a morphological idea of the city base on the notions of street layout, plot and urban fabric [...]*” (Levy 1999, p.50-51).

*The Second-Age City* found its origins in the progressist urban theoretic works of the CIAM in the 1920s, which differently by previous urbanists (like Cerda or Garnier) are organised at the international level with collective action. In 1933, with the Athen Chart, CIAM architects and urban planners completely revolutionised the way to conceive a city. Aesthetical principles of pre-modern urbanism are abandoned and the emphasis is given to a rational approach in the process of planning the city (*la ville machine, la ville fonctionelle, la ville reseau*). The street is a channel which should be organised and planned for the vehicle-based movement; the traditional notion of the closed street block is substituted by the new notion of the open block. However different it might seem, the suburban single-house residential sprawl of the second after-war period is morphologically more similar to the CIAM functionalist, open-block and car-dependent development than to the traditional models of garden cities.

It is evident that when studying human movement and socioeconomic behaviour and phenomena (such as in retail distribution), the distinction between *First and Second-Age cities* might be a key descriptor. For example, when studying the historic centres of Saint-Paul-de-Vence and of the Vieux Nice, knowing that the former is found as an isolated centre in the suburb while the second is integrated into a wider compact area, might result in different effects on both movement and retail distribution.

Beyond the nine urban fabrics detected at the neighbourhood scale, we need to define urban morphological macro-regions at a wider scale and define the boundaries of the urban spaces.

In the last few decades, several studies have investigated indicators and techniques for measuring characteristics of settlements, such as sprawl, compactness, and contiguity. Nevertheless, a general consensus does not exist for the definition of urban boundaries and morphological agglomerations (Parr 2007, Chaudhry and Mackaness 2008, Tannier 2011). Methods based on population distribution and density (Rozenfeld et al. 2008) or socio-economical administrative partitions (Parr 2007), mainly consider functional factors. If urban morphology is considered independently, other procedures have been proposed: a dilation-erosion procedure applied to the building footprint (Chaudhry and Mackaness 2008) or fractal approaches applied to both raster and vector representation of the building surfaces (Longely et al 1991, 1992). A dilation method is the starting point of both these approaches: Minkowski dilation for Tannier et al. (2011); morphological mathematics in Benediktsson (2003), Chaudhry and Mackaness (2008). The former investigates the variation in the number of aggregated elements at each dilation step in the overall space study, while the latter combines the dilation procedure with subsequent erosion. Both methods are applied to Euclidean space and to elementary components (buildings, built-up pixels, etc.) without any semantic characterisation of their urban content.

We propose here the application of the aforementioned dilation-erosion procedure to MFA probabilistic outcomes while respecting network contiguities. MFA results are used for the delineation of the boundaries of the traditional compact city. We will first use the semantic information of our urban fabric types and select clusters 1, 2, and 3 as sharing similar characteristics specific to the traditional urban compact environment. We will filter PBs for which  $p(UF1)+p(UF2)+p(UF3) \geq 0.8$  to focus on those elements that were assigned to the compact city with relatively low uncertainty. Applying topological dilation-erosion to networks of selected spatial units (lag of 3 topological steps), we can identify urban spaces where compact city characteristics might be considered as belonging to the same urban continuum, thus defining compact city borders (Fig.95). Repeating the same procedure with the remaining street segment where  $p(UF8)+p(UF9) \geq 0.8$  allows to isolate those natural areas characterised by a continuum of empty spaces. Finally, the remaining street segments define the Second-Age City region.

With this wider aggregation, the surrounding morphological context of every street segment is now described at two scales: the urban fabric and the morphological macro-region (first-era city, second-era city and natural space). Differently by the methods previously proposed, the delineation of the city borders is here approached from a street-based morphological point of view.

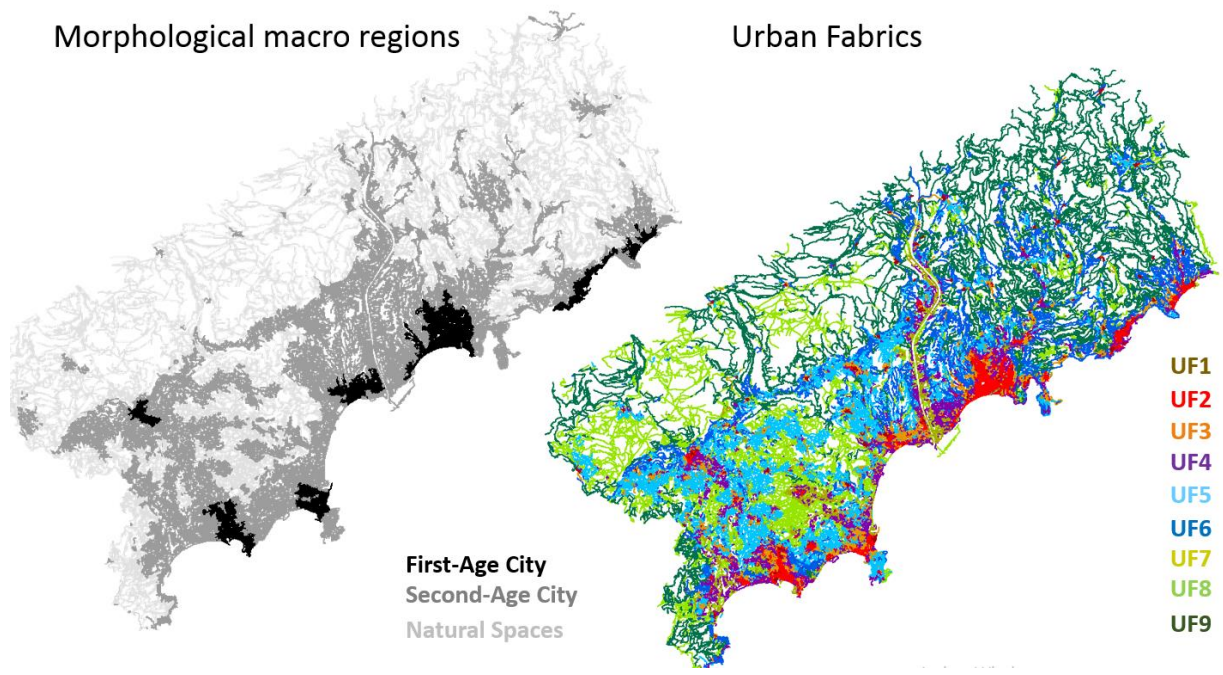


Figure 95 The urban morphological context at the meso- and macro-scales: projection of the nine Urban Fabrics and First/Second-Age Cities regions.

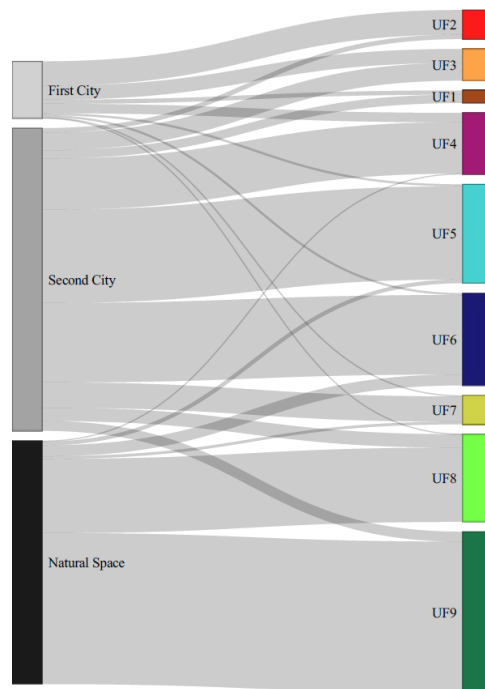


Figure 96 The urban morphological context at the meso- and macro-scales: the distribution of the street within the nine Urban Fabrics and First/Second-Age Cities regions and natural spaces.

*In the Third Part of this dissertation, we focused on the analysis of the physical components of the urban form. When reviewing the quantitative computer-aided procedures developed in the last decades, three main approaches have been identified.*

*In Section 6.1, we presented the street-network configurational approaches firstly developed in the seminal works of Hillier, named Space Syntax theory and successively enriched by a wider number of combinations of modelling methods. In Section 6.2, we reviewed quantitative computer-aided approaches for the description of the morphological properties of the urban fabrics. Nonetheless, typo-morphological analysis traditionally privileges block or building-based partition for the analysis of the urban form; in order to consider the street point of view in morphometric studies, urban design works need to be included. Recent works in this specific domain of studies propose geoprocessing protocols for the intrinsic compositional description of the skeletal streetscapes. Configurational, contextual and compositional procedures (the three C's) represent three approaches able to describe different facets of the urban physical space. As discussed in Section 6.3, these approaches have not been developed independently: more recently several works have been implementing and integrating these different procedures with the ultimate goal of providing a whole comprehensive vision of the multifaceted urban morphology.*

*The definition of the street segment as spatial unit led to the decision to use the specific configurational approach the Multiple Centrality Assessment: in Chapter 7, the traditional four MCA indicators were described and implemented at different scales. Through the 2SFCA approach, a normalisation version of the same indicators is proposed.*

*In Chapter 8, we presented two geoprocessing protocols for the quantitative description of the compositional characteristics of the skeletal streetscape. The former focuses on facades distribution similarly to the urban design tradition, while the latter proposes a description of the morphometric characteristics reinterpreting the notions traditionally developed by typo-morphology schools, but integrating the pedestrian point of view.*

*Chapter 9 presented the Multiple Fabric Assessment as a new protocol for the bottom-up identification of urban fabrics, specifically developed to answer to the goals of this dissertation. Spatial network-constrained patterns of streetscape descriptors were identified (9.1) and combined (9.2). In Section 9.3, a detailed description of MFA outcomes on the French Riviera was presented. Based on the MFA outcomes, further procedures allowed us to outline also a spatial morphological partition at the macro scale: First-, Second- Age Cities and Natural spaces.*

*In this Part we discussed, proposed and developed traditional and innovative computer-aided procedures for the quantitative description of different, yet complementary aspect of the urban physical form. These procedures have been applied to our specific case study of the French Riviera metropolitan area.*

*From the geographical projection of these variables, we have been able to better understand the characteristics of the urban system of the French Riviera. In the following part, the relationship between these aspects and the retail distribution is investigated with robust statistic modelling procedures.*

*Dans cette troisième partie de la thèse, nous nous sommes concentrés sur l'analyse des composants physiques de la forme urbaine. Lors de l'examen des procédures quantitatives assistées par ordinateur développées au cours des dernières décennies, trois approches principales ont été identifiées.*

*Dans la Section 6.1, nous avons présenté les approches de configuration du réseau de rues développées dans les travaux précurseurs de Hillier, appelées théorie de la syntaxe spatiale et successivement enrichies par un plus grand nombre de combinaisons de méthodes de modélisation. Dans la Section 6.2, nous avons passé en revue les approches quantitatives assistées par ordinateur pour la description des propriétés morphologiques des tissus urbains. Néanmoins, l'analyse typo-morphologique privilégie traditionnellement le découpage de l'espace parcellaire et bâti pour l'analyse de la forme urbaine; afin de prendre en compte le point de vue de la rue dans les études morphométriques, des travaux issus du design urbain doivent être inclus. De récents travaux dans ce domaine d'étude proposent des protocoles de géotraitement pour la description des caractéristiques géométriques des paysages bâtis le long des rues. Les procédures d'analyse configurationnelles, contextuelles et de composition (les trois C) représentent trois approches capables de décrire différentes facettes de l'espace physique urbain. Comme indiqué à la Section 6.3, ces approches n'ont pas été développées de manière indépendante: plusieurs travaux ont récemment mis en œuvre et intégré ces différentes procédures dans le but ultime de fournir une vision globale des multiples facettes de la morphologie urbaine.*

*La définition du segment de rue comme unité spatiale a conduit à la décision d'utiliser l'approche de la Multiple Centrality Assessment (MCA): Ainsi au Chapitre 7, les quatre indicateurs traditionnels de cette approche ont été décrits et mis en œuvre à différentes échelles. A travers la méthode 2SFCA, une version normalisée de ces mêmes indicateurs est proposée.*

*Au Chapitre 8, nous avons présenté deux protocoles de traitement géographique pour la description quantitative des caractéristiques géométriques des paysages bâtis le long des rues. Le premier propose des indicateurs descriptifs de la disposition spatiale des façades en respectant la tradition du design urbain, tandis que le second propose une description des caractéristiques morphométriques en ré-interprétant les notions développées traditionnellement par les écoles de typo-morphologie, en se focalisant cependant sur l'espace situé le long d'une rue.*

*Dans le Chapitre 9 nous avons présenté la Multiple Fabric Assessment, un nouveau protocole d'identification ascendante (bottom-up) des tissus urbains, spécifiquement développé pour les objectifs de cette dissertation. Des modèles de descripteurs de paysage de rue contraints par un réseau spatial ont été identifiés (9.1) et combinés (9.2).*

*La Section 9.3 présente une description détaillée des résultats de l'MFA sur la Côte d'Azur. Sur la base des résultats de la MFA, d'autres procédures nous ont permis de dessiner également une partition morphologique à l'échelle macro: Ville âge I, Ville âge II et espaces naturels.*

*Dans cette partie, nous avons discuté, proposé et développé des procédures traditionnelles et innovantes assistées par ordinateur pour la description quantitative d'aspects différents (mais*

*cependant complémentaires) de la forme physique urbaine. Ces procédures ont été appliquées à la région métropolitaine de la Côte d'Azur.*

*À partir de la projection géographique de ces variables, nous avons pu mieux comprendre les caractéristiques du système urbain de la Côte d'Azur. Dans la partie suivante, la relation de ces aspects avec la distribution commerciale est étudiée à l'aide de procédures de modélisation statistique.*





**Part 4**  
*Models for Retail Distribution in  
Urban Space*

*In Part 2 and 3, we separately analysed the retail distribution and the urban form. Specific procedures allowed us to outline different aspects of the two systems; nonetheless, in order to answer our research questions, we still have to explore how these two systems are related.*

*When reviewing the preceding scientific literature investigating the relationship between urban morphology and retail distribution, correlation analyses are traditionally implemented. The goal of these works is to provide statistical evidence about the association between specific morphological properties individually considered and retail distribution.*

*The simple bivariate analysis between retail and each indicator does not allow for us to understand the more complex relationship co-occurring between the morphological variables under investigation. On the contrary, modelling procedures might provide a more appropriate solution for this task. Only more recent works have been implementing Multivariate Linear Regressions (MLR) for the study of the spatial distribution of retail and the urban form (Sevtsuk 2010, Scoppa 2013, etc.).*

*However, several assumptions underlying MLR approaches are not met when considering our case study.*

*Firstly, the homoscedasticity assumption of the error is often violated when implementing linear regression on count data; even if we describe our store distribution in terms of densities, we still only observe a highly skewed distribution.*

*Secondly, our variables might be affected by high multicollinearity explained by the presence of several different metrics describing similar aspects of the urban morphological system. Since regression approaches are based on the assumption of statistical independence of the regressors, multicollinearity might preclude the implementation of regression techniques.*

*Independently from the specific modelling approach, other methodological questions specific to our case study need to be addressed; for example, how might we integrate the over-representation of zeros (store absence) in our dataset? How might we model retail fabric typologies?*

*In this fourth part, we will discuss which statistical modelling approach is most suited for our case study by considering both the research questioning of this dissertation and the specificities of our dataset. Finding the most adapted modelling approach is, in fact, the most important step in order for us to assure the consistency and coherence of our analysis and results.*

*Once the methodological framework is described, the selected modelling approaches are implemented and the outcomes discussed.*

*Dans les Parties 2 et 3, nous avons analysé séparément le commerce et la forme urbaine. Des procédures spécifiques nous ont permis de définir différents aspects de chacun; néanmoins, afin de répondre à nos questions de recherche, nous devons encore explorer les relations existantes entre ces deux systèmes.*

*Dans la littérature scientifique existante et traitant de la relation entre morphologie urbaine et distribution des commerces, des analyses de corrélation sont traditionnellement mises en œuvre. L'objectif de ces travaux est de fournir des preuves statistiques sur l'association entre des propriétés morphologiques spécifiques considérées individuellement et la distribution du commerce.*

*La simple analyse bivariée entre variables de commerce et chaque indicateur morphologique ne permet pas de comprendre l'ensemble des relations complexes qui sont à l'œuvre entre les différentes variables morphologiques étudiées. Au contraire, les procédures de modélisation pourraient fournir une solution plus appropriée pour cette tâche. Seuls des travaux plus récents ont mis en œuvre des régressions linéaires multivariées (MLR) pour l'étude de la distribution spatiale du commerce et de la forme urbaine (Sevtsuk 2010, Scoppa 2013, etc.).*

*Cependant, plusieurs hypothèses sous-jacentes aux approches de MLR ne sont pas satisfaites lors de l'examen de notre cas d'étude.*

*Premièrement, l'hypothèse d'homoscédasticité de l'erreur est souvent violée lors de la mise en œuvre d'une régression linéaire sur les données de comptage. Même si nous décrivons la distribution de nos magasins en termes de densités, nous observons toujours une distribution très asymétrique.*

*Deuxièmement, nos nombreuses variables pourraient être affectées par une multicolinéarité élevée, expliquée par la présence de plusieurs métriques différentes décrivant des aspects similaires du système morphologique urbain. Comme les approches de régression sont basées sur l'hypothèse d'indépendance statistique des régresseurs, cette multicolinéarité pourrait empêcher la mise en œuvre des dites techniques de régression.*

*Indépendamment de l'approche de modélisation, d'autres questions méthodologiques spécifiques à notre cas d'étude doivent être discutées. Par exemple, comment intégrer la surreprésentation de zéros (absence de magasin) dans notre jeu de données ? Comment modéliser des typologies de tissus commerciaux (retail fabric) ?*

*Dans cette quatrième partie, nous examinerons quelle approche de modélisation statistique convient le mieux aux caractéristiques de notre cas d'étude, en considérant à la fois les questions de recherche de cette thèse et les spécificités de notre ensemble de données. Trouver l'approche de modélisation la plus adaptée représente l'étape fondamentale permettant d'assurer la cohérence de nos analyses et résultats.*

*Enfin, une fois le cadre méthodologique décrit, les approches de modélisation sélectionnées sont mises en œuvre et les résultats discutés.*

## 10. Retail distribution and Urban Form:

### 10.1 Overall Framework: the street as a keystone

In the two previous parts, we focused our attention on the characteristics of the retail distribution and on the description of urban form.

In Part 2, retail has been described with both traditional counting/density measures and with an innovative bottom-up, location-based procedure (RFA, Chapter 5), with the ability to identify morpho-functional retail agglomeration typologies. Every street segment in our study area has been therefore associated with the number of stores lying along its edges; from this measure, we can describe the retail offer both in terms of **absence/presence** (Boolean) and of **count** (integer). The **retail agglomeration typology** (categorical) has also been associated to each street segment using the probabilistic outcomes of the Bayesian classification (highest weighted average probability with both the number of stores and their relative frequencies).

In Part 3, we discussed and implemented three street-based procedures for the analysis of the urban form following the traditional approaches developed for the study of the physical city. First, the street network **configurational** approach was implemented on a primal representation of the urban street network in order to characterise the relative position of retail in the network: each street segment is therefore described by a set of forty configurational indicators (Tab.8, Chapter 7), and each one is analysed in relation to a specific combination of scale and impedance to movement. Similarly, thirty-six indicators of **directional configuration** or accessibility (Tab.12, Chapter 8) consider the relative disposition of each street segment to specific urban morphological features (coastline and squares) and retail anchor stores. Second, from urban design and urban morphological approaches, we implemented twenty-eight indicators describing the **compositional or geometrical** skeletal streetscape, representing the hyper-local context of retail activity (Chapter 8). Third, based on urban typo-morphology theories, a specific procedure has been developed (MFA, Chapter 9) for the identification of **contextual or typo-morphological** urban fabrics, which are a mesoscale context for retail: each street segment is associated with nine probability values describing their association to the nine classes of urban fabrics detected in the overall study area. The aggregation of these nine classes allowed for the definition of three main subsystems or **macro-morphological regions** of our study area (First/Second-Age Cities, as well as natural spaces as defined in Section 9.3).

For each aspect discussed in Part 1 and represented in Fig.16 (Section 2.5), we have been able to associate a specific quantitative measure grounded on the respective theories of retail geography, network configurational, urban typo-morphology and urban design (Fig.97). All these measures share the same spatial unit definition and the street-based point of view of the urban form.

The more or less specific adaptation and development of quantitative protocols proposed in the previous parts allowed us to respect one of the main methodological challenges of this

dissertation: street-based measures of urban form (both morphological and functional). This approach should not be considered only as a procedural sophistication: it provides a coherent framework for the analysis of the relationship between form and function of the city at the street level introducing a common “behavioural-based” analytical approach, reducing important statistical biases traditionally overlooked in previous urban studies.

Indeed, the much-claimed need of convergence/bridging between urban morphological analytical approaches has not always come with theoretical coherence and methodological rigour. The need to bridge urban form methodologies firstly discussed by Kropf (2009, 2017) has been followed in the last ten years by a large number of works mixing or combining together several well-known procedures. Nonetheless, these approaches have been originally developed with different spatial unit definitions, goals, and theoretical frameworks; a simple combination of these procedures might raise both theoretical and statistical issues.

The combination of different approaches is a crucial phase of the research which should be integrated within the overall analysis framework with a specific theoretical and methodological reconceptualisation of the traditional analysis, done coherently with the phenomena under investigation.

The procedures developed in this dissertation allow for a **straightforward cross-analysis between form and functional characteristics** (but also only among morphological descriptors) **sharing the same perspective and spatial unit** without injecting further statistical biases from additional and unnecessary data manipulation. **The street segment becomes both the geometrical** (streetscape measures, the geometry of retail agglomerations, etc.) **and metrical reference** (local configurational properties, local morphological patterns, retail agglomeration size, etc.) (Batty 2005). Moreover, each indicator has an easily interpretable meaning directly associable with urban form perceived measures. The street assumes in this work not only the role of *keystone* among urban physical analytical approaches (as presented in Kropf 2009, 2017) but also between form and function.

The methodological framework presents a further advantage: its modularity (each variable is individually identified and implemented) provides the overall procedure flexibility and adaptability. Several case studies might be addressed with this same approach and at the same time each variable could be individually improved, added or removed (i.e. modifying spatial scales, adding/excluding indicators, substituting and comparing configurational approaches, etc.). As we will further discuss in the next sections, this framework is independent of the subsequent statistical/modelling choices (further discussed in this part): bivariate correlations, multilevel modelling, statistical regression, data-mining modelling, etc. might be all easily implemented, and their results compared.

Moreover, another advantage we might already highlight from this framework is its reproducibility: all the variables here implemented have been obtained from few vectorial data (as described in Section 3.4) already available and of easy access at the national scale in France and in several other European countries. The same procedures with minor modifications might be also easily implemented using OSM and VGI data.

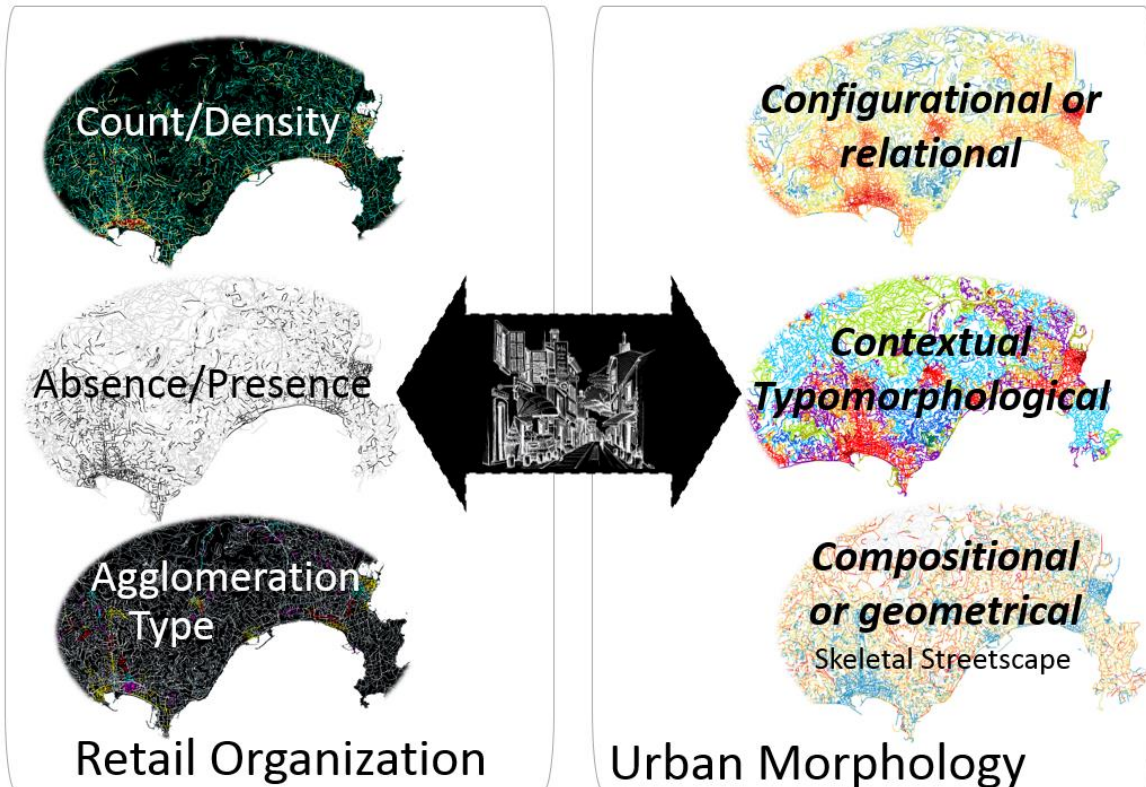


Figure 97 The street as a keystone: bridging urban form (*Configurational, Contextual, Compositional*) and retail measures (*magnitude, presence/absence and agglomeration typology*) at the street level.

Before discussing the specific modelling procedures developed in this dissertation, we should remind the reader the research questions of this dissertation. The underlying goal of this work is to understand the relationship between morphological characteristics of the urban form and the retail distribution; more specifically, as discussed in Part 1, we are interested in understanding which combination of morphological characteristics is observed in correlation with different aspects of retail distribution.

Beyond the traditional relationship between retail density and configurational properties, as originally described in the MET, two main developments are here proposed.

The first innovative aspect considers the integration of morphological descriptors of the streetscape at different levels, from the individual street scale to the contextual variables as defined by MFA. This innovation should enable one to verify whether specific combinations of 'network' and 'streetscape effect' within different morphological regions might be more or less correlated to retail distribution. The second aspect proposes to assess not only the retail count/density but also its absence (traditionally overlooked, inducing what is called *survivorship bias*) and the agglomeration typology as defined by the RFA procedure.

As well summarised in Fig.97, our goal is to identify which combination of the three C's *Configurational, Contextual, Compositional* is best correlated with retail absence/presence, magnitude and agglomeration typology.

We should also remind that, despite the importance of socioeconomic or accessibility descriptors (such as population density, incomes, proximity to the public transport, etc.)

highlighted in previous works (i.e. Sevtsuk 2010, 2014; Scoppa 2013, etc.), these variables are not included in this dissertation since our main focus is studying the role of the physical form of the city. As we will discuss in the conclusion, future works might include socioeconomic descriptors.

Before describing the specific procedures developed in this dissertation, we propose a brief review of the statistical/modelling approaches traditionally implemented in urban studies in the last twenty years previously discussed in Tab.2.

## 10.2 Statistical and modelling procedures: a literature review.

In this section, we propose an overview of the traditional statistical procedures implemented in the specific literature of urban studies where the relationship between street configurational properties and retail distribution is under investigation. When observing the selected works of Table 2, few simple statistical procedures have been implemented; only in more recent years, more sophisticated modelling procedures have been proposed.

In the review discussed in this section, those works proposing simple visual analysis exploring the spatial co-occurrence of retail and morphological pattern (such as Van Nes 2005, 2012; Van Nes and Lopez 2009) are excluded. Despite the interesting theoretical observations discussed in these studies, these notions are not corroborated by statistical evidence limiting any possibility to reproduce, compare, and discuss the outcomes of the analyses.

Of the remaining twenty-one works, six studies implement simple Pearson's correlation between each configurational variable individually considered and the retail density or count, often interpreted as dependent variable (Cui and Han 2015; Porta 2006, 2012; Saraiva 2019; Wang et al. 2014, 2018). Pearson correlation analysis is a statistical approach which allows one to evaluate the linear relationship between two continuous variables. When positive and statistically significant correlation values are identified, the two variables under analysis are described by a linear relationship: a change in one variable is associated with a proportional change in the other. Nonetheless, some assumptions underlying Pearson's correlation do not seem to be respected when considering retail distribution measures—in particular, the absence of outliers, normality of variables, linearity, and homoscedasticity. Indeed, the retail distribution, both measured as count or density, shows a strong positive/right skewness (in our case study, for example, the observed measure of skewness is equal to 31.8 and 201.1 when considering respectively retail count and density). The normality assumption is not verified since retail measures (independently by the choice of spatial unit) could only assume positive measures in the interval  $[0, +\infty)$ . Moreover, the presence of outliers explained in the MET as the result of the “multiplicative effect” and the high presence of zeroes (absence of retail) might further increase the skewness of the distribution. Similarly, the non-respect of these assumptions might be also observed in the statistical distribution of several morphological indicators (i.e. Betweenness centralities, building typology prevalence, etc.).

The observed Pearson correlation coefficients are not wrong per se: while the notion of a linear relationship between two variables still holds even when assumptions are not met, what might not be corroborated is the statistical significance of the correlation test which might result in biased results.

Moreover, several manipulations of the original data are traditionally proposed to meet the normality assumption: smoothing approaches such as KDE is the most implemented procedure, but also, the omission of spatial units with zeroes values and/or log-normalisation<sup>86</sup>. As discussed in 2.4, (and also highlighted by Scoppa 2013) these procedures might overstate the relationship between the variables under analysis and valuable and detailed information might be diluted through these manipulations.

To partially overcome these limitations, other *rank correlation approaches* (such as Spearman, Kendall rank, etc.) might be better suited for the same task. Rank correlation approaches evaluate the monotonic relationship between two continuous or ordinal variables based on their ranked values rather than the raw data; consequently, these approaches are more robust to the presence of outliers. Differently from the linear assumption of Pearson, rank approaches verify whether variables tend to change together—not necessarily at a constant rate.

Cutini (2001, Cutini et al. 2018) proposes a non-linear relationship between retail and configurational properties: exponential fit is investigated. Nonetheless, the exponential correlation coefficients are measured on a small sample of streets randomly selected among the overall study areas (respectively 100 in Grosseto and 30 in Milan) and statistical significance tests are not provided. What we should notice is that even when an exponential association is presupposed, a linear correlation test might be still implemented through a logarithmic transformation of the dependent variable, allowing for a simple assessment of the statistical significance of the exponential relationship. As previously observed, this solution requires the association of an arbitrary positive value to those spatial units where no stores are observed or, alternatively, their exclusion from the statistical analysis; in both cases, the statistical biases previously discussed would persist.

Independently from the specific function and statistical test chosen, the eight aforementioned studies propose only simple bivariate analysis (also known in statistics as Zero-order correlation) between each morphological variable and retail measures. These approaches measure the strength and direction of the relationship between two variables regardless of the effect of any other variable. However, multivariate effect occurs at the same time; occurrences of a specific phenomenon might be detected in correspondence of a specific combination of variables. The simple bivariate correlation should not be implemented when the goal of the study (as in this dissertation) is the evaluation of the combined importance of a set of explanatory variables. Given the underlying assumption of morphological/configurational factors being the underlying explanation of the observed retail distribution, the passage from bivariate to multivariate analysis is performed within the

---

<sup>86</sup> With the consequent need of association of a positive number, traditionally 1 unit, when zeroes are not excluded.



framework of regression models, where a single descriptor of retail activity becomes the dependent variable.

To account for the combinatory effect, Multiple Linear Regressions (MLR) is traditionally considered as a simple and powerful approach able to account for both the partial and semi partial correlation of the variables under investigation. MLR allows to assess both the combined effects of the independent variables and to disentangle and examine their separate effects.

Several works have proposed the implementation of MLR such as Omer and Goldblatt (2015), Sevtsuk (2010 and 2014), Scoppa (2013); Scoppa and Peponis (2015), etc. The implementation of MLR in these more recent works might be also seen as a response to the need to overcome the overly simplistic approaches traditionally implemented in this specific sub-domain of research (Sevtsuk 2010).

In particular, in the two PhD dissertations of Sevtsuk (2010) and Scoppa (2013), the authors proposed both to investigate the relationship between retail distribution and street network configuration properties through MLR. Several peculiarities from of these works have become two important references for our dissertation.

Sevtsuk (2010, 2014), as previously introduced, focused on the study area of Cambridge and Somerville (USA); in his works, the dependent variable is a Boolean value describing the absence/presence of retail in a specific building (the only case study among works of Tab.1-2). Independent variables describe both *weighted configurational properties*, compositional indicators (intrinsic morphometric descriptors around each building), and socioeconomic and accessibility (subway and bus-stops) characteristics surrounding each building. The morphological homogeneity of the study area does not require the integration of a specific context variable. Bivariate Pearson's correlations are firstly implemented as EDA step between each independent variable and retail presence/absence at the building level.

Although the same author suggests the use of probit or logit models, MLR is implemented at the building level. In this work, the overrepresentation of zeros is manually reduced: the author implemented MLR on a subset of 27 thousand buildings selected among those buildings within 100-meter from every store (this sub-selection is justified by the author as the need to consider the strong influence of zoning regulation in the North American context, which limits the retail presence in some urban subareas).

A first linear regression (Ordinary Least Square method OLS) is implemented and used to test the statistical significance of each predictor; moreover, the identification of spatial dependence both in the dependent variable and in the error term is observed using the Lagrange Multiplier diagnostic (Anselin 1988). Subsequently, a second OLS model includes the spatial lag of the dependent variable (100 metres) and the error effects. The statistical significance parameters of the first model are used as a threshold for the manual selection of a subset of variables (similarly to the principle underlying backwards stepwise multiple regression). Finally, the same procedure is implemented on eight different models considering different retail categories.

The works of Scoppa (2013) and Scoppa and Peponis (2015) focus on the South American case study of Buenos Aires (Argentina). Here again, an alternative measure of the dependent variable is proposed. Despite using the street segment as a spatial unit (23.7 thousand street segments), retail is not described with the traditional measures of count or density; the dependent variable in this work is given by the ratio between the length of building frontage of commercial lots and the street length. The resulting ratio might vary between 0, absence of retail, and 2, continuous commercial facades on both street sides. This choice is an interesting solution and proposes a proxy measure of retail magnitude, derived from the available dataset (land use data), which does not allow the traditional measure store number or density for each street.

The independent variables implemented in these works are: configurational properties, accessibility properties (distance from metro and train station), socioeconomic data, street width (intrinsic descriptor) and distance from the CB, and zoning district typology. These measures are both proposed at the local level and at the neighbourhood level (variables are implemented within a buffer of 800 metres/0.5 miles from each street).

Here again, after bivariate correlation, between each independent variable and the retail measure, Multiple Linear Regression is implemented at three levels: the street segment, street name (1.2 long-named streets) and network buffer scales. Variance Inflation Factors (VIF, see further) are measured in each model in order to control for the possible presence of strong collinearity among variables.

The 17 zoning typologies of Buenos Aires might be thought of as a morpho-functional definition of sub-spaces describing the context of each street. This variable is combined with street network configurational properties for the identification of retail intensity in subspaces. More particularly, Scoppa (2013, p.135) shows how the introduction of a zoning variable in the model increases significantly the coefficient of determination. The author also implements a Least Square Means approach, evaluating the aggregate average value of retail frontage within zoning regions (see Section 11.3 for further discussion about Least Squares Means).

In a similar work, Scoppa and Peponis (2015) used the same variables collected for Buenos Aires (configurational socioeconomic and distance from centre/metro/station) for further investigations: PCA is firstly implemented with the final goal of reducing multicollinearity between variables (all except street width, betweenness centrality, and distance from train stations considered as independent descriptors). Four factors are outlined: i) socioeconomic and distance from CBD, metro station; ii) network reach (describing the network density), iii) direct access (directional reach), and iv) choice and integration indicators. These four factors and some independent descriptors are investigated in a second phase with MLR approach.

MLR is also implemented by Omer and Goldblatt (2015): the traditional SSx street network configurational properties are used as independent variables while the dependent variable is represented by the number of stores within a 100-meter buffer from each axial line. As discussed in Chapter 2, the specific focus of this work is the assessment of the network effect within urban settlements with different planning developments. While statistical issues such as spatial autocorrelation or multicollinearity are not discussed, the absence of retail is

tackled: Pearson correlations are provided both on the whole datasets and only on axial lines with a retail count different from zero.

Despite MLR being considered a powerful technique, it might not always represent the best solution. Similarly to what has been previously discussed about statistical correlation approaches, strong assumptions are required for the implementation of MLR (such as linear relationship, homoscedasticity, absence of outliers, variance in predictors, multicollinearity, and normality in distributed residuals). When violated, MLR might result in biased and/or inefficient estimates.

More advanced techniques available to researchers might correct these violations. Nonetheless, before discussing possible violations and sophisticated solutions for the reduction of statistical biases, a simpler and much more effective aspect should be considered by the analyst before the choice of the specific statistical modelling approach: the distribution of the dependent variable.

When studying the retail distribution, each store represents a spatial discrete event; in most of the cases, the analysts disposes of a point pattern where each occurrence is associated with a single individual<sup>87</sup>. Independently of the spatial unit partition, each region under analysis might be simply described by the number of stores within each region (or, similarly, along a street). As we might observe, the dependent variable can only take discrete non-negative values and it does not necessarily follow a normal distribution. In such cases, a conventional linear regression cannot be applied (Cameron and Trivedi 2013, Hilbe 2011 etc.); instead of proposing ad hoc transformations, count regression approaches should be preferred.

Count regression approaches, as we further discuss in the next chapter, have been developed since the end of the '80s in different domains of studies, including retail geography. Nonetheless, among those works summarised in Tab.2, only one recent work proposed count regression: Ye et al. (2017) implemented negative binomial regression to study the relationship between urban block density and typology and urban vitality. The spatial unit in this work is represented by the street-block. The number of small catering businesses represents the dependent variable while the independent variables are measured as a set of accessibility measures (from the combination of SSx configurational properties), building typology, and functional mix. Although we share with this work the interest to the same modelling approach, the procedural procedures developed in Ye et al. (2017) completely diverge from the works previously discussed and from those notions underlying this dissertation (such as the street-base perspective).

---

<sup>87</sup> Differently from bioclimatic or geological data where each punctual observation represents a sampling value of an unknown and spatially continuous function; in this case, points are used for the estimation of the underlying variable/phenomena. When investigating micro-retail distribution, the analyst might discuss the presence of an 'economic/retail' potential function continuously distributed. Nonetheless, this function is only an artificial conjecture of the analyst which do not correspond to any real observed phenomena.

As discussed in this section, the simple bivariate correlation analysis between our variables is the most adopted approach in urban studies when the single effect of a morphometric variable on retail is under investigation. Nonetheless, when the goal of the investigation is the relationship between micro-retail distribution and the combined effect of urban morphological measures, bivariate correlations become inadequate. To overcome these limitations, MLR has been proposed as a better alternative. The intrinsic statistical characteristics of our variables (counting data with excess of zeroes) might represent an important limitation affecting the underlying assumptions of MLR and, consequently, the validity of the outcomes. In this dissertation, we, therefore, propose further development with the exploration and implementation of count regression approaches. In the next chapter (Chapter 11), we further discuss the theoretical explanations which underlie this modelling approach as well as the description of the specific procedures applied in this dissertation. The results of these procedures are presented in Chapter 12 where the advantages of count regression are clearly shown.

Before proceeding with the discussion of statistical modelling approaches, we think it might be worthwhile to re-present the store distribution on the French Riviera, considering the new spatial characteristics of the space study described in Part 2 and Part 3. This description provides the reader with a better understanding of the spatial and statistical characteristics of our variables and the reasons which motivated our search for a new modelling protocol, better fitted to our data.

### 10.3 Exploratory data analysis

Differently from the general presentation of Section 3.4, where the spatial heterogeneity was broadly described dividing the geographical space in littoral/sublittoral and east/west, we can now rely on the outcomes of the spatial analyses previously implemented. The superposition of the urban morphological partitions (from MFA) with the spatial distribution of stores and their agglomeration types (from RFA) allows highlighting the spatial co-occurrences and the heterogeneous distribution of stores.

In Table 15, we measured the following for each morphological region (from the top to the bottom): the number of streets, the ratio of streets where at least one store is observed, the average of the retail count, the number of stores and their share, and the overall density<sup>88</sup>.

We might observe how:

1) While in the overall study area only 12.7% of the street segments contain at least one store, when considering the nine urban fabrics previously described this ratio shows important differences: in descending order we find *traditional urban fabric of the plain with adjoining buildings* (UF2) where 43.3% of streets have at least one store, *old constrained fabric of town-houses* (UF1) with 23.2%, *discontinuous urban fabric with houses and buildings* (UF3) with 16.7% and *modern discontinuous urban fabric with big and medium-sized buildings* (UF4) with

---

<sup>88</sup> We excluded the Principality of Monaco and those streets within 600 metres from its border.

14.6%. Streets within suburban, natural, and connective fabrics have ratios which are lower than the average observed in the whole region (respectively 9.6% in UF5, 8% in UF6, 6.2% in UF7, 2.5% in UF8 and 1.4% in UF9).

2) While we previously observed that 76.6% of the stores are distributed in coastal municipalities and the 23.4% in sub-coastal municipalities, the urban fabrics partition further clarifies the retail distribution: in descending order, retail is mostly found in UF2 with 54.3%, UF4 16.1%, UF1 9%, UF3 and UF5 about 6%, UF6 and UF7 about 3%, UF8 and UF9 less than 1%.

3) While the store density in the overall study area of about 3.55 Stores/Street Length[km] with a higher concentration in the littoral municipalities 8.49 Stores/Street Length[km], important differences might be again observed when considering store differences within the nine urban fabrics: high densities are observed in UF2 (37.9 Stores/Street Length[km]) and UF1 (14.6 Stores/Street Length[km]), moderate densities in UF4, UF3, UF7 (respectively with 5.9, 4.18 and 3.8 Stores/Street Length[km]) while low densities in UF5 (1.45 Stores/Street Length[km]) and UF6,UF8 and UF9 (<1 Stores/Street Length[km]).

		UF1	UF2	UF3	UF4	UF5	UF6	UF7	UF8	UF9	1st ERA	2nd ERA	Natural
All Streets (98467)	N° streets	5935	9165	7996	15093	19588	12196	7326	10530	10638	16579	60635	21253
	[%] Streets with Retail	23,2	43,3	16,7	14,6	9,6	8,0	6,2	2,5	1,4	32	11	2
	Average Retail Street Count	0,65	2,55	0,35	0,51	0,15	0,12	0,32	0,05	0,02	1,59	0,26	0,03
	N° Retailis	3851	23086	2739	6861	2861	1353	1304	335	126	26172	15906	438
	[%] Length[km]	9,06	54,30	6,44	16,14	6,73	3,18	3,07	0,79	0,30	61,56	37,41	1,03
Streets with buildings (63071)	N° streets with build.	5523	8030	7210	10310	16528	9032	2018	2543	1877	14143	43969	4959
	% streets UF(i)	93,1	87,6	90,2	68,3	84,4	74,1	27,5	24,2	17,6	85,3	72,5	23,3
	[%] Streets with Street Count	25	48	18	19	11	10	14	7	5	36,9	14,3	5,5
		0,69	2,8	0,37	0,63	0,17	0,14	0,64	0,13	0,06	1,81	0,35	0,08

Table 15 The distribution of streets, stores and the relative store density for each morphological regions.

The measures in the lower part of Tables are implemented only considering those streets where the coverage ratio of the 50-metres PB is different from zero. The exclusion of those streets without built form on their edges allows reducing the excessive presence of zeroes of retail count associated with absence of built form (and null skeletal streetscape values), which represent trivial information. From a more theoretical point of view, we might consider streets with buildings as the real spatial domain where retailers might choose to locate and therefore the statistical domain on which we will implement our statistical analyses.

For each morphological region, we provide: the number of selected streets, the ratio of the streets with at least one store and the average of the retail count. What we might notice is that selecting streets with built forms, the number of features in UF1-6 is reduced to less than 26%, while on the contrary less than the 27% of the original streets are considered for UF7-9. Fig. 99 visually compares the repartition of streets with buildings in each urban fabric (the width of the columns is proportional to the number of selected streets) and the ratio of streets with/without stores is described on the vertical axes. A too small number of features with stores is observed in UF7-9, preventing us from the implementation of statistical analysis within these three specific sub-spaces.

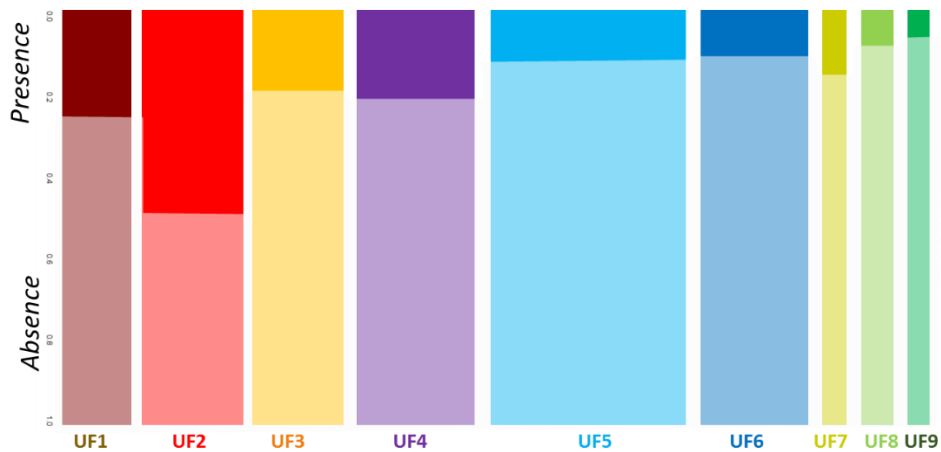


Figure 98 Schematic visualization of store presence/absence repartition in the French Riviera. Vertical axes: absence/presence of stores. Horizontal axes: Urban Fabric classes (the width of each column corresponds to the number of streets with built form on their relative 50-metres proximity bands).

Table 16 (upper part) shows the ratio of street segments of a given UF (in column) which are found in correspondence with different RFs (in rows). Inversely, Table 16 (lower part) shows the percentage of street segments of a given RF (in rows) which are found in correspondence with different UFs (in columns). This simple representation allows us to observe how the retail agglomeration types and the urban fabric families overlap: in correspondence of UF1 and UF2, where the highest concentration values are observed, we find pervasive retail agglomeration typologies. UF4 and UF7 are characterised by specialised retail concentrations. Retail in remaining UFs assumes a scattered or isolated form. Similarly, when considering the First/Second-Age cities as macro-scale morphological contexts, pervasive retail fabrics (RF1-4) are found within the former, while specialised (RF6), nucleated retail fabrics (RF8) and scattered/isolated stores (RF10) are found within the latter.

	UF1	UF2	UF3	UF4	UF5	UF6	UF7	UF8	UF9	1st ERA	2nd ERA	Natural
	Percentage of Urban Fabric <i>j</i> within Retail Fabric <i>i</i>									Perc. of Morph.region <i>j</i> within RF <i>i</i>		
RF 1.a	9	5	0	0	0	0	1	0	0	6	0	0
RF 1.b	8	2	0	0	0	0	0	0	0	4	0	0
RF 2.a	0	6	0	0	0	0	0	0	0	5	0	0
RF 2.b	0	2	0	0	0	0	0	0	0	2	0	0
RF 2.c	0	3	0	0	0	0	0	0	0	2	0	0
RF 3.a	1	4	0	0	0	0	0	0	0	3	0	0
RF 3.b	0	4	0	0	0	0	0	0	0	3	0	0
RF 4.a	0	9	0	0	0	0	0	0	0	7	0	0
RF 4.b	2	12	0	0	0	0	0	0	0	9	0	0
RF 5.a	2	11	3	2	0	0	1	0	0	8	2	0
RF 5.b	2	12	6	5	0	0	2	0	0	10	2	0
RF 6.a	3	2	10	19	6	2	30	6	1	3	11	3
RF 6.b	0	0	1	6	0	0	7	0	0	0	3	0
RF 7	9	3	0	0	0	0	0	0	0	3	1	0
RF 8	23	1	1	0	0	0	0	0	0	0	5	0
RF 9.a	9	11	5	4	0	0	1	0	0	10	3	0
RF 9.b	15	8	19	17	2	1	14	1	0	10	11	0
RF 10	16	3	54	46	91	97	44	92	99	14	61	96

	UF1	UF2	UF3	UF4	UF5	UF6	UF7	UF8	UF9	1st ERA	2nd ERA	Natural
	Percentage of Retail Fabric <i>i</i> within Urban Fabric <i>j</i>									Perc. of RF <i>i</i> in Morph.region <i>j</i>		
RF 1.a	38	59	1	1	0	0	1	0	0	100	0	0
RF 1.b	52	43	0	4	0	0	0	0	0	93	7	0
RF 2.a	0	98	0	1	0	0	1	0	0	100	0	0
RF 2.b	1	98	0	1	0	0	0	0	0	100	0	0
RF 2.c	0	100	0	0	0	0	0	0	0	100	0	0
RF 3.a	5	95	0	0	0	0	0	0	0	100	0	0
RF 3.b	0	99	0	1	0	0	0	0	0	100	0	0
RF 4.a	0	100	0	0	0	0	0	0	0	100	0	0
RF 4.b	5	92	0	1	0	0	0	0	0	98	2	0
RF 5.a	5	81	6	7	0	0	1	0	0	81	19	0
RF 5.b	5	66	12	16	0	0	1	0	0	77	23	0
RF 6.a	4	8	14	44	12	2	14	2	0	18	81	1
RF 6.b	0	8	6	68	3	0	16	1	0	10	90	0
RF 7	49	50	0	0	0	0	0	0	0	62	38	0
RF 8	81	14	2	2	0	1	0	0	0	6	94	0
RF 9.a	18	59	10	12	0	0	1	0	0	73	27	0
RF 9.b	16	26	20	29	3	1	5	0	0	41	59	0
RF 10	4	2	14	19	32	18	4	5	3	14	79	7

Table 16 Repartition of retail agglomeration distribution (RFA typology, Section 5.3) within urban morphological regions (Urban Fabrics and Morphological Macro Regions from MFA Section 9.3). Upper part: proportion of Urban Fabric *j* with Retail Fabric *i*; Lower part: proportion of Retail Fabric *i* within Urban Fabric *j*.

In Fig.99, we plotted the histograms of the retail count (upper part) and density (lower part) of the streets with built form in the overall study area; on the right side we plotted the same values excluding street without stores (logarithmic scale for the y-axes). As we might notice the distribution is highly skewed, characterised by a strong over-dispersion of the count part and an over-representation of zeros. Similar shapes are also observed when considering individual subsystems of Urban Fabrics and Morphological regions, with each one associated with different values of skewness, zero inflation and over-dispersion.

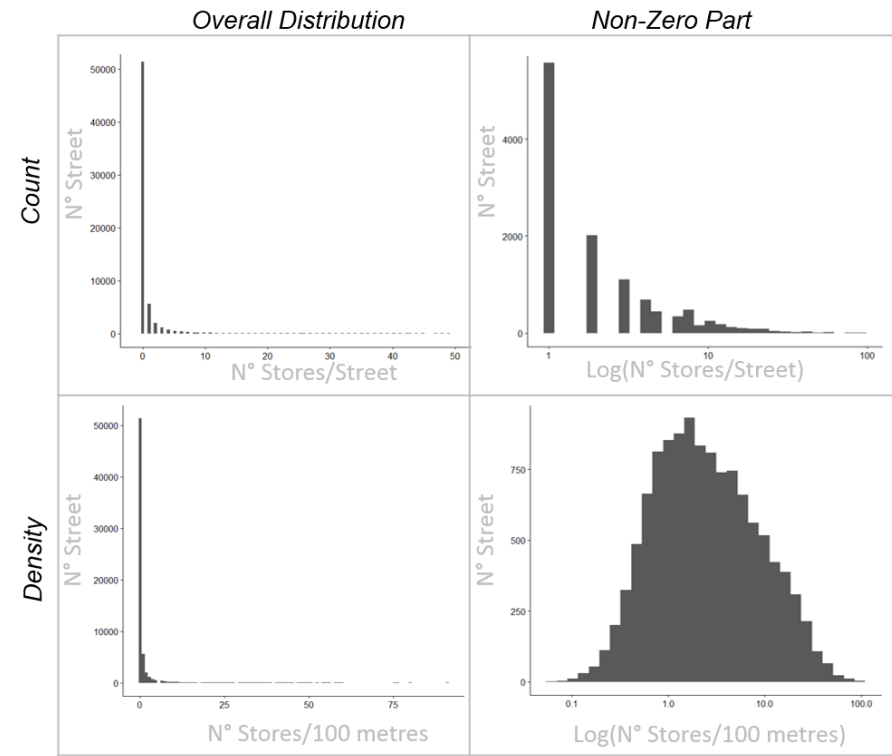


Figure 99 Histograms of the retail count (upper part) and density (lower part) for the streets with buildings in the overall study area.

As we might see from these graphical representations (Tab.15-16, Fig.98-99), retail distribution is characterised by strong heterogeneity, an over-representation of zeros and overdispersion of values. Although reduced, these characteristics persist when considering urban morphological subspaces. The subdivision of the study area in subspaces might help in excluding subspaces where the phenomenon is almost absent (UF7-8-9) allowing to focus on more central and peripheral areas. As previously discussed, in order to consider all these aspects without artificial manipulation (of the raw data of their statistical distribution), specific modelling procedures need to be implemented.

### **The spatial structure of retail distribution: the spatial autocorrelation**

As we previously discussed, the retail spatial distribution is not uniformly distributed in the urban space and strong spatial heterogeneities might be detected; in our study area, we already showed how stores are located on less than 20% of the street segments. Moreover, as observed in the histograms of Fig. 100, among those streets containing retail, only a small number of streets present high store counts/densities. These high concentrations have been explained by retail location theories as the interdependence of the location choices of different stores. MET describes this phenomenon being due to both the spatial distribution of specific configurational properties and the *multiplicative effect* of stores increasing the original attractiveness and economic potential of an urban space generated by the simple configuration of the urban space (Section 2.2).

This effect is what in spatial statistics is traditionally seen as spatial autocorrelation (SA) and it is defined as the relation among values of a variable attributable to the spatial arrangement



of the spatial units. Specific procedures for the detection of SA have been widely studied: these measures allow us to describe the nature and the strength of spatial interdependence of a variable under investigation. When similar values are spatially adjacent, we generally talk about positive SA; conversely, negative SA describes adjacent dissimilar values, null SA being associated with a random distribution of values (Griffith, 1987).

The notion of autocorrelation has already been introduced in the previous Parts of this dissertation, and, more specifically, autocorrelation analysis procedures have been described in Section 9.3 with the goal of identifying local patterns of morphological indicators. In this section, we focus our attention on the identification and description of the spatial autocorrelation affecting the retail distribution.

One of the main reasons why spatial autocorrelation is important is because several statistics procedures rely on the assumption that observations are independent of one another. If SA exists in our variables; then the assumption of independence of observations is violated and specific autoregressive approaches would be required.

For example, in the aforementioned work of Sevtsuk (2010), the author developed a spatially lagged OLS considering retail presence in buildings within 100 metres with the final goal being to explain the retail spatial organisation. This solution allows us to account for the spatial distribution and to reduce the possible statistical bias influencing the correlation coefficients of the model. Nonetheless, with the exception of Sevtsuk's work, urban studies investigating the relationship between urban morphological properties and retail distribution (Tab.2), traditionally overlook the role of SA.

As previously discussed in Part 1, the goal of this dissertation, differently from Sevtsuk, is not the modelling of the spatial distribution of retail *per se*, rather the investigation of the relationship between store distribution and the physical properties of the urban form.

In this work, we assume that the urban morphological variables and their spatial structure (made explicit by the morphological regions) might partially explain the autocorrelation observed in the retail system. Based on this observation, instead of proposing a specific modelling approach accounting for the SA such as autoregressive models, where retail distribution is used as explanatory factor, we propose, to measure and describe the magnitude of the spatial autocorrelation of the original retail distribution and of the model residuals. By comparing the value of the autocorrelation of these two distributions (observation and residuals), we will be able to identify the residual spatial autocorrelation associated with the phenomenon under study. Consequently, what has been explained by the model might be attributed to the overall organisation of the urban form. In other terms, the hypothesis underlying this procedure is that, of the overall SA affecting retail distribution, only a part is actually attributable to retail phenomena, while the rest should be considered as the effect of the spatial structure of the underlying urban form structure (described by our morphometric variables implemented in Part 3).

In order to assess spatial autocorrelation, two procedures have been developed. Correlography and variography represent two traditional quantitative approaches applied in geology and ecology to analyse the spatial organisation of natural phenomena (Legendre and

Legendre 1998). The former investigates and tests the presence of significant spatial autocorrelation in the dataset. The latter allows us to describe its spatial organisation and to relate the observed structure to hypothesised generation processes.

Within the Regionalized Variables Theory, empirical semi-variograms (called variograms for simplicity) are geostatistical functions conceived by Matheron (1965) to describe spatial phenomena combining the structural factor with a random noise component. Fluctuations of spatial distributions are considered in this theory as a characteristic of the phenomena with a specific structure at a given scale, rather than statistical noise. Variograms quantify the similarity between the value observed in a point  $i$  and those observed in its surrounding area at increasing distances. The final aim of these functions is to mathematise the zone of influence of a phenomenon. Variogram function  $\gamma(d)$  is defined by the following equation:

$$\gamma(d) = \frac{1}{2W} \sum_{h=1}^n \sum_{i=1}^n w_{hi} (y_n - y_i)^2 \quad (10.1)$$

$\gamma(d)$  decomposes the spatial variability of the observed variable among distances  $d$ .  $W$  is the matrix of weight describing the spatial relation between values  $y_i$  and the surrounding values  $y_h$  at distance  $d$ .

The application of variograms could reveal the spatial structures of the study field, the presence of nested structures, discontinuities, anisotropies and the size of influence zone of the phenomena (Dauphiné and Voiron-Canicio 1988).

From empirical variograms, three main pieces of information can be observed: the nugget effect identifies if sources of spatial variation might be associated at distances smaller than the sampling interval or to the presence of autocorrelation at smaller scales than the one of the spatial partition (in our case the street segment). The range is the spatial length at which  $\gamma(d)$  flatten and it describes the size of the spatial structure. Finally, the sill is calculated as the limit of the variogram tending to infinity distances,  $\gamma(d)$  informs us about the intensity of the spatial autocorrelation. Other features which might be outlined by the shape assumed by the variogram between the origin and infinity. We might usually identify i) Nested structures, allowing to describe the superposition of different scales of variation and ii) "Hole" effect by the presence of one or several bumps on the variogram. The hole effect denotes a tendency for areas of high values to be surrounded by areas of low values, or "holes".

Spatial correlograms, similarly to variograms, allows us to investigate spatial autocorrelation in the phenomena under study. Correlograms describe how the autocorrelation, measured with statistical approaches such as Moran's  $I$  or Geary's  $c$ , varies for increasing values of distance.

Network-constrained topological variogram and correlogram are here implemented considering the number of stores  $y_i$  associated to each street  $j$ , while the weight matrix  $W$  is determined by the network topological contiguities. In Figure 100, the two curves are plotted, providing a graphical description of the spatial structure and importance of the spatial autocorrelation of retail distribution on the French Riviera.

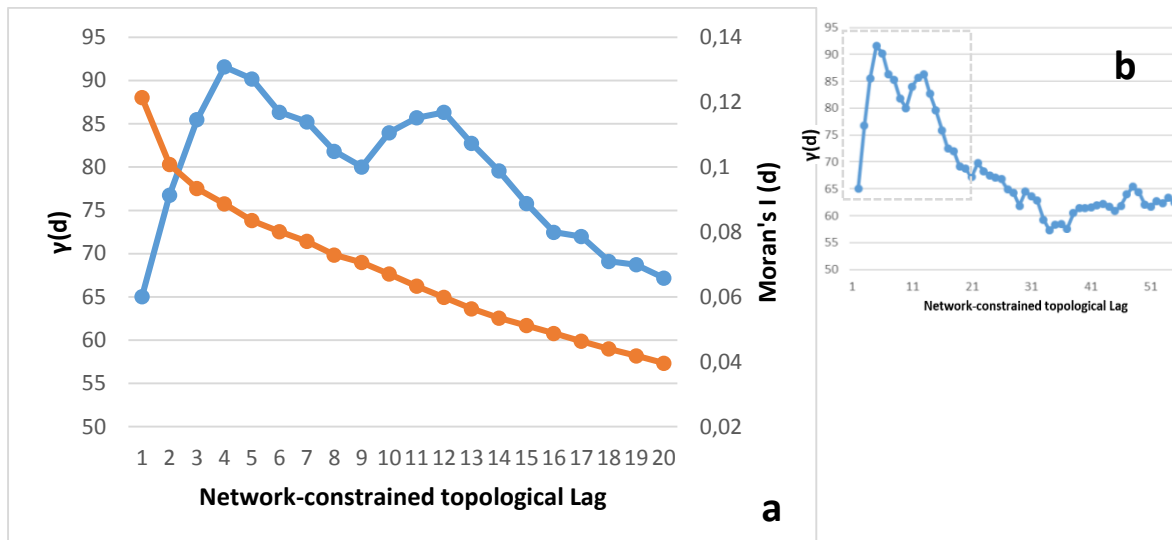


Figure 100 Network-constrained topological variogram (blue) and correlogram (orange) describing the spatial structure and autocorrelation of the retail distribution on the French Riviera

From the analysis of the variogram, we might deduce that: i) as expected, retail distribution might be characterised by spatial structures and interrelation intervening at hyperlocal scales below the street level considered in this work (nugget effect); ii) the variability of  $\gamma$  is strongly reduced for topological distances higher than 16; iii) the variogram computed up to step 51 (Fig.100.b) reveals a spatial structure that extends from step 4 to step 12, followed by a large hole effect. The curve thus reflects the presence of a major contrast in the distribution of stores: areas with a high number of shops, and separated by an average gap of 12 topological steps, are alternated with relatively low concentrations. Another smaller-size hole effect is embedded in this major structure, which reveals the existence of contrasting levels of retail clusters within the same areas of commercial concentration. In other terms, we might describe the spatial distribution of retail as influenced by a double structure intervening at a local-scale, a second spatial structure at mesoscale.

The same information about the maximal extension of the spatial autocorrelation is also described in the correlogram: considering 0.05 as threshold between positive and null spatial autocorrelation, the positive spatial autocorrelation of retail on the French Riviera is detected for a maximal distance of 15-16 topological steps.

In Chapter 12, we will discuss if and how the urban morphological form can partially explain the retail spatial structure isolating the real autocorrelation component attributable to the only retail component.

# 11. How to model the relationship between retail and urban form?

## 11.1 Modelling micro-retail distribution: from linear to count regression approaches

An explanatory model is a highly simplified representation of the reality which helps to better understand the underlying relations occurring between two or more variables through a mathematical equation, the model, involving quantities called parameters. This relation is traditionally described by the general expression:

$$y = f(x, \theta) \quad (11.1)$$

Where  $y$  is the dependent variable that we want to explain through a function  $f$  of the independent variable(s)  $x$ , also called explanatory variables, covariates or regressors. The unknown parameters  $\theta$  need to be estimated through inferential statistic procedures.

*“A function is a set of ordered pairs of numbers  $(x, y)$  such that to each value of the first variable  $(x)$  there corresponds a unique value of the second variable  $(y)$ ”* (Thomas, 1960).

Different algebraic conditions may be assumed by this function, allowing the mathematical modelling of the relationship between variables. The degree to which a specific model fits the observed data is assessed by specific statistical procedures which consider the distance between the observed and the predicted values of  $y$ , named as residuals, in each pair of observations  $(x, y)$ .

The choice of the model function  $f()$  is not always straightforward and it is intimately tied to the specific question under investigation, the shape and the relationship of the variable involved in the modelling procedure. Moreover, each statistical modelling approach has its own assumptions that should be respected to avoid the introduction of bias in statistical outcomes leading to misleading and equivocal conclusions about the phenomenon under study.

As we discussed in the previous chapter, when considering our specific case study, MLR has been proposed as a more sophisticated and robust approach (compared to simple bivariate correlation analysis) allowing the investigation of the combined effect of different variables. MLR assumes the following form:

$$y_i = b_0 + \sum_{j=1}^p b_j x_{ji} + e_i \quad (11.2)$$

Where  $y_i$  is the real-valued response for the  $i$ -th observation of  $\mathbf{y} = y_0, y_1, \dots, y_p$ ;  $b_0$  is the regression intercept;  $b_j$  is the  $j$ -th predictor's regression slope;  $x_{ji}$  is the  $j$ -th predictor for the  $i$ -th observation;  $e_i$  is a Gaussian error term.

Multiple specifies that the number of predictors  $p$  is higher than one (when  $p$  is equal to one, MLR is represented by a simple linear regression); the model is linear since each observation

of  $\mathbf{y} = y_0, y_1, \dots, y_i$  is expressed as a linear function of the parameters  $\hat{\mathbf{b}} = b_0, b_1, \dots, b_p$ ; *regression* expresses the fact that we are modelling the response variable  $Y$  as a function of predictor variables  $\mathbf{X} = X_0, X_1, \dots, X_p$ . In MLR, the traditional approach for the estimation of the parameters  $\hat{\mathbf{b}}$  is traditionally obtained with the Ordinary Least Square (OLS) approach which minimises the sum of the square root of the residuals. OLS solution has the following form:

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (11.3)$$

As previously introduced, before discussing the possible violations of the assumptions underlying the multiple linear regression approach, we should discuss some considerations about the characteristics of the response data we are going to model.

When reviewing works in retail geography, several measures have been proposed as response variables, such as store floor space, sales, workforce, etc. In the study of the relationship between retail and urban form, academics have mainly measured the distribution of stores as number, density (along streets), and retail surface or frontage length (Tab.2). As discussed in Part 2, despite this large variety of measures, when studying the spatial distribution of stores, the simple number of establishments within a given region represents the most adopted solution independently of the specific scientific domain of study. Individual stores represent the natural level of the analysis as well as the legal and functional unit for most businesses (Pipkin 1993, Lebrun 2002, etc.). Moreover, retailers in our original dataset are represented by a point pattern where detailed information about shape, exposition and surface (below 300m<sup>2</sup>) are not provided. Beyond our specific case study, counting values are the natural, obvious, and meaningful scale to describe discrete occurrences/distribution, *“and one should retain these virtues if possible”* (Garner et al. 1995).

The first main observation is that while in MLR the response variable can in principle take any real values on  $R$ , store count can only assume positive discrete values or zeroes on  $Z_0^+$ . As observed by Garner et al. (1995): *“on the one hand, the linear model relating the expected count to the predictors is likely to produce nonsensical, negative predicted values [negative value of retail count/density are theoretically impossible]. On the other hand, the validity of hypothesis tests in linear regression depends on assumptions about the variance of scores that are unlikely to be met in count data”*.

Indeed, the main assumption of linear models is that the residuals follow a normal distribution. When a response variable is skewed, such as the store count/density per street as previously observed (Section 10.3), this assumption is not respected.

Heteroscedasticity in linear regression is a problem *“since the absolute values of the residuals almost always correlate positively with the predictors and, consequently, the estimated standard errors of the regression coefficients are smaller than their true values. This, in turn, inflates the values of statistical tests concerning the values of these coefficients. However, heteroscedasticity is a problem with the model, not the data, because it makes perfect sense that the variance of a count should increase with its expected value.”* (Garner et al. 1995).

Transformations of the response variable are traditionally proposed with the final goal to produce errors that are approximately normal, addressing problems of abnormality and heterogeneity of variance, frequently appearing together (Snedecor 1956). Box and Cox (1964) proposed an exploratory procedure; nonetheless, these transformations have sensible interpretations because the log-transformation implies that mechanisms are multiplicative on the scale of the raw data. Moreover, the log transformation needs an artificial manipulation of zero values (in order to bypass  $\log(0) = -\infty$ ) without any theoretical explanation.

O'Hara and Kotze (2010) demonstrated with an example of ecological count data, that beyond the conceptual meaning of the log transformations, this procedure poorly performs when compared with counting data regression methods. In conclusion, when dealing with a skewed, discrete (but also categorical) response variable with a high number of zeroes, transformations are unlikely to produce normally distributed errors.

All these observations are completely absent within the specific literature studying the relationship between retail and urban morphological descriptors (Section 16.2). On the contrary, suboptimal transformation strategies of the original distribution have been proposed to meet the required assumptions (such as the log transformations, exclusion of zeroes and smoothing techniques which partially mask the real nature and characteristics of the response variable).

Instead of transforming the original observed variable (trying) to meet the assumptions of a MLR, specific approaches for count data have been widely developed in the last fifty years (Nelder and Wedderburn 1972; Cameron and Trivedi 1998, 2013, etc.): count regression approaches are able to consider the discrete nature of the data, the skewness of its shape, and the overrepresentation of zeroes without any artificial manipulation and consequently remove one of the main origins of statistical biases resulting from the improper use of MLR.

When looking at the scientific literature in geography, in 1989, Lovett and Flowerdew listed only 15 papers implementing count regression analysis (Poisson model). The same authors highlighted how, at that time, count methods were "*not very widely known in geography, social and environmental sciences generally*" despite a great number of dependent variables in these domains typically collected under the form of counts. Starting from the '90s, and more vigorously over the last two decades, we are witnessing a general movement from *data transformation* toward *model reformation* (St-Pierre et al. 2017). Count data modelling approaches have been acknowledged and integrated into varying degrees in many research fields: among others, health studies, biology, and ecology observe a wide number of applications (i.e. Neelon and al. 2016, St-Pierre et al. 2017, etc.). In the more specific domain of urban geography, count regression approaches are commonly applied in mobility and transport-related studies: number of car accidents (Kim et al. 2006), transport demand (Karathodorou 2012), etc. This last domain of research has been recently stimulated by the massive interest in the effect of urban form on active transport, walkability, and physical activity (walking commuters by county in Zahran 2008, counts of pedestrians in Rodriguez et al. 2009, etc.).

Within retail geography, count regression approaches have been initially implemented by authors such as Guy (1987, 1991), Davies and Guy (1987), Pipkin (1993), Shonkwiler and Harris (1993, 1996), Harris and Shonkwiler (1993), etc. More recently in retail geography, the use of count regression has become more commonly used especially in those works investigating the frequency of shopping and sales data (i.e. Lundevaller 2009, Verstraete et al. 2019, etc.).

Among the first academics proposing the implementation of count regression, Pipkin (1993) observes: *“a generally linear relationship between population and number of establishments has been identified (Berry 1967, p. 38) and repeatedly confirmed over the past twenty-five years. But residuals of the relationship have not been adequately accounted for. [...] Regression-oriented testing procedures tend to dismiss variations in the number of establishments across places of constant size as uninteresting residuals”*.

Despite the well-developed and acknowledged superiority of count regression methods in several domains of studies on social sciences, works related to retail and urban morphology seems to have been considering these approaches only in more recent years (Ye et al. 2017).

When considering our domain of study, Pipkin’s aforementioned observation, after almost thirty years, is still valid. With few modifications, we might re-propose the same observation as follows: *“A generally linear relationship between network configurational indicators and number of establishments has been identified (Hillier 1999) and repeatedly confirmed over the past twenty years (see Tab.2). But residuals of the relationship have not been adequately accounted for. Linear correlation and regression procedures tend to interpret variations in the number of establishments across streets as a multiplicative effect.”*<sup>89</sup>

This delay is probably related to the different scientific interests of the two disciplines (already discussed in Part 1): the urban morphology community has been developing a wide range of indicators and approaches for the analysis of different aspects of the urban form and its different facets (Part3). Nonetheless, statistical methods *“have often remained overly simplistic”* (Sevtsuk 2010). On the contrary, (retail) geography and spatial statistics, despite being less familiar with those notions developed in urban morphological and planning studies, provide more sophisticated spatial and statistical modelling approaches.

As this dissertation is at the crossway between urban morphology and retail geography, the specificities of these two domains are considered in this work. While the former guided us in the definition of the model descriptors (Part3), the latter provided more adapted solutions for both the response variables (Part 2) and the modelling approaches. Count regression retained our attention: their presentation and implementation are thus discussed in this work.

---

<sup>89</sup> With the exception of the two works of Cutini (2001) and Cutini et al. (2018) where an exponential fit is discussed.

## **Generalised Linear Models: count regression approaches**

We have seen that the traditional (Multiple) Linear Model is based on a strong hypothesis: the term error follows a normal distribution with the same variance. Nonetheless, variables measured as count (as in our case study) do not follow this prerequisite.

To overcome this critical issue, the most widespread approach is to implement a mathematical transformation of the variable trying to make it normal with the final goal to normalise its error<sup>90</sup> and to stabilise the variances. Nonetheless, we saw how these approaches are not all effective, and their normalising effect is sometimes difficult to estimate. The use of the original data should be preferred rather than their transformed values.

It is in this context that **Generalised Linear Models** (GLM) have been developed (McCullagh and Nelder 1989; Agresti 2011, etc.).

The idea developed in GLM is still to use a mathematical transformation on the variable to explain taking into account the *true distribution of errors* (for example, a law of Poisson in the counting case, Binomial in the case of ratios/percentages etc.). The dependent variable assumes a distribution from an exponential family (i.e. binomial, Poisson, multinomial, normal, etc.) and it does not need to be normally distributed. The linear relationship is assumed between the independent variable and the transformed response rather than its raw values.

Consequently, the estimation of  $\hat{\mathbf{b}}$  as in the traditional linear approach based on the minimisation of the least-squares errors (OLS) is no longer possible: an alternative approach is therefore proposed based on the **Maximum Likelihood Estimation (MLE)** procedure.

OLS is a method for approximately determining the unknown parameters of a linear regression model based on the differences between fitted values and observed values regardless of the form of the distribution of the errors. MLE is a technique used to estimate the most likely values of our parameters  $\hat{\mathbf{b}}$  of a statistical model and for fitting a statistical model given the data that we observe.

GLM are based on three primary features:

- *Random Component*: refers to the probability distribution of the dependent variable (and its errors)  $Pr(Y = y|\mu)$  with expected value  $\mu$ ; normal distribution in the linear regression, or binomial distribution in the binary logistic regression, etc. Also called a *noise model* or *error model*.
- *Systematic Component*: it specifies the explanatory variables ( $\mathbf{X} = X_0, X_1, \dots, X_p$ ) in the model, and the linear combination in creating the linear predictor (similarly to the traditional linear regression).
- *Link Function*: represents the mathematical function used for the transformation of the dependent variable which *linearises* the expected value of  $y$ ; it specifies the *link* between random and systematic components. In other words, the link function shows

---

<sup>90</sup> It is important to highlight that in the implementation of MLR what should be normal is the distribution of the error and not (necessarily) the distribution of the response variable.



how the expected value of the response variable is related to the linear predictor of explanatory variables.

While the systematic component is given by the explanatory variables as described in the previous chapters, several random components and link functions can be used and combined depending on the real distribution of the variables under investigation (and its error). Different models might be then specified.

**Linear Regression** is implemented when the distribution of the dependent variable follows a normal distribution and where the errors assume  $e_i \sim N(0, \sigma^2)$ ; the random component is therefore defined as:

$$Pr(Y = y|\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \text{ for } y \in R \quad (11.5)$$

Where  $\mu$  is the average and  $\sigma$  is the standard deviation. Since we are modelling the mean directly, the link function is  $f(y) = y$ , the identity, and we obtain the same linear model of 11.2 and the GLM results in the same estimates of the traditional MLR with OLS (Kutner et al. 2004).

When the variable to be analysed is represented by a count variable, the random component assumes the form of a Poisson distribution able to describe both discrete positive values and zeroes. The probability of observing a value  $Y$  is therefore defined as:

$$Pr(Y = y|\mu) = \frac{\exp(-\mu)\mu^y}{y!} \text{ for } y = 1,2,3 \dots \quad (11.6)$$

Where  $\mu$  is both the average and the variance; for low values of  $\mu$  the Poisson distribution shows a right-skewed shape while for higher values the Poisson distribution tends to the Gaussian distribution 11.5.

In this case, the link function used to analyse a variable following a Poisson distribution is usually a log function where  $f(y) = \log(y)$ ; the resulting model is also named **log-linear or Poisson regression model**. The linear regression of 11.2 is thus rewritten as:

$$E(y) = e^{b_0 + \sum_{j=1}^p b_j x_{ji}} \quad (11.7)$$

In the case of a count that can only take positive values, this link function is a logical function; it allows to have values predicted by the model that extends from  $-\infty$  to  $+\infty$ , which remains interpretable contrary to what it would have been with a traditional linear model.

When implementing a GLM, the meaning of the regression coefficients is different from the one defined by OLS of traditional MLR. In OLS, a one-unit increase in  $x$  adds  $f(y)$  units to the expected value of the dependent variable. In the Poisson regression model, the relationship between the dependent and independent variables is defined as in the 11.7, so a one-unit increase in  $x_{ji}$  multiplies the expected incidents by a factor of  $\exp(b_j x_{ji})$ , while a one-unit decrease divides the expected incidents by the same amount. Similar to traditional linear regression, an increase of the explanatory variable is associated with an increase/decrease in predicted  $y$  when  $b_j$  is respectively  $>$  and  $<$  0 while  $b_j = 0$  means that the predictor does not affect the expected count.

The Poisson regression model represents the benchmark model for count data, is rather simple to use, and has good robustness properties against potential model miss-specification (Gourieroux et al. 1984, Cameron and Trivedi 1998, 2013, etc.). Despite the simplicity of the GLM Poisson model, it should be implemented only when the assumption  $\mu = \sigma$  is respected; nonetheless, heterogeneity in the dependent variable might be at the origin of  $\mu \ll \sigma$ . If this extremely restrictive assumption is not met, the Poisson model produces incorrect estimates of its variance terms and misleading inferences about the regression. Several reasons might explain the heterogeneity, such as unobserved predictors, spatial/temporal dependency, and/or high variation in the observed dependent variable.

As alternative to the Poisson model, the Negative Binomial (NB) might be considered; this specific form of the random component provides an inbuilt solution to account for over-dispersion. The NB is expressed as:

$$Pr(Y = y|\mu, \phi) = \frac{\Gamma(y+\phi^{-1})}{y!\Gamma(\phi^{-1})} \left(\frac{\phi^{-1}}{\phi^{-1}+\mu}\right)^{\phi^{-1}} \left(\frac{\mu}{\phi^{-1}+\mu}\right)^y \quad (11.8)$$

Where  $\mu$  is the expected value (and its variance is given by  $\mu(1 + \mu)/n$ ) while  $\phi$  is equal to  $\mu/\sigma$  and it represents the parameter accounting for the overdispersion. As we might observe, differently from the Poisson distribution which is a one-parameter distribution, the NB is characterised by two parameters. When  $\phi$  is equal to 0, the binomial distribution in 11.5 is the same as a Poisson distribution 11.6.

The link function for the NB distribution is, again a logarithmic function  $f(y) = \log(y)$  and the linear relationship is described by the 11.7. The resulting model is also named **Negative Binomial regression**.

Similarly to what is observed for the 11.7, this link function is a logical function; predicted values assumed by the model range from 0 to  $+\infty$  which, again, are interpretable differently from what it would have been for simple linear models.

As we might observe, when considering our specific case study count regressions (Poisson and NB) represent two interesting alternatives to the classical MLR; the restrictive assumption of homoscedasticity which might be at the origin of serious statistical biases in the definition of traditional MLR coefficients is overcome by GLM considering the *true distribution of errors*.

Nonetheless, as we saw in the literature review, retail is also frequently measured as a density (Tab.2). One might argue that the raw count of stores might be strongly biased by the size of the underlying spatial unit, and, consequently, retail density should be preferred as independent variable of our models.

A specific approach to handle density variables is possible when implementing GLM. Density might be seen as a rate between a count value and the underlying spatial unit size, also named *exposure variable*; in our case study, these two quantities are, respectively, the number of stores and the street length. GLM handle exposure variables using simple algebra changing the dependent variable from a rate into a count by simply multiplying both sides of the equation by the exposure and moving it to the right side of the equation. In the final model,

the exposure variable becomes a term of the regression coefficients, also called **offset variable**.

With this solution, GLM allows us to preserve the natural form of the counting data (and its error) with the inclusion of an offset variable accounting for the possible variability determined by the underlying spatial unit dimension.

### **Count regression and zero-inflation**

So far we discussed how GLM might represent a well-adapted approach for the study of the retail distribution considering the original nature of the data. Despite being able to handle a discrete distribution of non-negative values, a further characteristic should be considered.

As described at several points in the previous sections, the distribution of retail shows a high concentration on few streets while the majority of the study area is characterised by the absence of stores. Moreover, we observed in section 10.3 how, even when specific sub-spaces (urban fabrics) are considered, the excessive presence of zeroes is still elevated.

The Poisson or the NB model previously described might not be appropriate if the overdispersion is due to excess zeroes; heuristic rules suggest a presence of zeroes not higher than 20% of the expected values (far less than what observed in our variable). Overlooking this specific aspect might cause the underestimation of zeros and consequently the variability present in the outcomes. In such situations, alternative models such as Zero Inflated and Zero Alternated models provide solutions to account for overdispersion due to excess zeros. Again, the GLMs modelling approach proposes different solutions able to handle the excessive presence of zeroes (Lambert 1992; Sileshi et al. 2009; Zuur et al. 2009, etc.).

In Zero-Inflated regression models, originally proposed by Lambert (1992), zeros are originated by two simultaneous processes: **structural zeros** (or true zeros) and **random zeros** (or false zeros). In our specific case study, the difference between structural and random zeros might be interpreted as follows: the former represent those street elements characterised by morphological attributes *hostile* for the presence of stores (streets with low probability to *receive* stores); the latter, on the contrary, represent those streets described by favourable morphological conditions but where stores, for other unknown (not investigated) reasons, are not present (streets with higher probability to *receive* stores)<sup>91</sup>.

The probability distribution of the Zero Inflated random component is therefore defined as:

$$\Pr(Y_i = y_i) = \begin{cases} \pi_i + (1 - \pi_i)f(y_i = 0) & \text{if } y_i = 0 \\ (1 - \pi_i)f(y_i) & \text{if } y_i > 0 \end{cases} \quad (11.9)$$

---

<sup>91</sup> This difference between structural/random zeros, or similarly low/high probability is traditionally used in medical and health studies to distinguish between those patients having higher and lower predisposition to certain illnesses (i.e. Dwivedi et al. 2010, etc.).

Where  $\pi_i$  is the logistic link function  $f(y) = \log(y/(1 - y))$  modelling the structural zeros; the count part is modelled by the function  $f(\dots)$  which might assume a Poisson or a Negative Binomial form as respectively described in 11.6 and 11.8.

The resulting models are named **Zero Inflated Poisson (ZIP)** and a **Zero Inflated Negative Binomial (ZINB)** regression models.

Zero Alternated (or hurdle) models originally proposed by Mullahy (1986) and King (1989) model all zeros as one part, while the non-zero part is modelled with zero-truncated count regression; differently from Zero Inflated, these approaches do not distinguish between structural and random zeros. In our specific case study, Zero Alternated models might separately identify those urban morphological characteristics describing absence/presence and magnitude of retail distribution. This second approach goes in the same direction of that common practice of academics in urban studies where spatial units with a non-zero count are isolated, and, consequently, statistical analysis is applied only on a zero-truncated distribution of values (i.e. Porta 2006, etc.).

The probability distribution of the Zero Alternated random component is therefore defined as:

$$\Pr(Y_i = y_i) = \begin{cases} (1 - \pi_i) & \text{if } y_i = 0 \\ (1 - \pi_i)f(y_i) & \text{if } y_i > 0 \end{cases} \quad (11.10)$$

Where, similarly to the Zero Inflated case,  $\pi_i$  is the logistic link function  $f(y) = \log(y/(1 - y))$  while  $f(\dots)$  might assume both a Poisson or a Negative Binomial as respectively described in 11.6 and 11.8. The resulting models are named **Zero Alternated Poisson (ZAP)** and a **Zero Alternated Negative Binomial (ZANB)** regression models.

Implementing ZI and ZA models allow us to explore the possibility that two processes might determine the observed zero and non-zero values instead of considering that these values come from the same data-generating process. On one side, the count process controls for the presence (magnitude), while a second process determines the excessive presence of zeros. Both ZI and ZA are typically described by the combination of logistic regression for the zero part, whereas non-zeroes are described through Poisson (ZIP-ZAP) or Negative Binomial (ZINB-ZANB). The main difference between ZI and ZA is that while the former considers the observed distribution of values as the result of the combined processes with the possibility to distinguish between structural and random zeros, the latter supposes that two separated generating-processes produce zero and nonzero values. Finally, the possibility to use Poisson and Negative Binomial both in ZI and ZA allows us to control for the combined overdispersion of count and zero parts.

The implementation and investigation of Zero Inflated and Zero Alternated models allow us to investigate and answer the specific research question of Section 2.5 and understand if specific combinations of morphometric indicators might explain the absence of retail along streets.

To the three models previously described (G, P, NB), four additional models might be implemented and evaluated, combining un-/truncated counting models with logistic regression (ZIP, ZAP, ZINB, ZANB).

GLM is a powerful technique which enables a wide number of modelling approaches going beyond the traditional MLR and allowing to investigate different aspects of the retail distribution.

While the implementation and comparison of these approaches have been already discussed in several disciplines, the same has never been discussed in the specific case of micro-retail distribution in urban studies.

For this reason, following the methodological approach of Yang et al. (2017), instead of manually choosing a specific model, in the next section we implemented and compare the seven alternatives described above: G, P, NB, ZIP, ZINB, ZAP, ZANB. AIC, Vuong Statistic Test (Vuong 1989) and further performance measures (Section 11.4) are implemented to support and help us in the evaluation of the different modelling approaches and in identifying those models which best fit our data for each system.

### **Modelling retail agglomeration type with multinomial regression**

So far we discussed how GLM could provide a better solution for the modelling of discrete distribution; moreover, GLM allows the implementation of mixture models where the overrepresentation of zeros is evaluated. The implementation of these approaches allows to answer several questions underlying this dissertation, which are related to both the presence/absence and the magnitude of the retail distribution. The explanation of the presence of different retail agglomeration categories have not been yet considered.

We discussed how some of the traditional works in urban studies of Tab.2, propose statistical modelling approaches both for the relationship between urban morphometric variables and retail count/density, but also considering each retail typology count/density (Porta 2012, Sevtsuk 2010).

As discussed in Section 2.4, the simple enumeration of stores and businesses within a street or a given urban space “does not tell the whole story” (Greenberg 1995); moreover, we believe that the separate evaluation of each retail category is not appropriate since the interaction between retail categories is overlooked. The presence of a specific store is also influenced by the typology of the overall retail agglomeration within which it is located.

In order to consider the interaction and spatial co-occurrences of retail typologies in a same location, a specific procedure (RFA) has been developed in Part 2 for the identification of morpho-functional retail agglomeration types (Retail Fabrics). The modal value of RF, obtained as a weighted average of the probabilities of assignment of each store in the street, has been associated with each street segment. Instead of using individual retail categories, we might then use the typologies of retail agglomerations and their relative probabilities.

Two main methodological issues might be associated with this procedural choice: firstly, the separation of retail agglomeration typologies would engender an excessive multiplication of models<sup>92</sup> and the complexification of the interpretation of the results. Secondly, and more importantly, the retail agglomeration typology is a categorical variable where each street might be associated to one of the  $n$  retail agglomeration types: a separate-fitting would result in a less efficient set of models with larger standard errors (Agresti 2002). On the contrary, multinomial regression allows one to simultaneously fit all the logit functions associated to each category of our variable.

GLM is an appropriate model for the explanation of retail agglomerations, too. Indeed, GLM allows for not only continuous and count variables, but also categorical measures: multinomial regression approaches are traditionally implemented for analysing interrelated categorical data (Agresti 2002). The multinomial logit link function is an extension of the binomial version  $f(y) = \log(y/(1 - y))$  where the denominator is defined as the sum of the probabilities of each outcome's class, and, consequently, the probability function might be written as:

$$Pr[(Y_1 = n_1, Y_2 = n_2, \dots, Y_c = n_c) | \sum Y_j = n] = \frac{\exp(\beta_{K-1} \mathbf{X}_i)}{1 + \sum_{k=1}^{K-1} \exp(\beta_k \mathbf{X}_i)} \quad (11.11)$$

Where  $c$  are the classes of the categorical variable to predict,  $\beta_K$  is the set of regression coefficient associated with outcome  $k$ ,  $\mathbf{X}_i$  is the set of the explanatory variables.

Here again, the absence of retail might represent a specific additional class which should be considered in the multinomial GLM. Nonetheless, the zero inflation discussed in the previous section would result in an unbalanced distribution of probabilities, where the class associated with the retail absence would represent more than 80% of the streets under analysis. The implementation of Multinomial GLM would focus on the class associated with retail absence overlooking the importance of the different retail fabrics.

Since ZI and ZA regression models previously discussed already deal with overdispersion and zero inflation of the retail distribution, the multinomial GLM model will be applied here only on those street segments where retail is different from zero. The interpretation of the outcomes would inform us about the urban morphological characteristics of the physical city which is associated with each retail agglomeration typology when retail is observed along a street.

The exclusion of streets should not be considered as an omission of an important part of the dependent variable at the same level of what has been previously discussed regarding the counting models: if our research question is *which urban form characteristics are associated with retails of each agglomeration type*, the absence of retail is not part of the possible answer. The retail presence defines the natural spatial domain on which we might investigate the typology of retail agglomeration. For these reasons, the outcomes of the multinomial GLM model would need to be read and interpreted considering those of the count models.

---

<sup>92</sup> At least 10 retail agglomeration typologies for each of the 7 urban fabrics and 2 morphological regions which results in, at least, 90 models.

In this chapter, we focused on the definition of different models which might be implemented in our case study, and considered specific forms of the dependent variable and its error distribution, allowing to overcome the heteroscedasticity assumption. Nonetheless, before discussing the application and outcomes of these methods, as we will see in the following sections, three more questions need to be discussed and defined.

Firstly, we will see how multicollinearity might affect our models and we review methods that deal with multicollinearity detection. As we will see, this issue is strictly linked with feature selection problems. As we discussed previously, one of the goals of this work is to identify and compare the most relevant urban form indicators that could explain presence, volume and typology of retail patterns. Methodological solutions that could help in this specification are reviewed.

Secondly, we still have to define how the contextual variable (the morphological regions) might be considered in our modelling approach. Different solutions might be implemented considering both the probabilistic and the categorical outcomes of MFA.

Thirdly, we should discuss and establish a common procedure to assess the performance of the models; the final goal is to provide the analyst with the right set of measures supporting the tasks of model and variable evaluation and selection. As we previously introduced, when implementing GLM, the traditional  $R^2$  coefficient of determination used in MLR for the evaluation of the goodness of the model is no longer available. This aspect might prevent analysts to implement count regression with GLM (for instance, this reason motivates the modelling choice of Bielik et al. 2018 where counting methods are suggested but traditional linear regression is implemented). In section 11.4, alternative measures are discussed.

Although these three issues are disentangled and separately discussed for the sake of clarity, issues related to multicollinearity/feature selection, model selection, and spatial heterogeneity represent highly interrelated aspects. Methodological solutions presented in these three sections might be not exhaustive considering the whole corpus of the scientific literature. Moreover, what might be considered the “best” solution could potentially result from the combination of various sophisticated approaches with the risk of excessively increasing the number of parameters required, which results in a not-implementable model when considering regular computing capacities. Beyond these technical considerations, the complexity might also induce conceptual/theoretical issues about the principle of simplicity of a model (*Occam razors, simple is better*). The solutions proposed in the next sections propose an overall good balance between a too complex and an overly simplistic approach.

## 11.2 Multicollinearity and feature selection.

While GLM approaches can process a heteroscedasticity assumption, a second important assumption underlies both MLR and GLM procedure: the independence of explanatory variables  $\mathbf{X} = X_0, X_1, \dots, X_p$ . When explanatory variables are correlated, multicollinearity issues might arise.

Non-experimental studies are nearly always characterised by the presence of multicollinearity. This is even more true in this work where different facets and metrics of the same phenomenon, the urban physical form, are studied and combined. It is then clear how multicollinearity represents one of the major difficulties that we need to consider throughout this section, independently from the form of the model that we want to fit our data.

Multicollinearity might be distinguished in two categories: the former, **structural multicollinearity**, occurs when we create an explanatory variable using the mathematical combination of other variables. In other words, multicollinearity is generated as a by-product of the model rather than being present in the data itself. On the contrary, **data multicollinearity** is present in the explanatory variables rather than being an artefact of the model. Observational experiments, such as our analysis, are more likely to exhibit data multicollinearity even if the distinction between the two is arguable since the variables considered are artefacts specifically conceived for the detection of different still correlated aspects of the urban physical form.

In general terms, Multicollinearity has been defined as a condition in which two or more independent variables are perfectly or nearly perfectly correlated. In this case, one independent variable can be expressed as a linear function of other explanatory variables. This situation can be mathematically illustrated as follow:

$$X_2 = a + bX_1 \quad (11.12)$$

Which means that an explanatory variable  $X_2$  might be expressed as a linear combination of a second explanatory variable  $X_1$ .

In case of **quasi-perfect multi-collinearity** in a general linear model  $y = \mathbf{X}\beta + \varepsilon$  one or more eigenvalues of the moment matrix  $\mathbf{X}^T\mathbf{X}$  is close to zero and consequently the eigenvalue of  $(\mathbf{X}^T\mathbf{X})^{-1}$  is very large. This creates unstable regression estimates and large standard errors. Despite its presence, multicollinearity does not impact the overall fit of regression: the predictive power and the reliability of a model, such as  $R^2$  or  $F$  ratios, are not influenced. Of course, this is not true in the extreme case of perfect multicollinearity, where eigenvalues of the moment matrix are equal to zero, hindering the implementation of the regression analysis ( $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$  does not exist).

Perfect multicollinearity might occur, for example with categorical variables: when transforming  $n$  factors of a categorical variable in  $n$  dummies variables, each of them can be expressed as a linear function of the others. In our problem configuration, this might be the case, for instance, when considering Urban Fabric classes (outcome of the MFA, Chapter 9) or the building typology prevalence (Chapter 8). Quasi-perfect multicollinearity occurs, for



example, when including configurational indicators representing variations of the same measure of centrality (as observed for Reach and Straightness indicators where correlation coefficients are higher than 0.8, see Section 12.3).

Why is multicollinearity so important and how it affects the results of our model? Although multicollinearity does not influence the global model precision, its main consequences concern the analysis and interpretation of individual regression coefficients, and when attempts are made to isolate the individual contribution of one explanatory variable (Makridakis et al. 2008, Kutner et al. 2005, Cohen et al.2003, etc.). The unique effect of individual predictors is estimated by holding all predictors constant, thus ignoring any shared variance between predictors. If multicollinearity is present, each data feature contributes less information to the estimates of individual effects that it does to the overall analysis. This situation may lead to unstable and biased coefficients and standard errors.

Even if increasing dataset size might reduce statistical issues, producing more accurate estimates of each variable, a theoretical issue still persists. Estimates of each variable describe the unique effect for each variable without considering the relationships with the other dependent factors. Consequently, it is impossible to infer each regression predictor as if it could be modified individually, without any impact on the other variables. This means, for example, inferring the sole importance of coverage ratio of the build surface on retail presence without considering that the corridor effect indicator might vary proportionally. Multicollinearity prevents any discussion about the importance of each component included in a regression model and possible comparison with other factors.

In order to detect and reduce multicollinearity issues, several approaches have been proposed and traditionally applied, proposing the reduction of the dimensionality of the problem.

The simple pairwise correlation between two variables might help in multicollinearity reduction. Bivariate correlation coefficients help the researcher in identifying all those pairs of indicators which create multicollinearity issues. The choice of which variables need to be dropped out from the model between each pair of highly correlated variables is left to the researcher's expert knowledge. Justifying each variable choice might not always be easy and it could influence the final results of the analysis. Applying a pairwise correlation and manually selecting and justifying each choice would be highly expensive, or even more if we consider that we need to reproduce the variable selection method for the global model and for each spatial subsystem where pairwise correlation assumes different values every time. Moreover, this manual approach does not provide any solution when considering the combined effect of sub-groups of variables, which is not detectable by simple pairwise correlation analysis. We need, then, a more robust procedure to answer to this task, which would assure reproducibility and reduce the possible arbitrary contribution of an expert-based choice.

Factor analysis, Linear Discriminant Analysis (LDA), principal component analysis (PCA) and similar approaches of *dimensionality reduction* might represent possible solutions to multicollinearity issues. Applying these statistical procedures to a high number of correlated variables enables the identification of a lower number of unobserved variables called factors. The original variables are then expressed as a linear combination of the factors. In urban

morphology and retail geography, PCA has been successfully implemented by Vialard and Carpentier (2015), Bernabé et al. (2015) and in the specific case of retail distribution by Scoppa and Peponis (2015). Factor analysis is considered to be an efficient dimensionality-reduction technique for multivariate analysis (Dunteman 1989); despite the mathematical similarity between several available techniques, different results might arise with possible complications in the interpretation of each factor (Lee and Moudon 2006). The interpretation of the original features and their relation to the original system variables is completely lost in most cases. Factor interpretability is even more evident when comparing results from the analysis of different spatial partitions of our study area, hindering comparative intra and inter level analysis (between partitions from a hierarchical nested structure). For these reasons, factor analysis approaches do not fit the needs of this work. This study focuses on objectively measured, individually observable and disaggregated measures of urban form instead of subjective/latent/composite factors, in order to facilitate both the interpretation of each individual indicator and to allow comparative analysis through the hierarchical partition given by the spatial partition of MFA analysis.

In order to deal with multicollinearity without collapsing individual variables, other methodologies of variable selection should be investigated. In statistical literature, variable selection is one of the most active research areas.

*Tolerance* diagnostic and the Variance Inflation Factor (VIF) represent two valid variable selection techniques: they measure the proportion of the unique information that a predictor provides in the regression. Tolerance is calculated as one minus the proportion of predictor variance that overlaps with the other predictors. Its relation with VIF is given by the formula:

$$VIF_i = \frac{1}{(1-R_i^2)} = \frac{1}{tolerance} \quad (11.13)$$

Where  $R_i^2$  is the coefficient of determination of the regression equation of  $X_i$  on the others  $X$ 's, and it can be applied for each independent variable analysed  $X_i$ . The squared root of VIF informs us about how much the standard error of a variable is greater than the standard error of the same variable if no collinearity with other variables were detected. When VIF values exceed a certain value, a problem of multicollinearity is detected. In the scientific literature, there is no consensus on a precise cut-off value of VIF: depending on the reference, this threshold may vary between 4 and 10 (respectively *tolerance* of 0.18 and 0.10) (Neter et al. 1989; Menard 1995; Hair et al. 2010; etc.). VIF criterion is mostly used as a method to evaluate the final results of a model, as, for example, implemented in Tsou and Cheng (2013) in the specific case of the spatial distribution of retailers. Stepwise routines might be implemented for variable selection where at each step, firstly,  $VIF_i$  is evaluated for every variable of the dataset, and, secondly, the variable described by the highest  $VIF_i$  is excluded from the dataset (Craney and Surlis 2002). This procedure is repeated until the highest  $VIF_i$  of the remaining variables does not exceed the selected cut-off. The outcome of VIF stepwise selection is represented by a subset of independent variables.

Differently from factor analysis, VIF feature selection allows for the removal of multicollinearity problems without collapsing individual variables. Nevertheless, VIF selection does not consider the dependent variable as a target of the selection process, exploring the

intercorrelation only between independent variables (Maddala et al. 2009); moreover, a change or transformation of one independent variable might change the correlations with the other variables and thus the VIFs levels.

Others' stepwise variables selection methods include forward and backward stepwise. Both these procedures reduce the regressors of a model by considering the less highly correlated variables to the target variable. Starting from an empty model, forward stepwise methods add explanatory variables to the model one at the time. At each step, the variables producing the highest increment of  $R^2$  is selected (other criteria might be used for inclusion/exclusion of variables as AIC, BIC). This procedure is implemented until the increment of  $R^2$  is smaller than a given cut-off value. Similarly, backward stepwise selection method deletes iteratively the lowest correlated variable. The Forward/Backward method combines the two approaches. In urban morphology and retail geography, examples of application of forward and backward selection processes can be found related to walkability in Lee and Moudon (2006), building density and frontage in Vialard and Carpentier (2015) and to the specific case of retail distribution in Davies (1973).

Several criticisms have been highlighted in different scientific disciplines. The most evident is that stepwise is sensitive to small perturbation in the data, thus it does not necessarily produce the best model if there are redundant/collinear predictors (Judd et al. 2011). Suppose, for example, that we have two highly correlated independent variables with an almost identical correlation value with the dependent variable; only one of them will enter the regression equation. A small modification of the initial sampling data (or errors in the measurements) might easily reverse the results with possible divergence in the following selection steps of the forward selection process. This divergence might occur at any step of the procedure. Divergence of the resulting models might also be observed when (ex)-including independent variables at the beginning of the process. A possible solution might be a cross-validation approach, which analyses the results from several applications of forward/backward procedures to different subsets of both features and independent variables from the original dataset and compares the variables' coefficients from the  $n$  resulting models. Other criticisms consider the falsely narrow confidence intervals, too small  $p$  values and inflated  $R^2$  (Altman and Andersen 1989), the biased regression coefficients of the remaining variables (Tibshirani 1996), and the bias induced by the application of tests on the same dataset without any cross-validation (Wilkinson and Dallal 1981).

The research of solutions for multicollinearity within the domain of study of statistics has converged in the last decades with the one of *feature selection* research from the machine learning research community. The attention to these methods has been driven by an enthusiastic growing research in data mining processes stimulated by the explosion of technologies and data storage capacities, which provided a high dimensional dataset. Processing and analysing data under these circumstances create issues of efficiency for predictive models and computational costs.

The traditional aforementioned approaches of dimension reduction (or feature extraction) are often associated with inefficiency because they involve the analysis of the whole original dataset, including possible irrelevant variables and variables which may lead to overfitting, causing a loss in both modelling and computational efficiency. The same criticism prevents the application of VIF and stepwise selection methods: the need to reduce the bias in the resulting parameters would require the consideration of all possible combinations of a feature subset which grows exponentially with the original number of features.

New approaches of feature selection have been investigated, and all share a common goal: to propose computational efficient procedures able to identify the most significant subset of features to a targeted variable by removing those features characterised by low relevance and high redundancy (Wei and Billings 2007). Reviewing an all-inclusive panorama of the scientific literature of feature selection algorithms goes beyond the scope of this section (a synthetic review of feature selection approaches as proposed in Heinze et al. 2018). Since the main modelling approach developed in this thesis is regression modelling for count data as discussed in the previous section, we will focus on a specific category of feature selection, namely Penalised Regression for Generalised Linear Models.

The idea underlying methods such as forward and backward selection is to remove predictors from the model which are not considered significant; this operation can be seen as settings their regressor coefficients to zero. The notion underlying Penalised Regression, on the contrary, is to penalise them toward zero, without forcing them to be exactly zero. For these reasons, these methods are also known as shrinkage or regularisation methods. In this way, the complexity of the model is reduced while keeping all or part of the variables in the model. The coefficient associated with the variables with a minor contribution to the outcome of the model is close to zero. These penalised regressions traditionally require the choice of a shrinkage value of lambda to define the magnitude of the penalisation. Three main penalised regression procedures are the most commonly used approaches: Ridge, on Least Absolute Shrinkage Selection Operator (LASSO) and Elastic Net (Enet) penalised regressions.

In Ridge regression when considering a regular Linear regression with OLS procedure, the OLS loss function is augmented to minimise the sum of the squared residuals but also to account and penalise for the size of the parameter estimates with the final goal to shrink them toward zero. The loss function to minimise is:

$$L_{Ridge}(\hat{\beta}) = \sum_{i=1}^n (y_i - x'_i \hat{\beta})^2 + \lambda \sum_{j=1}^m \hat{\beta}_j^2 \quad (11.14)$$

And the 11.3 might be rewritten as:

$$\hat{\beta}_{Ridge} = (\mathbf{X}'\mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}'\mathbf{y} \quad (11.15)$$

Where  $\mathbf{I}$  is the identity matrix while the parameter  $\lambda$  is the regularisation penalty; if  $\lambda = 0$  the shrinkage has no effect and the 11.5 and the 11.3 are equivalent while for  $\lambda \rightarrow \infty$ ,  $\hat{\beta}_{Ridge} \rightarrow 0$ . Ridge regressions show good level of performances even when the predictors outnumber the observations; despite these advantages, Ridge regression never sets the variable coefficient equal to zero and consequently, the final model will contain the original number of variables.

A second approach, LASSO (Tibshirani, 1996) firstly proposed the original LASSO method, which includes both shrinkage and selection method for linear regression. In the case of the LASSO regression, the loss function to minimise is:

$$L_{LASSO}(\hat{\beta}) = \sum_{i=1}^n (y_i - x'_i \hat{\beta})^2 + \lambda \sum_{j=1}^m |\hat{\beta}_j| \quad (11.16)$$

Where differently from the 11.14 the penalisation component is proportional to the sum of the absolute coefficients. In LASSO procedure, differently from the Ridge, several values of lambda are explored allowing the regression coefficient to be shrunk toward zero and, those with a minor contribution might be forced to be exactly equal to zero. For this reason, LASSO might be seen as an alternative variable selection approach which, similarly to Stepwise regression approaches produces a final reduced model encompassing only the most significant variables. When considering the overall capacity of the two penalized regression approaches, both show similar performance values; while Ridge seems to be more adapted when coefficient parameters are of a similar size, LASSO regression is more adapted when the model presents a subset of variables with high coefficient parameters while the remaining have very small coefficients (James et al. 2014).

Finally, Elastic net regression combines both Ridge, and LASSO penalisation approaches allowing to both shrink the coefficient toward zero and to set some of them equal to zero.

$$L_{enet}(\hat{\beta}) = \frac{\sum_{i=1}^n (y_i - x'_i \hat{\beta})^2}{2n} + \lambda \left( \frac{1-\alpha}{2} \sum_{j=1}^m \hat{\beta}_j^2 + \alpha \sum_{j=1}^m |\hat{\beta}_j| \right) \quad (11.17)$$

A second parameter  $\alpha$  allows exploring different combinations of penalisation approaches. When  $\alpha = 0$  we the 11.17 is equal to the Ridge Regression while on the contrary for  $\alpha = 1$  Enet regression is equal to the LASSO regression.

Enet, combining both the advantages of Ridge and Lasso procedures, allows producing simpler and more interpretable models. Implementing Enet regression to our case study will enable to outline the subset of urban morphological variables most related to the spatial distribution of retail.

As we might see in 11.17, two parameters need to be defined in Enet approach,  $\alpha$  and  $\lambda$ . In order to find the optimal values of these parameters, specific iterative processes are implemented allowing to choose  $\alpha$  and  $\lambda$  from a large number of possibilities and based on optimisation procedure based on information criterion, such as AIC or BIC (see further in Section 11.4 for the definition of these parameters).

Penalised regression has been recently applied in retail studies and, more specifically, for the modelling of sales/count data as in the recent works of Roth Tran (2016), Vakhutinsky et al. (2019), Verstraete et al. (2019). These last two works successfully implement penalised regression approaches combined with count regression models.

In conclusion, as we have seen in this section, there is not one specific technique which can truly “solve” the problem of multicollinearity. The approach to implement to deal with multicollinearity should be chosen on the specific case study under analysis; multiple methods

should be employed if possible to help identify the underlying structure within the predictor data set (Dormann et al. 2013).

In this dissertation, we implement both Backward and the Enet feature selection procedures; as previously discussed, the former has already been applied in urban morphology studies while the second has been recently implemented in retail economic studies. In this way, we will be able to assess the performance of the two approaches in our empirical case study, providing valuable information and support to the choice of one of the two methods in future studies.

### 11.3 How to introduce spatiality: the contextual variable

As discussed in the previous sections, the retail distribution is characterised by strong heterogeneities with different degrees of presence/absence and concentration/clustering in different subsystems of our study area. One research hypothesis of this dissertation is that different combinations of urban form properties might be associated with the retail presence, magnitude, and agglomeration typologies within different urban morphological contexts.

The contextual effect also called neighbourhood, integral, or landscape effect, depending on the discipline (Mayer 1990, Mayer and Jenks 1990), is traditionally overlooked by those configurational studies reviewed in Table 2. Only a few authors have tackled the importance of the urban morphological context. For example, Cutini 2001 analyses the relationship between retail and (SSx) configurational properties within six sub-spaces of the city of Grosseto (Italy); Scoppa (2013) uses the official land-use zoning definition of the city of Buenos Aires (Argentina); Omer and Goldblatt (2015) consider eight Israeli cities with distinct historical planning evolutions; Saraiva (2017) shows differences between city centres and peripheries of four medium-sized Portuguese cities. All these works, as well as this dissertation, share the common idea that network properties (as well as other morphological properties) might differently influence retail distribution (as well as other urban functioning phenomena) depending on the urban morphological context (traditionally, between spontaneous and planned city neighbourhoods).

The main idea underlying these works is that spatial context is a fundamental component of the urban system under study, even more when studying large spaces, encompassing a large variety of urban forms<sup>93</sup>. The observed distribution of the response variable does not only result from the spatial relationship with one variable (or a subset of variables). It might be also caused by its dependence on several explanatory variables which are themselves spatially structured (Legendre and Legendre, 1998). The context of the surrounding environment (i.e. neighbourhood in sociology, habitat in biology, etc.) might be considered as a result of the combination of spatial patterns of independent variables. Context variable might also be considered as a confounding variable; with the term confounding are designated those “extra”

---

<sup>93</sup> All the aforementioned works focus on single cities or the comparison of individual cities, while in this dissertation we consider an (almost) entire French department encompassing different cities and towns making up the large urban system of the French Riviera conurbation.

variables that are not measured in a model, although they have a simultaneous effect on both dependent and independent variables. Not considering confounding variables might cause confounding bias in model outcomes with a systematic effect of over/underestimation (called positive or negative confounding bias) and spurious correlations between the dependent and independent variables and a consequent incorrect rejection of the null hypothesis (Leamer 1983, Kaufman and Cooper 1999, Greenland and Morgenstern 2001).

This aspect has been largely studied in ecology (Delgado de la Flor et al. 2017), educational studies (Raudenbush and Bryk 1986, Raudenbush and Whillms 1995, etc.) and epidemiology (i.e. Macyntyre et al. 1993, McMichales 1999, Lawson 2001, etc.). For example, in these studies, the aim is to understand if there are specific neighbourhoods which provide comparative advantages in the outcomes of a specific phenomenon (clean safe and less stressful environment on health) or how/if the socio-economic status of the neighbourhood has a significant incidence on diseases (Diez-Roux et al. 2001, Oakes 2014 etc.). The common idea of all these works is that the contextual effect might change the outcomes of a given phenomenon under study.

In urban studies, the most evident example is the one of the hedonistic model where three groups of variables are traditionally considered (Rosen 1974): structural attributes describing the intrinsic properties of a house, locational attribute describing the accessibility and proximity to services and transport, and the neighbouring contextual attribute.

In our case study, we might find some analogies with hedonic models: indeed, we describe the economic value of a street segment (in terms of quantity and typology of retail presence), focusing on three groups of morphometric properties of the urban form, which describe compositional, configurational and contextual characteristics (as described in Part 3). In our case study, the MFA procedure allowed us to quantify the contextual variable expressed in terms of urban fabric families and morphological regions (Chapter 9).

The question of this section is then how to consider and integrate the contextual variable in our statistical modelling analyses. In scientific literature, four main approaches have been developed:

- 1) We might implement *aggregated data analysis* such as correlations regression or analysis of the variance (ANOVA or Fisher analysis of variance) or covariance (ANCOVA); these statistical techniques assess latent differences in a scale-level dependent variable by a nominal-level variable having two or more categories. These approaches are able to handle large amounts of data summarising and the comparison of variables through summary statistics.

In our case study, we might implement ANOVA, or similar approaches, considering the average and dispersion statistics of the variables within the nine typologies of urban fabrics (and the three morphological regions). This approach has been implemented, for example, in Scoppa (2013): the Least Squares Means approach allows one to consider the statistical relationship between the summary statistics of independent and dependent variables in the 16 zoning areas of Buenos Aires.

These approaches, similarly to what has been discussed for correlation and linear regression analysis, assume that variable distribution within groups are normally distributed and have equal standard deviations. Beyond these assumptions, a more theoretical issue might arise: these “aggregation” techniques might help in identifying whether the contextual partition is significantly correlated to a specific variable; nonetheless, aggregating the data would describe the characteristics of the partition and not about the individual belonging to each region. We lose the fine grain description of each individual, and, more important, it might lead to what is known as *the ecological fallacy* (Robinson 1950), which consists of thinking that relationships observed for groups necessarily hold for individuals. Inferences made at a group level may be correct also at the individual level, but this evidence is not supported by the aggregate data analysis approaches.

- 2) A second approach, in contrast to the previous one, considers *disaggregation* techniques (Nezlek 2011); each individual feature under analysis is labelled by  $n$  dummy variables, which describe the association to the region in which the individual is located, (the group membership but also disaggregate observations measured at the group level). Methods such as Least Square Dummy Variables (LSDV) attributes to the traditional Linear regression  $n-1$  dummy variables ( $y = \alpha + X\beta + g_1 + g_2 + \dots + g_{n-1}$ ). The effect of adding dummy variables does not affect the linear relationship  $\beta$  between  $y$  and  $x$  which remain the same; this approach only allows to adjust the intercept value (shift). When integrating dummy variables in a (generalised) linear regression model, we are assuming that the relationship between individual level variables is the same in every group.
- 3) A third approach is to analyse individual level variables within each subgroup such as regression by groups and group level correlation approaches. Differently from the previous techniques, regression by subgroups is a *separation approach*, which allows the analysis of the variables at the individual element level and allows for the function between variables to vary between groups (not only the intercept but also the slope, differently from the previous approach). The downside of a separation approach is that error estimation might be biased since inter-class variability is overlooked (the models are separately studied).
- 4) Multilevel Linear Modelling (MLM) is a fourth approach which is largely used in social studies when variables are provided at different aggregation levels from the individual level to different aggregate levels (often administrative units) and data presents a nested structure. The advantage of MLM is that they are able to simultaneously consider the variability at the individual and at higher aggregative levels. The application of MLM allows one to consider separate intercept and slope for each spatial context: this would mean to overcome the problem of different retail concentrations within different neighbourhoods, as well as the investigation of different effects of the estimates.



Multilevel theories and more sophisticated statistical models have been proposed in the last two decades with a prevalent application in social, health, and epidemiology studies (Oakes 2004).

When considering urban studies, this modelling approach has been recently proposed as a possible solution to fill the gap between urban morphology and social/ behavioural studies in Liebst (2015); the same author implemented MLM combining configurational analysis (SSx) with social descriptors and survey data in 26 locations of Copenhagen to study the spatial distribution of the “collective effervescence” (Liebst 2019).

Considering that the goal of our analysis is the investigation of the relationship at the street-level of the urban morphometric descriptors and the retail presence, it is clear how the first two approaches are not well-fitted for our research goals. On the contrary, regression by groups and MLM are both good candidates: MLM might be considered the most adapted approach when considering the hierarchical nested structure of our data (individual street nested into urban fabric typologies, on their turn nested in urban morphological regions). Nonetheless, our design research is different from those of traditional social and behavioural studies: our compositional and configurational variables are defined at the individual level of the street segment, while contextual descriptors have been obtained from a specific aggregative procedure that considers the spatial association of the lower (individual) level. In other terms, our variables are not defined at different levels needing a specific methodological procedure to simultaneously evaluate the intra and inter group variability. Moreover, the use of Bayesian classification approaches in MFA allows one to describe the spatial partition in probabilistic terms, which, consequently, might be used as a continuous (instead of dummy) variable in our model.

Finally, but not less important, as discussed in the previous parts, our focus is on the identification and selection of the subset of variables associated with retail distribution and retail agglomeration typologies; specific variable selection approaches (such as penalised regression models) are not compatible with MLM.<sup>94</sup>

For these reasons, the methodological choice in this dissertation is to implement the following: firstly, penalised regression at the global level using contextual variables in terms of probabilities; secondly, (penalised) regression by groups for each subspace identified by MFA procedures (both macro regions and urban fabrics).

---

<sup>94</sup> The ideal model considering all the aspects discussed so far would be a *Penalised Multilevel Zero Inflated Generalised Linear Model*, which is able to consider the dependent variable as a count variable, the nested structure of the data, and to select the most significant regressors for each region with shrinkage regression approaches.

## 11.4 Goodness of fit measures

In explanatory modelling, evaluating a model performance means to assess the strength of the relationship between dependent and independent variables through the function  $f$ . Different parameters might be implemented to indicate the level of goodness of fit of a model. Defining a common procedure for the assessment and comparison of different models represent a task of paramount importance: fitting a wide number of models to our data without being able to assess their appropriateness and their relative goodness would hinder a reasonable interpretation of our modelling outcomes.

In order to answer to this need, the coefficient of determination  $R^2$  is traditionally implemented in OLS regressions as the proportion of the variance of the dependent variable, which is predictable from the independent variables.  $R^2$  coefficient of determination also quantifies, through its square root, the degree of linear correlation between observed and predicted  $y$ .

The most general definition of the determination coefficient is obtained as one minus the ratio between the residual and the total sum of squares (respectively  $SS_{Res}$  and  $SS_{Tot}$ ) as in the following equation:

$$R^2 = 1 - \frac{SS_{Res}}{SS_{Tot}} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (11.18)$$

where  $N$  is the number of observations,  $y_i$  is the dependent variable,  $\hat{y}_i$  the value predicted by the model and  $\bar{y}_i$  is the mean of the  $y_i$  values.  $R^2$  might vary in the range  $[0, 1]$ ; values close to 1 (or 0) indicates that the regression prediction perfectly (or poorly) fit the data.

In OLS, regressions predictions are calculated through the minimisation of  $SS_{Res}$  and  $R^2$  might be consequently implemented as expressed in 11.18. This version of the coefficient of determination is strictly increasing with the number of variables included in the model. In order to consider the complexity of the model, the number of parameter  $p$  might be included as a penalisation factor in the 11.18 providing the adjusted  $R^2$  ( $R_{adj}^2$ )

$$R_{adj}^2 = 1 - \frac{SS_{Res}/(N-1)}{SS_{Tot}/(N-p-1)} \quad (11.19)$$

When evaluating and comparing the fit of alternative models (both at the stage of model and feature selection)  $R_{adj}^2$  should be preferred to  $R^2$ .

However, as reviewed in Section 11.1, regression models with categorical dependent variables GLMs (including ZI/ZA models) are usually fitted by the method of MLE through an iterative process. In these cases, the traditional measure of  $R_{adj}^2$  does not exist and the coefficient  $R^2$  of determination might be obtained only from the raw predictions. Moreover,  $R^2$  is a measure of fit for linear regression with a homoscedastic distribution of error; when implementing the same measure to count regression outcomes, as we discussed in the previous section, the assumption of homoscedasticity is not respected and  $R^2$  might also assume negative values.

The traditional  $R^2$  measure needs to be redefined for the specific case of count regression.

Within the scientific literature, two groups of Pseudo  $R^2$  have been proposed. The first suggests measures of pseudo  $R^2$  based on: *likelihood or log-likelihood* as in McFadden's (1974), adjusted McFadden's, Cox and Snell (1989), Nagelkerke (1992), Crag and Uhler (1970); *estimated probabilities* in Efron (1970); *variance decomposition* (for binary and ordinal logit models) in McKelvey and Zavoina (1975). A second group considers functions of fitted and observed values. Count and adjusted-Count pseudo  $R^2$  are implemented as the proportion of correct predictions on the overall number of predictions (Gujarati 2009). Other methods are discussed in Mittlbock and Schemper (1996) and Menard (2000). Most of these methods are presented and discussed in Long and Farese (2003) and Hardin and Hilbe (2001). While an extensive literature has focused on pseudo- Pseudo  $R^2$  for logistic and count data regression models (Cameron and Windmeijer 1995, Cameron and Trivedi 2005), goodness of fit for ZI regressions has been only recently addressed by Martin and Hall (2016), who conceived a specific pseudo  $R^2_{ZIP/ZINB}$  based on deviances. As this large variety of works on pseudo  $R^2$  suggests, there is no general consensus on which measure should be preferred and each choice leads to drawbacks (Tjur 2009). Moreover, most of these measures have been specifically conceived to fit particular type of GLMs regressions (Binomial, Poisson, ZI, etc.), hindering any possibility to apply them to a large variety of models with the final goal of supporting the model selection phase.

A different approach to assess the goodness of fit of a model, which has become increasingly popular, is represented by information criteria (IC) measures. The notion behind IC approaches is the need to find a compromise between two contrasting measures: on one side, the likelihood maximisation, while on the other is the principle of parsimony (or Occam's razor principle) which favours simpler models (Cameron and Trivedi 2005). Akaike Information Criterion (AIC, Akaike 1973), proposes a new way to compare different models on a given outcome. The choice of the *best* model is guided by the AIC score, implemented as:

$$AIC = 2K - 2\log(L(\hat{\theta}|y, M)) \quad (11.20)$$

where  $K$  is the number of estimable parameters which correspond to the degree of freedom, while  $L(\hat{\theta}|y, M)$  is the maximum value of likelihood function for the model  $M$ . AIC score is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model; lower AICs are associated to models which are closer to the reality. It is important to highlight that AIC scores do not have a specific meaning when independently considered: it is only the comparison of different AIC scores that helps the analyst in ranking and selecting the best model between a finite set of models.

Another method that might be applied to the model selection task similarly to the AIC approach is the Bayesian Information Criterion (BIC) or Schwarz Information Criterion (SIC) developed by Schwarz (1978) and defined as:

$$BIC = K \ln(n) - 2\log(L(\hat{\theta}|y, M)) \quad (11.21)$$

Where again,  $K$  is the degree of freedom,  $L(\hat{\theta}|y, M)$  is the log-likelihood and  $n$  the sample size.

The main difference between the two ICs is that the score resulting from BIC is more penalized by the number of parameters  $K$  than the score obtained from AIC. Therefore, the latter is preferred if more importance is given to model parsimony in model selection task (Cameron and Trivedi 2005).

Both these methods might be applied to several statistical methods (not limited to regressions) and they allow to compare non-nested models, which ordinary statistical tests cannot do. Moreover, differently from Pseudo  $R^2$  measures, ICs penalise the inclusion of variables that do not significantly increase the fit.

As a rule of thumb, Burnham and Anderson (2004) suggest that an AIC difference greater than 10 favours the model with the smaller IC. Another way to interpret ICs considers the relative probability that the  $i$ -th model minimises the information loss as:

$$e^{(IC_{min}-IC_i)/2} \quad (11.22)$$

where  $IC_{min}$  denotes the model with the minimum value of IC and  $IC_i$  the IC of the  $i$ -th compared model.

Likelihood ratio-based tests (LR test) might be implemented to help the analyst understanding if a statistically significant difference between model IC scores is detected. LR tests null hypothesis is that both the compared models are equally close to the true model. If the null hypothesis is not verified, one of the two models should be considered as better performing. Specific LR tests should be considered when comparing nested or non-nested models: the former is defined when “*one model is a special case of the other*”, which means that all terms of the smaller model occur in the larger one. On the contrary, non-nested models are defined when “*neither can be represented as a special case of the other*” (Cameron and Trivedi 2005, p.278).

Vuong (1989) provided a LR test for non-nested models which represents, so far, the most applied LR test in different domains of the scientific literature without any restrictions on GLMs regressions. However, when comparing ZI models with non-ZI models, boundary issues in the parameter space generates questions about the definition of whether they might be considered non-nested models or not. Wilson (2015) has recently highlighted how Vuong non-nested test assumptions are violated when considering the specific case of ZI models. The same author proposes a graphical method where fitted probabilities resulting from ZI and the relative non-ZI model are plotted similarly to what is traditionally done with “normal Q-Q plot” in order to visually assess whether or not zero inflation is appropriate. The same issue is addressed in Neelon et al. (2016) who mention Todem et al. (2012) and Cao et al. (2014) as possible solutions.

In this work, we choose to use AIC and BIC scores in order to quantify and rank the performance of our models. For the sake of comparability, this choice would require to implement linear regression through the GLM approach (with a Gaussian link function); as discussed in section 11.1, the multiple linear models implemented through the OLS or MLE approach lead to equivalent outcomes. Non-nested Vuong test is applied in order to guide the

model selection. Aware of the possible biases when considering ZI models (Wilson 2015), other measures and graphical methods are also implemented in order to assess our models.

From this brief review, it could be inferred that the problem of model evaluation mainly surrounds the research of the good parameter of the goodness of fit (i.e.  $R^2$ , pseudo  $R^2$ , ICs). These parameters provide a single piece of information summarising the general behaviour of a regression model without any additional indication about the possible source of fitting problems (Kleiber and Zeileis 2016).

Other parameters should be investigated in order to describe different aspects of the model outcome distributions and their relationships with observed values, which would bring new useful evidence that might contribute to the task of interpretation and model selection.

More precisely, three aspects of the goodness of fit need to be further explored: overdispersion, over/underfitting, and, most importantly, the zero inflation. Researchers demonstrated how the presence of a large number of zero observations (10-80% of the overall values) should not be necessarily considered as inflation of zero (Warton 2005). Regular P/NB binomial might already account for the excess of zeros without an unnecessary complexification of the model. The compulsory implementation of ZI and ZA regression approaches is not recommended and different goodness-of-fit measures should carefully evaluate this specific aspect of the models.

For these reasons, beyond AIC and graphical approaches, further measures should be considered. In the following part of this section, a specific set of measures and graphical representations are proposed in order to support the process of model outcome analysis and comparison included in the process of model selection.

An additional graphical tool that we use to assess the goodness of fit of our models is the *rootgram*. From some early works of Tukey (1977), Kleiber and Zeileis (2016) recently extended and developed this specific method to the specific case of counting data regression models. Instead of looking at the Q-Q plot, rootgram plot the distribution of observed and expected frequencies on a square root scale rootgrams; this graphic representation allows analysts to easily assess the noise and the fit of the model separately for high and low values. As illustrated by its authors, rootgrams reveal their efficacy when evaluating presence of zero inflation, over/underfitting and over-dispersion. Although this method allows for easy visual evaluation and identification of possible fitting issues, more accurate methods need to be explored. In Figure 101-left we might observe an example of rootgram where expected counts, given the model, are shown by the red line while observed counts are shown as bars; in Fig.101-right the bars correspond to the residuals (difference between observed and expected counts). Considering the 0-axis as a reference line, if a bar is over the line then the model over predicts a particular count value while if the bar exceeds the zero line it under predicts.

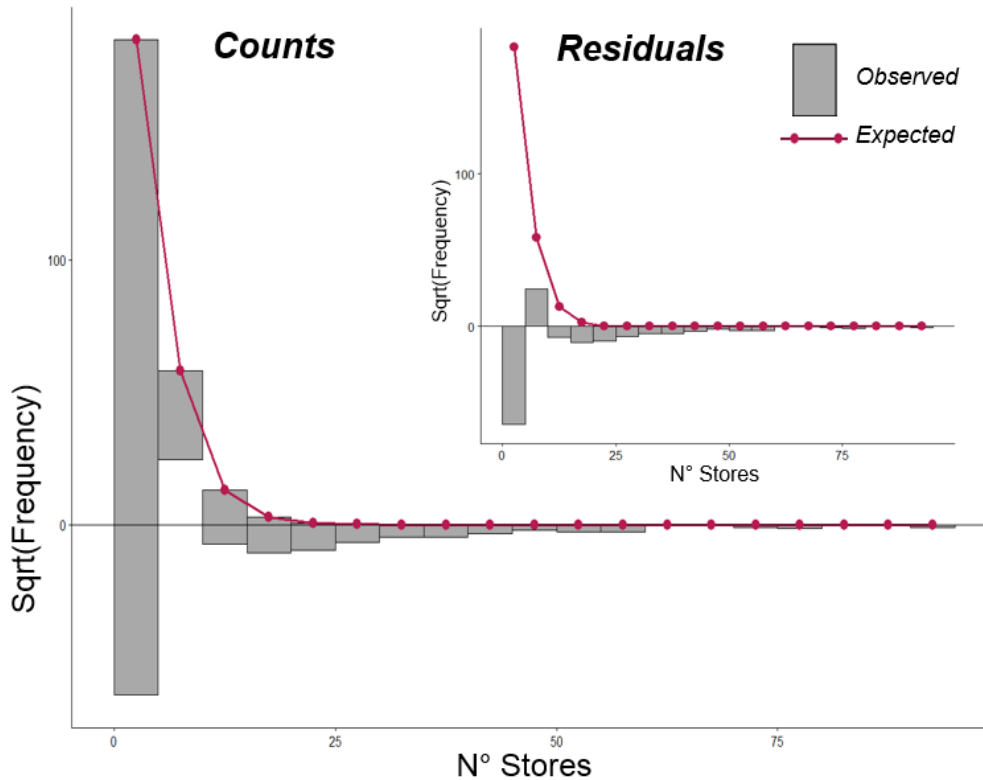


Figure 101 Rootgram : a graphical approach to assess the fit of count regression models.

**Additional performance measures for count regression models: Count REC curves**

Another relevant issue in our case study is related to the ability of AIC BIC and loglikelihood measures to allow comparison of models only if those have the same underlying dataset (same variables and same features). How might one compare the result from the model outcomes on the overall space study with the composite outcomes obtained from the n models implemented in each subspace? In other words, how does one evaluate if the regression by groups performs better than a global model? The only solution consists in comparing a series of indicators based on the prediction and the errors of the models.

Beyond different possibilities of measures of the goodness of fit, we could also observe and compare the traditional measures of dispersion of the residuals (mean absolute and standard deviation) for each model as well as, separately for its zero and count part, respectively implemented as:

$$Mean\ abs\ error = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{n} ; St.\ Dev.\ Error = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{n - 1}} \quad (11.23)$$

But the most straightforward question the reader might want to know about each model performance is: *What proportion of the observations is correctly estimated by the model?* Which, in our case study might be reformulated as “*What proportion of the streets are predicted by the model with exactly the right (observed) amount of stores?*”

The answer to this question might be found in the aforementioned Count pseudo  $R^2$  measure which informs us about the proportion of correctly estimated observations.

If we adopt the traditional approach of model performance evaluation in multi-class classification problems, we could summarise the results of our regression models through confusion matrix CM (or error matrix). The square matrix  $N \times N$  consists of columns and rows that list the number of instances as absolute (or relative) observed values (on the top axis) vs. predicted values (on the left axis).

$$CM = \begin{matrix} & \text{Observed } j \\ \text{Fitted } i & \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix} \end{matrix} \quad (11.24)$$

With  $N = \max(\text{observed values } j, \text{fitted values } i)$

Each element  $c_{ij}$  represents the number of elements (streets) with  $j$  observation value (number of stores observed) and  $i$  predicted values (number of stores predicted by the model). We can thus observe that the Count pseudo  $R^2$  ( $C$ ) corresponds to the performance indicator called *Accuracy* as the ratio between the sum of the diagonal elements and the total number of elements in CM (Figure 103).

$$C = \text{Accuracy} = \frac{\sum_{i,j=0}^N c_{i=j}}{\sum_{i,j=0}^N c_{ij}} \quad (11.25)$$

From the same approach, two other measures of model performance might be implemented.

*Sensitivity* (or *Recall*) is the true positive rate and answers the question: *What is the proportion of streets correctly captured by the model?*

$$S_j = \text{Sensitivity } j = \frac{c_{i=j}}{\sum_{i=0}^N c_{ij}} \quad (11.26)$$

*Precision* represents the correct positive prediction divided by the total number of predictions and answers to the question: *What is the proportion of correctly predicted streets? Or How well is the model predicting store counts?*

$$P_i = \text{Precision } i = \frac{c_{i=j}}{\sum_{j=0}^N c_{ij}} \quad (11.27)$$

Considering that one of the main aspects we are interested in analysing is the ability of each modelling approach to capture the overdispersion and the zero-inflation behaviours, we can focus our attention on the model performance when considering zeroes and counting parts individually (graphical illustration in Figure 103.c). From the 11.26 and 11.27, we redefine Precision and Sensitivity of zeroes predictions ( $S_z, P_z$ ) as:

$$S_z = \frac{c_{00}}{\sum_{i=0}^N c_{i0}} \quad P_z = \frac{c_{00}}{\sum_{j=0}^N c_{0j}} \quad (11.28)$$

And similarly, Precision and Sensitivity of the counting part of the model predictions ( $S_c, P_c$ ) as:

$$S_c = \frac{\sum_{i,j=1}^N c_{i=j}}{\sum_{i=0,j=1}^N c_{ij}} \quad P_c = \frac{\sum_{i,j=1}^N c_{j=i}}{\sum_{i=1,j=0}^N c_{ij}} \quad (11.29)$$

A simpler way to visualize how to obtain  $C, S_z, P_z, S_c, P_c$  is shown in the following schematic representation.

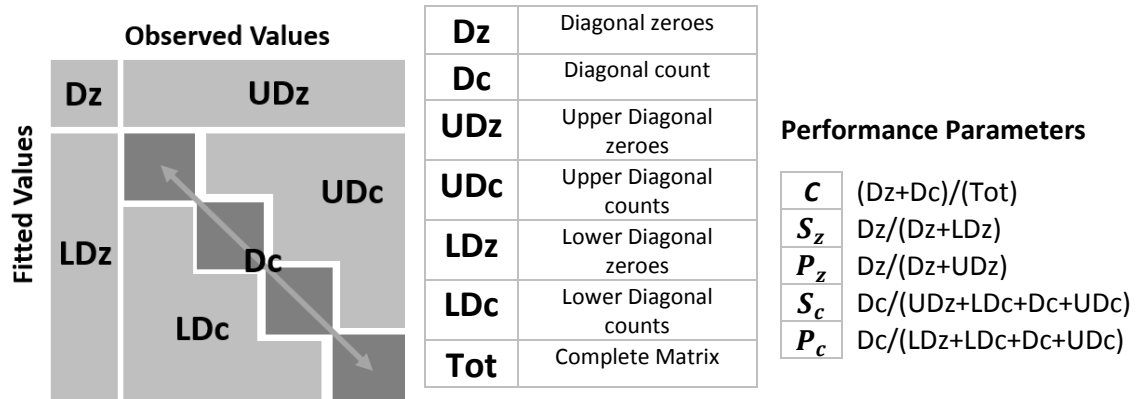


Figure 102 A simple schematic representation for the implementation of Performance Parameters using the Confusion Matrix representation of a regression model outcomes.

From the aforementioned measures, we might also implement the F1 score formula which has the advantage of simultaneously evaluating both precision and sensitivity between count and zero parts as:

$$F_1 = 2 * \frac{S*P}{S+P} \quad (11.30)$$

Measures of Accuracy, Sensitivity and Precision, are traditionally conceived for the analysis of the values disposed exclusively on the CM diagonal, while errors of misclassification are equally considered independently by their importance. In regressive models, differently from classificatory approaches, different importance is given to prediction errors. In the CM representation, values with small errors are disposed close to the diagonal while values with high errors are more distant. When analysing regressive models outcomes, we could study how performance parameters vary as function of the absolute deviation  $|e|$ . In other terms this means to include the information about the dispersion of occurrences relatively to the CM diagonal within performance parameters (as illustrated in Figure 103.d).

From 11.25-29 we then obtain:

$$C(|e|) = \frac{\sum_{i,j=0}^N c_{i=j\pm|e|}}{\sum_{i,j=0}^N c_{ij}} \quad (11.31)$$

$$S_z(|e|) = \frac{\sum_{i,j=1}^N c_{i=j\pm|e|}}{\sum_{i=0}^N c_{i0}} \quad P_z(|e|) = \frac{c_{00}}{\sum_{j=0}^N c_{0j}} \quad (11.32)$$



$$S_c(|e|) = \frac{\sum_{i,j=1}^N c_{i=j\pm|e|}}{\sum_{i=0,j=1}^N c_{ij}} \quad P_c(|e|) = \frac{\sum_{i,j=1}^N c_{j=i\pm|e|}}{\sum_{i=1,j=0}^N c_{ij}} \quad (11.33)$$

Plotting the function  $C(|e|)$  versus the error tolerance  $|e|$  on the x-axis, we obtain what Bi and Bennet (2003) called the Regression Error Characteristic curve. The REC curve represents a development of the Receiver Operating Characteristic curve (ROC, Egan 1975) for regression models which estimates the cumulative distribution function (CDF) of the error. Considering that every model is characterised by a certain degree of inaccuracy, REC curves helps in visualizing the trade-off between accuracy and error (or tolerance): for values of tolerance close to zero, we obtain the proportion of occurrences perfectly predicted while for values close to N-1 (the maximum tolerance), all occurrences are considered accurate (Bi and Bennet 2003). The same notions might be implemented considering  $S_z(|e|)$ ,  $P_z(|e|)$ ,  $S_c(|e|)$ ,  $P_c(|e|)$  functions.

Among the several advantages reported by the same authors, REC facilitates the visual comparison of the expected performances of different regression model approaches, revealing additional information for the model selection task.

To our knowledge, this work represents the first application of the REC curve approach to the specific case of Count Regression modelling. We name Count-REC the set of functions composed by the traditional accuracy function  $C(|e|)$  together with the four additional functions  $S_z(|e|)$ ,  $P_z(|e|)$ ,  $S_c(|e|)$ ,  $P_c(|e|)$ .

In the specific context of this research, these curves might answer to the following questions: *What is the accuracy (or sensitivity, or precision) of the overall (or zero part, or count part) model M in predicting the exact number of store streets with a tolerance level of e?*

Count-REC curves are not only an extension of REC: this approach reveals its novelty and usefulness when comparing ZI and non-ZI models. The Precision and Sensitivity of both zero and count parts might be individually visualized as functions of the tolerance for different count data modelling approaches (namely P, NB, ZIP, ZINB, ZAP, ZANB). Considering AIC and BIC, it is evident how Count-REC curves allow a finer appreciation and comprehension of the benefits (downsides) of choosing a ZI or non-ZI model.

While zero inflation and overdispersion have been addressed by the aforementioned approaches, a measure of over/under dispersion is still missing. Another synthetic approach might be implemented considering the ration between the Accuracy Count-REC curve (from 17.30) of only positive and negative tolerance levels as follow:

$$Ov./Un. Dispersion Ratio = \frac{C(|e^+|)}{C(|e^-|)} = \frac{\sum_{i,j=0}^N c_{i=j-|e|}}{\sum_{i,j=0}^N c_{i=j+|e|}} \quad (11.34)$$

Where the numerator and denominator of the 11.34 correspond to the accuracy at different tolerance thresholds applied to the CM diagonal plus the CM lower and upper diagonal (see Figure 104 for a schematic representation). When this ratio is more than one, overdispersion

of the model error is detected; on the contrary, under dispersion is characterised by values less than one. As for the Count-REC curves, for each model M involved in the model selection task, we can graphically plot this ratio as a function of the tolerance, allowing a simple comparative visualisation tool.

Fig.103 shows some representation of Count-REC curves; Fig.104 proposes a schematic illustration of model performance parameters derived from a confusion matrix representation of its outcomes.

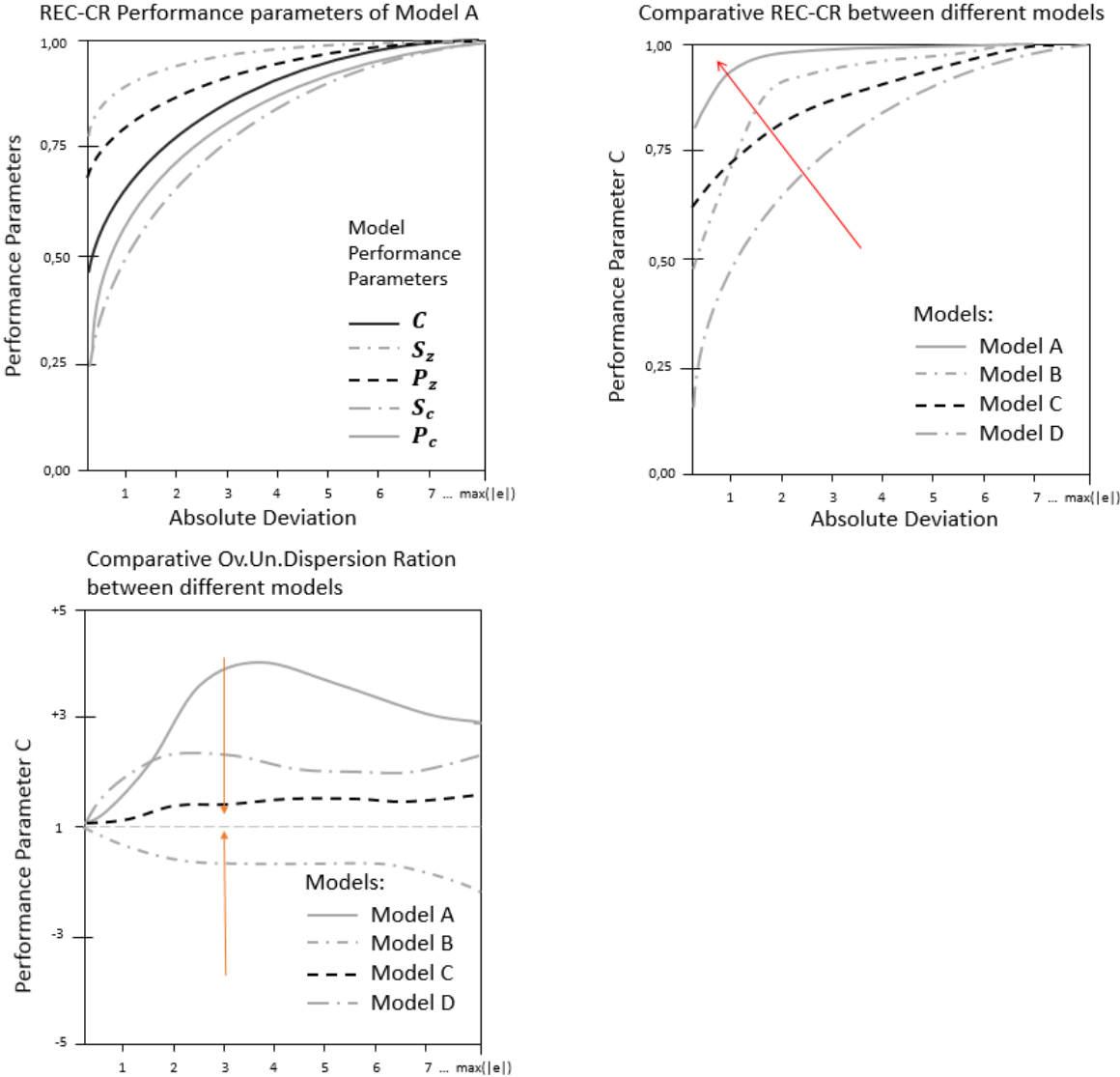


Figure 103 Schematic example of the REC-CR curves implement for the evaluation of the goodness of fit of the models.

Summarising, in this dissertation, the model assessment and selection process is based on both numerical measures and graphical approaches: 1) ICs values and loglikelihood tests; 2) a vector of twelve measures of performance describing each model outcomes (Mean Abs Error,

Standard Error, C, P, S, F1 score); 2) rootgrams; 3) a set of Count-REC curves able to assess the dispersion of the predictions as a function of the error.

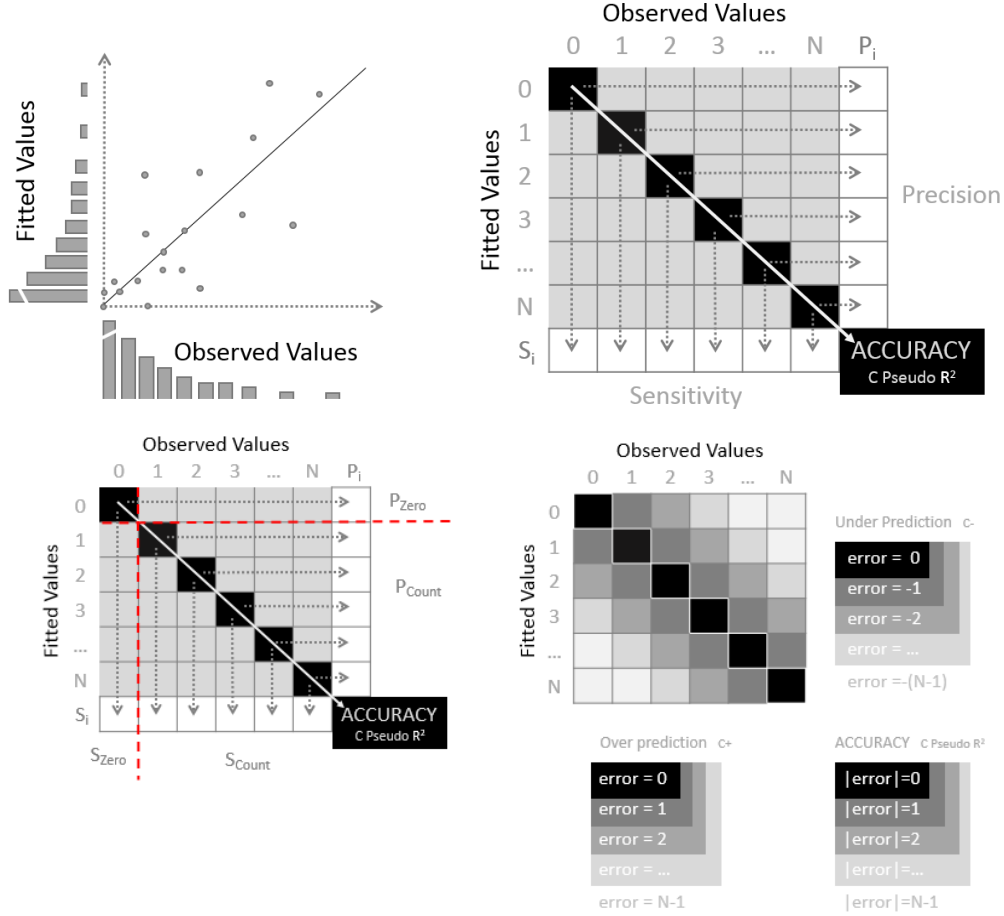


Figure 104 Schematic illustration of model performance parameters derived from a confusion matrix representation of its outcomes.

## 12. Models of Retail Distribution on the French Riviera

### 12.1 Counting model selection

The statistical procedures discussed and identified in Chapter 11 have been implemented in the study area of the French Riviera. As presented in the previous sections, our study region encompasses eighty-eight municipalities, which encompass five large and average-sized cities and a large number of small cities and villages. The contemporary evaluation of a large number of cities in the same statistical analysis allows us to overcome the risk of idiosyncratic outcomes describing the specific characteristics of individual cities; conversely, all the results discussed in this chapter might be seen as common characteristics associated with several urban areas within the context of the Mediterranean coastal region of the French Riviera.

In a first phase of the analysis, we focused on the model selection task with the final goal being to understand which regression model is the most appropriate to describe the relationship between retail and urban form variables on the French Riviera. The seven regression models previously described (G, P, NB, ZIP, ZINB, ZAP, ZANB) are implemented on the overall space, on the two morphological sub-regions (First and Second-Age city), and six of the nine urban fabrics (UF1-6) for a total of sixty-three models.

In this section, we temporarily overlook the role of individual regressors; as previously discussed, the inevitable multicollinearity in our models might influence the coefficients associated with each explanatory variable, preventing their accurate interpretation. However, the overall model selection criteria and predictions are not influenced by multicollinearity problems. For these reasons, in this section we focus only on the assessment of models, measuring their quality level as explained in Section 17.3; once the best models are identified and selected, in the next section we further analyse the role of regressors.

In the next three tables (Tab.17-19), we provide to the reader an overview of the criteria used for the model selection analysis.

In Table 17, we reported the descriptors of the sixty-three models; for each of them, the following information is provided: the **AIC**, **-2loglikelihood**, the number of features  $n$  (street) and the number of parameters  $c$  (number of variables + number of parameters of the model). The *best model* might be found in the correspondence of the lowest AIC value (AIC scores are highlighted in red-high values to green-low values).

We remind the reader that the implementation of the linear regression through the GLM approach with Gaussian function (instead of the traditional OLS approach) allows for obtaining the correspondent loglikelihood and therefore AIC criteria for a straightforward comparative analysis with the other count regression models.

The first observation is that the highest values of AIC are found, as expected, in correspondence of linear (G) and Poissonian (P, ZIP, ZAP) regression models. The inappropriateness of these models (respectively explained by the non-normality and to the

over-dispersion of the dependent variable, Section 16.3) is here empirically confirmed by our data independently of the spatial region under analysis.

Lower AIC scores are found for NB, ZINB and ZANB regressions; in the overall study area and in every sub-space, ZINB always shows the lowest AIC values: depending on the specific region, we might observe AIC scores for ZINB from 0.64 to 3.75% lower than ZANB, and from 0.54 to 3.29% lower than NB.

Non-nested Vuong tests were used to further compare the models and assess the statistical significance of the apparent superiority of ZINB model (Section 11.4). For each spatial region, we implemented the Vuong statistical test starting from the model with the lowest AIC value towards the ones with highest values; the results are shown in Table 18. As we might observe, the statistically significant superiority of ZINB is confirmed in every region with the exception of UF5 and UF6. In these two specific cases,  $p > 0.05$  when comparing AIC values of ZINB and ZANB (if we consider  $p > 0.01$  as a threshold the same outcome might be outlined for UF4).

These observations are further confirmed when plotting the relative rootgrams. In Fig.105 we might observe rootgrams for the seven models implemented on the overall study area; similar behaviours are observed for every morphological sub-region. From Fig.105, we might observe how linear (G) and Poisson regression models (P, ZIP, ZAP) do not account for overdispersion, contrary to all Negative Binomial regression models (NB, ZINB ZANB).

Model selection								
		G	P	NB	ZIP*	ZINB**	ZAP*	ZANB*
Global	AIC	284770,9	120492	90024,06	100454,9	87373,54	101170,5	88247,1
	-2log-likelihood	284558	120282	89812	100034	86956	100750	87824
	n	63071	63071	63071	63071	63071	63071	63071
	c	105+2	105+1	105+2	210+2	210+3	210+2	210+3
					**	**	*	*
First	AIC	77041	47955,88	36477,71	41594	35451	41729	35869
	-2log-likelihood	76828	47696	36264	41150	35036	41280	35450
	n	14143	14143	14143	14143	14143	14143	14143
	c	101+2	101+1	101+2	202+2	202+3	202+2	202+3
					*	**	*	*
Second	AIC	171475	68000	49808	54969	48841	55431	49198
	-2log-likelihood	171260	67862	49616	54522	48416	54976	48746
	n	43969	43969	43969	43969	43969	43969	43969
	c	102+2	102+1	102+2	204+2	204+3	204+2	204+2
					**	*	*	*
UF1	AIC	22254	10816	9405	9883	9137	9968	9480
	-2log-likelihood	22078	10774	9280	9580	8968	9662	9164
	n	5506	5506	5506	5506	5506	5506	5506
	c	93+2	93+1	93+2	186+2	186+3	186+2	186+3
					**	**	*	*
UF2	AIC	47155	37145	27049	31622	26158	31721	26352
	-2log-likelihood	46948	36914	26866	31218	25786	31308	25976
	n	7954	7954	7954	7954	7954	7954	7954
	c	94+2	94+	94+2	188+2	188+3	188+2	188+3
					*	*	*	*
UF3	AIC	22097	9976	9025	9161	8909	9267	9058
	-log-likelihood	21890	9766	8830	8760	8522	8890	8662
	n	7108	7108	7108	7108	7108	7108	7108
	c	93+2	93+1	93+2	186+2	186+3	186+2	186+3
					*	**	*	*
UF4	AIC	48306	25602	16408	19036	16023	19154	16135
	-2log-likelihood	48108	25340	16186	18598	15622	9357	15730
	n	10259	10259	10259	10259	10259	10259	10259
	c	94+2	94+	94+2	188+2	188+3	188+2	188+3
					*	*	*	*
UF5	AIC	36232	15636	13765	14156	13691	14267	13778
	-2log-likelihood	36138	15462	13596	13800	13316	13946	13458
	n	16453	16453	16453	16453	16453	16453	16453
	c	86+2	86+1	86+2	86+2	86+3	86+2	86+3
					*	*	*	*
UF6	AIC	16590	7186	6721	6793	6638	6883	6800
	-2log-likelihood	17878	7358	6720	6714	6586	6792	6658
	n	8889	8889	8889	8889	8889	8889	8889
	c	86+2	86+1	86+2	86+2	86+3	86+2	86+3
					*	*	*	*

Table 1 Model selection: AIC-Akaike Information Criteria; c-number of parameters in the model; n-feature number; G-Gaussian; P-Poisson; NB-Negative Binomial; ZI-Zero Inflated; \* and \*\* indicate that AIC has been obtained with, respectively, logit and probit functions.

	G vs.P	P vs. NB	NB vs.ZIP	ZIP vs.ZINB	ZINB vs.ZAP	ZINB vs.ZANB	ZINB* vs.ZINB**	
<b>Global</b>	Vuong test statistic	-89,546	-21,39	13,692	-16,887	17,724	9,313	w2=0
	p	< 0.22e-15	< 0.22e-15	< 0.2e-15	< 0.22e-15	< 0.22e-15	< 0.2e-15	0,5
	Preferable Model	P	NB	ZIP	ZINB	ZINB	ZINB	-
<b>First</b>	Vuong test statistic	-37,51	-19,416	11,817	-18,483/-15,244***	15,491	7,93	0
	p	< 0.22e-15	< 0.2e-15	< 0.2e-15	< 0.22e-15	< 0.2e-15	< 0,713e-12	0,5
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB	-
<b>Second</b>	Vuong test statistic	-66,592	-14,391	8,313/9,084***	-14,731/-11,954***	11,29	4,46	-8,854
	p	< 0.22e-15	< 0.22e-15	< 0.2e-15	< 0.74e-8	< 0.22e-15	< 1,296e-6	< 0.22e-15
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB	-
<b>UF1</b>	Vuong test statistic	-40,545	-9,674	3,179/5,30***	-9,66/-3,49***	7,186	4,663	0
	p	< 0.22e-15	< 0.22e-15	0,7e-3/0,5e-7	< 0.2e-15/0,24e-3	3,34E-14	1,56E-06	0,5
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB	/
<b>UF2</b>	Vuong test statistic	-13,339	-17,263	10,96/11048***	-16,99/13,79***	14,338	5,633	0
	p	< 0.22e-15	< 0.22e-15	< 0.22e-15	< 0.22e-15/1,1e-3	< 0.22e-15	8,85E-09	0,5
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB	/
<b>UF3</b>	Vuong test statistic	-49,992	-9,183	-1,04/15,72	-5,59	6,9	4,357	0,385
	p	< 0.22e-15	< 0.22e-15	0,0856/ < 0.22e-15	< 0.22e-15	4,42E-12	2,18E-05	0,5
	Preferable Model	P	NB	ZIP/NB	ZINB	ZINB	ZINB	/
<b>UF4</b>	Vuong test statistic	-22092	-10,486	6,835/7,436	-8,246	8,558	2,103	0,001
	p	< 0.22e-15	< 0.22e-15	0,4e-12/0,5e-15	< 0.22e-15	< 0.22e-15	0,0178	0,987
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB	/
<b>UF5</b>	Vuong test statistic	-52	-8,314	2,19/11,74***	-9,16	6,505	1,97	0
	p	< 0.22e-15	< 0.22e-15	0,0153/0,45e-5	< 0.22e-15	3,87E-11	0,9756	0,5
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB/ZANB	/
<b>UF6</b>	Vuong test statistic	-43,096	-6,158	-0,187/15,13	-5,8/15,64	1,95/15,64	1,85	1,2
	p	< 0.22e-15	3,69E-10	0,425/0,22e-15	2,06e-9/0,22e-15	2,3e-2/0,22e-15	0,107	0,107
	Preferable Model	P	NB	NB	ZINB	ZINB	ZINB/ZANB	/

Table 2 Results of the Vuong test between our seven models for the overall study area and each morphological region.

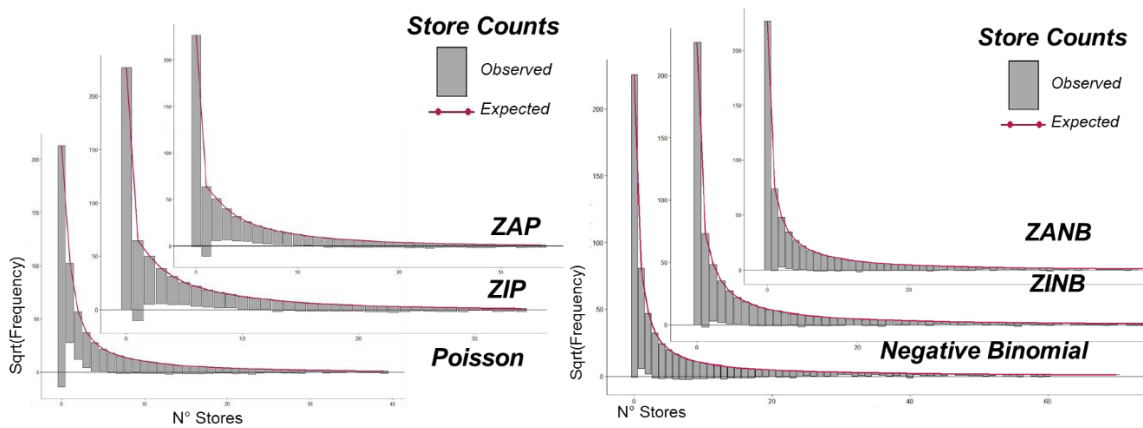
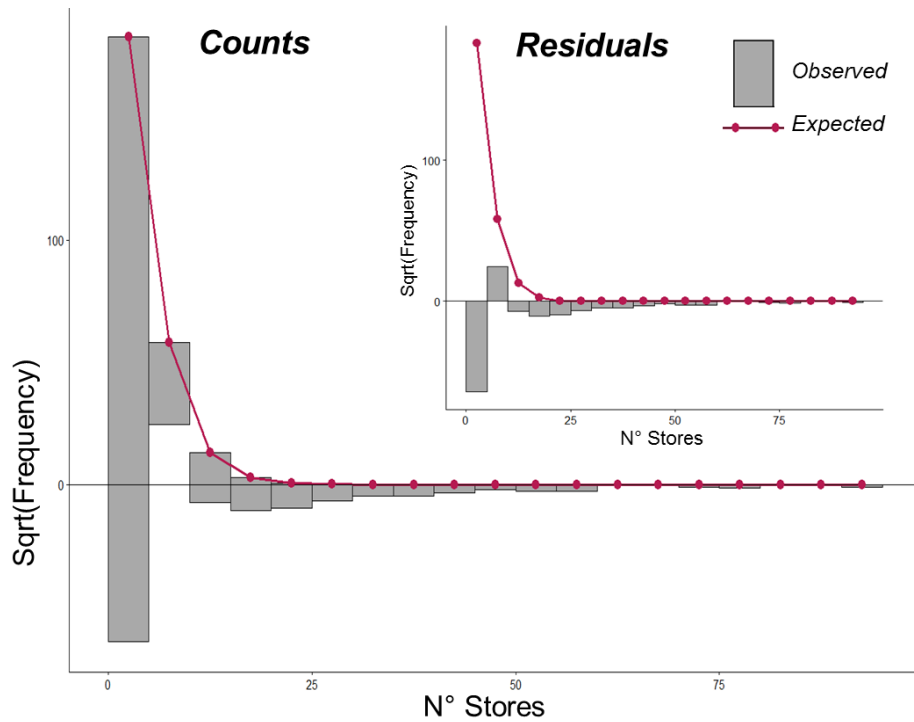


Figure 105 Rootgram of the Global model (overall space study and variables) for: Linear model (top), Poissonian Regressions P-ZIP-ZAP (bottom-left) and Negative Binomial regressions NB-ZINB-ZANB (bottom-right). The last group of models shows a better fit between expected and observed values.

Before proceeding with the phase of variable selection, we might already evaluate how the separation of the study area in subspaces might improve the fit values compared to the overall model. As previously discussed, AIC cannot be implemented for this task since different numbers of features underlie each modelled region. For this reason, we use the set of thirteen measures previously described.

In Table 19, we reported the 13 variables implemented for each study region; in the top part are the results for NB while in the lowest part are the results for ZINB.

The regression on the overall study area with all the morphological variables (with a number of regressors varying from a maximum of 105 to a minimum of 86) allows us to identify with



exact precision the retail distribution of  $C=70.17$  and  $C=71.16\%$  of the overall streets respectively with the NB and ZINB model. Higher values of Sensitivity and Precision are observed for those streets without stores, respectively  $S_0=84.3\%$  and  $P_0=90.4\%$  in NB and  $S_0=83.9\%$  and  $P_0=90.7\%$  in ZINB. When considering streets with non-zero count of stores, sensitivity and precision assume  $S_c=36.4$  and  $P_c=28.1\%$  in NB while  $S_c=35.6\%$  and  $P_c=26.8\%$  in ZINB. F1 score is 0.531 for NB and 0.536 for ZINB. The average overall error shows lower values for ZINB than NB. These same measures are reported for every subspace.

What we might observe is that the lowest accuracy and sensitivity values are found in more central compact fabrics, while more peripheral urban fabrics show higher values; on the contrary, the precision of the count part of the models drops in peripheral urban fabrics.

When implementing the same models separately on First and Second-Age city, we improve the accuracy of the overall model of only  $+0.38\%$  and  $+0.98\%$ ; while using the 6 UFs, it improves of  $+0.52\%$  in ZINB, while a reduction of  $-0.15\%$  in NB is detected. Despite the decomposition of the overall study area, it does not seem to affect the overall predictability of the model; what we should observe is that different values are observed within each sub-regions: while UF1 and UF2 and UF3 show higher improvement values with respectively  $+4.66\%$ ,  $+11.03\%$  and  $+2.55\%$ , on the contrary, UF4 and UF5 shows an almost null variation, while a negative value of  $-0.36\%$  is associated to UF6. When observing the separate sensitivity and precision values for count and zero parts, we might observe an overall growth in precision and a loss in sensitivity in the counting parts (especially in ZINB models). Moreover, both precision and sensitivity for the zero parts improve with the exception of UF6 and 7.

We might conclude that the decomposition of the study area in morphological subspaces (both UF and morphological macro-regions) improve the goodness of fit for central traditional areas, while, on the contrary, semi-peripheral and peripheral regions seem to be penalised. Since street elements in peripheral regions outnumber those of more central areas, when evaluating the combined results of sub-models (First/Second-Age City and UF1-6), the model improvements achieved in central areas are diluted and reduced to modest values.

NB	C		Sz		Sc		Pz		Pc		F1		E(T)		E(Tz)		E(Tc)		Sd(T)		Sd(Tz)		Sd(Tc)		
	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val	±[%]	val
Global	0.717		0.843		0.171		0.904		0.132		0.531		0.629		0.221		2,414		2,040		0.723		4,015		
Global*	0.698		0.828		0.174		0.899		0.134		0.536		0.675		0.231		2,435		2,104		0.720		4,055		
Global(F+S)	0.701	0.38	0.831	0.39	0.175	0.25	0.899	0.03	0.134	1.27	0.539	0.50	0.670	-0.68	0.237	-1.50	2,426	-0.36	2,118	0.64	0.724	0.59	4,089	0.85	
First(Global)	0.460		0.620		0.185		0.869		0.124		0.675		1.506		0.617		3,028		3,085		1.229		4,420		
First	0.470	2.23	0.633	2.09	0.191	3.01	0.874	0.55	0.130	4.26	0.683	1.17	1.471	-2.40	0.617	-6.07	2,992	-1.18	3,102	0.54	1.165	-5.50	4,481	1.36	
Second(Global)	0.775		0.877		0.165		0.904		0.140		0.403		0.407		0.152		1,941		1,580		0.491		3,650		
Second	0.775	0.01	0.877	0.10	0.161	-2.67	0.904	-0.06	0.137	-1.90	0.402	-0.46	0.412	1.28	0.156	2.53	1,955	0.69	1,600	1.26	0.540	9.08	3,666	0.43	
Global*	0.705		0.836		0.173		0.899		0.135		0.543		0.658		0.229		2,399		2,026		0.704		3,869		
MFA(6)	0.711	0.84	0.843	0.78	0.178	2.66	0.902	0.29	0.141	4.11	0.555	2.22	0.652	-0.90	0.222	-3.24	2,394	-0.17	2,083	2.71	0.740	4.83	3,987	2.95	
UF1(Global)	0.619		0.756		0.211		0.879		0.148		0.570		0.720		0.331		1,883		1,559		0.730		2,510		
UF1	0.647	4.32	0.785	3.68	0.236	10.74	0.888	0.93	0.175	15.34	0.599	4.82	0.663	-8.60	0.273	-21.21	1,829	-2.97	1,550	-0.57	0.634	-15.09	2,563	2.09	
UF2(Global)	0.288		0.392		0.176		0.872		0.111		0.725		2.263		1.079		3,532		3,867		1.593		5,016		
UF2	0.304	5.46	0.425	7.73	0.175	-0.45	0.875	0.34	0.113	1.99	0.733	1.13	2.224	-1.75	0.982	-9.82	3,555	0.64	4,063	4.81	1.468	-8.48	5,336	6.00	
UF3(Global)	0.753		0.876		0.199		0.881		0.193		0.461		0.385		0.132		1,526		1,003		0.368		1,831		
UF3	0.761	1.02	0.884	0.91	0.205	3.03	0.886	0.55	0.203	4.57	0.485	4.91	0.387	0.51	0.145	8.98	1,477	3.26	1,106	9.28	0.708	47.99	1,742	-5.14	
UF4(Global)	0.610		0.702		0.228		0.897		0.120		0.456		0.779		0.371		2,481		2,319		0.712		4,701		
UF4	0.619	1.48	0.715	1.88	0.220	-3.90	0.896	-0.05	0.119	-0.41	0.461	1.00	0.779	-0.01	0.373	0.62	2,471	-0.41	2,268	-2.23	0.834	14.65	4,489	-4.72	
UF5(Global)	0.873		0.967		0.096		0.906		0.217		0.230		0.185		0.035		1,444		0.686		0.190		1,523		
UF5	0.869	-0.49	0.962	-0.53	0.099	2.86	0.908	0.15	0.197	-10.29	0.244	5.80	0.194	4.57	0.045	23.26	1,439	-0.32	0.737	6.92	0.356	46.66	1,510	-0.90	
UF6(Global)	0.888		0.975		0.092		0.915		0.230		0.241		0.154		0.026		1,336		0.547		0.163		1,118		
UF6	0.889	0.09	0.972	-0.26	0.123	25.23	0.919	0.40	0.265	13.20	0.287	15.99	0.152	-1.26	0.028	9.69	1,291	-3.46	0.552	0.97	0.171	4.54	1,182	5.38	
ZINB	0.716		0.839		0.178		0.907		0.134		0.536		0.624		0.228		2,350		1,997		0.707		3,943		
Global	0.697		0.824		0.181		0.902		0.134		0.541		0.668		0.249		2,370		2,067		0.727		3,981		
Global*	0.704	0.98	0.834	1.15	0.177	-2.21	0.902	-0.03	0.136	1.13	0.547	1.06	0.662	-0.89	0.239	-4.02	2,380	0.39	2,091	1.15	0.751	3.16	4,018	0.93	
Global(F+S)	0.456		0.610		0.192		0.878		0.126		0.678		1.476		0.639		2,908		2,995		1.231		4,298		
First(Global)	0.484	5.78	0.654	6.78	0.192	-0.10	0.882	0.37	0.133	5.27	0.696	2.58	1.434	-2.93	0.574	-11.34	2,906	-0.09	2,998	0.08	1.248	1.32	4,273	-0.58	
First	0.475		0.643		0.172		0.906		0.143		0.411		0.408		0.156		1,922		1,575		0.501		3,637		
Second(Global)	0.775	0.01	0.876	0.13	0.165	-3.86	0.905	-0.07	0.139	-2.84	0.409	-0.56	0.414	1.42	0.160	2.08	1,944	1.10	1,624	3.03	0.535	6.43	3,750	3.02	
Second	0.703		0.831		0.180		0.902		0.137		0.547		0.652		0.237		2,332		1,985		0.706		3,792		
Global*	0.714	1.65	0.845	1.60	0.186	3.24	0.907	0.50	0.146	6.34	0.569	3.90	0.634	-2.80	0.224	-6.02	2,293	-1.68	1,967	-0.88	0.725	2.57	3,739	-1.41	
MFA(6)	0.628		0.756		0.246		0.888		0.171		0.584		0.685		0.301		1,833		1,512		0.615		2,496		
UF1	0.659	4.66	0.793	4.62	0.259	5.03	0.903	1.74	0.190	10.08	0.631	7.38	0.650	-5.42	0.261	-15.21	1,811	-1.20	1,649	8.35	0.621	0.95	2,812	11.22	
UF1(Global)	0.288		0.400		0.167		0.876		0.106		0.728		2.209		1.102		3,395		3,747		1.563		4,874		
UF2	0.328	12.23	0.479	16.48	0.166	-0.94	0.885	1.02	0.111	4.84	0.475	2.89	2.113	-4.55	0.995	-10.80	3,312	-2.53	3,651	-2.65	1.652	5.38	4,681	-4.12	
UF2(Global)	0.743	2.56	0.881	2.63	0.225	1.36	0.892	0.40	0.216	9.91	0.503	5.50	0.394	-3.44	0.151	-3.78	1,443	-3.28	1,000	-1.27	0.510	23.76	1,722	-8.76	
UF3(Global)	0.611		0.702		0.230		0.900		0.120		0.463		0.787		0.389		2,448		2,319		0.778		4,677		
UF4	0.615	0.76	0.707	0.62	0.236	2.56	0.902	0.23	0.124	3.20	0.469	1.24	0.777	-1.25	0.379	-2.65	2,439	-0.35	2,319	-0.02	0.756	-2.87	4,690	0.28	
UF4(Global)	0.867		0.958		0.108		0.907		0.204		0.239		0.191		0.043		1,427		0.686		0.206		1,527		
UF5	0.868	0.13	0.959	0.09	0.112	3.05	0.910	0.27	0.204	0.05	0.266	10.34	0.192	0.41	0.045	4.70	1,417	-0.72	0.690	0.51	0.241	14.31	1,510	-1.15	
UF5(Global)	0.886		0.971		0.100		0.916		0.225		0.248		0.156		0.029		1,326		0.546		0.169		1,119		
UF6	0.885	-0.03	0.967	-0.48	0.140	28.69	0.922	0.66	0.253	10.95	0.317	21.91	0.156	-0.14	0.035	16.85	1,270	-4.42	0.583	6.37	0.193	42.18	1,320	15.25	

Table 3 Comparison of results of NB and ZINB models when global and sub-region are evaluated on the same subgroup of features. For each goodness of fit measure the raw values and the percent change between the modelsimplified on the overall space study each specific subregion.

(C-Accuracy; Sc-Sensitivity count part; Sz- Sensitivity zero part; Pc-Precision count part; Pz-Precision zero part; F1 score; E(T)-Average Tolerance; E(Tz)-Average Tolerance Zero Part; E(Tc)-Average Tolerance count part; Sd(T)-Standard Deviation Tolerance, Sd(Tz)-Standard Deviation Tolerance Zero Part; Sd(Tc)-Standard Deviation Tolerance Count Part)

Projecting the model (and sub-model) outcomes in the geographical space might allow the analyst to observe the spatial characteristics of the predictive power of the models: without exploring these spatial representations of the model outcomes, we might not be able to detect possible limitations of the models precluding important observations for future research developments. For example, we might detect specific patterns of residuals suggesting the omission of important variables. Moreover, this task becomes even more important in our case study since different models have been studied for each morphological subregion.

In the following figures, we illustrate in the geographical space (zoomed in on the city of Nice) the observed distribution of retail, the predicted values and their error (Fig.106); from the left to the right we might observe the global model, First/Second-Age Cities composite model, and the Urban Fabrics composite models.

Despite the set of goodness-of-fit measures previously described were indicating higher performance values associated with the combined models when compared to the original global model, these differences are hardly detectable from Fig.106. We might observe a general underprediction on the hilly neighbourhoods surrounding the city centre and along the coast, despite the inclusion of a specific configurational indicator describing the directional configurational centrality toward the coastline.

Only expert-based knowledge of the study area might allow us to better understand and explain the underlying reasons for specific over/under prediction values. For instance, we might observe in every model underprediction for the pedestrian area of Rue Massena and the antique shop cluster along Rue Catherine Ségurane on the west side of the Port Lympia (highlighted in Fig.106 as respectively point *a* and *b*). The first example might suggest the need to further consider, in future studies, the role of pedestrianization<sup>95</sup>, while the second might support the notion of historical inertia of retail agglomeration where streets which were suitable for retail activity in the past might still be active nowadays despite the disadvantaged location of the contemporary conditions (Sevtsuk 2010, p.120).

An example of overprediction in every model might be observed along the streets surrounding the Esplanade Kennedy (point *c* in Fig.106). Despite the proximity to open spaces being traditionally an attractive characteristic for retail distribution (as we further discuss in the 12.2, 12.3), Esplanade Kennedy is located at the crossing point of several main axes and urban routes, transforming the esplanade into an oversized roundabout. This observation might suggest the need to include further indicators describing the repulsive power of large transport infrastructures and/or high-traffic areas.

These specific examples demonstrate the usefulness of the geographical projection of model outcomes not only to better understand the model and the area under study, but also as a starting point to elaborate new working hypotheses to develop in future research.

---

<sup>95</sup> Pedestrianization is an urban design solution which is traditionally implemented in urban spaces where a naturally high concentration of stores is already observed and explained by other locational factors; nonetheless, as highlighted by our models, we might suggest that pedestrianization stores could further increase the economic potential of a street. Specific attention should be given to this aspect, considering the diachronical evolution of retail presence before and after the pedestrianization of a street.

The aforementioned examples have been selected from those streets where the same under/over prediction outcomes have been detected independently of the specific model: global, First/Second, and Urban Fabric composite models were always over/under predicting the retail presence. Nonetheless, we have not been able yet to identify the overall differences between these models; one possible solution is to compare the differences between the predicted values of the three models. In Figure 107, we plot in the geographical space the differences between: a) Global and First/Second-Age City composite model, b) Global and Urban Fabric composite model, c) First/Second- Age city and Urban Fabric composite model. These maps allow us to better visualise the impacts of the introduction of morphological spatial partitions in our models: First/Second-era city partition triggers important modifications of prediction in central areas, while adjustments interest both central and peripheral areas when the Urban Fabric partition is considered. Nonetheless, from these maps, nothing can be asserted about the correctness of these predictions (although we already discussed in the previous part the improvement of the overall goodness-of-fit parameters).

The last aspect of these models to be considered is the analysis of the residual autocorrelation. As discussed in Section 10.3, the analysis of the spatial autocorrelation of our model residuals might allow us to distinguish between the autocorrelation properly associable to the retail phenomena and the part explained by the structure of the underlying urban form.

The graph in Fig.108 shows the correlogram of the retail distribution (from Fig.100, Section 10.3) and of the residuals of the Global model, First/Second, and Urban Fabric composite models; as we might see, the value of the spatial correlation are reduced, with an increasing reduction for higher topological lags. If we consider the same threshold of 0.05 for the distinction between positive and null spatial autocorrelation, a statistically significant and positive autocorrelation of the residual is observed for a topological radius of: i) four, when considering the global model, ii) three, when considering the First/Second composite model, iii) two, when considering the urban fabric composite model. For the three models, independently of the specific numerical threshold, the three correlograms show an elbow at a topological distance of three topological steps. The residual spatial autocorrelation described by the model errors corresponds (approximately) to the small-scale structure described by the variogram of Fig.100 (Section10.3).

From these observations, we might deduce that: i) as expected, retail spatial distribution is characterised by positive autocorrelation effects; ii) the overall spatial autocorrelation of retail might be divided in *intrinsic* autocorrelation, related to the phenomena itself, and the *systemic* autocorrelation which is explained by the underlying structure of the urban morphological system; introducing in the model the information about the spatial structure of the underlying urban system (the morphological region partitions), which allows us to isolate the intrinsic autocorrelation. iii) In the case of the French Riviera retail system, the intrinsic autocorrelation is positive and significant for a network-constrained distance between 2 and 4 topological steps; when considering 77 metres as the median length of the street segment in the overall

study regions, the radius of the autocorrelation in metrical terms corresponds to a length of 150-300 metres.

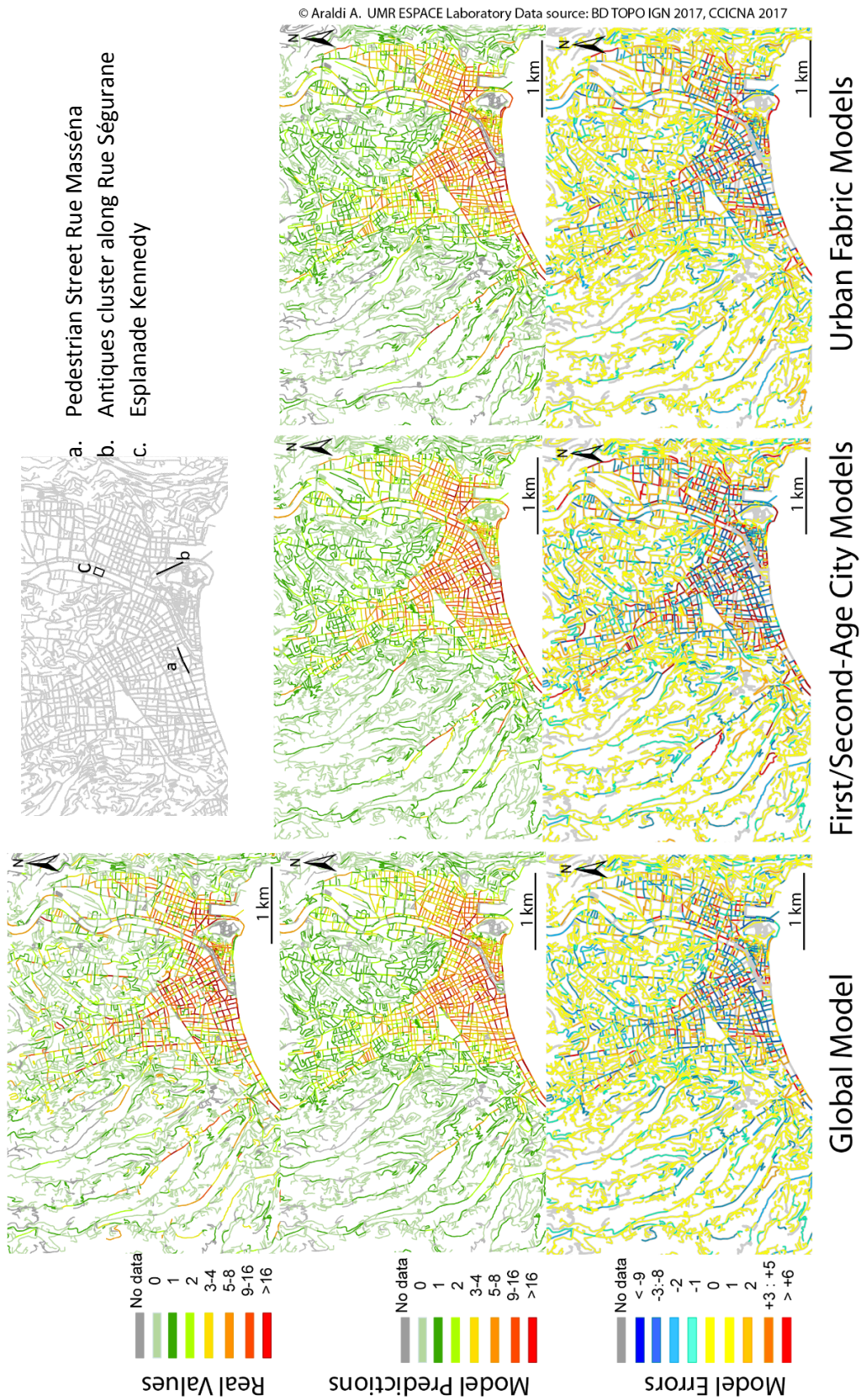
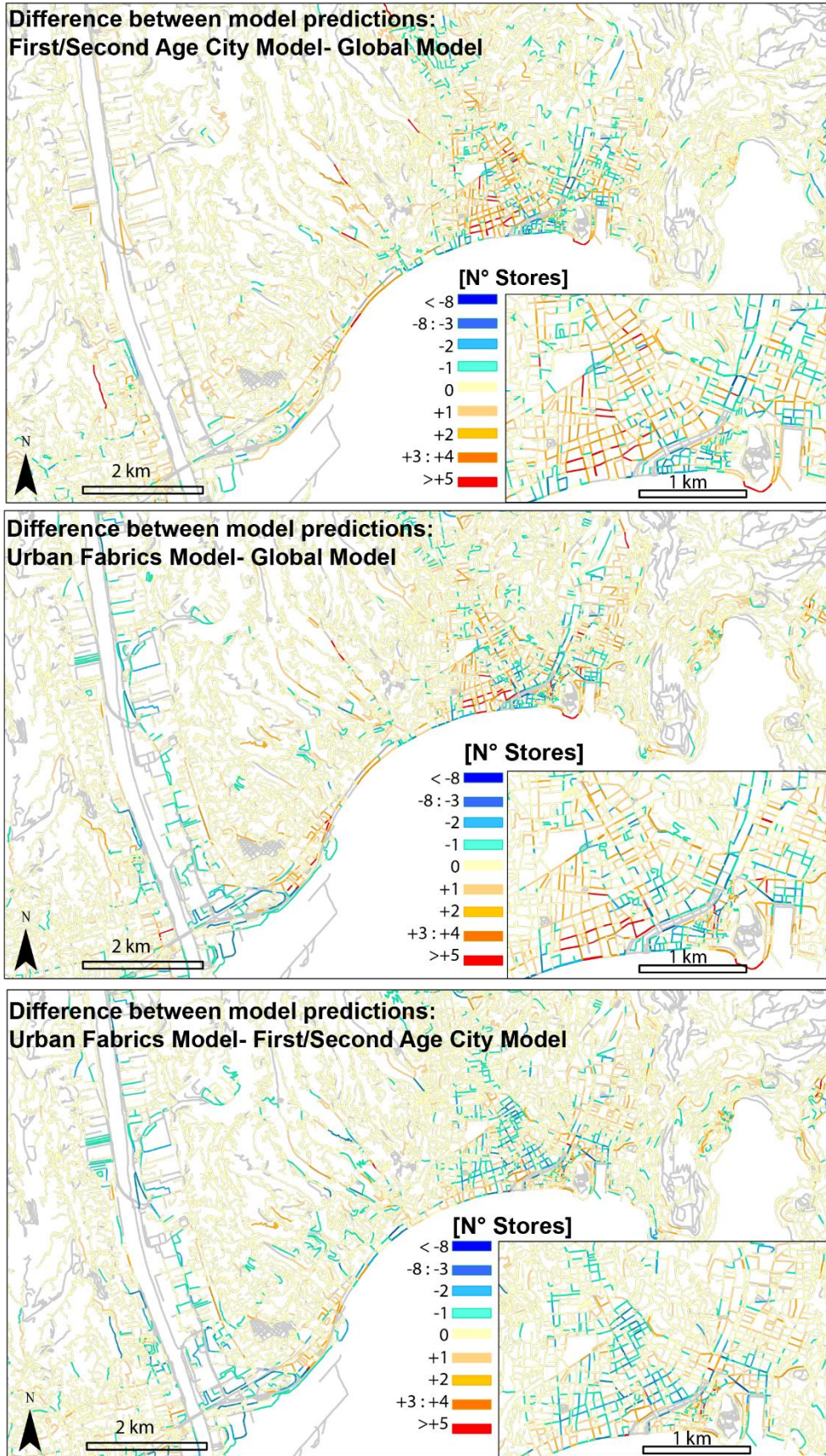


Figure 106 Projection in the geographical space of the observed, predicted and residual values for the Global model, First/Second-Age and Urban Fabric composite models.



© Araldi A. UMR ESPACE Lab. Data source: BD TOPO IGN 2017, CCINCA 2017.

Figure 107 Projection in the geographical space of the model prediction differences. From the top to the bottom: (First/Second-age-City – Global); (Urban Fabric – Global); ( First/Second-age-City –Urban Fabrics)

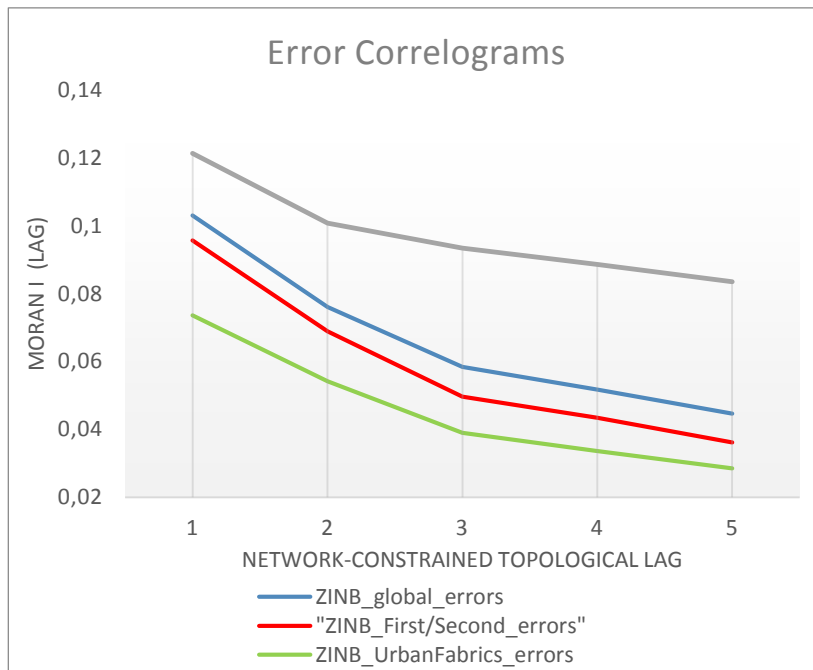


Figure 108 Correlogram of model errors and of store counts. The spatial structure of the urban morphological system allows explaining the spatial organisation of retail at the mesoscale while residual autocorrelation is detected at the local-scale (about 3 topological steps)

Finally, the visual inspection of the REC-curves, both for the different count models and for each region, inform us how the low values of Accuracy, Precision and Sensitivity observed for all our models increase rapidly when considering small tolerance values. For example, when considering the Global ZINB model, the REC curves in Fig.109 inform us how with one unit of tolerance, the model is able to predict almost the 90% of the count values; more specifically, precision and accuracy multiply their values by three (from respectively 0.13 and 0.17 to 0.67 and 0.58). Nonetheless, few differences are detected among REC curves obtained from different models, with the exclusion of the Gaussian model for which the lowest values are observed.

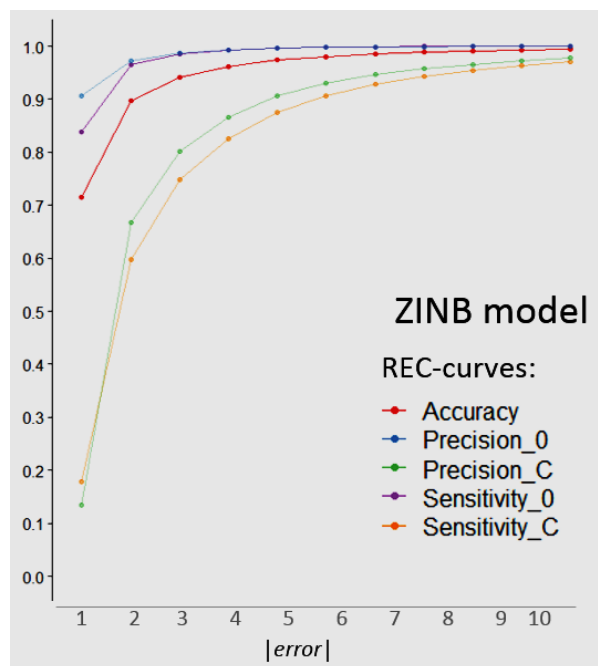


Figure 109 REC-Curves for the Global ZINB model.

## 12.2 Variable selection for count regression models

While in the previous section we presented and discussed the outcomes of models applied on the whole set of explanatory variables, in this section we focus on the outcomes of the variable selection procedures. This allows us to identify and describe which combination of morphometric indicators underlie the spatial distribution of retail in the global study area as well as in each morphological region.

As presented in Section 17.2, two methodological procedures have been retained and implemented for the variable selection phase of our research: Backward Regression (BR) and Elastic Net penalised regression (Enet)<sup>96</sup>. The former is one of the most adopted solutions in urban studies (automatically or manually implemented), while the latter represents a more sophisticated procedure which has been recently implemented in economic and retail geography studies. In the scientific literature, there is no evidence comparing and assessing the two procedures, although statistic and data mining literature seem to favour penalised regression approaches (Harrell 2001).

For these reasons, the two procedures have been implemented in our case study.

BR approach removes the indicators below a certain significance threshold separately for the count and the zero parts. In our application, the threshold implemented is 0.001; consequently, *by construction*, all the selected variables which we are going to discuss in this section show a high level of statistical significance and therefore a strong relationship with the retail distribution. Enet removes regressors with a penalised regression procedure; for each study region, we ask the Enet algorithm to explore 20 values of lambda<sup>97</sup>. The regression coefficients reported in this section correspond to the penalised model for which the lowest BIC score is observed.

Before describing and commenting on the results of the application of these two procedures to each study area and subspace, we should anticipate some general observations about BR and Enet outcomes constantly found in the following applications:

1) The two variable selection procedures have been implemented on the overall study area and on each subspace, each time considering the same dataset composed by the same number of features and regressors. The morphological variables independently selected by these two procedures partially overlap for all our models. The common morphological indicators identified represent a sort of *double selection* and *double endorsement* of the relevance of those morphological indicators in relationship to retail distribution.

---

<sup>96</sup> These procedures have been implemented using *mpath* R package (Wang et al. 2015) for both BR and Enet Penalised Regressions with ZINB models.

<sup>97</sup> 100 is the default value implemented in recent studies. High values of lambda allows a higher precision in the selection of regressors; nonetheless, the implementation of these algorithms with a high number of features and variables (between 86 and 105 indicators) with a regular computer requires long computational time. Using 20 as a lambda value, the computational time varied between nine hours and almost a week (in the case of the global model). The use of more powerful computer systems and future algorithmic developments might reduce the time of implementation and the adoption of higher lambda values. On the contrary, the implementation of backward regression procedures does not require important implementation time and it is limited to a maximum of five hours.



2) For each region under investigation, BR and Enet procedures select a similar number of features for the count part of the original ZINB model. Conversely, the zero part is differently handled by the two procedures: BR maintains a wider number of variables while Enet is more selective, and, in some cases, it completely erases the zero part proposing a final NB model. The reasons for this difference can most likely be found in the model selection procedure of Enet algorithms, which is based on the minimisation of BIC; as previously discussed, this information criterion considers a higher number of parameters as a penalising factor. These results might support the idea that a NB model is, in certain cases, a simpler and more efficient solution, which is more adapted for the study of the retail distribution in urban systems; on the contrary, the ZINB approach, despite being the *best solution* when the full model is studied (Section 18.1), becomes too complex when a smaller number of variables is investigated and selected.

3) Differently from Enet, BR overlooks the complexity of the model: only the significance of the coefficients is used as a selection criterion. Smaller significance thresholds do not converge the original model to a simpler solution<sup>98</sup>, especially when a dataset with a large number of features is investigated. As previously discussed, in statistical modelling, one solution to obtain significant regression coefficients is to increase the number of observations; in our case study, we already dispose of a large dataset providing highly significant coefficients.

4) As we further describe, the BR selected regressors show apparent inconsistencies and contradictory results for two main reasons. Firstly, selected indicators describing different characteristics of the same morphological aspect are associated with regression coefficients with opposite signs: we observe for instance a positive/negative coefficient for Reach/Straightness centralities, Regular/Normalised centralities, Average/Standard Deviation of morphometric descriptors, etc. Secondly, the resulting odd ratios from the BR procedure assume, in many cases, excessive positive/negative values which are indicative of a statistical issue affecting our results. One possible reason behind these results might be found in the multicollinearity of our indicators: when implementing VIF to the BR selected regressors, residual multicollinearity is still present, especially among those regressors showing *contradictory* results (in the following parts we use the terminology *low-modest-high* VIF for respectively values of  $VIF \leq 4$ ,  $4 < VIF < 10$  and  $VIF \geq 10$  since suggested VIF thresholds in the statistical literature vary between 4 and 10). On the contrary, these inconsistencies are not observed in Enet selected variables since penalised regression procedures handle multicollinearity with specific shrinkage procedures (as discussed in 17.2).

These four observations are aligned with Harrell's conclusion (Harrell 2001) about the inadequateness of stepwise regression approaches, especially the two observations about inflated parameter estimates and the higher complexity of the resulting models. We might support the conclusion that Enet penalised regressions outperform the BR approach, but at the price of higher model sophistication and higher computational power/time required.

---

<sup>98</sup> Several significance thresholds have been tested: between 0.5 and 0.001 a significant reduction of the number of selected variables is observed. Between 0.001 and 0.0001 the number of selected features is almost equivalent (with the exception of UF5 and UF6).

Nonetheless, further developments are required to endorse our empirical observations and demonstrate the superiority of one of the two approaches (for instance, exploring higher ranges of values of lambda, here reduced to a small value for computational power restrictions). For these reasons, in the following part of this section, although discussing the results of Enet penalised regression, we chose to also describe (in Text-Boxes 4,5 and 6), for each study area, the results of both procedures. For the same aforementioned reasons, the specific numerical coefficient (odd ratios) associated to each indicator will not be attentively considered; we limit our discussion only to the consideration of the positive/negative value that each regressor shows in relation to retail distribution.

Before describing these results, we remind the reader that 1) all the selected indicators in BR are significant by construction, as a result of the selection procedures while the significance is assured by the persistence of the regressors at the shrinkage. 2) the coefficient outcomes from count regressions are expressed as Odds Ratios (OR) or similarly Incidence Rate Ratios (IRR). An odds ratio is the ratio of the event or the outcome occurring in the exposed group divided by the event occurring in the non-exposed group. In probabilistic terms, OR is the probability to observe an event divided by the probability that the event is not observed. Odds ratios in our case study might be interpreted as follow: streets with a higher presence of stores are OR times more likely to be observed when the higher values of the given morphological characteristics under analysis are detected. When OR values are higher than one, the effect is positive, while, on the contrary, values lower than one describe an inverse relationship between the morphological variable and the retail presence. Finally, when OR values are equal to one, regressors and independent variables show no correlation.

The description and discussion of the variable selection analysis are organised as follows. Firstly, we describe in the Text-Boxes the indicators selected by the two procedures starting from the subset of indicators commonly identified. Moreover, the main variations in the model goodness-of-fit measures are highlighted (Appendix X). Secondly, we discuss the thematic importance of the selected variables considering exclusively the Enet penalised regression outcomes (for the aforementioned reasons).

The description and discussion of the results consider first the global model, secondly, the First/Second-age Cities partition, and, finally, the six Urban Fabric partition.

## **Retail distribution in the French Riviera**

### **TEXT-BOX 4. Variable Selection outcomes for the overall study region**

When considering the overall study area, BR and Enet select respectively 31 and 19 variables in the count parts; 10 indicators are found in common between the two variable selection procedures, respectively nine with a positive relationship and two with a negative one.

The shared morphological variables identified by the two procedures with a positive relationship are i) three configurational descriptors: 1200-metres Straightness, 20-minutes Reach, 5-minutes normalised Betweenness ii) six morphometric variables: 50-metres PBs coverage ratio, Building Fragmentation, street corridor effect, Average Building Height, Street Length and ii) one-directional centrality measure, the 1200-metres Betweenness towards anchor stores. The three indicators selected by both BR and Enet regressions and negatively related to retail presence consider only morphometric descriptors: Street Acclivity and average Street Openness.

Two additional aspects of the urban form with positive correlation with retail distribution are similarly identified by both BR and Enet with analogous but not equal indicators: 20-minute Straightness (regular in Enet and normalised in BR) and 1200-metre Directional Centrality measures toward the coast (Straightness and Reach respectively for BR and Enet).

Additional regressors with a positive relationship with retail distribution are selected by the two procedures: i) BR also considers 1200-meter normalised closeness, 5-minutes normalised reach, 600-metres directional betweenness toward the coastline and prevalence of specialised buildings. ii) Enet also select 600-metres and 1200-metres betweenness centrality, UF4 and UF7.

BR selects additional regressors with a negative relationship with retail: five streetscape intrinsic descriptors such as Street Windigness, Building Misalignment, Cross-Sectional Ratio, Prevalence of Small Buildings (C1[<150m<sup>2</sup>] and C2[150-250 m<sup>2</sup>]) but also five configurational indicators such as 600-metres Closeness, 5-minutes Reach, 20-minutes Normalised Reach and Straightness. This second group of indicators seems to be in contrast with those variables conjointly selected by the two procedures; as previously discussed, multicollinearity problem might be at the origin of these observations. The VIF values for the selected variables by the BR procedure is moderate for 5-minutes Reach, 1200-metres Straightness but high for 5 and 20-minutes Reach and Betweenness, both regular and normalised. Beyond the problem of multicollinearity detected by the high values of VIF, further observations let us doubt about the validity of the BR procedure: as illustrated in Fig. 110, some estimates present inflated OR (in zero part), while standard errors of the parameter estimates are too small, with, consequently, too narrow confidence intervals around the parameter estimates (similarly to what is observed in Harrell 2001).

The zero part of the model is maintained only by the BR procedure while no variables are selected by the Enet reducing the original ZINB model to a simpler NB model. The variables selected by the BR approach with positive correlation with over-representation of zeroes among which the 600-metres Reach centrality, Large Building Prevalence (>1000 m<sup>2</sup>),

Prevalence of Building Specialisation, average Open Space, 2400-metres Directional Reach toward coastline and 600-metres Straightness towards squares. On the contrary, thirteen indicators are negatively correlated with the over-representation of zeroes, among which the most important are: Street Length, First/Second-Age City (compared to natural), 600-metres Straightness, and Building Fragmentation.

When observing the goodness-of-fit measures for the two reduced models, we might detect important differences. In the case of the BR, all parameters are only slightly reduced; for example, the overall accuracy of the model falls from 0.715 to 0.701. The main variation is detected in the mean and standard deviation of the absolute error of zeroes ( $E(T_z)=0.26$ , +16%;  $Sd(T_z)=0.89$ , +26). These small variations of the overall performance of the BR reduced models (compared to the original model) are explained by the high number of selected variables, 26 for the count part and 31 for the zero part (half of the original number of regressors). On the contrary, when considering the Enet reduced model, only 19 variables are selected for both zero and count parts, since the original ZINB model is reduced to a simpler NB model. Consequently, substantial modification of the goodness-of-fit measures are detected. The overall accuracy of the model increases from 0.715 to 0.78; nonetheless, this overall increase is explained by a higher performance of the model to predict zeroes, and, on the contrary, lower performances for the count part (average absolute error of the zero-part  $E(T_z)=0.081$ , -64%; average absolute error of the count-part  $E(T_c)=3.327$ , +42%).

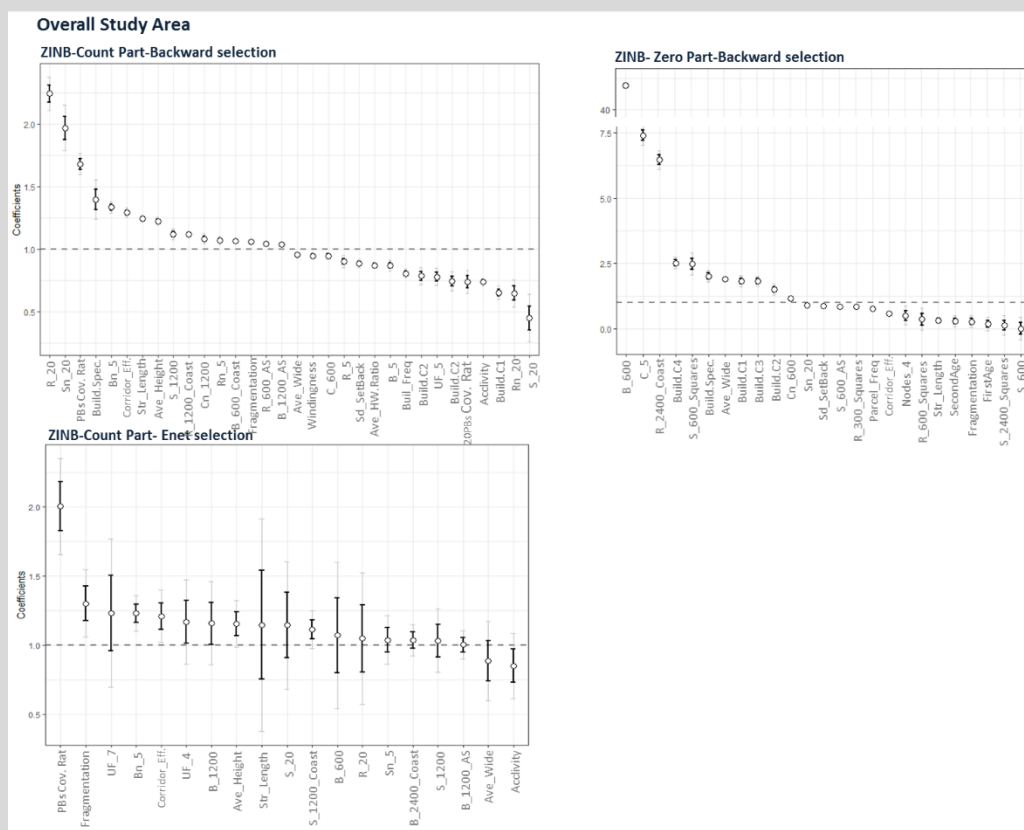


Figure 110 Selected variables from the Backward Regression and Enet Penalised Regression to the Global model

Observing the selected variables obtained by the Enet penalised regression procedure on the overall study area of the French Riviera, several observations on the relationship between retail distribution and urban form might be formulated. Retail distribution is associated with a combination of both configurational and compositional descriptors of the streetscape, selected in equal number among the original subset of regressors.

As already demonstrated in most of the works in Tab.1-2, we also identify the through-movement centrality (betweenness) as one of the most important retail locational factors among street-network configurational properties. Both pedestrian-oriented and vehicle-oriented through-movement are selected, at the meso- and macro-scales (5 and 20 minutes) for the latter, at the two radii of the meso-scale (600 and 1200 metres) for the former. Two additional configurational indicators consider vehicle-oriented centrality (both reach and straightness) at the macro-scale (20 min) and the pedestrian straightness at the meso-scale (1200 metres).

However, configurational measures are not the variables with the highest OR values with retail presence; on the contrary compositional descriptors of coverage ratio and built-up unit fragmentation of the streetscape are found with the highest OR values. The combination of these two variables with Street Corridor Effect and Average Building Height represents the main morphological characteristics associated with dense and compact urban areas.

Modern discontinuous urban fabric UF4 and connective fabric UF7 are positively related to retail, describing those areas in semiperipheral and suburban space where large specialized buildings allow the presence of mid- to big-sized stores. The street length is positively related to retail presence, which indicates that longer segments are associated with the presence of retail. Two directional centrality measures consider the proximity to anchor stores and direction towards the coastline.

Beyond the description of the individual and relative importance of each indicator and/or category of indicator, what is more important is to identify how these variables are combined together, providing the favourable urban morphological conditions for the location of stores.

Indeed, from the variables selected in the overall area reduced model, we might detect the presence of two sub-groups of locational factors. A first group of indicators identifies those streets crossing urban areas with higher built-up density values, where several contiguous high buildings are close and orderly disposed along the street edges. These traditionally central and compact areas distributed along the coast are also associated with pedestrian-oriented centrality at the mesoscale. A second group of indicators describes those long streets crossing modern and connective fabrics with large-sized buildings; completing the description, we might associate the high vehicle-oriented centrality and through-movement centrality measures and the influence of large anchor stores.

As we might see, despite the application of one model to the overall study area encompassing both dense urban areas and semiperipheral suburban spaces, the indicators identified highlight the presence of two different locational processes. These outcomes corroborate the necessity to investigate central and peripheral areas with two distinct models. In the following section we will observe the results of the separate models.

## **Retail distribution in the First- and Second-Age Cities**

### **TEXT-BOX 5. Variable Selection outcomes, First- and Second-Age Cities**

The BR and Enet procedures applied to the subsystem of the **First-Age City** identify respectively 37 and 26 indicators for the count part and 8 and 2 indicators for the zero part.

Here again, we might observe an important number of indicators which are consistently identified by both the variable selection approaches. Thirteen regressors show a positive relationship with retail distribution: i) four configurational measures: 5-minutes normalised Betweenness, Reach/Straightness (respectively in BR and Enet), 1200-metre Straightness, 300-metres normalised Straightness; ii) five morphometric descriptors: 20 and 50-metre Coverage Ratio, average Building Height, street corridor effect, Building Fragmentation. iii) four directional centrality measures: 2400-metres Betweenness centrality towards coastline, 1200-metres Reach/Straightness towards coast (respectively in BR and Enet), 1200-metres betweenness centrality toward anchor stores, 300-metres Straightness centrality towards squares. When considering indicators with a negative relationship with retail distribution, six regressors are coherently selected by the two procedures, here again only among morphometric variables: Street Acclivity, Open Space and Setback standard deviations, Prevalence of Small-Sized Buildings (<125m<sup>2</sup>), residential urban fabric of small houses (UF5), standard deviation of Cross-Sectional ratio (respectively in BR and Enet).

Additional indicators in the case of the BR are 2400-metres Straightness towards coastline, 20-minutes normalised Closeness with a positive statistical relationship; on the contrary, a negative relationship is found with average Cross-Sectional Ratio and Setback, 1200-metres Reach and normalised Straightness, 20-minutes Reach and Closeness, 5-minutes normalised Straightness, 300 and 1200-metres normalised Reach, the 2400-metres Reach and 1200-metres straightness toward coastline. Similarly to when observed in the global model, here as well multicollinearity might be at the origin of the apparent contradictions among some of the selected indicators with BR regression; VIF values are always low except for moderate values in correspondence with 50-metres coverage ratio, average Setback and Cross-Sectional ratio, Reach and Straightness at 300 and 1200 metres, 5 and 20-minutes and Directional centrality measures towards coasts at both 1200 and 2400-metres radius. Inflated estimates both in the count and zero parts and narrow confidence intervals corroborate, here again, the inadequateness of the BR procedure.

Enet regression, on the contrary, selects fewer additional indicators which strengthen the notion of centrality at 1200-metres (Betweenness and normalised Straightness) with a positive relationship with retail distribution; a negative relationship is detected in correspondences of semi-peripheral urban fabrics UF3, UF4 and connective fabric UF7.

When considering the zero parts, BR and Enet both select Building Fragmentation with a negative relationship with an over-representation of zeroes. Additional indicators are, in the case of Enet, regression Parcel Frequency; for BR 1200-metres Betweenness, 300-metres normalised Betweenness, 5-minutes normalised Reach with a positive relationship; negative

values in correspondence with 5-minutes Reach and normalised Straightness, 1200-normalised Betweenness.

When observing the goodness-of-fit measures for the two reduced models, important differences might be outlined. In the case of the First-Age City, we detect a significant reduction of the overall accuracy of the BR model which falls from 0.484 to 0.385. The important drop of the model performance is associated with the intrinsic difficulty in predicting correctly the precise number of stores of the commercial streets of the traditional compact city, as well as with the lower presence of zeroes, which normally boosts model accuracy. Here again, the BR reduced model penalises the prediction of zeroes where both mean and standard deviation of the absolute error of zeroes almost double compared to the full model ( $E(Tz)=1.10$ , +92%;  $Sd(Tz)=2.71$ , +118%). When considering the Enet reduced model, similarly to what was previously observed in the global model, we observe an increase of the overall accuracy of the model from 0.484 to 0.573; here again the overall increase is attributable to higher performance of the reduced model to predict zeroes (increase in both sensitivity and precision), and, on the contrary lower performances for the count-part (average absolute error of the zero-part  $E(Tz)=0.262$ , -54%; average absolute error of the count-part  $E(Tc)=4.42$ , +52%).

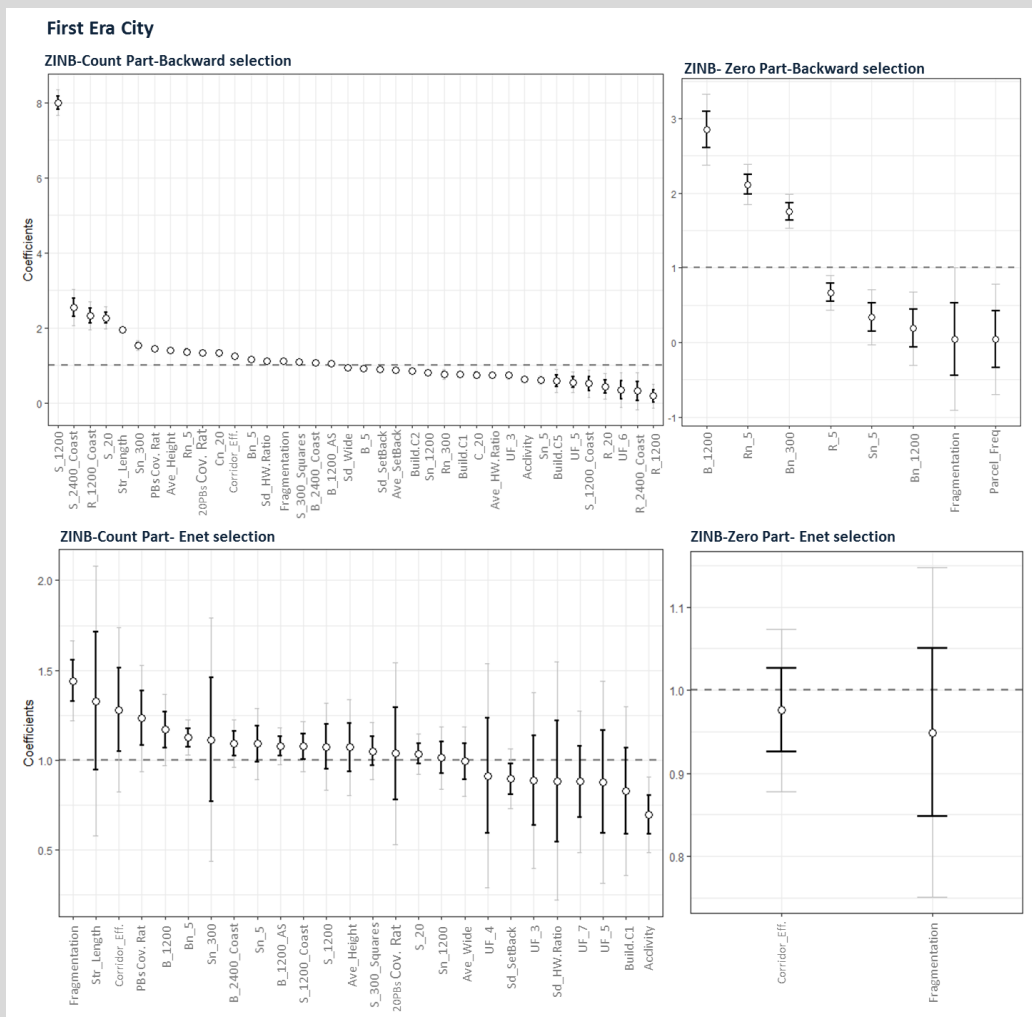


Figure 111 Selected variables from the Backward Regression and Enet Penalised Regression to the First-Age City Model

The BR and Enet regressions applied to the subsystem of the **Second-Age City** identify respectively 25 and 21 indicators for the count part and 15 and 2 indicators for the zero part.

The two count parts both select 12 indicators, 9 with a positive relationship with retail distribution, 3 with a negative one. In the first group we find i) two configurational indicators: 5-minutes normalised Betweenness and the 20-minutes Reach, ii) five morphometric descriptors: 50-metres PBs Coverage Ratio, Street Length, average Building Height, residential and connective Urban Fabrics (UF4 and UF7) and iii) two Directional configurational measures, the 600-metres Betweenness towards the coast and the 1200-metres Betweenness towards anchor stores. When observing the common selected indicators with a negative relationship with retail distribution, Street Acclivity and Prevalence of Small Buildings (C1 and C2, < 250 m<sup>2</sup>) are identified. BR regression additionally selects 1200-metres Reach and 20-minutes Straightness (both associated with high VIF values) while Enet considers the average Open Space width.

Additional indicators with a positive relationship with retail distribution are, in the case of the BR, Parallel Façades, 5-minutes Closeness, 20-minutes normalised Straightness and UF2. The Enet regression, in its turn, considers several centrality measures at 1200-metres radius, such as normalised Straightness, Betweenness towards the coastline, Straightness towards anchor stores, and two morphological measures, Average-Sized Buildings (C3, 250-1000m<sup>2</sup>) and Building Specialisation Prevalence.

Only one indicator, the Building Frequency, seems to have a divergent outcome: a positive relationship in Enet regression and negative in BR.

When looking at the zero parts Parcel Frequency and 20-metres PBs coverage ratio are found with a negative relationship with excess of zeroes by both the Enet and Bw regression approaches. The latter also add thirteen additional variables, among which, the highest values are found in correspondence with large buildings (C5, > 4000 meters) and 2400- metres Reach towards the coast with a positive relationship.

Nonetheless, here again, correlation coefficients of BR regression might be biased by the presence of multicollinearity. VIF is always low except for moderate values for 5-minutes Betweenness (regular and normalised), Prevalence of Large Buildings (C4 and C5, >1000 m<sup>2</sup>), 50-metres PBs Coverage Ratios, average Open Space and 1200/600-metres Betweenness towards coastline; high VIF values are found for same-distance Reach and Straightness pairs (1200 metres and 20 minutes). High VIF values in the zero part are found only for 2400-metres reach and straightness towards coastline; furthermore, inflated estimates in the zero part and narrow confidence intervals support, here again, the inadequateness of the BR procedure.

Changes in the goodness-of-fit measures for the two reduced models of the Second-Age City show similar values to what was previously highlighted for the First-Age City models: BR is characterised by a significant reduction of the overall accuracy of the model, which falls from 0.774 to 0.626. The prediction capacity of zeroes is penalised: both mean and standard deviation of the absolute error of zeroes almost double compared to the full model (E(Tz)=0.429, +169%; Sd(Tz)=1.023, +91%). When considering the Enet reduced model, there is a milder increase of the overall accuracy of the model from 0.774 to 0.82; the overall



increase is associated with higher performance of the reduced model to predict zeroes (increase in sensitivity), and, on the contrary, lower performances for the count-part (average absolute error of the zero-part  $E(T_z)=0.111$ , -31%; average absolute error of the count-part  $E(T_c)=2.376$ , +22%). One specific characteristic differentiates Enet regression application to the Second-Age City from the First-Age City: despite the observed reduction of the absolute error of the model for the zero part, its standard deviation doubles ( $Sd(T_z)=1.116$ , +108%), describing a higher difficulty to differentiate between absence and presence of stores in modern car-oriented regions surrounding the compact cities.

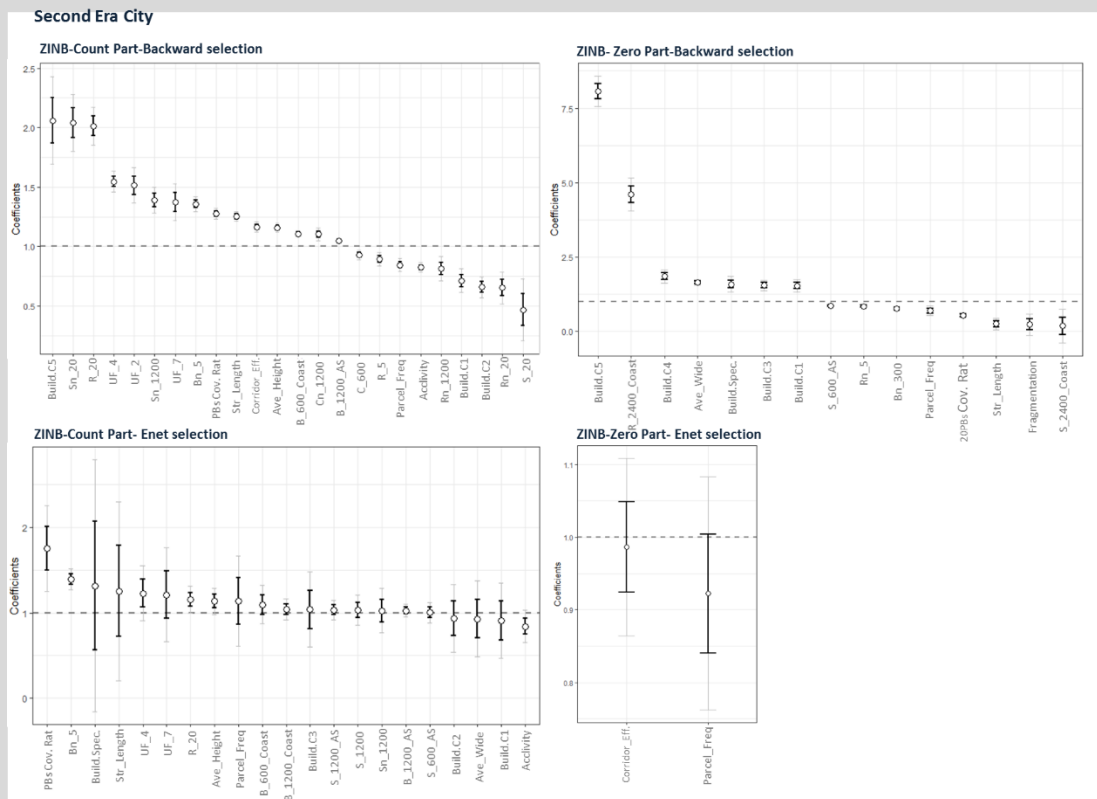


Figure 112 Selected variables from the Backward Regression and Enet Penalised Regression to the Second-Age City Model

Before analysing the selected indicators, we should remind to the reader the main characteristics which differentiate these two sub-systems identified as the First and Second-Age Cities, as highlighted in the previous part of the dissertation. The First/Second-Age Cities partition describes two urban systems encompassing different proportions and spatial arrangements of urban fabrics. More particularly, when modelling the First-Age City, we are considering five different central cities, four distributed along the coast (Nice, Antibes with Juan-les-Pins, Cannes with Le-Cannet, Menton, Cagnes-sur-mer with Saint-Laurent-du-Var) and only one in the hinterland (Grasse). Within these urban compact areas, we observe a higher representation of compact urban fabrics (UF1 and UF2) in their core, which gradually leave the place to residential and irregular urban fabrics of small houses (UF3) and larger buildings in modern urban fabrics (UF4) in more semiperipheral and peripheral parts. Some small fragments of suburban fabrics (UF5 and 6) might also be observed as a result of the

specific generative process of the agglomerations of Antibes with Juan-les-Pins, Cannes with Le-Cannet and Cagnes-sur-mer with Saint-Laurent-du-Var, where development from different urban cores has generated unique urban areas which include fragments of suburbia. The prevalence of compact fabric within First-Age Cities is also associated with pedestrian-oriented mobility. Finally, we should remind the reader how within these urban areas we found 61% of all small and average-sized stores of the French Riviera.

The Second-Age City system, on the contrary, is characterised by a unique interconnected and almost continuous common matrix. Scattered within this matrix, we might recognise: i) several isolated historical centres (UF1), some of them surrounded by small planned developments (UF2); the too small extension of these compact centres and the high distances between them limit the development of pedestrian-oriented mobility. ii) modern and connective urban fabrics composed by large buildings both residential and specialised forms; these urban fabrics might be found both independently and combined within more complex morphological regions (as previously described for the Var Valley and Sophia Antipolis technological park; Text-Box 3, Section 9.3). Vehicle-oriented mobility is traditionally associated with the Second-Age City, as suggested by Fusco (2016).

Applying variable selection procedures to the First/Second-Age Cities partition allows us to focus on the specific internal characteristics of each subsystem corresponding to specific statistical variability distributions. Consequently, more detailed descriptions of the morphological characteristics and of the pedestrian/vehicle-oriented centrality measures might be combined and associated with retail presence.

In the case of the **First-Age City**, the coverage ratio still represents an important characteristic of the urban morphology positively correlated with retail distribution; nonetheless, a relatively lower value of OR is observed. On the contrary, higher importance is given to the built-up unit fragmentation and the street corridor effect. As we previously highlighted in the global model, within dense morphological regions, retail distribution is associated with the combination of building density together with pedestrian-oriented through-movement centrality at the meso-scale (as for the global model, here again the 1200-meter Betweenness is selected).

Moreover, while the global model considers the proximity to the coastline as positively related to store distribution, when observing the specific results for the First-Age City systems (which are always located along the coast), Enet regression further specifies that positive influence is associated with straightest and shortest paths towards the coast (directional Betweenness and Straightness).

Several morphometric indicators are identified as negatively correlated to retail distribution: at the street scale, the irregularity of the streetscape and the presence of small houses are repulsive factors (standard deviation of both setback and crosssectional ratios); at the neighbourhood scale, residential fabrics (both of small houses UF3 and UF5 and of large buildings UF4) and connective fabrics (UF7) are also negatively related to store distribution. As observed for the global model, street acclivity reduces the commercial potential of a street also when considering First-Age cities.

Here again, we might identify two specific combinations of urban morphological parameters explaining the retail distribution in First-Age Cities. On one side we have denser planned urban areas closer to the coast with a regular distribution of fragmented built-up units. Within this context, the presence of streets with irregular streetscape, small buildings or higher acclivity represent hyper-local characteristics which reduce the probability to observe retail.

On the other side, residential and connective urban fabrics are usually negatively correlated to retail in the overall space of the First-Age city, unless hyper-local streetscape conditions with fragmented built-up units are close and parallel to the street edges are found. This specific exception is highlighted by the zero part of the model, where built-up unit fragmentation and street corridor effects are negatively related to the presence of zeroes.

In the case of the **Second-Age Cities**, building density is still one of the most important indicators able to identify those small isolated and scattered concentrations (typically of UF1 and UF2). Vehicle-oriented configurational centralities are also selected (Betweenness 5 minutes and reach 20 minutes) confirming the car-oriented mobility logic dominating these peripheral areas.

Modern and connective urban fabrics (UF4 and UF7) appear here again but with a positive relationship with retail. This observation confirms the double role of these urban fabrics suggested by the global model. Depending on the overall macroregion context, these two retail fabrics might assume contrasting roles.

The positive effect of the built-up unit fragmentation is replaced in the Second-Age City by the indicator of parcel frequency. The reason underlying this change is probably due to the lower contiguity of the built forms typical of the suburban spaces. For the same reason, the acclivity, the presence of small buildings, individual homes and villas represent the local intrinsic conditions associated with residential functions and therefore improper for retail activities.

Similarly to the case of the First-Age City reduced model, even in the case of the Second-Age City, the zero part highlights how the parcel frequency and the presence of street corridor effect (even if observed at wider distances, 20-metres corridor effect) might provide the hyper-local conditions for enhancing the presence of retail activities.

As discussed in this part, the application of variable selection models on the separate spaces of the First- and Second-Age Cities allows us to better understand the urban structure of the specific area under study, in this dissertation, the conurbation of the French Riviera. More in particular, specific combinations of configurational and geometrical descriptors (both at the street and neighbourhood scales) allowed us to better recognise the locational factors underlying the retail distribution within these two systems.

In the following part of this Chapter, we investigate the results of the variable selection procedures applied to the finer spatial partition made of 6 Urban Fabrics identified from the MFA procedures. Natural and connective urban fabrics (UF7, UF8, UF9) are excluded from this

analysis: the small number of streets with buildings and stores (as discussed in section 16.3) does not allow us to obtain statistically significant outcomes. Since urban fabrics are characterised by a higher homogeneity compared to the larger First-/Second-Age Cities, the statistical distribution of morphological indicators is strongly reduced, and, consequently, a lower number of regressors is selected by the variable selection procedures.

### **Retail distribution within specific Urban Fabrics**

#### **TEXT-BOX 6. Variable Selection outcomes, the Urban Fabrics**

##### **UF1: Urban Fabric of old constrained fabric of town-houses**

The BR and Enet regressions applied to the old constrained fabric of town-houses (UF1) identify respectively 18 and 13 indicators for the count part and 8 and 0 indicators for the zero part.

The two count parts share the same information about 8 indicators, 6 with a positive relationship with retail distribution, 2 with a negative one. In the first group we find i) two morphometric descriptors: Street corridor effect and average Building Height ii) three configurational indicators: 20-minutes Reach, 600-metres Betweenness and iii) one Directional configurational measure, 1200-metre Reach towards the coast. Negative relationship is found in the correspondence of Street Acclivity and Open Space standard deviation.

Additional indicators selected by the Bw regression approach are 300-metres Betweenness, Building Frequency, 20-minutes Straightness, 2400-metres Reach toward the coast all with a negative relationship with retail distribution. VIF values are always low except for moderate values in correspondence of 1200/2400-metres Reach toward the coastline, 300/600-metres Betweenness toward the coastline, 20-minutes Reach and Straightness and a 20-minutes Closeness and Normalised Reach.

The indicators additionally considered by Enet are 5-minutes normalised Straightness, 600-metres Betweenness towards the coastline, 1200-metres Betweenness and Betweenness towards anchor stores with a positive relationship; 600-metres normalised Closeness with a negative relationship.

When considering the zero part, we might observe how only BR procedure persist while Enet reduces the ZINB model to a simple NB model. The same problems relatively to BR previously highlighted for the global and the First/Second-Era partition, even if reduced, are still detected when considering the fined partition of the Urban Fabrics.

Every goodness-of-fit measures of the BR reduced UF1-model shows significant reductions when compared to its relative original full-model. The overall accuracy falls from 0.659 to 0.598. The prediction capacity of count values, and, more strongly, zeroes-values are both penalised, ( $E(Tz)=0.377$ , +44%;  $E(Tc)=1.936$ , +7%). When considering the Enet reduced model a slighter increase of the overall accuracy of the model from 0.659 to 0.705; as previously observed for the global model, the overall increase of the model accuracy is explained by the reduction of the ZINB model to a simpler NB model. Consequently, we observe higher performance of the Enet reduced model to predict zeroes (average absolute error of the zero-

part  $E(T_z)=0.0667$ , -64%) and, on the contrary, lower performances for the count-part (average absolute error of the count-part  $E(T_c)=2.596$ , +43%).

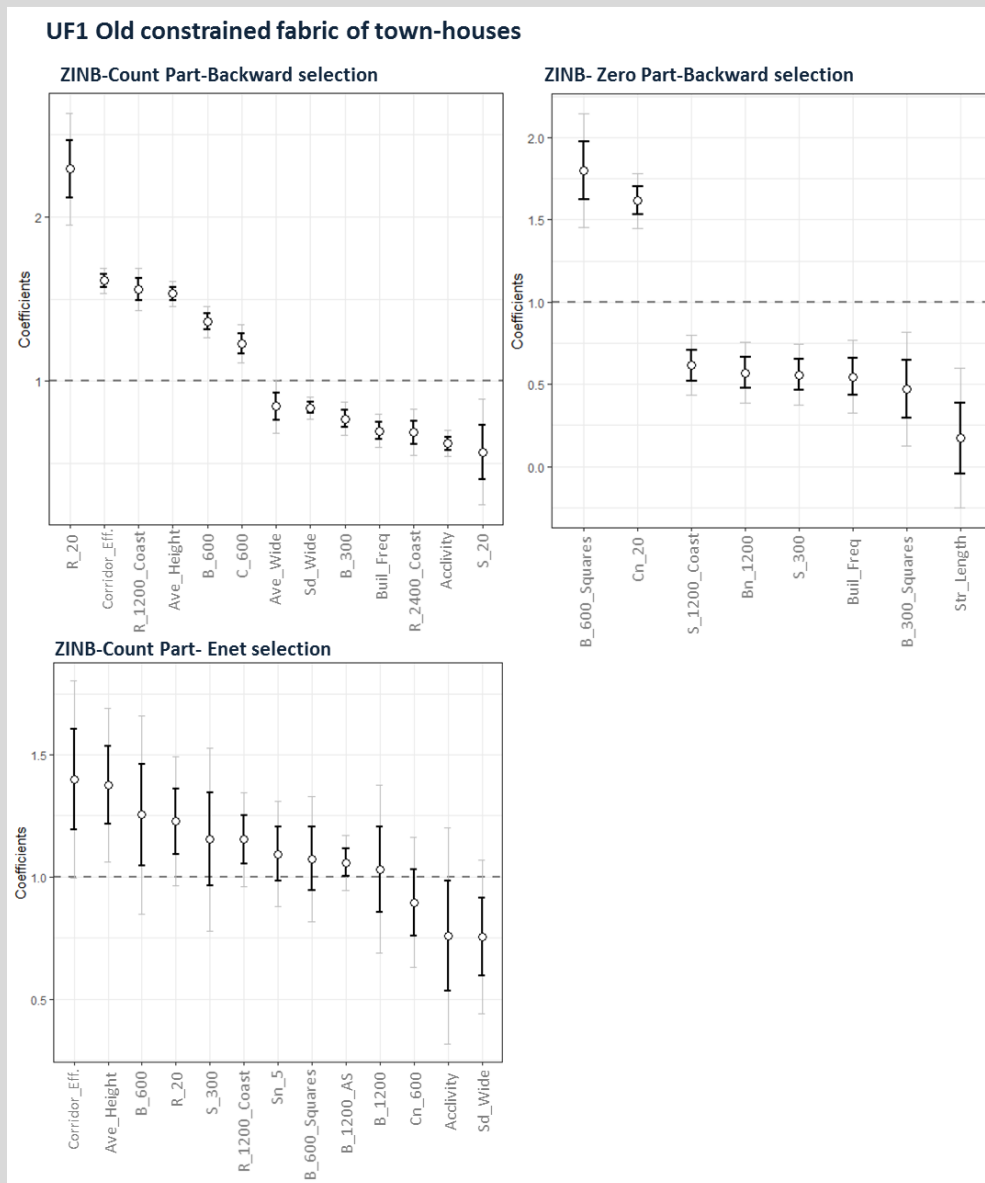


Figure 113 Selected variables from the Backward Regression and Enet Penalised Regression to the UF1 Model

## **UF2: traditional urban fabric of the plain with adjoining buildings**

The BR and Enet regressions applied to traditional urban fabric of the plain with adjoining buildings (UF2) identify respectively 17 and 18 indicators for the count part and 12 and 4 indicators for the zero part. The two count parts consistently identify ten indicators, seven with a positive relationship and three with a negative one. Within the first group we find i) four streetscape descriptors: 50-metres PB Coverage Ratio, Street corridor effect, Nodes Degree Four, Building Fragmentation; ii) two configurational indicators, the 5-minutes normalised Betweenness and 20-minutes Straightness and one 1200-metres

Reach/Straightness towards the coastline (respectively BR and Enet). A negative relationship is found only with streetscape indicators: Street Acclivity, Standard Deviation of Open Space and of Cross Sectional Ratio (H-W).

Additional indicators selected by the BR procedure are 300-metres normalised Straightness and normalised Reach with a positive relationship with retail distribution; Parcel Frequency, on the contrary, show a negative relationship. Enet regression selects 600-metre Betweenness with a positive relationship with retail distribution while Small Building Prevalence (C1 and C2, < 250 m<sup>2</sup>) and standard deviation of the Setback with a negative relationship.

Here again, we should consider that some of the selected indicators by the BR procedure might suffer from multicollinearity problems: VIF is high in correspondence of 300-metres normalised Reach and a moderate VIF for 20-minutes Closeness and normalised Reach.

Two indicators seem to assume a conflicting role between BR and Enet outcomes: average Building Height assume a negative and positive relationship with retail distribution in respectively Enet and Bw. Contrarily, 1200-meter Betweenness towards places a assume negative and positive relationship with retail distribution in respectively Bw and Enet.

When considering the zero part, all of the four indicators selected from the Enet procedure are also found among those identified by the BR. Three streetscape descriptors—Street corridor effect, Parcel Frequency, Built-up-units Fragmentation—and one configurational indicator, the 5-minutes Reach, are all negatively correlated with the excessive presence of zeroes. BR selects eight additional indicators: 20-metres PB Coverage Ratio, 20-minutes normalised Closeness and normalised Reach and Street Length with a negative correlation with retail absence. On the contrary, a positive relationship is found in correspondence of Large Building Prevalence (C5, >4000 m<sup>2</sup>), the standard deviation of Setback and Building Height and 300-metres normalised Closeness.

The overall accuracy of the BR reduced UF2-model shows the highest reduction among every BR implementation; it falls from 0.327 to 0.254. The prediction capacity of count and zero values both penalised, ( $E(T_z)=0.377$ , +32%;  $E(T_c)=1.936$ , +15%). When considering the Enet reduced UF2-model a slighter increase of the overall accuracy of the model from 0.327 to 0.444; the overall increase of the model accuracy, hide a strong increase of the predictive power for zeroes (average absolute error of the zero-part  $E(T_z)=0.417$ , -58%) and, on the contrary, a strong reduction of the predictive performance for the count-part (average absolute error of the count-part  $E(T_c)= 5.158$ , +56%). The high statistical dispersion of the count part within UF2 morphological regions, with the presence of the most commercial streets of the French Riviera, explains the highest average errors and standard deviation observed among our models when trying to predict the exact number of stores on each street segment.

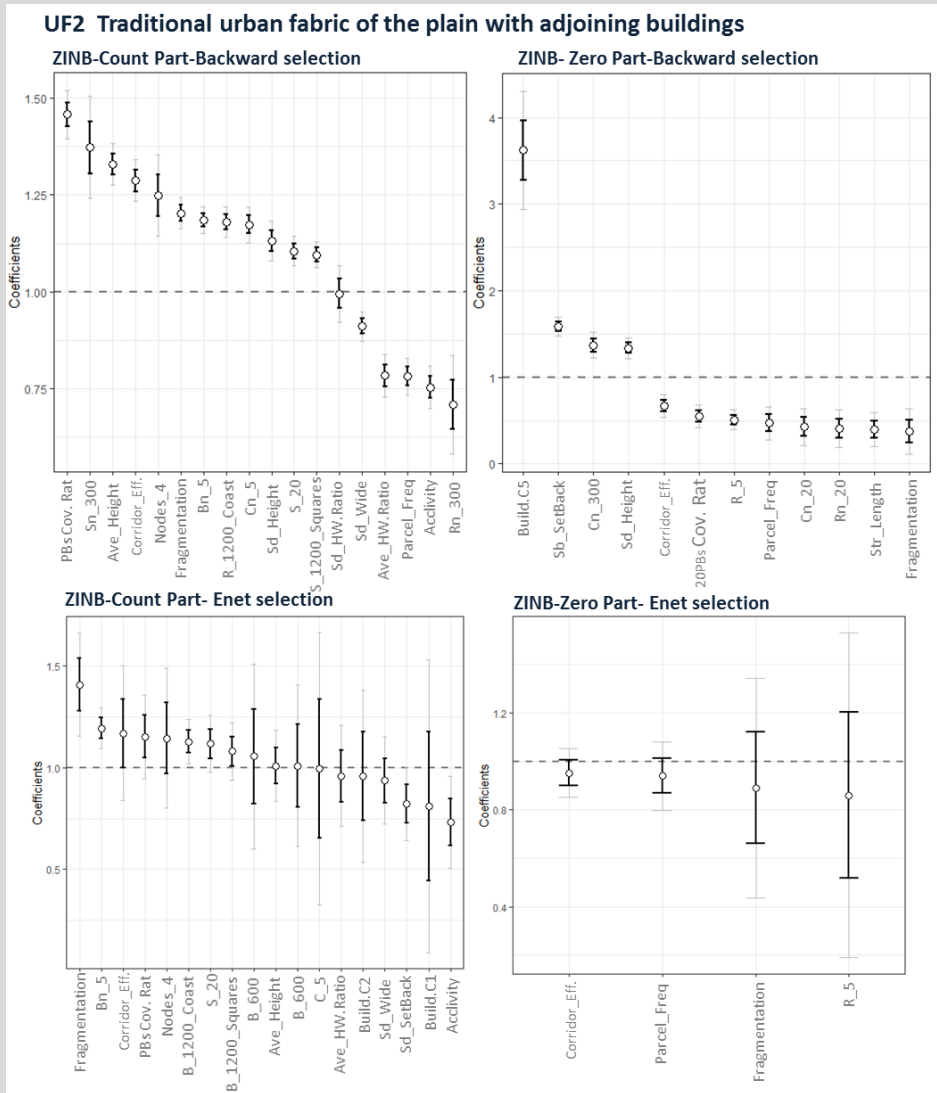


Figure 112 Selected variables from the Backward Regression and Enet Penalised Regression to the UF1 Model

**UF3: Discontinuous UF with houses and buildings**

The BR and Enet regressions applied to discontinuous urban fabric with houses and buildings (UF3) identify respectively 13 and 15 indicators for the count part and 4 and 2 indicators for the zero part. The two count parts consistently identify 10 indicators, 6 with a positive relationship and 4 with a negative one. Within the first group we find i) three streetscape descriptors: 50-metres PB Coverage Ratio, average and standard deviation of Building Height, ii) two configurational indicators: 5-minutes normalised Reach, 1200-metres Betweenness and iii) one directional configurational indicator, the 2400-metres Betweenness towards the coastline. A negative relationship is found in correspondence with two streetscape indicators, Street Acclivity and average Open Space and two configurational indicators, 1200 and 600-metres Reach.

Two configurational properties are similarly identified by the two procedures: 600-metres Straightness with a positive relationship with retail distribution and 300-metres Betweenness

with a negative one (in both cases a regular centrality indicator for BR and the normalised version for Enet).

Additional selected indicators are: Street Length (positive relationship) and 1200-metres Straightness for BR. Enet selects Parcel Frequency (positive relationship) and 300-metres directional Straightness towards places (negative relationship). In UF3, VIF values are high only for 600-metres Reach and Straightness centrality.

When considering the zero part, only one indicator is selected by both procedures: the 1200-metres Betweenness centrality with a negative relationship with store absence. BR also selects Street Length and 5-minutes Straightness with a negative relationship while 300-metres normalised Closeness with a positive relationship. Enet selects the Street corridor effect indicator with a negative relationship with retail absence.

Every goodness-of-fit measure associated with the resulting model from the BR procedure shows small negative reductions. The overall accuracy falls from 0.762 to 0.736. The prediction capacity of count and zero values both penalised, ( $E(T_z)=0.179$ , +23%;  $E(T_c)=1.522$ , +6%). When considering the Enet reduced UF3-model, this is a slighter increase of the overall accuracy of the model from 0.762 to 0.782; the overall increase of the model accuracy hides a strong increase of the predictive power for zeroes (average absolute error of the zero-part  $E(T_z)=0.0789$ , -46%), and, on the contrary, a strong reduction of the predictive performance for the count-part (average absolute error of the count-part  $E(T_c)= 1.834$ , +27%).

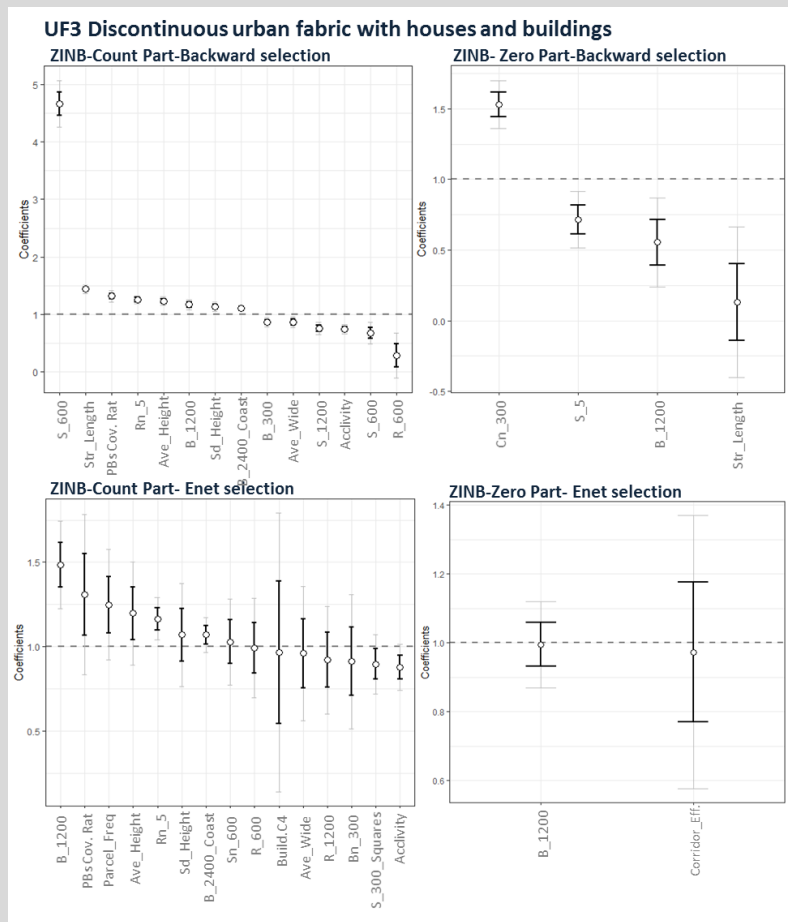


Figure 114 Selected variables from the Backward Regression and Enet Penalised Regression to the UF3 Model



#### **UF4: Modern discontinuous UF with big/medium-sized buildings**

The BR and Enet procedures applied to modern discontinuous UF with big/medium-sized buildings (UF4) identify respectively 10 and 17 indicators for the count part and 19 and 2 indicators for the zero part. The two count parts consistently identify 7 indicators, 5 with a positive relationship and 2 with a negative one. Within the first group we find i) three streetscape descriptors: 50-metres PB Coverage Ratio, Built-up-Unit Fragmentation, Street Length, average Setback and ii) one directional configurational indicator, the 600-metres Betweenness towards the coastline. A negative relationship is found with the two streetscape indicators of Street Acclivity and Setback standard deviation.

As we might notice, this is the only subspace for which the morphological indicators commonly selected by BR and Enet do not consider any configurational indicator. Only street properties (length and acclivity) and the relative disposition of building (Setback measures) represent the most relevant characteristics related to retail presence.

Configurational properties are not completely absent: different indicators are selected by the two procedures. BR considers 5-minutes Reach and normalised Betweenness with a positive relationship, while a negative relationship is found in correspondence with small buildings and 20-metres PB Coverage Ratio. Finally, 600-metres normalised Straightness and normalised Reach are found with, respectively, a positive and negative relationship with retail distribution. Here again, as previously discussed, one of the possible causes might be found in the multicollinearity among variables: VIF is moderate for 600-metres normalised Betweenness and high for 600-metres Reach and Straightness.

Several additional indicators are also selected by the Enet procedures. A positive relationship is found in association with 300 and 1200-metres normalised Straightness, 5-minutes Straightness, 1200-metres normalised Betweenness and 600-metres directional Betweenness towards the coastline. A negative relationship is associated with average Open Space. A positive but weak relationship is found with proximity directional measures towards anchor stores.

When considering the zero part, only two indicators are found in common between the two procedures, Parcel Frequency and 1200-metres Betweenness, both negatively correlated with retail absence. Several additional indicators are selected from the BR procedure. Nonetheless, important inconsistencies are found: for example, all Building Prevalences have a positive relationship with retail absence (higher values for larger surfaces, C4 and C5); 600-metres Reach/Straightness and, similarly, 2400-metres directional Reach/Straightness towards coastline, both with, respectively, a positive/negative relationship with retail absence.

The overall accuracy falls from 0.615 to 0.574. The prediction capacity of count and zero values are both penalised, ( $E(T_z)=0.445$ , +17%;  $E(T_c)=2.535$ , +4%). When considering the Enet reduced UF4-model, an important increase of the overall accuracy of the model from 0.6155 to 0.734 is observed; the overall increase of the model accuracy, similarly for the previous urban fabric UF3, hides a strong increase of the predictive power for zeroes (average absolute error of the zero-part  $E(T_z)=0.207$ , -45%) and, on the contrary, a strong reduction of the predictive performance for the count-part (average absolute error of the count-part  $E(T_c)=$

3.060, +25%). Similarly to what is observed for UF2, the higher presence and overdispersion of stores in UF4 might be at the origin of the higher errors observed within these specific morphological regions.

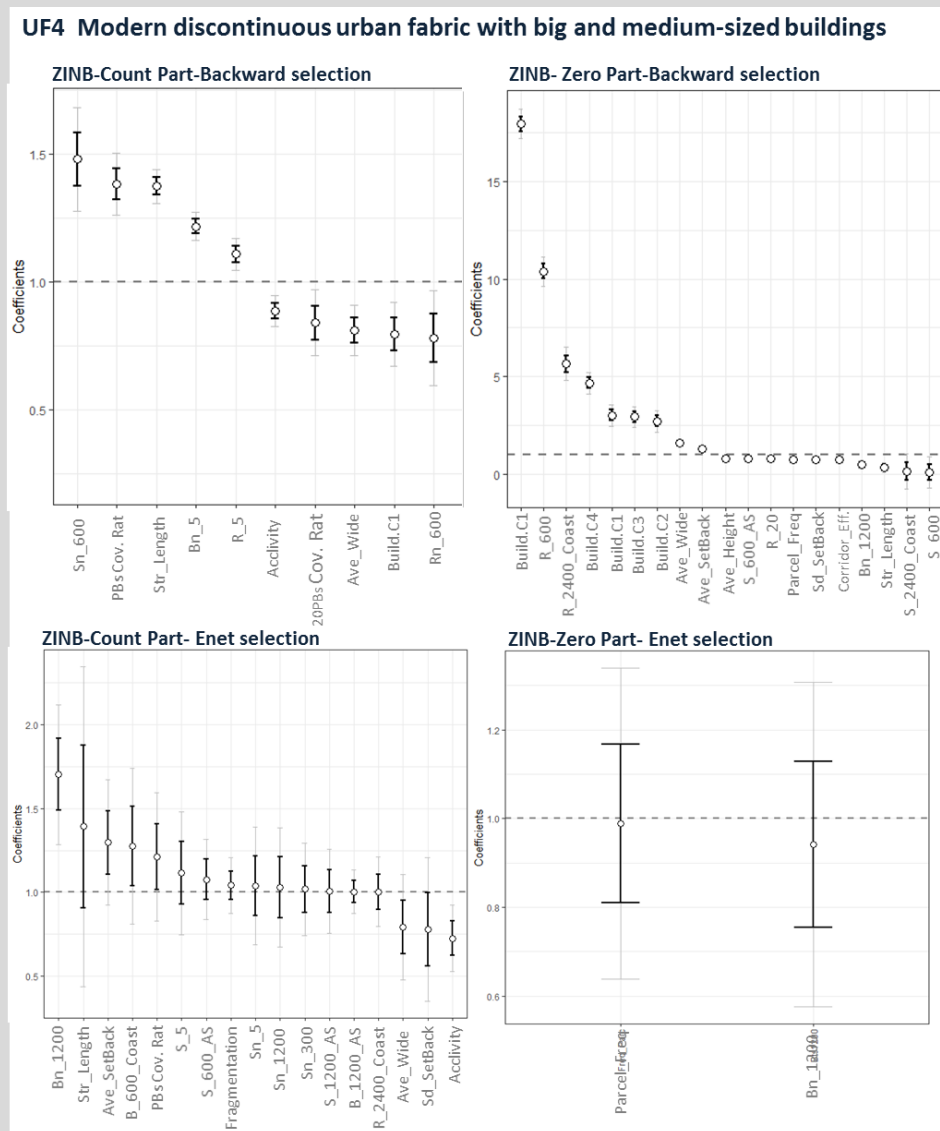


Figure 115 Selected variables from the Backward Regression and Enet Penalised Regression to the UF4 Model

### **UF5: Suburban residential suburban fabric in hills or plain**

Both the BR and Enet regressions applied to suburban residential suburban fabric in hills or plain (UF5) identify 11 indicators for the count part, while respectively 5 and 0 indicators are selected for the zero part. The two count parts consistently identify 5 indicators, 4 with a positive relationship and 1 with a negative one. Within the first group we find i) two streetscape descriptors, the 50-metres PB Coverage Ratio and the average Open Space and ii) three configurational measures, the 600-metres normalised Straightness and the 5-minutes regular and normalised Reach. A negative relationship is found in association with the Street Acclivity.

Additional indicators are selected by the BR are 5-minutes Reach and Street Length with a positive relationship, while Prevalence of Small Buildings (C1) and 600-metres normalised Reach with a negative relationship. As we might observe, no inconsistencies are detected in UF5 subsystem; VIF is always low within UF5 subsystem.

When considering the additional variables selected by Enet, we find 1200-metres Betweenness, standard deviation Building Height, and Parcel Frequency with a positive relationship; a negative relationship is associated with 300-metres Betweenness.

When considering the zero part, Enet reduces the original ZINB model to a simple NB. On the contrary, BR selects two variables positively related to retail absence, 20-minutes normalised Reach and normalised Closeness and three variables negatively correlated with the absence of stores, 5-minutes Betweenness, 20-minutes normalised Straightness and Street Length.

Minor changes in goodness-of-fit measures are observed when implementing both BR and Enet procedures to UF5. The overall accuracy of the two models slightly grows from 0.8682 to 0.872. The prediction capacity of count and zero values are both penalised: both increase the predictive power for zeroes values (average absolute error of the zero-parts  $E(T_z)=0.039$ , -13% and  $E(T_z)=0.034$ , -23% for respectively BR and Enet procedure); a minor reduction of the predictive performance for the count-part (average absolute error of the zero-parts  $E(T_c)=1.453$ , +3% and  $E(T_c)=1.480$ , +4% for, respectively, BR and Enet procedure).

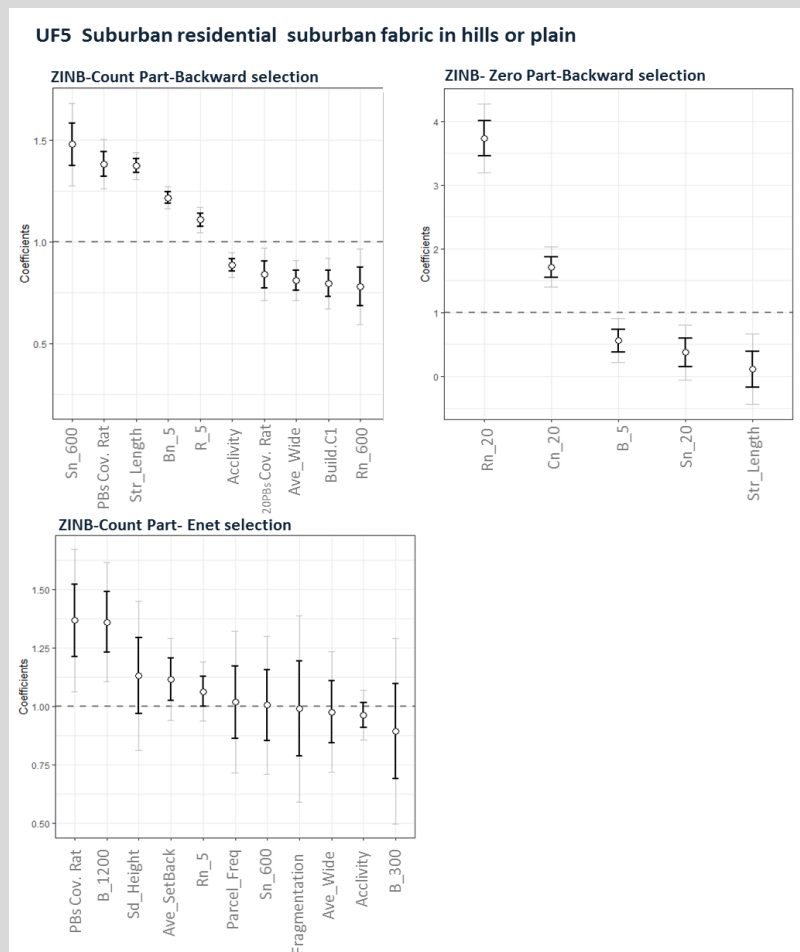


Figure 116 Selected variables from the Backward Regression and Enet Penalised Regression to the UF5 Model

### **UF6: Small house constrained suburban fabric**

The BR and Enet regressions applied to Small house constrained suburban fabric (UF6) respectively identify 5 and 12 indicators for the count part, while respectively 2 and 0 indicators are selected for the zero part. Differently from the previous subsystems, no indicators are found in common between the two procedures. The odd ratios in both cases are close to one describing weak relationships between morphological indicators and retail distribution.

BR select four indicators with a positive relationship with retail distribution: one morphometric indicator, the Prevalence of Averaged-sized Buildings (C3) and three configurational indicators, the 20-minutes Reach, 600-metres normalised Betweenness, and 5-minutes Betweenness. One configurational indicator is found negatively correlated with retail distribution, which is the 300-metres normalised Reach. In the zero part, we find two indicators negatively related to retail absence: the 300-metres normalised Straightness and the Street Length.

Enet selects six indicators with a positive relationship and six with a negative one. In the former group, we find the 50-metres PB Coverage Ratio, the standard deviation and the average of the Building Height, the average Setback and the 300-metres normalised Betweenness. The latter encompasses 600-metres Reach, normalised reach and Closeness, Parcel Frequency, Street Acclivity, and average Open Space.

Similarly to what is observed with the outcomes of UF5, even for UF6, slight changes in goodness-of-fit measures are observed when implementing both BR and Enet procedures to UF5. The overall accuracy of the two models slightly grows from 0.885 to, respectively, 0.886 and 0.9. The prediction capacity of count and zero values are both penalised: the procedures increase the predictive power for zeroes values (average absolute error of the zero-parts  $E(T_z)=0.021$ , -40% and  $E(T_z)=0.002$ , -94% for respectively BR and Enet procedure); a minor reduction of the predictive performance for the count part (average absolute error of the zero-parts  $E(T_c)=1.435$ , +13% and  $E(T_c)=1.509$ , +19% for, respectively, BR and Enet procedure).

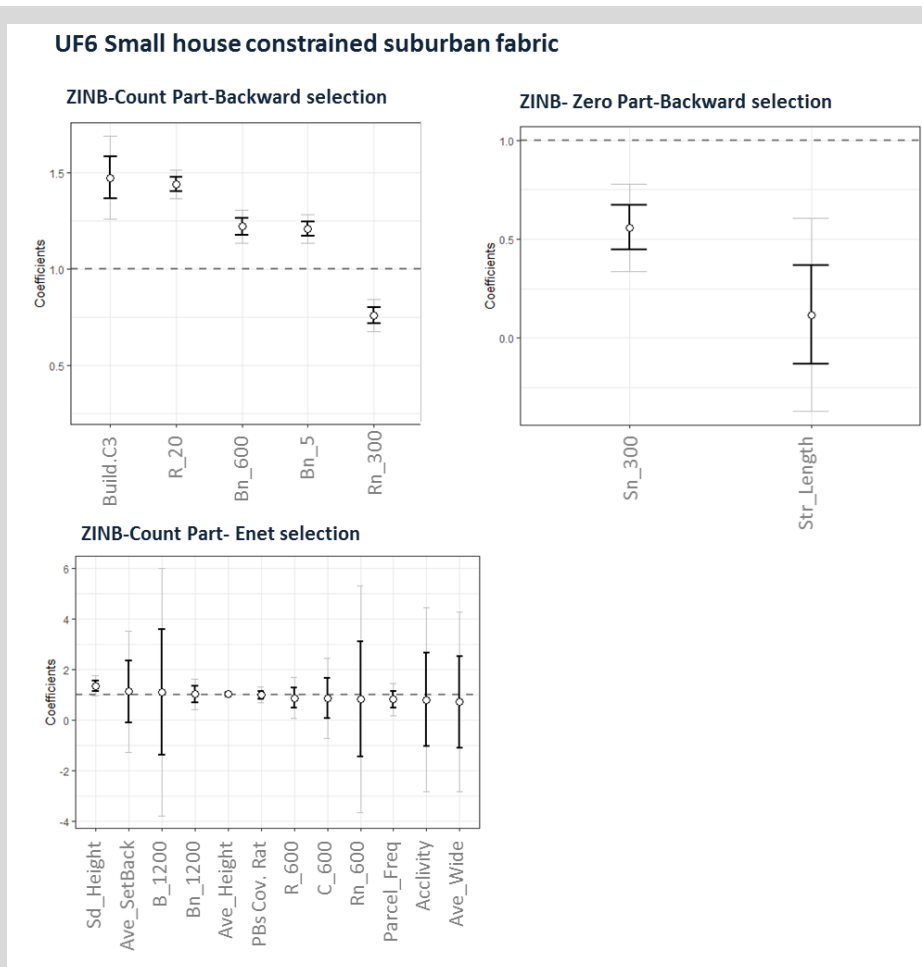


Figure 117 Selected variables from the Backward Regression and Enet Penalised Regression to the UF6 Model

### **Comparing the internal logics of urban fabrics**

When implementing the Enet penalised regression procedure to the individual urban fabrics, as in the previous case of the First-/Second-Age City partition, what we might expect is the identification of specific urban morphological parameters related to the retail distribution which show high statistical variability. Nonetheless, the different spatial scales at which the morphological regions are defined bring with them important theoretical differences.

First-/Second-Age City partition identifies two different subsystems of the French Riviera metropolitan area with different morphological and mobility characteristics; on the contrary, when decomposing the model considering individual urban fabrics, we are focusing on more homogeneous urban sub-spaces at the neighbourhood scale, losing the information of the spatial continuity and contiguity at the macro-scale. Indeed, when modelling each UF, we are generally considering a large number of small morphological regions characterised by a given urban fabric type and scattered in the overall study region. These urban fragments might be found both in the First- and in the Second-Age City, for which we already observed two different mobility systems. Moreover, the information about the spatial contiguity of different urban fabrics is not included. For example, historical centres (UF1) might be found both within

the First-Age City, surrounded by large planned urban developments of the XIX<sup>th</sup> and early XX<sup>th</sup> centuries, but also in isolated positions, and dispersed within the Second-Age City.

What we might expect from the implementation of Enet penalised regression to each urban fabric type is the identification of specific combinations of configurational and compositional urban parameters both at the intrinsic hyper-local scale (streetscape) and at the macro-scale (macro-contextual regions).

The main advantage of this application is that the selection of homogeneous urban areas reduces the variability of underlying street-scale compositional/geometrical descriptors. In other terms, a relevant part of the information about the urban physical form is already defined by the selected spatial area under study.

Within the **old constrained urban fabric of town-houses (UF1)**, the main morphological indicator related to retail presence is the street corridor effect and the building height: these indicators identify those streets bordered by aligned, continuous, and high facades. While the building coverage ratio and the built-up unit fragmentation characterised the models both at the global and at the First-/Second-Age City level, when focusing on UF1 neighbourhoods, the same indicators lose their importance. The reason underlying this observation is that the majority of streets in historical centres, having undergone a complete burgage cycle of densification (Conzen 1960), are surrounded by similarly high values of coverage ratio and built-up-unit fragmentation both observed in association with presence and absence of stores. Consequently, the streetscape descriptors which add a marginal attractive effect for retail presence is represented by the alignment of high facades along the street edges. Street acclivity and open space are negatively related to retail distribution. In Fig.X, we might observe some examples of the effect of these streetscape indicators with different observed retail presences.

The selection of specific configurational indicators allows us to further specify the relative position of these streetscapes within their local surrounding street-network. As previously discussed, these historical centres might be located both in the First- and Second-Age City macro-context. In historic centres within First-Age City, higher retail concentrations are observed along streets with strong pedestrian-oriented through-movement centralities at the meso- and macro-scales (600 and 1200 metres Betweenness). The combination of local-scale pedestrian-oriented centrality with large-scale vehicle-oriented centrality (300-metres and 20 minutes Reach) allow us to identify those small isolated historical centres surrounded by vast vehicle-oriented suburban areas of the Second-Age City.

Finally, the directional proximity and through-movement towards the coast contributes to the higher retail presence potential, even along those streets without the aforementioned streetscape properties (see example in Fig. 118).



Figure 118 A-B: example of streetscapes in historical centres with strong street corridor effect and high building façade values, positively related to retail presence. C-D: examples of streetscapes with streetscape conditions negatively correlated to retail activities: lower corridor effect (interruptions observed with small empty plot -C- and presence of specialized building -D-), higher activities, and lower building heights. E-F: examples of streetscape where despite the low corridor effect (buildings along only one side of the street), low building height (in E) and high activities (in F) they still have high retail densities due to their proximity to the coastline.

Within the **traditional urban fabric of the plain with adjoining buildings (UF2)**, retail activity is found with higher concentrations in connection to those street segments where fragmented large built-up units, street corridor effect and higher building heights are observed. While the street corridor effect and the building height are streetscape characteristics which UF2 shares with the historical centres (UF1), built-up unit fragmentation is a third descriptor exclusively selected for UF2 (and for which the highest OR value are observed).

Concerning built-up unit fragmentation, two observations should be made. Firstly, fragmented built-up units might also be expressed in terms of building contiguity: those streetscapes where a higher number of contiguous buildings is found provide an attractive urban morphological environment for higher retail presence. This characteristic is independently considered from the size/typology of the buildings composing the streetscape. Indeed, the negative effect of the presence of large-sized buildings (such as city theatres, hospitals, stadiums, etc.) is individually accounted for by a further specific indicator (Prevalence of C5 buildings whose surface is more than 4000 m<sup>2</sup>). This indicator identifies, for example, those streets both in central and semiperipheral areas of UF2 where independent large-sized buildings are built respecting the relationship of regularity with the surrounding

streets but detached from other buildings. The highest presence of retail activity is thus associated with differently-sized contiguous buildings.

Secondly, we should specify how small values of built-up-unit fragmentation is associated with a negative effect on small and average retail counts; conversely, these spaces might offer the larger surface sizes required by others urban functions among which large anchor stores (and which, in their turn, might attract small retailers in their surroundings). In this case, from a pedestrian perspective, we might still observe a streetscape characterised by a high street corridor effect with several shop windows, although only one (or few) large-surface stores are found. These two observations describe how low fragmentation values might be observed both associated with retail absence and with the presence of large commercial activities (Fig.119.C-D).



Figure 119 Image A-B) High Built-up-unit fragmentation associated with high number of stores. Image C) In the same street, with the same values of configurational and contextual descriptors, we might observe: on the left hand-side, high building fragmentation associated with eight small-sized stores at the ground floor. On the right hand-side, a lower built-up-unit fragmentation (only two contiguous building) with only 3 average-sized stores. Image D) Example of street with presence of exceptional large surface buildings with specialized function, the National Theatre of Nice. Image E) Example of small built-up-unit fragmentation associated with lower retail concentrations: individual buildings (Historical Palaces and Hotels) along the Promenade des Anglais.



When considering the configurational properties associated with the aforementioned streetscape descriptors, we might observe both pedestrian-oriented through-movement centrality at the mesoscale (600-metres betweenness) and large scale vehicle oriented centrality (20-minutes betweenness). Proximity to the coastline and squares (1200-metres) further increment the retail attractiveness of a street.

Among configurational indicators, straightness is often preferred over reach centrality; this observation and the specific selection of Nodes of degree 4 indicator highlight how the overall street pattern regularity is a positive factor in explaining retail distribution. When the planned regularity of UF2 becomes weaker, fewer retailers are observed: indicators such as building misalignment and open space variability are negatively related to store counts. The irregularity of the street network and of the streetscape might be observed in semiperipheral areas where, for different reasons, the original plans of the XIX<sup>th</sup> and early XX<sup>th</sup> century have not been fully implemented or modified since the second after-war period. Similarly, the urban morphological irregularity might be explained by the underlying irregularity of the site morphology (street acclivity with negative relationship with retail).

Finally, while for UF1, Enet penalised regression reduced the model from a ZINB to a simpler NB model, in the application to UF2, the ZINB model is preserved. The zero part of the model shows that retail absence is negatively linked to the street corridor effect, building fragmentation and parcel frequency. We might then infer that even if more irregular fragments of UF2 (with lower values of through-movement, density and building heights) are generally associated with lower retail counts, the presence of few contiguous or independent buildings close and aligned to the street edges might augment the probability to detect at least some retail activity.

As we might observe from the results of the Enet regression models on compact urban fabrics (UF1 and UF2) beyond the through-movement centrality (Betweenness centrality) identified in those traditional studied of Tab.2, a strong influence on retail distribution is also determined by a specific combination of streetscape descriptors (more than two thirds of the selected indicators). While the site and the street-network properties describe an overall and broad commercial potential of central and compact urban spaces, it is only considering the specific combination of streetscape measures which allows us to better comprehend the different urban morphological conditions related to the presence, the absence, and the magnitude of store distribution.

When considering the **discontinuous and irregular urban fabric with houses and buildings (UF3)**, the morphological indicators positively correlated to retail activity are the coverage ratio, the average building height, and the parcel frequency. Similar to what has previously been described for the First-/Second-Age Cities, when analysing semiperipheral and suburban areas, mostly composed by individual buildings and houses, the built-up-fragmentation is always low; consequently, the built-up-unit fragmentation is substituted by the parcel frequency. As already discussed in Chapter 8, the choice of the parcel frequency instead of the building frequency might be explained by the choice to consider the parcel and buildings within the first 50 metres from the street edges: the latter might not detect the presence of houses and buildings located beyond this proximity threshold. Differently from the previous

urban fabrics, the most important urban characteristics related to retail distribution within UF3 is the high meso-scale pedestrian-oriented through-movement centrality. Betweenness centrality at 1200 metres is found both related to retail count and negatively related to the absence of retail. From the zero part of the UF3 Enet-reduced model, we might infer that independently of the specific morphological properties of the streets within UF3, the only presence of high values of through-movement centrality, together with street corridor effect, provide the minimal conditions for stores to locate.

Several indicators are further identified by the Enet procedure, such as the presence of large buildings, average street openness, and pedestrian centrality, but their OR are slightly different from one with a relatively large confidence interval. As a consequence, we cannot conclude anything on the role of these indicators. Conversely, acclivity is still found negatively related to store distribution even within the UF3 context.

As previously highlighted in the model for the overall study area and for the First-/Second-Age City partition, **modern discontinuous urban fabric with big/medium-sized buildings (UF4)** plays a different role in the explanation of the retail activity, depending on the surrounding context within which this type of urban fabric is located. When implementing the Enet procedure to this specific urban fabric, we find both common characteristics, and, here again, two sub-groups of parameters endorsing the results about the double role of UF4 detected by the previous models.

Urban fragments classified as UF4 are traditionally composed by few main axes connecting the neighbourhood to the rest of the urban system and a larger number of short segments, cul-de-sacs, or longer and tortuous paths connecting sparse buildings to the main collector streets. The windiness of these residential streets are even more emphasised when considering the hilly residential suburbs scattered throughout the French Riviera. The combination of configurational parameters, such as micro- and meso-scale pedestrian straightness centrality (300 and 1200 metres), together with street length and street acclivity, allows us to discriminate between residential paths, loops, and cul-de-sacs from the main axes.

When these neighbourhoods are found more isolated within suburbs or grafted on connective fabrics UF7 (as described in Text-Box.6), vehicle-oriented mobility prevails. Consequently, retail is more probably found along those axes with the highest vehicle efficiency at the meso-scale (5 minutes). In these vehicle-oriented subspaces, stores are found in larger buildings located far from the street edges: car parks are traditionally interposed between buildings and street edges. Large anchor stores might also attract in their proximity small- and mid-sized stores (600-1200 metres straightness directional centrality toward anchor stores).

When UF4 neighbourhoods are found in a closer position or included within more central areas, pedestrian betweenness identifies the main paths connecting these residential neighbourhoods to the more mixed-use and compact areas. Fragmentation, similarly to the close central areas, becomes again an important factor for retail activity. Moreover, pedestrian betweenness and parcel frequency are also identified by the zero part of the model as negatively linked to the absence of retail.

Finally, when implementing Enet regression procedure to **residential suburban fabric on hills or in the plain (UF5)** and to **constrained suburban fabric of small houses (UF6)** similar outcomes might be outlined. As described in Text-Box.4, the penalized procedure reduces the original full ZINB models to simpler NB models in both cases. 11 and 12 variables are selected for UF5 and UF6, respectively. Despite the drastic reductions in the number of selected variables, the overall goodness-of-fit measures remain almost unaltered (Text-Box. 6). In both models, the meso-scale pedestrian-oriented through-movement centrality (1200 Betweenness), and the building coverage ratio are positively linked to retail. On the contrary, negative relations are found for street acclivity and average street openness. Other indicators are also selected, but their OR values are close to 1 in both models and also present large confidence intervals, providing weak evidence for the relationship between retail and morphological descriptors in suburbs.

## 12.3 A model for the typology of retail agglomerations

So far, we focused our attention on the outcomes of count regression models describing the spatial distribution of stores in terms of presence/absence and quantity. While Section 12.1 focused on the counting model selection phase, Section 12.2 described and discussed the outcomes of the variable selection procedures implemented considering different morphologically-based partitions of the French Riviera metropolitan area. An important aspect is still missing in our analysis: the retail fabric typologies identified in Part 3 have yet to be considered. In this section, we will present the results of the multinomial regression model explaining the spatial organisation of retail agglomeration types and completing the understanding of the relationship between retail distribution and urban form on the French Riviera.

As discussed in Section 11.1, a multinomial regression model is implemented in order to investigate if and how urban morphological properties might explain the typology of retail agglomerations associated with the street segments.

Since the model described in the previous sections provided us with the information about the retail presence/absence and magnitude, the multinomial model has been implemented only on the subset of the original street segments where retail is detected (about 11.6 thousand of street segments).

A first multinomial model is implemented, considering as a dependent variable the 10-levels Retail Fabric typologies (RFA procedure, Chapter 9), and, as independent variables, all 105 urban morphological variables at our disposal (as described in Chapter 11). The confusion matrix in Tab.20.a allows us to visualise the predicted and observed values for each RF; an overall accuracy of 75% is observed. The sensitivity and precision for each Retail Fabric typology are described in respectively Table 20.b and c respectively.

From these outcomes, we can deduce that in the French Riviera metropolitan area, a strong statistical relationship is observed between urban morphological descriptors and store agglomeration types. For almost all retail types, we observe values of sensitivity and precision

higher than 73 and 75% respectively; lower values are observed for three retail fabric families: car-oriented ribbons (RF6), pedestrian-oriented ribbons (RF5), and the fragmented pedestrian ribbons (RF9). Misclassification errors seem to occur especially between these two last retail fabrics.

To control for over-fitting problems, which might be linked to the small size of the dataset under analysis, the original dataset (of 11.6 thousand street segments) was divided in a training and a test subset (containing respectively 70% and 30% of the original features, randomly selected). We then implemented a second multinomial model on the training test and the resulting model was applied on the test set; an accuracy of 72.56 % on the test set suggests the absence of over-fitting problems in our original model.

Figure 122 allows us to visually compare the observed and predicted values of RFA in the geographical space of the French Riviera: a limited number of misclassifications might be observed along those streets with an intermediary position among two or more retail fabric types.

		Observed										a									
		RF 1.a	RF 1.b	RF 2.a	RF 2.b	RF 2.c	RF 3.a	RF 3.b	RF 4.a	RF 4.b	RF 5.a	RF 5.b	RF 6.a	RF 6.b	RF 7	RF 8	RF 9.a	RF 9.b	RF 10		
Predicted	RF 1.a	418	5	5	27	5	0	2	5	17	0										484
	RF 1.b																				
	RF 2.a	22	352	21	33	5	0	0	0	2	3										438
	RF 2.b																				
	RF 2.c																				
	RF 3.a	16	22	268	42	1	0	0	0	5	1										355
	RF 3.b																				
	RF 4.a	33	27	41	593	46	2	4	5	31	1										783
	RF 4.b																				
	RF 5.a	8	11	0	63	690	35	28	2	256	8										1101
RF 5.b																					
RF 6.a	0	0	0	1	45	426	0	1	92	134										699	
RF 6.b																					
RF 7	8	0	0	15	10	1	192	7	3	0										236	
RF 8	7	0	0	1	6	1	5	299	44	11										374	
RF 9.a	10	4	1	41	342	186	17	37	1038	243										1919	
RF 9.b																					
RF 10	0	0	0	1	35	357	0	13	386	4440										5232	
		<b>522</b>	<b>421</b>	<b>336</b>	<b>817</b>	<b>1185</b>	<b>1008</b>	<b>248</b>	<b>369</b>	<b>1874</b>	<b>4841</b>										<b>11621</b>

		Observed										b									
		RF 1.a	RF 1.b	RF 2.a	RF 2.b	RF 2.c	RF 3.a	RF 3.b	RF 4.a	RF 4.b	RF 5.a	RF 5.b	RF 6.a	RF 6.b	RF 7	RF 8	RF 9.a	RF 9.b	RF 10		
Predicted	RF 1.a	80%	1%	1%	3%	0%	0%	1%	1%	1%	0%										0%
	RF 1.b																				
	RF 2.a	4%	84%	6%	4%	0%	0%	0%	0%	0%	0%										0%
	RF 2.b																				
	RF 2.c																				
	RF 3.a	3%	5%	80%	5%	0%	0%	0%	0%	0%	0%										0%
	RF 3.b																				
	RF 4.a	6%	3%	12%	73%	4%	0%	2%	1%	2%	0%										0%
	RF 4.b																				
	RF 5.a	2%	3%	0%	8%	58%	3%	11%	1%	14%	0%										0%
RF 5.b																					
RF 6.a	0%	0%	0%	0%	4%	42%	0%	0%	5%	3%										0%	
RF 6.b																					
RF 7	2%	0%	0%	2%	1%	0%	77%	2%	0%	0%										0%	
RF 8	1%	0%	0%	0%	1%	0%	2%	81%	2%	0%										0%	
RF 9.a	2%	1%	0%	5%	29%	18%	7%	10%	55%	5%										0%	
RF 9.b																					
RF 10	0%	0%	0%	0%	3%	35%	0%	4%	21%	92%										0%	

		Observed										c									
		RF 1.a	RF 1.b	RF 2.a	RF 2.b	RF 2.c	RF 3.a	RF 3.b	RF 4.a	RF 4.b	RF 5.a	RF 5.b	RF 6.a	RF 6.b	RF 7	RF 8	RF 9.a	RF 9.b	RF 10		
Predicted	RF 1.a	86%	1%	1%	6%	1%	0%	0%	1%	0%										0%	
	RF 1.b																				
	RF 2.a	5%	80%	5%	8%	1%	0%	0%	0%	0%	1%										1%
	RF 2.b																				
	RF 2.c																				
	RF 3.a	5%	6%	75%	12%	0%	0%	0%	0%	1%	0%										0%
	RF 3.b																				
	RF 4.a	4%	3%	5%	76%	6%	0%	1%	1%	4%	0%										0%
	RF 4.b																				
	RF 5.a	1%	1%	0%	6%	63%	3%	3%	0%	23%	1%										1%
RF 5.b																					
RF 6.a	0%	0%	0%	0%	6%	61%	0%	0%	13%	19%										0%	
RF 6.b																					
RF 7	3%	0%	0%	6%	4%	0%	81%	3%	1%	0%										0%	
RF 8	2%	0%	0%	0%	3%	0%	1%	80%	12%	3%										0%	
RF 9.a	1%	0%	0%	2%	18%	10%	1%	2%	54%	13%										0%	
RF 9.b																					
RF 10	0%	0%	0%	0%	1%	7%	0%	0%	7%	85%										0%	

Table 20 Confusion matrix displaying the results of the multinomial model for the prediction of the retail fabrics (a), its sensitivity (b) and precision (c).

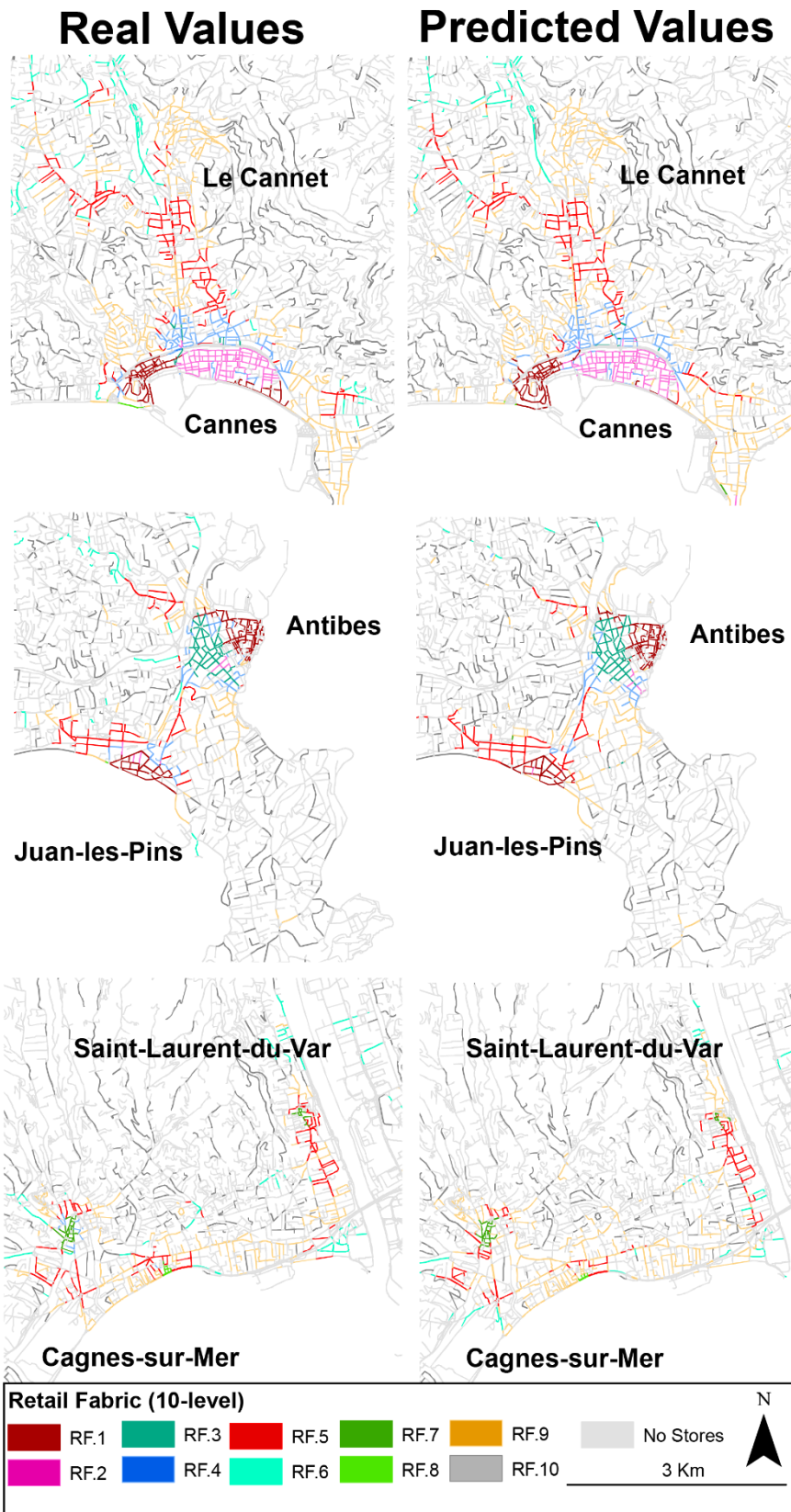


Figure 120 Projection in the geographical space of the observed values (left) and of the predictions of the multinomial model for the retail fabrics typology

As for the count regression models, here again, we implemented a penalised Elastic Net multinomial regression in order to define the key subset of morphological indicators able to identify the localisation of the retail agglomeration types. The smaller size of the dataset allows us to explore 100 values of lambda. Similar to the previous implementation of the Enet procedure, the best solution of lambda is automatically selected in correspondence to the model with the smallest BIC value.

From the original 105 variables, the reduced model selected a total of 36 regressors among which include: i) 9 streetscape descriptors, ii) 6 meso- and macro-scale contextual variables; iii) 10 configurational and 12 directional configurational properties. An overall accuracy of 72.45% is observed for the reduced model. In Table.21, we summarised the selected variables for each retail fabric and the corresponding ORs.

	RF1	RF2	RF3	RF4	RF5	RF6	RF7	RF8	RF9	RF10
	a	b	a b c	a b	a b	a b	a b	a b	a b	RF10
R_300	1,776									
S_300										0,902
R_600						0,778				0,740
S_600										0,276
S_1200				1,430						
S_5				1,116						
Sn_5									0,855	0,823
B_20						1,020				
Rn_20						1,027				
Sn_20					2,227					0,860
R_300_Squares							1,374			
B_300_Squares							1,108			
S_600_Squares	1,017								0,824	
R_600_Squares					0,849					
S_1200_Squares		1,133		1,102					0,986	
R_600_Coast	1,666									
S_600_Coast										0,995
S_2400_Coast		1,242								0,958
R_600_AS			1,414							
S_600_AS		1,221								
R_1200_AS			1,122						0,999	
S_1200_AS									0,844	
Build. Heigh Misalignement					1,049					
Average Set Back										1,022
Street acclivity										1,224
Pre. Specialised Build.						1,708				
Average Open Space wide						0,980				1,069
Parallel Façades						0,853				
Average Height						0,933				0,905
PB Coverage Ratio									0,980	0,708
UF1								6,443		
UF2				1,306	1,366	0,667				0,444
UF5										1,906
UF4						1,691				
UF7						1,206				
First-Era City						0,650		0,821		

Table 21 Odds Ratios of the ElasticNet penalised regression procedure applied to the multinomial model describing the relationship between retail agglomeration types (10-levels RFs) and urban morphological indicators. A total of 36 regressors are selected.

With the underlying condition that retail is present (which follows the considerations discussed in the previous sections), we now analyse the outcomes of the reduced model, which help us to better understand the morphological characteristics associated to each Retail Fabric type.

RF1, touristic, and leisure-oriented agglomerations are most likely to be found in those urban spaces characterised by high network densities at small scales (300 metres), close to the coast, and to public open spaces (both accessible within 600 metres). While the first condition identifies almost all the historical centres of the French Riviera, both along the coastline and the inland, the second morphological condition allows us to differentiate between the two groups. The presence/accessibility towards squares and public spaces further differentiates between larger historical centres (such as the historic centre of Nice, Antibes, Cannes etc.) and the small fishermen villages (Cros-de-Cagnes, Villefranche-sur-mer, etc.) where a lower number of public spaces is observed. As we already observed in Section 5.5, RF1 are not only observed in historical city centres, but also in more recent dense and planned areas in the proximity of the city centre such as the case of Juan-les-Pins or the lower part of the Le-Cannet (in Cannes). These specific spaces, classified as more regular and planned (UF2), still present high values of network density.

RF2 encompasses three variations of clothing and leisure-oriented retailscapes. As already observed in Section 5.5, these retail agglomeration types are found in the city centres of Nice and Cannes, and, with a smaller spatial extent, in Menton and Antibes. The reduced multinomial model further details the specific morphological characteristics of the city centres where RF2 is found: three directional configurational centralities considering, more specifically, the straightness proximity towards the coastline, anchor stores, and squares. The specific selection of straightness centrality measures is associated with the underlying regular urban grid typical of the planned urban fabric extensions of the XIX<sup>th</sup> and early XX<sup>th</sup> centuries over the historical centres where RF2 is found.

As we might observe, the proximity to the coastline and squares is a common locational factor of both leisure-oriented RF1 and RF2, the former characterised by smaller scales (600 metres) while the latter by meso-scales distances (600-1200-2400 metres respectively to anchor stores, squares and coast). Beyond the differences in the spatial scales and the network regularity of the underlying street network, an important additional consideration considering the network configurational properties might explain the differences in location between these two retail fabrics.

As we observed, high street network densities are associated with RF1; this morphological characteristic is also often associated with higher plot and building fragmentations. All these characteristics strongly limit the locational potentials for large anchor stores (more specifically, department stores particularly present in RF2), requiring larger commercial surfaces, and, consequently, the surrounding retail functional mix of these urban spaces. As for most of the studies in urban morphology, we also confirm that high street network densities promote functional mix. This relationship is twofold: on the one hand, street network density provides the morphological condition which attracts functional mix, an observation already made by Jane Jacobs (Jacobs 1961); on the other, it has a repulsive effect

on those department stores (especially clothing stores) which bring with them functional specialisation.

RF3 is associated with the proximity to anchor stores, too. Nonetheless, differently from RF2, these spaces are located further away from the coastline and squares. These two morphological features represent two important elements able to differentiate between leisure-oriented and service-oriented functional polarisation within dense and pervasive retail fabrics.

RF4 is the fourth retail agglomeration type characterised by high store densities and spatial pervasiveness. Differently from the previous RFs, this specific class is associated with more residential compact areas. When observing the multinomial regression outcomes, meso-scale straightness indicators are selected (1200 metres and 5-minutes) together with the straightness proximity to squares.

As for the previous retail fabric, the straightness centralities are selected above the regular reach centrality highlighting the importance of the regularity of the underlying street network. In fact, as already observed in Section 10.3, RF2, RF3 and RF4 are all found in UF2, the urban fabric type characterised by a regular street grid.

In Section 9.5, we observed how rooted in core areas and town centres, pedestrian ribbons (RF5.a) are radially disposed along the main boulevards and avenues. These main streets play the important role of collector of the surrounding residential-oriented neighbourhoods where a more scattered, out-of-ribbon retailscape is observed (RF.b). Further pieces of information are provided by the reduced multinomial model: firstly, as we might observe in Tab.21, RF5 is associated with macro-scale vehicle-oriented straightness centralities (20 minutes); this indicator, as discussed in Chapter 7, corresponds to a measure of network vehicle efficiency. Moreover, these concentrations are most likely to be found within the traditional urban fabric of the plain with adjoining buildings UF2 (as already observed in Tab.15 Section 10.3). A third morphological aspect associated with RF5 is the negative correlation with square accessibility; finally, a positive relationship is found with building height misalignment.

The combination of these four indicators reveals how ribbon retail agglomerations are found along and close to those street collectors within an overall area with dense built-up characteristics in semiperipheral areas. These longitudinal axes, from the core centre of cities towards more residential areas are those urban spaces where pedestrian and vehicle-oriented mobility models coexist. The building misalignment reveals also a gradual transformation from core compact central areas towards more residential irregular urban fabrics (mainly UF3 but also UF4, 5 and 6).

The transformation from pedestrian to vehicle-oriented urban spaces becomes definitive in those peripheral areas where vehicle-oriented retail ribbons and specialized retail agglomerations RF6 are observed. RF6 is, indeed, positively correlated with large-scale centrality and through-movement measures (20 minutes) and negatively related to meso-scale pedestrian-oriented centrality (600 metres). Nonetheless, the simple network configuration centrality is not a sufficient condition to observe a specialised retail agglomeration; morphometric descriptors at the local, meso- and macro- scales should also



be considered. At the macro-scale, RF6 is most likely to be found in a Second-Age City subsystem, and, more specifically in correspondence with modern discontinuous urban fabrics (UF4) or connective artificial fabrics (UF7). Further streetscape morphometric descriptors at the hyper-local scale define the streetscape intrinsic conditions: buildings along these streets are most likely to assume specialised forms, with low heights, closer proximity to the street centreline, although not disposed with regularity along the street edges.

So far, we have described the urban morphological characteristics associated with retail fabric types located within core areas of large cities and in suburban areas. RF7 is a mixed isolated retail cluster with anchor stores, which has been found in several average-sized cities and four sub-centres included within the municipality of Nice; here again, ribbon agglomerations are grafted to these clusters. The multinomial model associated RF7 with two directional configurational variables: proximity to squares, and, more specifically, directional through-movement toward squares. These two indicators describe how, in average-sized cities, the role of the central square becomes primordial as the originating core from which the socioeconomic vivacity might be spread in the surrounding spaces following the main collector axes. As discussed in the introduction, average-sized cities are those urban spaces which most suffer from the current transformations of the retail system, as in the case of the city of Grasse. The outcomes of our model disclose important suggestions about the possible strategies for the commercial development of these average-sized cities: the focus of the requalification activity should start from the central square, in which the socio-economic core of these cities are located, and, following the main axes toward more residential areas, help structure pedestrian retail ribbons.

RF8 corresponds to the retail agglomeration type found in correspondence with isolated, small inland villages and fisherman villages along the coast. As discussed in Section 9.5, these retail agglomerations are characterised by a leisure/touristic functionality characteristics influenced by their historical heritage and/or environmental characteristics. Their location is, consequently, most likely to be defined by morphological characteristics, rather than accessibility/centrality measures: these small historical centres and villages are rarely found along main pedestrian/vehicle axes but rather located on peculiar natural configurations (such as the hill-top for Eze, Saint-Paul-de-Vence, etc. or the sea-side for the village of Saint-Jean-Cap-Ferrat, Villefranche-sur-mer, etc.). For these reasons, no configurational or directional indicator is relevant for the definition of the localisation of these retail agglomerations. Their localisation, on the contrary, is well explained by the combination of two morphometric descriptors at the meso- and macro- scales: RF8 are found in old constrained fabric of town-houses (UF1) which are not inserted in a First-Age City system.

RF9 corresponds to fragmented pedestrian-oriented retail ribbons and small isolated concentrations in suburbs. Differently from the other retail fabric types, the outcomes of the multinomial model for RF9 only select urban morphological variables with a negative relationship: PB coverage ratio, meso-scale vehicle-oriented straightness, straightness proximity towards squares and anchor stores at the large scale (1200 metres). For this retail fabric class, after RF6, we found the lowest sensitivity and precision values (respectively, 56

and 54%); misclassification errors are classified as pedestrian or vehicle-oriented retail, or, eventually, isolated stores. These observations confirm how RF9 might be considered to be a transitioning agglomeration type between the three aforementioned retail fabric types.

Finally, when considering isolated stores, RF10, our multinomial models, selects fifteen variables. These indicators describe a negative relationship with: i) pedestrian and vehicle-oriented centrality values at almost every scale (300 and 600 metres, 5 and 20 minutes); ii) proximity toward the coast; iii) coverage ratio and building height. These very characteristics were positively related to retail activity in the count regression models, confirming how their presence is more favourable for more retail-intensive RFs than for isolated stores. Along streets with higher values of activity, building setback and openness, we are more likely to observe isolated stores. Two urban fabric types are also observed in correspondence with this store distribution: a positive relationship with suburban residential fabrics (UF5) and a negative relationship with traditional planned urban fabrics (UF2).

To conclude, we might observe how from the most compact areas towards the most peripheral and low-density areas, the reduced multinomial model requires an increasing number of variables to discriminate each retail fabric typology: i) few indicators of proximity to squares, coast and/or anchor stores are sufficient when considering core areas of large cities and isolated average-sized cities; ii) specific combinations of configurational and morphometric indicators when investigating semiperipheral and suburban areas; iii) morphometric, configurational and proximity all intervene when considering isolated stores and isolated centralities. Finally, it is particularly challenging to predict the presence of isolated stores in suburban areas.

*In the fourth part of this dissertation, we focused on the analysis of the statistical relationship between retail spatial distribution and urban morphological descriptors.*

*In Chapter 10, we reviewed all the urban and retail form descriptors elaborated in the previous parts, describing how through the street segment we have been able to describe the retail and urban systems with a common spatial unit and a coherent methodological framework.*

*In Chapter 10.2, we reviewed the statistical approaches implemented in previous works investigating the relationship between retail and urban form. The development of more sophisticated procedure is therefore outlined as an essential aspect to overcome the overly simplistic procedures traditionally implemented.*

*Chapter 11 describes the modelling procedure implemented in this dissertation: we firstly discuss in Section 11.1 how generalised linear regression models allow us to consider the specific nature of our data (count), but also the retail absence (with Zero Inflated and Zero Alternated approaches) and of retail agglomeration typologies (with multinomial regressions). In section 11.2, the multicollinearity problem is discussed; among the different procedures available, penalised regressions retained our attention. Section 11.3 describes how to introduce contextual variables within regression analysis, while Section 11.4 defines the set of measures and functions implemented for the evaluation of our models.*

*In Chapter 12, we described and discussed the outcomes of our models. Firstly, in Section 12.1, the model selection procedure demonstrated how on the French Riviera, ZINB models seem to better perform when modelling the retail distribution. Similar results are also obtained when considering several urban morphological sub-regions. The projection in the geographical space of the model outcomes and errors allowed us to observe how models perform in different urban areas. Once we defined the most adapted model, Section 12.2 presented the results of the penalised regression procedures for each morphological region: from these outcomes several insights about the process relating the spatial distribution of stores and the urban morphological properties is described. Finally, in 12.3, we described the outcomes of the multinomial regression analysis modelling the relationship between urban morphological characteristics and retail agglomeration typology.*

*The rich and detailed information provided by these models allows us, in the conclusion section, to answer to our initial research questions.*

*Dans la quatrième partie de cette thèse, nous nous sommes concentrés sur l'analyse de la relation statistique entre la distribution spatiale de la distribution et les descripteurs morphologiques urbains.*

*Au Chapitre 10, nous avons passé en revue tous les descripteurs de formes urbaines et commerciales élaborés dans les parties précédentes, expliquant comment nous avons pu décrire à travers le segment de la rue les systèmes urbains et commerciales avec une unité spatiale commune et un cadre méthodologique cohérent.*

*Au Chapitre 10.2, nous avons passé en revue les approches statistiques mises au point dans des travaux antérieurs sur la relation entre le commerce de détail et la forme urbaine. Le développement de procédures plus sophistiquées est donc présenté comme un aspect essentiel pour surmonter les limitations des procédures trop simplistes traditionnellement employées.*

*Le Chapitre 11 décrit la procédure de modélisation mise en œuvre dans cette thèse: nous discutons tout d'abord dans la section 11.1 comment les modèles de régression linéaire généralisée permettent d'envisager la spécificité de nos données (comptage), mais aussi l'absence de commerces de détail (avec des approches par excès de zéros et zéros alternés) et les typologies d'agglomération du commerce de détail (avec régressions multinomiales). Dans la Section 11.2, le problème de multicolinéarité est discuté; parmi les différentes procédures disponibles, les régressions pénalisées ont retenu notre attention. La section 11.3 décrit comment introduire des variables contextuelles dans l'analyse de régression, tandis que la Section 11.4 définit l'ensemble des mesures et fonctions implémentées pour l'évaluation de nos modèles.*

*Au Chapitre 12, nous avons ensuite décrit et discuté les résultats de nos modèles. Tout d'abord, dans la Section 12.1, la procédure de sélection des modèles a montré comment, sur la Côte d'Azur, les modèles ZINB semblent mieux fonctionner lors de la modélisation de la distribution du commerce. Des résultats similaires sont également obtenus, en considérant plusieurs sous-régions morphologiques urbaines. La projection dans l'espace géographique des résultats et des erreurs des modèles nous permet d'observer les performances des modèles dans différentes zones urbaines. Une fois le modèle le plus adapté identifié, la Section 12.2 présente pour chaque région morphologique les résultats des procédures de régression pénalisée: à partir de ces résultats, plusieurs informations sur le processus reliant la distribution spatiale des magasins et les propriétés morphologiques urbaines sont décrites. Au paragraphe 12.3, nous décrivons les résultats de l'analyse de régression multinomiale, modélisant la relation entre les caractéristiques morphologiques urbaines et la typologie de l'agglomération de magasins.*

*Enfin, les informations riches et détaillées fournies par ces modèles nous permettent, dans la partie de conclusion, de répondre à nos questions de recherche initiales.*

# **Conclusions and Perspectives**

In this research, we investigated the relationship between the spatial patterns of retail activity and the properties of the form of the physical city following the line of research initiated twenty years ago by Hillier's Movement Economy Theory. Several theoretical and methodological developments have been proposed and implemented in this dissertation, allowing us to further understand the relationship between urban form and retail spatial patterns in the French Riviera metropolitan area.

On the one hand, based on retail geography literature, a new analytical protocol (RFA) has been specifically conceived for the identification of retail agglomeration types integrating functional and spatial characteristics of retail clusters. Subsequently, retail spatial distribution has been investigated considering three main pattern descriptors: the store presence/absence (Boolean), the magnitude (count measure) and the RFA agglomeration typology (categorical descriptor). On the other hand, urban form theories were at the base-level of a second set of measures proposed for the description of the form of the physical city. Beyond the traditional street-network configurational properties, a large number of streetscape descriptors have been proposed, inspired by recent urban design studies. From urban typo-morphological t, a specific protocol has been developed for the identification of the urban morphological context at the meso- and macro- scales within which each street segment is located (MFA protocol).

The two sets of retail and urban physical form measures have been finally investigated through the implementation of different count regression approaches. These modelling procedures, which represent a methodological innovation when considering available literature, allowed us to describe the overall strength of the statistical relationship between the urban form and retail. The implementation of penalised regression procedures allowed us to identify the specific combination of morphometric variables most related to each characteristic of the retail spatial patterns; these procedures have been implemented considering different morphological spatial partitions of the study area.

The methodological and thematic findings of this dissertation might be resumed as follow:

1) Two innovative network-constrained and location-based procedures have been proposed and implemented for the description of the retail and urban fabric typologies. Both the protocols consider: 1) the implementation of morphological (or morpho-functional as far as RFA is concerned) indicators through geoprocessing algorithms; 2) the identification of spatial and statistically significant patterns for each indicator; finally, 3) the combination of these morphological patterns with AI-based classification analysis. These three-step procedures provide a data-driven and bottom-up identification of urban and retail fabrics in large metropolitan areas. The methodological challenges to identify urban and retail morphological regions from the street point of view have been overcome. The street became our metrical, geometrical, and statistical reference for the implementation in the whole dissertation, allowing us to bridge retail and urban form measures, but also different urban form approaches which are traditionally independently analysed.

Secondly, we showed the superiority of the Zero Inflated Negative Binomial regression model on traditional linear modelling approaches adopted in most of the reviewed works. This

specific model overcomes the main problem of heteroscedasticity and over-dispersion of the retail count data without any ambiguous data transformation. Moreover, the zero inflation observed in retail distributions is specifically investigated with ZI and ZA procedures. When considering the full models, ZINB shows the best fitting scores.

Thirdly, the implementation of Backward and Penalised Regression procedures allowed us to empirically test these two variable selection approaches. The outcomes of these two procedures (both the overall goodness-of-fit measures, and the selected variables statistical parameters) suggest the superiority of Penalised Regression procedures. Penalised regression procedures allow, on the one hand, to identify the most significant subset of variables explaining the relationship between retail distribution and urban form for each morphological sub-region; on the other, they further specify the overall form of the models between ZINB and NB models (with the maintain/exclusion of the zero part).

The aforementioned methodological developments allowed us to investigate the thematic questions which have been proposed at the beginning of this work (Chapter 2) and propose our answers backed by empirical research on the case study of the French Riviera.

***(1) Can the contextual information of the urban form help in better understanding the 'network effect' on retail distribution? If yes, how?***

We will answer this question positively. The consideration of the contextual morphological regions and macro-regions allowed us to highlight how street network configurational properties might represent both general characteristics but also specificities of a given urban fabric or morphological region.

On one hand, we highlighted how the through-movement street-network centrality represents the most important street network property explaining retail distribution. Through-movement might be considered a *constant*, which is related to the retail distribution and which has always been selected among the most important explanatory variables. These outcomes are aligned with what has already been observed in previous studies investigating the relationship between retail and street-network configuration (Porta 2006, Peponis 2008, Fusco et al. 2017, etc.).

On the other hand, the separate analysis of morphological regions allowed us to further appreciate the specific characteristics of the network properties in relationship with retail distribution. When considering the through-movement properties, while the pedestrian meso-scale is the most important in unplanned organic forms (which on the French Riviera corresponds both to historical centres and most residential suburbs), vehicle through-movement at the meso-scale is associated with both traditional and modern planned fabrics. Moreover, new specific street-network properties are selected in certain subspaces, such as the straightness centrality in modern and traditional planned areas.

***(2) Beyond the network effect, are there other characteristics of the urban form inhibiting retail presence and development? If yes, which ones?***

Once again, our answer is positive. Network configurational properties are not the only morphological variables which might help in the identification of retail and often other streetscape descriptors might show a higher potential to influence the retail distribution. This information is even more true when considering compact and modern areas where specific combinations of configurational and compositional descriptors are found to be differently related to retail.

Some streetscape descriptors show an important role when studying the retail spatial distribution. More particularly, the building coverage ratio, built-up-unit fragmentation (contiguity), and street corridor effect, are frequently found positively related to retail distribution, while on the contrary, street acclivity and street openness are negatively correlated. Here again, considering the contextual information, allowed us to understand the 'streetscape effect' or, in other terms, how specific streetscape variables might assume different meanings in different subspaces. For example, we have observed how specific combinations of urban fabrics and morphological descriptors at the meso- and macro-scales might explain the retailscape.

***(3) Can specific urban morphological characteristics be linked to the spatial distribution of retail agglomeration types? If yes, which characteristics play the most important role?***

Even to this last question a positive answer is given. Specific urban morphometric variables are associated with different retail morphofunctional typologies. More specifically, we highlighted how specific attractors such as the coastline, anchor stores, and squares might influence the functional specialisation of the core areas of cities where pervasive retail fabrics are observed. Configurational and compositional descriptors become more important for the discrimination among semiperipheral retail fabrics. Finally, the morphological characteristics of streets and the vehicular street-network configurational centrality measures are fundamental to explain the localisation of specialised concentrations in peripheral areas.

In conclusion, this dissertation has demonstrated how beyond the network configurational properties, a more complex combination of urban design characteristics within different morphological contexts might be correlated to the retail distribution, and, consequently to the socio-economic attractiveness of urban spaces.

## **Limitations and future perspectives**

All throughout this research work, and more specifically in the modelling part, a large number of research ideas and questions have continuously emerged and stimulated my interest in different aspects of this dissertation. The hidden challenge during the whole doctoral research has been to keep focused on the main research questions without getting lost in interesting, but diverging issues for the overall economy of this dissertation. For this reason, several interesting thematic and procedural ideas have been left unexplored and/or only marginally discussed.



As introduced in the introduction of this work, the modularity of this dissertation allows us to propose several developments on individual sub-parts, still keeping the overall methodological framework which might allow comparing future improvement with the results presented in this dissertation.

When considering the two procedures developed in Part 2 and Part3, respectively the RFA and the MFA, many methodological improvements might be proposed. Despite the strong theoretical background on which each indicator has been specifically conceived and implemented, there is still an interesting work of sensitivity analysis which might consider different metrical dimensions, thresholds, and radii which have been chosen. Similarly, other IA-backed clustering procedures approaches might be investigated and implemented. Validation of the proposed protocols through the application of other case studies is also an important perspective.

This work is already ongoing: since MFA has been developed during the first half of the doctoral research activities<sup>99</sup>, in the last two years MFA has already been applied to other case studies, such as Osaka (Japan) and Marseille (France) metropolitan regions, Brussel (Belgium), and Izmir (Turkey)<sup>100</sup>. These implementations allowed us to test, develop, and validate the MFA procedure to other case studies beyond the French Riviera, with different data sources and different urban contexts. In Araldi et al. (2018), we specifically discussed how MFA might be easily implemented and adapted to different case studies. The outcomes of these new applications demonstrated not only the validity of MFA protocol but also its flexibility and adaptability to different data sources.

Beyond the technicalities of the discussion about the individual morpho-functional/morphometric indicators, the statistical procedures, or the classificatory approach, a particularly critical point is the automation of the overall procedure. Only by providing a unique and ergonomic tool (or a set of tools similar to the street network configurational analytical tools reviewed in Chapter 3 or in MORPHO by Olivera 2013, etc.) might the procedure be easily tested and applied to a larger number of case studies, and, more interestingly, allow both diachronic and synchronic comparative studies.

Differently from MFA, the RFA procedure has been developed in the second half of the doctoral research activities<sup>101</sup> and, consequently, it has still to be tested on other case studies.

Starting from October 2019, a one-year collaboration between ESPACE laboratory and the private company Kinaxia will allow us to automatize and optimise both MFA and RFA protocols, enabling new possibilities and the research towards new research questioning about both urban form and retail. Within this framework, all French large urban areas will be

---

<sup>99</sup> First diffusion of MFA procedure at ICCSA conference and result at ISUF conference in July 2017 (Araldi and Fusco 2017, Fusco and Araldi 2017).

<sup>100</sup> Within the framework of, respectively, Joan Perez post-doc *Coastal cities* project with Professor Takashi Fuse (The University of Tokyo), Madeleine Guyot Phd with Isabelle Thomas (Catholic University of Louvain) and Irem Erin PhD research project with Professor Ebru Cubukcu (Dokuz Eylul University). Papers with the presentation of these new MFA implementation are currently under edition.

<sup>101</sup> First diffusion of RFA procedure and outcomes at Spatial Analysis and Modeling conference, 8th-9th sept. Tokyo (Araldi and Fusco 2018)

used as a case study, allowing a comparison of the urban morphology at the national scale. Together with the theoretical research questioning which emerged within this context, several technical and methodological challenges will also arise from the larger amount of data to study. The outcomes of these analyses will be integrated by the private company Kinaxia in their City Scan<sup>102</sup> services, demonstrating the interest in the research activities developed in this dissertation which have also interested not only our current research partner, but also a private company. The application of the proposed protocols at the French national level will also allow us to generalize results found on the French Riviera and highlight local specificities when compared to other large French urban areas.

Methodological innovations might also be investigated in relation to the modelling phase of this dissertation (Part 4). Other penalised regression approaches might also be implemented on the same set of variables, further verifying the results highlighted in Section 12.2. Beyond penalised regression, Multi-Level Modelling or Network-Constrained Geographical Weighted Regressions represent two interesting methodological procedures. The former might allow for the inclusion of the hierarchical nested organisation of our dataset, and, therefore, allow us to consider the between-class variability, which (as previously highlighted) might be detected in our outcomes. The latter might inform us about the spatial variation of the statistical correlation coefficient between our morphometric variables and the retail distribution.

Considering the large size of the dataset, AI modelling procedures might also be considered both for the modelling and the variable selection phases. Some preliminary exploratory analysis has already been implemented using Bayesian Networks and Random Forests learning methods; nonetheless, the results are not yet conclusive and further analysis should be carried on.

Beyond the methodological limitations and perspective, this research leaves open a much wider range of thematic questions. Three main axes might be considered.

Firstly, some retail localisation factors have been deliberately set aside in this study in order to focus on the specific relationship between the properties of the physical city and retail distribution. If enlarging the research goal to the retail localisation problem, other registers of the urban form should be considered (Chapter 2): the inclusion of other variables such as socioeconomic and accessibility factors become of paramount importance (as demonstrated in the previous dissertation of Scoppa and Sevtsuk ). Population distribution, the socio-professional characteristics of neighbourhood, the consumption index (*Indices de Disparité des Dépenses de Consommation* CCI FRANCE, INSEE) and all the characteristics of the social register of the urban form might be easily integrated within this model. As suggested by Okabe (2012), the same street-based Thiessen tessellation at the origin of the MFA protocol might be used once again for the attribution of the social descriptors to each street element with simple proportional factors. Integrating these new variables within the same statistical modelling procedure developed in this dissertation would allow investigating further interesting questions, for instance: which is the marginal importance of the urban form

---

<sup>102</sup> <https://www.cityscan.fr/accueil>

characteristics on the retail distribution? What is the role of each urban form register? How might their combined effect explain retail distribution? What is the role of accessibility?

As discussed in the first part of this dissertation, the retail distribution is also considered in urban studies as a proxy of a place vibrancy/vivacity/attractiveness. Several other sources of data describing different aspects of the social attractiveness of streets might be modelled conjointly or independently from the retail distribution, providing new insights into the social and economic functioning of cities and their relationship with the urban morphology. For example, using the data from Airbnb and Twitter (which we have gathered in the last two years) as dependent variables in our modelling approach, it might allow us to investigate how leisure and touristic activities are related to the urban environment; this specific question is an interesting aspect when considering the French Riviera where the touristic sector represents one of the most important economic activities.

Secondly, the temporal dimension is one of the most important absent aspects which has not been investigated in this dissertation. The exclusion of this component is, here again, a specific methodological choice: this dissertation might be considered in a wider two-step research design considering longer-term goals. With this work, we specifically focused on the development of robust and reproducible protocols for the cross-section analysis of the urban morphology and retail distribution with a fine-grained and easily accessible data (at least in the French and most of the European countries or with open-source datasets). These protocols will represent, in a second phase, the cornerstone of future research aiming at introducing the temporal dimension in the study of retail and urban fabrics and their generative processes. On one hand, diachronic comparative analysis of the retail and urban morphology individually considered might reveal several important aspects of the transformation process and resilience properties of these two components. On the other, introducing the time-dimension within the models proposed in this dissertation might allow us to disclose further information about the relationship between the retail system and urban form. For example, we might study how the location factors might vary in time for the overall retail sector as well as for specific categories. Similarly, we might investigate which morphological characteristics are mostly associated with retail development and decay (as described, for example, by store vacancy).

Finally, another interesting research perspective should consider the potential of the Bayesian Network approaches. Firstly, we might further exploit the probabilistic content of the Bayesian Network classification procedures which implement both RFA and MFA, allowing for the evaluation of the uncertainty in the study of urban systems (Fusco 2018). Secondly, the implementation of probabilistic modelling approaches in urban morphology studies represents an interesting methodological solution which overcomes the architectural/morphological determinism criticisms traditionally highlighted in quantitative studies relating urban form and socioeconomic phenomena. Moreover, in the urban morphology community, we observe a strong separation between expert-based and a more

recent data-driven analytical approaches. The importance of bridging the two approaches and expertise is an important topic included in many discussions, especially at the recent ISUF 2017 and 2019 annual conferences to which we participated. The probabilistic framework of the Bayesian Networks approaches might allow for the transformation of the deterministic quantitative analysis and results in a more flexible tool for the support of more qualitative urban studies.

# **Conclusions et Perspectives**

Dans cette recherche, nous avons étudié la relation entre les schémas spatiaux de l'activité commerciale dans la ville et les propriétés de la forme de la ville physique dans la lignée des recherches initiées il y a vingt ans par la théorie de l'économie du mouvement de Hillier. Plusieurs développements théoriques et méthodologiques ont été proposés et mis en œuvre dans cette thèse, permettant de mieux comprendre la relation entre la forme urbaine et les schémas spatiaux de vente au détail dans la région métropolitaine de la Côte d'Azur.

D'un côté, basé sur la littérature de la géographie du commerce, un nouveau protocole analytique (RFA) a été spécifiquement conçu pour l'identification des types d'agglomération de magasins intégrant les caractéristiques fonctionnelles et spatiales des clusters de magasins. Par la suite, la distribution spatiale du commerce a été étudiée en tenant compte de trois principaux descripteurs : la présence / absence de magasin (booléen), la magnitude (nombre de magasins) et la typologie d'agglomération de RFA (descripteur catégorique). D'un autre côté, les théories de la forme urbaine sont à la base d'un deuxième ensemble de mesures proposées pour décrire la forme de la ville physique. Au-delà des propriétés de configuration traditionnelles du réseau de rues, un grand nombre de descripteurs de paysage à l'échelle de la rue ont été proposés, inspirés d'études récentes sur le design urbain. À partir des principes d'analyse typo-morphologique urbaine, un protocole spécifique a été développé pour l'identification du contexte morphologique urbain aux échelles méso et macro autour de chaque segment de rue (protocole MFA).

Les deux ensembles de mesures de la forme physique et commerciale de l'espace urbain ont finalement été étudiés par la mise en œuvre de différentes approches de régression basées sur des modèles de Poisson. Ces procédures de modélisation, qui représentent une innovation méthodologique compte tenu de la littérature existante, nous ont permis de décrire la force générale de la relation statistique entre la forme urbaine et l'activité commerciale. La mise en œuvre de procédures de régression pénalisées nous a permis d'identifier la combinaison spécifique de variables morphométriques les plus liées à chaque caractéristique de la configuration spatiale des détaillants ; ces procédures ont été mises en œuvre en tenant compte de différentes partitions spatiales morphologiques de la zone d'étude.

Les conclusions méthodologiques et thématiques de cette thèse pourraient être résumées comme suit :

- 1) Deux procédures innovantes, orientées-rue et basées sur la localisation, ont été proposées et mises en œuvre pour la description des typologies de tissu commercial et urbain. Les deux protocoles comprennent : 1) la mise en œuvre d'indicateurs morphologiques (ou morpho-fonctionnels en ce qui concerne RFA) au moyen d'algorithmes de géotraitement; 2) l'identification de motifs spatiaux statistiquement significatifs pour chaque indicateur; enfin 3) la combinaison de ces modèles morphologiques avec une analyse de classification basée sur l'IA. Ces procédures en trois étapes data-driven et bottom-up, fournissent la possibilité d'identifier les tissus urbains et commerciaux dans des grandes régions métropolitaines. Les défis méthodologiques de cette thèse d'identification des régions morphologiques urbaines et commerciales du point de vue de la rue ont été surmontés. La rue est devenue notre

référence métrique, géométrique et statistique pour la mise en œuvre au long de toute la thèse, nous permettant ainsi de faire le pont entre les différentes mesures de formes commerciales et urbaines, mais également entre les différentes approches de formes urbaines traditionnellement analysées de manière indépendante.

Deuxièmement, nous avons montré la supériorité du modèle de régression binomiale négative à inflation zéro sur les approches de modélisation linéaire traditionnelles adoptées dans la plupart des travaux examinés. Ce modèle spécifique résout le problème principal de l'hétéroscédasticité et de la surdispersion des données de comptage des magasins sans transformation artificielles de la distribution de la donnée originale. En outre, l'excès de zéros observés dans les distributions des commerces est spécifiquement étudiée avec les procédures ZI et ZA. Lors de l'examen des modèles complets, ZINB affiche les meilleurs résultats de prédictions.

Troisièmement, la mise en œuvre de régressions à la fois par étapes inversées et pénalisées nous a permis de tester empiriquement ces deux approches de sélection de variables. Les résultats de ces deux procédures (les mesures globales de la qualité de l'ajustement et les paramètres statistiques des variables sélectionnées) suggèrent la supériorité des procédures de régression pénalisée. Les procédures de régression pénalisées permettent, d'un côté, d'identifier le sous-ensemble le plus significatif de variables expliquant la relation entre la distribution de détail et la forme urbaine pour chaque sous-région morphologique; de l'autre côté, elles précisent également la forme générale des modèles entre les modèles ZINB et NB (avec maintien / exclusion de la partie zéro).

Les développements méthodologiques susmentionnés nous ont permis d'enquêter sur les questions thématiques proposées au début de ce travail (Chapitre 2) et de proposer nos réponses appuyées par des recherches empiriques sur le cas d'étude de la Côte d'Azur.

***(1) Les informations contextuelles de la forme urbaine peuvent-elles aider à mieux comprendre «l'effet de réseau» sur la distribution de détail? Si oui comment?***

Nous répondrons à cette question de manière positive. La prise en compte des régions morphologiques et des macro-régions contextuelles nous a permis de mettre en évidence la manière dont les propriétés de configuration du réseau de rues peuvent représenter à la fois des caractéristiques générales et des spécificités d'un tissu urbain ou d'une région morphologique donnée.

D'un côté, nous avons souligné à quel point la centralité d'intermédiation du réseau de rues représente la propriété la plus importante expliquant la distribution du commerce. La centralité d'intermédiation peut être considérée comme une *constante* liée à la distribution du commerce et qui a toujours été sélectionnée parmi les variables explicatives les plus importantes. Ces résultats sont alignés sur ce que l'on avait déjà observé dans des études précédentes sur la relation entre la configuration des réseaux de vente au détail et des réseaux de rues (Porta 2006, Peponis 2011, Fusco et al. 2015, etc.).

D'autre part, l'analyse séparée des régions morphologiques nous a permis d'apprécier davantage les caractéristiques spécifiques des propriétés du réseau en relation avec la

distribution de détail. En ce qui concerne les propriétés de centralité d'intermédiarité, alors que la méso-échelle piétonne est la plus importante dans les formes organiques non planifiées (qui sur la Côte d'Azur correspondent à la fois aux centres historiques et à la plupart des banlieues résidentielles), la centralité d'intermédiarité des véhicules à la méso-échelle est associée avec des tissus planifiés traditionnels et modernes. De plus, de nouvelles propriétés spécifiques du réseau de rues sont sélectionnées dans certains sous-espaces, tels que la centralité de rectitude dans les zones planifiées modernes et traditionnelles.

***(2) Au-delà de l'effet de réseau, existe-t-il d'autres caractéristiques de la forme urbaine qui inhibent la présence et le développement de la vente au détail? Si oui lesquels?***

Encore une fois, notre réponse est positive. Les propriétés de configuration du réseau ne sont pas les seules variables morphologiques qui pourraient aider à identifier le commerce de détail et, souvent, d'autres descripteurs de paysage à l'échelle de la rue pourraient présenter un potentiel plus élevé d'influence sur le commerce au détail. Cette information est encore plus vraie lorsque l'on considère des zones compactes et modernes où des combinaisons spécifiques de descripteurs de configuration et de composition se révèlent être différemment reliées à la distribution spatiale des activités commerciales

Certains descripteurs de paysage de rue jouent un rôle important dans l'étude de la distribution spatiale des activités commerciales. Plus particulièrement, le taux de couverture des bâtiments, la fragmentation des unités bâties (contiguïté) et l'effet de corridor de rue sont fréquemment associés de manière positive à la distribution de détail, tandis qu'au contraire, la pente et l'ouverture des rues sont négativement corrélées. Là encore, compte tenu des informations contextuelles, a permis de comprendre «l'effet de paysage de rue» ou, en d'autres termes, comment des variables de paysage de rue spécifiques pourraient prendre une signification différente dans des sous-espaces différents. Par exemple, nous avons observé comment la combinaison spécifique de tissus urbains et de descripteurs morphologiques aux échelles méso et macro pourrait expliquer le commerce de détail.

***(3) Peut-on associer des caractéristiques morphologiques urbaines spécifiques à la distribution spatiale des types d'agglomération de détail? Si oui, quelles caractéristiques jouent le rôle le plus important?***

En rapport à cette dernière question, une réponse positive est donnée là encore. Des variables morphométriques urbaines spécifiques sont associées à différentes typologies morpho fonctionnelles de détail. Plus spécifiquement, nous avons souligné comment des attracteurs spécifiques tels que le littoral, les grands magasins et les places pourraient influencer sur la spécialisation fonctionnelle des quartiers centraux des villes où l'on observe des tissus pervasifs pour la vente au détail. Les descripteurs de configuration et de composition deviennent plus importants pour la discrimination entre les tissus de détail semi-périphériques. Enfin, les caractéristiques morphologiques des rues et les mesures de



centralité de configuration du réseau de rues véhiculaires sont fondamentales pour expliquer la localisation de concentrations spécialisées dans les zones périphériques.

En conclusion, cette thèse a montré comment, au-delà des propriétés de configuration du réseau, une combinaison plus complexe de caractéristiques de design urbain dans différents contextes morphologiques pourrait être corrélée à la distribution de détail et, par conséquent, à l'attractivité socio-économique des espaces urbains.

## Limites et perspectives futures

Tout au long de ce travail de recherche, et plus particulièrement dans la partie modélisation, de nombreuses idées et questions de recherche ont constamment émergé, stimulant ainsi mon intérêt pour différents aspects spécifiques de cette thèse. Le défi caché pendant toute la durée de la recherche doctorale a été de rester concentré sur les principales questions de recherche sans perdre la piste avec des questions intéressantes mais divergentes pour l'économie globale de cette thèse. Pour cette raison, plusieurs idées thématiques et procédurales intéressantes ont été laissées inexploitées et / ou discutées de manière marginale.

Comme présenté dans l'introduction de ce travail, la modularité de cette thèse permet de proposer plusieurs développements sur des sous-parties individuelles, tout en conservant le cadre méthodologique global qui pourrait permettre de comparer les améliorations futures aux résultats présentés dans cette thèse.

Lors de l'examen des deux procédures développées dans les parties 2 et 3, respectivement le RFA et le MFA, de nombreuses améliorations méthodologiques pourraient être proposées. Malgré le solide arrière-plan théorique sur lequel chaque indicateur a été spécifiquement conçu et mis en œuvre, il reste un travail intéressant d'analyse de sensibilité qui pourrait prendre en compte différentes dimensions métriques, seuils et rayons choisis. De la même manière, d'autres approches de procédures de classification pourraient être étudiées et mises en œuvre. La validation des protocoles proposés via l'application à d'autres études de cas est également une perspective importante.

Ce travail est déjà en cours: depuis que la méthode MFA a été développée au cours de la première moitié des activités de recherche doctorale<sup>103</sup>, dans ces dernières années MFA a déjà été appliqué à d'autres études de cas telles que les régions métropolitaines d'Osaka (Japon) et de Marseille (France), Bruxelles (Belgique) et Izmir (Turquie)<sup>104</sup>. Ces implémentations nous ont permis de tester, développer et valider la procédure MFA auprès d'autres études de cas au-delà de la Côte d'Azur, avec différentes sources de données et différents contextes urbains.

---

<sup>103</sup> Première diffusion de la procédure MFA à la conférence ICCSA et résultat à la conférence ISUF en juillet 2017 (Araldi et Fusco 2017, Fusco et Araldi 2017).

<sup>104</sup> Dans le cadre, respectivement, du projet post-doc de Joan Perez *les villes côtières* avec le professeur Takashi Fuse (Université de Tokyo), du doctorat de Madeleine Guyot avec Isabelle Thomas (Université catholique de Louvain) et du projet de recherche de Irem Erin PhD avec le professeur Ebru Cubukcu (Dokuz Université Eylul). Les articles présentant la nouvelle mise en œuvre de l'AMF sont en cours d'édition.

Dans Araldi et al. (2018), nous avons spécifiquement discuté de la manière dont MFA pourrait être facilement mise en œuvre et adaptée à différentes études de cas. Les résultats de ces nouvelles applications ont démontré non seulement la validité du protocole MFA, mais également sa flexibilité et son adaptabilité à différentes sources de données.

Au-delà des aspects techniques de la discussion sur les différents indicateurs morpho-fonctionnels / morphométriques, les procédures statistiques ou l'approche classificatoire, l'automatisation de la procédure globale est un point particulièrement critique. Fournissant uniquement un outil unique et ergonomique (ou un ensemble de boîtes à outils similaire aux outils analytiques de configuration du réseau routier décrits au chapitre 3 ou dans MORPHO d'Olivera V. 2013, etc.), la procédure peut être facilement testée et appliquée à un plus grand nombre de cas plus intéressants, permettant des études comparatives diachroniques et synchroniques.

Contrairement à MFA, la procédure RFA a été développée dans la seconde moitié des activités de recherche doctorale<sup>105</sup> et, par conséquent, doivent encore être testées sur d'autres études de cas.

À partir d'octobre 2019, une collaboration d'un an entre le laboratoire ESPACE et la société privée Kinaxia nous permettra d'automatiser et d'optimiser les protocoles MFA et RFA, offrant ainsi de nouvelles possibilités et la recherche de nouvelles questions de recherche, tant sur la forme urbaine que sur la vente au détail. Dans ce cadre, toutes les grandes zones urbaines françaises serviront d'études de cas permettant de comparer la morphologie urbaine à l'échelle nationale. Outre le questionnement théorique qui émerge dans ce contexte de recherche, la quantité plus importante de données à étudier posera également plusieurs problèmes techniques et méthodologiques. Les résultats de cette analyse seront intégrés par la société privée Kinaxia dans leurs services CityScan<sup>106</sup>, démontrant ainsi l'intérêt des activités de recherche développées dans le cadre de cette thèse et qui ont également intéressé, outre notre partenaire de recherche actuel, une société privée. L'application des protocoles proposés au niveau national français permettra également de généraliser les résultats de la Côte d'Azur et de mettre en évidence les spécificités locales par rapport aux autres grandes zones urbaines françaises.

Des innovations méthodologiques pourraient également être examinées en relation avec la phase de modélisation de cette thèse (Partie 4). D'autres approches de régression pénalisée pourraient également être appliquées sur le même ensemble de variables, en vérifiant davantage les résultats soulignés dans la section 12.2. Au-delà de la régression pénalisée, la modélisation multiniveau ou les régressions pondérées géographiques en réseau constituent deux procédures méthodologiques intéressantes. Le premier pourrait permettre l'inclusion de l'organisation hiérarchique imbriquée de notre ensemble de données et donc permettre de prendre en compte la variabilité entre les classes qui (comme souligné précédemment) pourrait être détectée dans nos résultats. Ce dernier pourrait nous renseigner sur la variation

---

<sup>105</sup> Première diffusion de la procédure de RFA et des résultats lors de la conférence Spatial Analysis and Modeling, les 8 et 9 septembre. Tokyo (Araldi et Fusco 2018)

<sup>106</sup> <https://www.cityscan.fr/accueil>

spatiale du coefficient de corrélation statistique entre nos variables morphométriques et la distribution de détail.

Compte tenu de la taille importante de l'ensemble de données, les procédures de modélisation de l'intelligence artificielle pourraient également être envisagées à la fois pour les phases de modélisation et de sélection des variables. Certaines analyses exploratoires préliminaires ont déjà été mises en œuvre à l'aide de réseaux bayésiens et de méthodes d'apprentissage basées sur les forêts d'arbres décisionnels, toutefois, les résultats ne sont pas encore concluants et il conviendrait de poursuivre les analyses.

Au-delà des limites méthodologiques et de la perspective, cette recherche laisse ouverte un éventail beaucoup plus vaste de questions thématiques. Trois axes principaux pourraient être envisagés.

Premièrement, certains facteurs de localisation des commerces de détail ont été délibérément mis de côté dans cette étude afin de mettre l'accent sur la relation spécifique entre les propriétés de la ville physique et la distribution de détail. Si l'on élargit l'objectif de la recherche au problème de la localisation du commerce de détail, il conviendrait alors d'envisager d'autres registres de forme urbaine (Chapitre 2): l'inclusion d'autres variables telles que les facteurs socioéconomiques et d'accessibilité devient d'une importance primordiale (comme démontré dans les thèses de Scoppa et Sevtsuk). La répartition de la population, les caractéristiques socioprofessionnelles du quartier, l'indice de consommation (*Indices de Disparité des Dépenses de Consommation* CCI FRANCE, INSEE) et toutes les caractéristiques du registre social de la forme urbaine pourraient être facilement intégrées dans ce modèle. Comme suggéré par Okabe (2012), la même partition de Thiessen basé sur la rue à l'origine du protocole MFA pourrait être utilisé, une fois encore pour l'attribution des descripteurs sociaux à chaque élément de rue avec des facteurs proportionnels simples. L'intégration de ces nouvelles variables dans la même procédure de modélisation statistique développée dans cette thèse permettrait d'enquêter sur d'autres questions intéressantes, par exemple : quelle est l'importance marginale des caractéristiques de la forme urbaine sur la distribution de détail? quel est le rôle de chaque registre de formes urbaines? comment leur effet combiné pourrait-il expliquer la distribution de détail? Quel est le rôle de l'accessibilité?

Comme discuté dans la première partie de cette thèse, la distribution de détail est également considérée dans les études urbaines comme un indicateur indirect du dynamisme / vivacité / attractivité du lieu. Plusieurs autres sources de données décrivant différents aspects de l'attractivité sociale des rues pourraient être modélisées conjointement ou indépendamment de la distribution de détail, ce qui permettrait de mieux comprendre le fonctionnement social et économique des villes et leurs relations avec la morphologie urbaine. Par exemple, en utilisant les données d'Airbnb et de Tweeter (que nous avons recueillies au cours des deux dernières années) comme variables dépendantes de notre approche de modélisation, cela pourrait nous permettre d'étudier le lien entre les activités de loisirs et touristiques et l'environnement urbain. Cette question spécifique est un aspect intéressant pour le cas

d'étude de la Côte d'Azur, où le secteur du tourisme représente l'une des activités économiques les plus importantes.

Deuxièmement, la dimension temporelle est l'un des aspects absents les plus importants qui n'a pas été étudié dans cette thèse. L'exclusion de cette composante est, là encore, un choix méthodologique spécifique : cette thèse pourrait être envisagée dans un schéma de recherche plus large en deux étapes tenant compte d'objectifs à plus long terme. Avec ce travail, nous sommes spécifiquement concentrés sur le développement de protocoles robustes et reproductibles pour l'analyse transversale de la morphologie urbaine et de la distribution de détail, avec des données précises et facilement accessibles (au moins dans les pays français et dans la plupart des pays européens ou avec données open source). Ces protocoles constitueront, dans un deuxième temps, la pierre angulaire des recherches futures visant à introduire la dimension temporelle dans l'étude des tissus pour la vente au détail et des tissus urbains et de leurs processus génératifs. D'un côté, l'analyse comparative diachronique de la morphologie urbaine et de la distribution commerciale individuellement considérées pourrait révéler plusieurs aspects importants concernant le processus de transformation et les propriétés de résilience de ces deux composants. D'autre part, l'introduction de la dimension temporelle dans les modèles proposés dans cette thèse pourrait permettre de divulguer des informations supplémentaires sur la relation entre le système de vente au détail et la forme urbaine. Par exemple, nous pourrions étudier la manière dont les facteurs d'emplacement peuvent varier dans le temps pour l'ensemble du secteur commercial ainsi que pour des catégories spécifiques. De même, nous pourrions rechercher quelles caractéristiques morphologiques sont principalement associées au développement de la vente au détail et à la décroissance (comme décrit, par exemple, par la vacance d'un magasin).

Enfin, une autre perspective de recherche intéressante devrait considérer le potentiel des approches de réseau bayésien. Premièrement, nous pourrions exploiter davantage le contenu probabiliste des procédures de classification du réseau bayésien mises en œuvre, à la fois par RFA et par MFA, permettant d'évaluer l'incertitude dans l'étude des systèmes urbains (Fusco 2018). Deuxièmement, la mise en œuvre d'approches de modélisation probabiliste dans les études de morphologie urbaine représente une solution méthodologique intéressante qui résout les critiques du déterminisme architectural / morphologique traditionnellement mises en avant dans les études quantitatives reliant forme urbaine et phénomènes socioéconomiques. De plus, dans la communauté de la morphologie urbaine, nous observons une forte séparation entre les approches analytiques basées sur la connaissance-experte et les approches analytiques plus récentes basées sur l'analyse des données. L'importance de faire le lien entre les deux approches et expertise est un sujet important qui fait l'objet de la plupart des discussions, en particulier lors des récentes conférences annuelles ISUF 2017 et 2019 auxquelles nous avons participé. Le cadre probabiliste des approches des réseaux bayésiens pourrait permettre de transformer l'analyse quantitative déterministe et ses résultats en un outil plus flexible permettant de soutenir des études urbaines plus qualitatives.

# References

- Adebayo, A. A., Greenhalgh, P., and Muldoon-Smith, K. (2019). Investigating retail property market dynamics through spatial accessibility measures. *Journal of European Real Estate Research*.
- Agresti, A. (2002). *Categorical Data Analysis*. John Wiley and Sons. Inc., Publication.
- Agresti, A., and Kateri, M. (2011). *Categorical data analysis* (206-208), Springer Berlin Heidelberg.
- Akaike, H. (1974). A new look at the statistical model identification. In *Selected Papers of Hirotugu Akaike* (215-222). Springer, New York, NY.
- Alexander, C. (1964). A city is not a tree. *1965*, 124.
- Alexander, C. (1977). *A pattern language: towns, buildings, construction*. Oxford university press.
- Alexander, E. R. (1993). Density measures: A review and analysis. *Journal of Architectural and Planning Research*, 181-202.
- Allain, R. (2004). *Morphologie urbaine: géographie, aménagement et architecture de la ville*. Paris: Armand Collin.
- Alonso, W. (1964). Location and land use. Toward a general theory of land rent. *Location and land use. Toward a general theory of land rent*.
- Altman, D. G., and Andersen, P. K. (1989). Bootstrap investigation of the stability of a Cox regression model. *Statistics in medicine*, 8(7), 771-783.
- Anselin, L. (1988). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical analysis*, 20(1), 1-17.
- Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), 93-115.
- Anselin, L., and Bera, A. K. (1998). Introduction to spatial econometrics. *Handbook of applied economic statistics*, 237.
- Araldi A, Fusco G (2019). From the built environment along the street to the metropolitan region. Human scale approach in urban fabric analysis, *Environment and Planning B: Urban Analytics and City Science*, 46(7), 1243–1263.
- Araldi A., Fusco G. (2016). Urban Form from the Pedestrian Point of View: Spatial Patterns on a Street Network. *International Conference on Innovation in Urban and Regional Planning- INPUT, Turin, Italy Sept. 2016*, Proceedings INPUT 2016 pp. 32-37. <https://hal.archives-ouvertes.fr/hal-01417484/document>
- Araldi A., Fusco G. (2017). Decomposing and Recomposing Urban Fabric: The City from the Pedestrian Point of View, in O. Gervasi et al. (Eds.) *Computational Science and Its Applications - ICCSA 2017*, Trieste, Italy Proc. Part IV, LCNS Springer, vol. 10407, p. 365-376. DOI: [https://doi.org/10.1007/978-3-319-62401-3\\_27](https://doi.org/10.1007/978-3-319-62401-3_27)
- Araldi A., Fusco G. (2018). Describing retail distribution in urban space: local fractal indicators for point pattern analysis, *Revue Internationale de Géomatique*, 28(4), 505-532.
- Araldi A., Fusco G. (2019). Retail Fabric Assessment: describing retail patterns within urban space”, *Cities*, 85(2019) 51-62.
- Araldi A., Perez J., Fusco G., Fuse T. (2018). Multiple Fabric Assessment: Focus on Method Versatility and Flexibility, in: O. Gervasi et al. (eds) *Computational Science and Its Applications – ICCSA2018*. Proceedings, Part III, Lecture Notes in Computer Science, vol.10962, pp.251-267. Springer, Berlin, p. 251-267.

- Arbia, G. (1989). *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*. Kluwer, Dordrecht.
- Asami, Y., Kubat, A. S., Kitagawa, K., and Iida, S. I. (2003). Introducing the third dimension on Space Syntax: Application on the historical Istanbul. In *Proc. 6th International Space Syntax Symposium* (48-1). ITU Faculty of Architecture.
- Assunção, R. M., and Reis, E. A. (1999). A new proposal to adjust Moran's I for population density. *Statistics in medicine*, 18(16), 2147-2162.
- Aversa, J., Doherty, S., and Hernandez, T. (2018). Big Data Analytics: The New Boundaries of Retail Location Decision Making. *Papers in Applied Geography*, 4(4), 390-408.
- Axenov, K. E., Bondarchuk, E., and Brade, I. (1997). The new retail trade and services and their emerging location pattern in St. Petersburg. *GeoJournal*, 42(4), 403-417.
- Bacon, E. (1965). *Design of cities*, Princeton, Princeton university Press.
- Badariotti, D., Banos, A., and Moreno, D. (2007). Conception d'un automate cellulaire non stationnaire à base de graphe pour modéliser la structure spatiale urbaine: le modèle Remus. *Cybergeo: European Journal of Geography*.
- Badariotti, D., Banos, A., and Moreno, D. (2009). Morphologie urbaine et réseau. Etude des discontinuités et des ruptures induites par le réseau de circulation à l'aide du modèle Remus. *Revue Internationale de Géomatique*, Lavoisier, 2009, 19 (1), 45-66.
- Baddeley, A., and Turner, R. (2004). Spatstat: An R package for analyzing spatial point patterns" 1-42.
- Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial point patterns: methodology and applications with R*. Chapman and Hall/CRC.
- Balsas, C. J. (2017). Retail planning in Japan: Implications for city centres. *Journal of Urban Regeneration and Renewal*, 10(4), 357-368.
- Barthélemy, M. (2018). *Spatial networks* (2872-2882). Springer New York.
- Batty, M. (2005). Agents, cells, and cities: new representational models for simulating multiscale urban dynamics. *Environment and Planning A*, 37(8), 1373-1394.
- Batty, M. (2007). Model cities. *Town Planning Review*, 78(2), 125-151.
- Batty, M. (2008). The size, scale, and shape of cities. *science*, 319(5864), 769-771.
- Batty, M., and Longley, P. A. (1994). *Fractal cities: a geometry of form and function*. Academic press.
- Bauer, S. (2016) *Indikatorengestützte Analyse der Einflussnahme von Investoren auf planende Gemeinden im Ansiedlungsprozess von Vorhaben des grossflächigen Einzelhandels. Ein deutsch-französischer Vergleich*, PhD dissertation, Technische Universität Darmstadt, Germany.
- Bauer, S. and Fusco, G. (2016) Indikatorengestützte Analyse der Einflussnahme von Investoren auf die kommunale Planung am Beispiel der stadtplanerischen Standorteignung, Flächenmanagement und Bodenordnung (FUB), 2/2016, ISSN: 1616-0991, p. 79-85.
- Bavelas, A. (1948). A mathematical model for group structures. *Human organization*, 7(3), 16.
- Beaujeu-Garnier, J. (1977). *Atlas et géographie de Paris et la région d'Ile-de-France: Paris, ma grande ville* (1). Flammarion.
- Beavon, K. S. O. (1977). *Central place theory: a reinterpretation* (73). London: Longman.
- Beirão, J. N., and Chaszar, A. (2014). Convex-and Solid-Void Models for Analysis and Classification of Public Spaces.

- Beirão, J. N., and Koltsova, A. (2015). The Effects of Territorial Depth on the Liveliness of Streets. *Nexus Network Journal*, 17(1), 73-102.
- Benedikt, M. L. (1979). To take hold of space: isovists and isovist fields. *Environment and Planning B: Planning and design*, 6(1), 47-65.
- Benediktsson, J. A., Pesaresi, M., and Amason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9), 1940-1949.
- Bennison, D. J. and Davies, R. L. (1977). *The Movement of S1ropper.s Within the Central Area of Newcastle-upon-Tyne*, Department of Geography. Seminar Papers No.34, Newcastle: University of Newcastle-upon-Tyne.
- Berghauser Pont, M., and Haupt, P. (2002). Spacemate: FSI-GSI-OSR als instrument voor verdichting en verdunning, PERMETA Architecten, Amsterdam.
- Berghauser Pont, M., and Haupt, P. (2004). *Spacemate: the spatial logic of urban density*. Delft, The Netherlands: Delft University Press Science.
- Berghauser Pont, M., and Haupt, P. (2005). The Spacemate: density and the typomorphology of the urban fabric, *Nordisk Arkitekturforskning* (4), 55–68.
- Berghauser Pont, M., Stavroulaki, G., Gil, J. et al. (2017). Quantitative comparison of cities: Distribution of street and building types based on density and centrality measures. Proceedings of the 11th Space Syntax Symposium Instituto Superior Técnico, University of Lisbon.
- Berghauser-Pont, M. Y., and Haupt, P. (2010). *Spacematrix: space, density and urban form*. NAI Publishers.
- Bernabé, A., Calmet, I., Musy, M., Bocher, E., and Andrieu, H. (2013). Classification automatique des tissus urbains par la méthode des nuées dynamiques. *31e Rencontres AUGC* (Cachan, France).
- Bernabé, A., Musy, M., Andrieu, H., and Calmet, I. (2015). Radiative properties of the urban fabric derived from surface form analysis: A simplified solar balance model. *Solar Energy*, 122, 156-168.
- Berry B.J.L. (1967). *Geography of market centers and retail distribution*, New Jersey, Prentice Hall.
- Berry, B. J. (1959). Ribbon developments in the urban business pattern. *Annals of the Association of American Geographers*, 49(2), 145-155.
- Berry, B. J. (1963). Commercial structure and commercial bright: retail patterns and process in the City of Chicago. *Research paper*, (85).
- Besag, J., and Newell, J. (1991). The detection of clusters in rare diseases. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 154(1), 143-155.
- Besicovitch, A. S. (1964). On Fundamental Geometric Properties of Plane Line-Sets. *Journal of the London Mathematical Society*, 1(1), 441-448.
- Bi, J., and Bennett, K. P. (2003). Regression error characteristic curves. In *Proceedings of the 20th international conference on machine learning (ICML-03)* ( 43-50).
- Bielik, M., König, R., Schneider, S., and Varoudis, T. (2018). Measuring the impact of street network configuration on the accessibility to people and walking attractors. *Networks and Spatial Economics*, 18(3), 657-676.
- Birnbaum, D., Jacquez, G. M., Grimson, R., Waller, L. A., and Wartenberg, D. (1996). The analysis of disease clusters, part II: introduction to techniques. *Infection Control and Hospital Epidemiology*, 17(6), 385-397.

- Black, W. R. (1991). Highway accidents: a spatial and temporal analysis. *Transportation Research Record*, 1318, 75-82.
- Black, W. R. (1992). Network autocorrelation in transport network and flow systems. *Geographical Analysis*, 24(3), 207-222.
- Blanchard, R. (1960). *Le Comté de Nice: étude géographique*. A. Fayard.
- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126-139.
- Bolón-Canedo, V., Sánchez-Marono, N., Alonso-Betanzos, A., Benítez, J. M., and Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences*, 282, 111-135.
- Borie, A., and Denieul, F. (1984). Méthode d'analyse morphologique des tissus urbains traditionnels, UNESCO, <http://unesdoc.unesco.org/images/0006/000623/062310fb.pdf>
- Borruso, G. (2005). Network density estimation: analysis of point patterns over a network. In *International Conference on Computational Science and Its Applications* (126-132). Springer, Berlin, Heidelberg.
- Borruso, G. (2008). Network density estimation: a GIS approach for analysing point patterns in a network space. *Transactions in GIS*, 12(3), 377-402.
- Box, G. E., and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211-243.
- Boyce, R. R., and Clark, W. A. (1964). The concept of shape in geography. *Geographical Review*, 54(4), 561-572.
- Boyer, M. (2001). *L'invention de la Côte d'Azur: l'hiver dans le Midi*. Editions de l'Aube.
- Brisco, N. A., Doran, J. J., and Edgerton, J. H. (1927). *Principles of retailing*. Prentice-Hall, Incorporated.
- Browet, A. (1997). L'apport de l'approche géographique dans le conseil quotidien aux stratégies du commerce de détail. *Revue belge de géographie*, 121, 283-290.
- Brown, S. (1989). *Retail location theory: The legacy of Harold Hotelling*.
- Brown, S. (1991). Retail location: the post hierarchical challenge. *International Review of Retail, Distribution and Consumer Research*, 1(3), 367-381.
- Brown, S. (1992a). Tenant mix, tenant placement and shopper behaviour in a planned shopping centre. *Service Industries Journal*, 12(3), 384-403.
- Brown, S. (1992b). The wheel of retail gravitation?. *Environment and Planning A*, 24(10), 1409-1429.
- Brown, S. (1992c). *Retail location: a micro-scale perspective*. Avebury, Aldershot.
- Brown, S. (1993a). Micro-scale retail location: Cinderella or ugly sister?. *International Journal of Retail and Distribution Management*, 21(7).
- Brown, S. (1993b). Retail location theory: evolution and evaluation. *International Review of Retail, Distribution and Consumer Research*, 3(2), 185-229.
- Brown, S. (1994). Retail location at the micro-scale: inventory and prospect. *Service Industries Journal*, 14(4), 542-576.
- Brown, S. (1995). Sources and status of marketing theory. In *Marketing Theory and Practice* (23-39). Palgrave, London.
- Bruno O.M. Costa, L.F. (2001). A parallel approach for a plant taxonomy system. In *Proceedings of the VI Ibero-American Symposium on Pattern Recognition*, Florianópolis, S.C., Brazil.



- Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods and research*, 33(2), 261-304.
- Cachinho, H. (2011). Urban retail dynamics: from shopping spaces to consumer places. *Retail planning for the resilient city*, 147-168.
- Cachinho, H. (2014). Consumerscapes and the resilience assessment of urban retail systems. *Cities*, 36, 131-144.
- Cameron, A. C., and Trivedi, P. (1998). *The Econometric Analysis of Count Data*. Wiley.
- Cameron, A. C., and Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.
- Cameron, A. C., and Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge university press.
- Cameron, A. C., and Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of econometrics*, 77(2), 329-342.
- Campagna, M., Floris, R., Massa, P., Girsheva, A., and Ivanov, K. (2015). The role of social media geographic information (SMGI) in spatial planning. In *Planning support systems and smart cities* ( 41-60). Springer, Cham.
- Caniggia, G., and Maffei, G.L. (1979). *Composizione architettonica e tipologia edilizia*. Venezia: Marsilio.
- Cao, G., Hsu, W. W., and Todem, D. (2014). A score-type test for heterogeneity in zero-inflated models in a stratified population. *Statistics in medicine*, 33(12), 2103-2114.
- Carol, H. (1960). The hierarchy of central functions within the city. *Annals of the Association of American Geographers*, 50(4), 419-438.
- Carré, J. P., and Rouleau, R. (1974). Le commerce dans le tissu urbain: le dynamisme du commerce de détail alimentaire au Mans. *L'Espace géographique*, 311-317.
- Carruthers, W. I. (1962). Service centres in greater London. *Town Planning Review*, 33(1), 5.
- Carter, H., and Rowley, G. (1966). The morphology of the central business district of Cardiff. *Transactions of the Institute of British Geographers*, 119-134.
- Castex, J., Céleste, P., Burlen, K., and Panerai, P. (1980). *Lecture d'une ville: Versailles*. Ed. du Moniteur.
- Chaudhry, O., and Mackaness, W. A. (2008). Automatic identification of urban settlement boundaries for multiple representation databases. *Computers, Environment and Urban Systems*, 32(2), 95-109.
- Chiaradia, A., Cooper, C., and Webster, C. (2012). sDNA a software for spatial design network analysis. *Accessed August, 21*.
- Chiaradia, A., Hillier, B., Barnes, Y., and Schwander, C. (2009). Residential property value patterns in London: Proceedings of the 7th International Space Syntax Symposium. 016 Stockholm: KTH.
- Christaller, W. (1933). *Die zentralen Orte in Süddeutschland*. Jena: Gustav Fischer Verlag; translated by C. W. Baskin as *Central Places in Southern Germany*. Englewood Cliffs, NJ: Prentice-Hall.
- Clark, P. J. (1956). Grouping in spatial distributions. *Science*, 123(3192), 373-374.
- Clark, P. J., and Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35(4), 445-453.
- Clarkson, R. M., Clarke-Hill, C. M., and Robinson, T. (1996). UK supermarket location assessment. *International Journal of Retail and Distribution Management*, 24(6), 22-33.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2005). *Applied linear statistical models* (Vol. 5). Boston: McGraw-Hill Irwin.

- Clifton, K., Ewing, R., Knaap, G. J., and Song, Y. (2008). Quantitative analysis of urban form: a multidisciplinary review. *Journal of Urbanism*, 1(1), 17-45.
- Cliquet G., (2000). Large format retailers : a French tradition despite reactions", *Journal of Retailing and Consumer Services*, Vol.7, 183-195. Cliquet G., 2002a, "La localisation commerciale: méthodes, stratégies et perspectives", in Cliquet G. and Josselin J. M. (eds.), *Stratégies de localisation des entreprises commerciales et industrielles*, Bruxelles, De Boeck, p. 39-67.
- Cliquet G., (2002). *Le géomarketing: méthodes et stratégies du marketing spatial*, Paris, LavoisierHermès.
- Coca-Stefaniak, A. (2013). A strategic approach to town centres—a must or a maybe?. *The Geographer*.
- Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). Applied multiple correlation/regression analysis for the social sciences.
- Coluccia, E., and Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of environmental psychology*, 24(3), 329-340.
- Conrady, S., and Jouffe, L. (2013). Introduction to Bayesian Networks and BayesiaLab. *Bayesia SAS, USA*.
- Converse, P. D. (1949). New laws of retail gravitation. *Journal of marketing*, 14(3), 379-384.
- Conzen, M. P. (2010). The elusive common denominator in understanding urban form. *Urban Morphology*, 14(1), 55-58.
- Conzen, M. P., Gu, K., and Whitehand, J. W. R. (2012). Comparing traditional urban form in China and Europe: A fringe-belt approach. *Urban Geography*, 33(1), 22-45.
- Conzen, M. R. G. (1960). Alnwick, Northumberland: a study in town-plan analysis. *Transactions and Papers (Institute of British Geographers)*, (27), iii-122.
- Cooper, C. H., Harvey, I., Orford, S., and Chiaradia, A. J. (2018). Testing the ability of Multivariate Hybrid Spatial Network Analysis to predict the effect of a major urban redevelopment on pedestrian flows. *arXiv preprint arXiv:1803.10500*.
- Cox, D. R., and Snell, E. J. (1989). Analysis of binary data (Vol. 32). *Monographs on Statistics and Applied Probability*.
- Cragg, J. G., and Uhler, R. S. (1970). The demand for automobiles. *The Canadian Journal of Economics/Revue Canadienne d'Economique*, 3(3), 386-406.
- Craney, T. A., and Surles, J. G. (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, 14(3), 391-403.
- Cui, C., and Han, Z. (2015). Spatial patterns of retail stores using POIs data in Zhengzhou, China. In *2015 2nd IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM)* (88-92). IEEE.
- Cullen, G. (1961). *Townscape*, Architectural press, Londres.
- Cutini, V. (2001). Centrality and land use: three case studies on the configurational hypothesis. *Cybergeo: European Journal of Geography*.
- Cutini, V. (2003). Lines and Squares: towards a configurational approach to the analysis of open spaces. In *4th Space Syntax Symposium 2*(49-1). University College of London.
- Cutini, V. (2014). Spatial analysis of urban squares. *Tema. Journal of Land Use, Mobility and Environment*.
- Cutini, V., Farese, D., and Rabino, G. (2018). Milan: The Configuration of a Metropolis. In *Smart Planning: Sustainability and Mobility in the Age of Change* (343-357). Springer, Cham.
- Cutini, V., Paio, A., and Pezzica, C. (2017). A square for people a visually communicative, configuration informed and social-oriented redevelopment project for Largo da Graça, Lisbon. In *11th International Space Syntax Symposium, SSS 2017* (184-1). Instituto Superior Técnico.

- Cutini, V., Petri, M., and Santucci, A. (2004). From Axial Maps to Mark Point Parameter Analysis (Ma. PPA)—A GIS Implemented Method to Automate Configurational Analysis. In *International Conference on Computational Science and Its Applications* (1107-1116). Springer, Berlin, Heidelberg.
- Cutini, V., Petri, M., and Santucci, A. (2005). Dalla Space Syntax alla Ma. PPA: un contributo per l'evoluzione dell'analisi configurazionale. In A. Bruzzo, S. Occelli (eds.) *Le relazioni tra conoscenza e innovazione nello sviluppo dei territori*. Milano: Franco Angeli, 351-370.
- Dagher, N. (1998). *Online grocery shopping*, INSEAD, Fontainebleau, France.
- Dagorne, A. (1995). La Côte d'Azur, un mur de béton ? Alpes-Maritimes et Monaco, *Recherches régionales, Alpes-Maritimes et contrées limitrophes, Conseil Général des Alpes-Maritimes*, 163 – 213.
- Dale, M. R., Dixon, P., Fortin, M. J., Legendre, P., Myers, D. E., and Rosenberg, M. S. (2002). Conceptual and mathematical relationships among methods for spatial analysis. *Ecography*, 25(5), 558-577.
- Dart, M., and Lewis, R. (2017). *Retail's Seismic Shift: How to Shift Faster, Respond Better, and Win Customer Loyalty*. St. Martin's Press.
- Dauphiné, A. (1990). *Nice: une eurocité méditerranéenne*. Nice, Serre.
- Dauphiné, A., and Voiron-Canicio, C. (1988). *Variogrammes et structures spatiales* (No. 12). Gip Reclus.
- Davies R. L. (1974). Nucleated and ribbon component of the urban retail system in Britain, in Davies R. L. (eds.), *Marketing geography, with special reference to retailing*, Corbridge, Retailing and Planning Associates.
- Davies R. L. (1976). *Marketing geography with special reference to retailing*, New York, Up.
- Davies, L. D. (1984). *Retail and commercial planning, Croom Helm*, London and Sydney, St.
- Davies, R. B., and Guy, C. M. (1987). The statistical modeling of flow data when the Poisson assumption is violated. *Geographical Analysis*, 19(4), 300-314.
- Davies, R. L. (1969). Effects of consumer income differences on shopping movement behavior. *Tijdschrift voor economische en sociale geografie*, 60(2), 111-121.
- Davies, R. L. (1973). Evaluation of retail store attributes and sales performance. *European Journal of Marketing*, 7(2), 89-102.
- Davies, R. L., and Bennison, D. J. (1978). Retailing in the city centre: the characters of shopping streets. *Tijdschrift voor economische en sociale geografie*, 69(5), 270-285.
- Davies, R.L. (1972a). Structural models of retail distribution: analogies with settlement and land use theories, *Transactions, Institute of British Geographers*, 57(1): 59-82.
- Davies, R.L. (1972b). The retail pattern of the central area of Coventry, *The Retail Structure of Cities*, London: Institute of British Geographers, Occasional Publication (1), 1-32.
- Davies, W. K. (1967). Centrality and the central place hierarchy. *Urban studies*, 4(1), 61-79.
- Dawson, J. A. (1969). Some early theories of settlement location and size. *Journal of town planning institute*, 55(10), 444-448.
- Dawson, J. A. and Kirby, D. A. (1980). Urban Retail Provision and Consumer Behaviour: Some Examples From Western Society, in D. '1'. Herbert and R. J. Johnston (eds), *Geography and the Urban Environment: Progress in Research and Applications*, vol. 3, London: John Wiley, 87- 132.
- Dawson, J. A., and Sparks, L. (1986). New technology in UK retailing: issues and responses. *Journal of Marketing Management*, 2(1), 7-29.

- Dawson, J.A. (1980). *Retail activity and public policy*, in J.A. Dawson (ed.) *Retail Geography*, London: Croom Helm, 193-235.
- Dawson, J.A. (1983). *Shopping Centre Development*, London: Longman.
- de Palma, A., Ginsburgh, V. and Thisse, J-F. (1987). On the existence of location equilibria in the 3-firm Hotelling problem, *Journal of Industrial Economics*, 36 (2): 245-52.
- Decoupigny, F., and Fusco, G. (2009). Mobilités potentielles et émergence de structures réticulaires en région Provence–Alpes–Côte d'Azur. *LEspace géographique*, 38(3), 267-282.
- Delage, M., and Fleury, A. (2011). Centralités commerciales et logique hiérarchique en milieu urbain dense : Une démarche renouvelée d'analyse spatiale appliquée au cas de Paris. *Bulletin de la Société Géographique de Liège, Société Géographique de Liège*, 2011, 56, 7-21.
- Delgado De La Flor, Y. A., Burkman, C. E., Eldredge, T. K., and Gardiner, M. M. (2017). Patch and landscape-scale variables influence the taxonomic and functional composition of beetles in urban greenspaces. *Ecosphere*, 8(11), e02007.
- DeLisle, J. R. (2005). US shopping center classifications: challenges and opportunities. *Research Review*, 12(2), 96-101.
- Dennis, C., Marsland, D., and Cockett, T. (2002). Central place practice: shopping centre attractiveness measures, hinterland boundaries and the UK retail hierarchy. *Journal of Retailing and Consumer Services*, 9(4), 185-199.
- Deprez, S., and Vidal P (2016). "Drives des villes versus drives des champs: les nouveaux territoires du commerce alimentaire." In Desse R.-P., Lestrade (2016) *S, Mutations commerciales et devenir de l'espace marchand*, Rennes: Presses Universitaires de Rennes: 29-44.
- Deprez S. (2017). Les drives: une proximité renforcée ou réinventée? Quand la distribution alimentaire connectée réécrit les territoires d'approvisionnement des consommateurs." *Flux 3* (2017) : 102-117.
- Desse, R. P. (2013). La fin de l'urbanisme commercial à la française. In *Les Annales de la recherche urbaine* (Vol. 108, No. 1, 4-15). Centre de Recherche d'Urbanisme.
- Desse, R. P., and Lestrade, S. (2016). *Mutations de l'espace marchand*. Presses universitaires de Rennes.
- Desyllas, J. (2000). The relationship between urban street, configuration and office rent patterns in Berlin. *PhD dissertation, Bartlett School of Graduate Studies, University College London, UK*.
- Diamond, D.R. (1962). The central business district of Glasgow, in K. Norborg (ed.) *The IGU Symposium in Urban Geography*, Lund 1960, Lund: Gleerup, 525-34.
- Dibble, J. L. (2016). *Urban morphometrics: towards a quantitative science of urban form*. PhD dissertation, University of Strathclyde, UK.
- Dirichlet, G. L. (1850). Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *J. reine angew. Math.*, 40, 209.
- Dolega, L., Pavlis, M., and Singleton, A. (2016). Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services*, 28, 78-90.
- Dolega, L., Reynolds, J., Singleton, A., and Pavlis, M. (2019). Beyond retail: New ways of classifying UK shopping and consumption spaces. *Environment and Planning B: Urban Analytics and City Science*.
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., and Münkemüller, T. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27-46.

- Dubey, A., Naik, N., Parikh, D., Raskar, R., and Hidalgo, C. A. (2016). Deep learning the city: Quantifying urban perception at a global scale. In *European conference on computer vision* (196-212). Springer, Cham.
- Duda, R. O., Hart, P. E., and Stork, D. G. (2001). Bayesian decision theory. *Pattern classification*, 11(4), 99-102.
- Dudey, M. (1990). Competition by choice: The effect of consumer search on firm location decisions. *The American Economic Review*, 1092-1104.
- Dunteman, G. H. (1989). *Principal components analysis* (No. 69). Sage.
- Durandy, D. (1920). Mon pays. Visages et paysages de la Riviera, *Librairie nationale d'art et d'histoire*, édition Van Oest, Paris Bruxelles.
- Duranton, G. and Overman, H.G. (2005). Testing for Localisation Using Micro-Geographic Data. *Review of Economic Studies*, 72(4), 1077–1106.
- Durkheim, E. 1960. *De la division du travail social*, Paris, puf.
- Eaton, B. C., and Lipsey, R. G. (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. *The Review of Economic Studies*, 42(1), 27-49.
- Eaton, B. C., and Lipsey, R. G. (1979). Comparison shopping and the clustering of homogeneous firms. *Journal of Regional Science*, 19(4), 421-435.
- Efron, R. (1970). Effect of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics*, 8(4), 231-234.
- Egan, J. P. (1975). *Signal detection theory and ROC-analysis*. Academic press.
- Ellis, L. (1988). Les Russes sur la Côte d'Azur. Nice, Serre.
- Enström, R., and Netzell, O. (2008). Can space syntax help us in understanding the intraurban office rent pattern? Accessibility and rents in downtown Stockholm. *The Journal of Real Estate Finance and Economics*, 36(3), 289-305.
- Erin I., Fusco G., Cubukcu E., Araldi A., (2017) Quantitative Methods of Urban Morphology in Urban Design and Environmental Psychology, *International Seminar on Urban Form- ISUF 2017, Valencia, Spain* DOI.org/10.4995/ISUF2017.2017.5732
- Escourrou, G. (1980). *Climat et environnement*, Paris, Masson.
- Ester, M., Kriegel, H. P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd* 96(34), 226-231.
- Ewing, R., and Clemente, O. (2013). *Measuring urban design: Metrics for livable places*. Island Press.
- Ewing, R., and Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban design*, 14(1), 65-84.
- Ewing, R., Handy, S., Brownson, R. C., Clemente, O., and Winston, E. (2006). Identifying and measuring urban design qualities related to walkability. *Journal of Physical Activity and Health*, 3(s1), S223-S240.
- Fabre, C. (2008). *Géographie du commerce et aménagement urbain dans les Alpes-Maritimes: acteurs et stratégies pour le schéma de développement commercial*, PhD dissertation, Paris Est, France.
- Fernandes, J. R., and Chamusca, P. (2014). Urban policies, planning and retail resilience. *Cities*, 36, 170-177.
- Figueiredo, L. (2009). Continuity Lines-An investigation of urban form through street networks. *PhD dissertation, University College London, London, UK*.

- Figueiredo, L., and Amorim, L. (2005). Continuity lines in the axial system. In *The Fifth Space Syntax International Symposium. Delft University of Technology, Delft, The Netherlands*.
- Fischer, J. H., and Harrington Jr, J. E. (1996). Product variety and firm agglomeration. *The RAND Journal of Economics*, 281-309.
- Flahaut, B., Mouchart, M., San Martin, E., and Thomas, I. (2003). The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. *Accident Analysis and Prevention*, 35(6), 991-1004.
- Fleischmann, M. (2017). *Measuring Urban Form. A Systematisation of Attributes for Quantitative Urban Morphology*, MSc Dissertation, University of Strathclyde-UK.
- Fleury, A. (2004). La rue: un objet géographique? *Tracés. Revue de Sciences humaines*, (5), 33-44.
- Fleury, A., Mathian, H., and Saint-Julien, T. (2012). Définir les centralités commerciales au cœur d'une grande métropole: le cas de Paris intra-muros. *Cybergeo: European Journal of Geography*.
- Floris, R., and Campagna, M. (2014). Social media geographic information in tourism planning. *Tema. Journal of Land Use, Mobility and Environment*.
- Fong, P. (2003). What makes big dumb bells a mega shopping mall. In *Proceedings of the 4th International Space Syntax Symposium, London* (10-1).
- Fong, P. (2005). A study of store location patterns inside enclosed shopping environments. In *5th International Space Syntax Symposium* (103-116).
- Forriez, M., and Nottale, L. (2010). Lois d'échelle et transitions fractal-non fractal en géographie. *LEspace géographique*, 39(2), 97-112.
- Fotheringham, A. S. (1989). Scale-independent spatial analysis. *Accuracy of spatial databases*, 221-228.
- Frankhauser, P. (1994). La fractalite des structures urbains. Pairs: Collection Villes. *Anthropos*.
- Frankhauser, P. (1998). The fractal approach. A new tool for the spatial analysis of urban agglomerations. *Population: an english selection*, 205-240.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 35-41.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3), 215-239.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1), 1.
- Friske, W., and Choi, S. (2013). Another Look at Retail Gravitation Theory: History, Analysis, and Future Considerations. *ABD Journal*, 5(1).
- Fujita, M. (1988). Urban land use theory in J.J. Gabszewicz, J-F. Thisse, M. Fujita and U. Schweizer, *Location Theory*, London: Harwood, 73-149.
- Fujita, M., and Smith, T. E. (1990). Additive-interaction models of spatial agglomeration. *Journal of Regional Science*, 30(1), 51-74.
- Fujita, M., Krugman, P. R., Venables, A. J., and Fujita, M. (1999). The spatial economy: cities, regions and international trade (Vol. 213). *Cambridge, MA: MIT press. doi, 10, 1467-9787*.
- Fusco G, Cagliioni M, Araldi A (2017). Street network morphology and retail locations. Application to the city of Nice (France) *European Colloquium on Theoretical and Quantitative Geography- ECTQG 2015 Bari, Italy. in Plurimondi VIII, 17, 2015, 15-22* <http://plurimondi.poliba.it/index.php/Plurimondi/article/view/26>
- Fusco G., Araldi A. (2016). Urban Fabrics with LINCS. Empirical Evidence from the French Riviera. *Spatial Analysis and GEOmatics - SAGEO 2016, Nice, France*. Proceedings SAGEO Dec. 2017 p. 229-244 <https://hal.archives-ouvertes.fr/hal-01424734v1>

- Fusco, G. (2003). *Un modèle systémique d'indicateurs pour la durabilité de la mobilité urbaine: les cas de Nice et Gênes dans une comparaison internationale*, PhD dissertation, Université de Nice, France.
- Fusco, G. (2016). Beyond the built-up form/mobility relationship: Spatial affordance and lifestyles. *Computers, Environment and Urban Systems*, 60, 50-66.
- Fusco, G. (2018). *Ville, complexité, incertitude. Enjeux de connaissance pour le géographe et l'urbaniste*. Habilitation à diriger des recherches.
- Fusco, G., and Araldi, A. (2017). Significant patterns in urban form-Spatial analysis of morphological indicators. *Revue Internationale de Géomatique*, 27(4), 455-479.
- Fusco, G., and Araldi, A. (2017). The Nine Forms of the French Riviera: Classifying Urban Fabrics from the Pedestrian Perspective. In *24th ISUF International Conference. Book of Papers*( 1313-1325). Editorial Universitat Politècnica de València.
- Fusco, G., and Tirico, M. (2016). Configurational Approaches to Urban Form: Empirical Test on the City of Nice (France). INPUT 2016: the 9th International Conference on Innovation in Urban and Regional Planning, SITI - ISMB - DIST, Sep 2016, Turin, Italy.
- Gardner, William, Edward P. Mulvey, and Esther C. Shaw. "Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models." *Psychological bulletin* 118.3 (1995): 392.
- Garner, B. J. (1966). *The internal structure of retail nucleations* (No. 12). Department of Geography, Northwestern University.
- Garrison, W. L. (1950). *The Business Structure of the Customer Tributary Area of the Fountain Square Major Outlying Business Center of Evanston, Illinois: A Study of the Dispersion and Agglomeration of Satellite Business Structure in a Portion of the Chicago Metropolitan Area*. Northwestern University.
- Garrocho-Rangel, C., Álvarez-Lobato, J. A., and Chávez, T. (2013). Calculating intraurban agglomeration of economic units with planar and network K-functions: a comparative analysis. *urban Geography*, 34(2), 261-286.
- Gehl J. (2010) *Cities for People*. New York: Island Press.
- Gehl, J. (1971). *Life between buildings: using public space* Copenhagen.
- Gehlke, C. E., and Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29(185A), 169-170.
- Gehrig, T. (1998). Competing markets. *European Economic Review*, 42(2), 277-310.
- Genre-Grandpierre, C. (1999). La desserte spatiale des réseaux de transport routier: une approche fractale. *FLUX Cahiers scientifiques internationaux Réseaux et Territoires*, 15(38), 56-68.
- Genre-Grandpierre, C., and Josselin, D. (2008). Dépendance à l'automobile, tension dans les mobilités et stratégies des ménages. *Cybergeo: European Journal of Geography*.
- Getis, A. (1984). Interaction modeling using second-order analysis. *Environment and Planning A*, 16(2), 173-183.
- Getis, A., and Franklin, J. (1987). Second-order neighborhood analysis of mapped point patterns. *Ecology*, 68(3), 473-477.
- Getis, A., and Getis, J. M. (1968). Retail store spatial affinities. *Urban Studies*, 5(3), 317-332.
- Gehl, J. (1971). *Life Between Buildings: Using Public Space*, translated by Jo Koch.
- Ghosh, A. and McLafferty, S. (1987) *Location Strategies for Retail and Service Firms*, Lexington: Lexington Books

- Giannopoulou, M., Vavatsikos, A. P. and Lykostratis, K. (2016). A process for defining relations between urban integration and residential market prices. *Procedia-Social and Behavioral Sciences*, 223, 153-159.
- Gil, A., Stutz, J. and Chiaradia, Alain J. (2007) Confeego: tool set for spatial configuration studies. Presented at: 6th International Space Syntax Symposium, Istanbul, Turkey, 12-15 June 2007. Published in: Turner, A. ed. *New Developments in Space Syntax Software*. Istanbul: ITU Faculty of Architecture, 15-22.
- Gil, J. (2017). Street network analysis “edge effects”: Examining the sensitivity of centrality measures to boundary conditions. *Environment and Planning B: Urban Analytics and City Science*, 44(5), 819-836.
- Gil, J., Beirão, J. N., Montenegro, N., and Duarte, J. P. (2012). On the discovery of urban typologies: data mining the many dimensions of urban form. *Urban morphology*, 16(1), 27.
- Goldstucker, J.L., Bellenger, D.N., Stanley, T.J. and Otte, R.L. (1978) *New Developments in Retail Area Analysis and Site Selection*, Atlanta: Georgia State University, College of Business Administration.
- Golledge, R.G., Rushton, G. and Clark, W.A.V. (1966) Some spatial characteristics of Iowa's dispersed farm population and their implications for the grouping of central place functions, *Economic Geography*, 42: 261-72.
- Gonzalez-Benito, O., Munoz-Gallego, P. A., and Kopalle, P. K. (2005). Asymmetric competition in retail store formats: Evaluating inter-and intra-format spatial effects. *Journal of Retailing*, 81(1), 59-73.
- Goodchild, M. F. (1986). *Spatial autocorrelation* (Vol. 47). Geo Books.
- Goodchild, M. F., and Mark, D. M. (1987). The fractal nature of geographic phenomena. *Annals of the Association of American Geographers*, 77(2), 265-278.
- Gourdon, J. L. (2001). *La rue: essai sur l'économie de la forme urbaine*. Editions de l'Aube.
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica: Journal of the Econometric Society*, 701-720.
- Gouyet, G.F. (192) *Physique et Structures Fractales*, Masson, Paris.
- Graff, Ph. (2000) *L'exception Urbaine*. Nice: de la Renaissance au Consiglio d'Ornato. (Parenthèses, Marseille).
- Graff, Ph. (2013). *Une ville d'exception*. Nice, dans l'effervescence du 20<sup>e</sup> siècle. (Serre, Nice).
- Greenberg, M. (1995). *The poetics of cities: Designing neighborhoods that work*. The Ohio State University Press.
- Greenland, S., and Morgenstern, H. (2001). Confounding in health research. *Annual review of public health*, 22(1), 189-212.
- Griffith, D. A. (1987). Spatial autocorrelation. *Resource publications in Geography*.
- Grimaldi, D., Fernandez, V., and Carrasco, C. (2019). Heuristic for the localization of new shops based on business and social criteria. *Technological Forecasting and Social Change*, 142, 249-257.
- Grimmeau, J. P., Wayens, B., and Hanson, E. (2011). Une macro-géographie du commerce de détail en Belgique. *EchoGéo*, (15).
- Grondeau, A. (2018). La compétitivité des territoires de l'innovation confrontés aux crises et à la démondialisation: Le cas de Sophia-Antipolis. In *Annales de géographie* (No. 5, 463-491). Armand Colin.
- Gujarati, D. N. (2009). *Basic econometrics*. Tata McGraw-Hill Education.



- Gujarati, D. N. (2009). *Basic econometrics*. Tata McGraw-Hill Education.
- Gutiérrez, J., García-Palomares, J. C., Romanillos, G., and Salas-Olmedo, M. H. (2017). The eruption of Airbnb in tourist cities: Comparing spatial patterns of hotels and peer-to-peer accommodation in Barcelona. *Tourism Management*, 62, 278-291.
- Guy, C. M. (1976). *The location of shops in the Reading area*. Department of Geography, University of Reading, UK.
- Guy, C. M. (1977). A method of examining and evaluating the impact of major retail developments upon existing shops and their users. *Environment and Planning A*, 9(5), 491-504.
- Guy, C. M. (1983). The assessment of access to local shopping opportunities: a comparison of accessibility measures. *Environment and Planning B: Planning and Design*, 10(2), 219-237. <https://doi.org/10.1068/b100219>
- Guy, C. M. (1983). The assessment of access to local shopping opportunities: a comparison of accessibility measures. *Environment and Planning B: Planning and Design*, 10(2), 219-237.
- Guy, C. M. (1984). The urban pattern of retailing B: within UK. *Store location and store assessment research.*, 73-90.
- Guy, C. M. (1987). Recent advances in spatial interaction modelling: an application to the forecasting of shopping travel. *Environment and Planning A*, 19(2), 173-186.
- Guy, C. M. (1991). Spatial interaction modelling in retail planning practice: the need for robust statistical methods. *Environment and Planning B: Planning and Design*, 18(2), 191-203.
- Guy, C. M. (1998). Classifications of retail stores and shopping centres: some methodological issues. *GeoJournal*, 45(4), 255-264.
- Haggett, P., and Chorley, R. (1969). 1969: Network analysis in geography. London: Edward Arnold.
- Haggett, P., Cliff, A. D., and Frey, A. (1977). Locational analysis in human geography. *Tijdschrift Voor Economische En Sociale Geografie*, 68(6).
- Haig, R. M. (1927). Regional survey of New York and its environs. *New York: New York City Planning Commission*.
- Hair, J. F., Jr., R. E. Anderson, R. L. Tatham, and W. C. Black. (2010) *Multivariate data analysis*, 7th ed. Prentice-Hall, Inc., Upper Saddle River, New Jersey.
- Halbwachs, M. (1928). *La population et les tracés de voirie à Paris*, Paris, Alcan.
- Hallowell, G. D., and Baran, P. K. (2013). Suburban change: A time series approach to measuring form and spatial configuration. *The Journal of Space Syntax*, 4(1), 74-91.
- Hallowell, G. D., and Baran, P. K. (2013). Suburban change: A time series approach to measuring form and spatial configuration. *The Journal of Space Syntax*, 4(1), 74-91.
- Hamaina, R., Leduc, T., and Moreau, G. (2014). A new method to characterize density adapted to a coarse city model. In *Information Fusion and Geographic Information Systems (IF AND GIS 2013)* ( 249-263). Springer, Berlin, Heidelberg.
- Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of planners*, 25(2), 73-76.
- Hanson, J. (1994). 'Deconstructing' architects' houses. *Environment and Planning B: Planning and Design*, 21(6), 675-704.
- Hardin, J. and J. Hilbe. (2001) *Generalized Linear Models and Extensions*. College Station, TX: Stata Press.
- Harrell, F. E. (2001), *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*, Springer-Verlag, New York.

- Harvey C.W. "Measuring Streetscape Design for Livability Using Spatial Data and Methods" (2014). Graduate College Dissertations and Theses. 268. <https://scholarworks.uvm.edu/graddis/268>
- Harvey, C., and Aultman-Hall, L. (2015). Urban streetscape design and crash severity. *Transportation Research Record*, 2500(1), 1-8.
- Harvey, C., and Aultman-Hall, L. (2016). Measuring urban streetscapes for livability: A review of approaches. *The Professional Geographer*, 68(1), 149-158.
- Harvey, C., Aultman-Hall, L., Hurley, S. E., and Troy, A. (2015). Effects of skeletal streetscape design on perceived safety. *Landscape and Urban Planning*, 142, 18-28.
- Harvey, C., Aultman-Hall, L., Troy, A., and Hurley, S. E. (2017). Streetscape skeleton measurement and classification. *Environment and Planning B: Urban Analytics and City Science*, 44(4), 668-692.
- Hausdorff, F. (1919). Dimension und äußeres Maß. *Mathematische Annalen*, 79(1-2), 157-179.
- Heinritz, G. (2001). Les changements des horaires de fermeture des magasins et leurs conséquences sur les structures du commerce de détail. *Revue Géographique de l'Est*, 41(1-2).
- Heinze, G., Wallisch, C., and Dunkler, D. (2018). Variable selection—a review and recommendations for the practicing statistician. *Biometrical Journal*, 60(3), 431-449.
- Hermenault, L. (2017). *La ville en mouvements. Circulations, échanges commerciaux et matérialité de la ville : pour une articulation systémique des facteurs d'évolution du tissu urbain parisien entre le XVe et le XIXe siècle*. PhD dissertation in Archéologie et Préhistoire, Université Panthéon-Sorbonne - Paris I, France.
- Hermosilla, T., Palomar-Vázquez, J., Balaguer-Beser, Á., Balsa-Barreiro, J., and Ruiz, L. A. (2014). Using street based metrics to characterize urban typologies. *Computers, Environment and Urban Systems*, 44, 68-79.
- Hernandez, T., and Bennison, D. (2000). The art and science of retail location decisions. *International Journal of Retail and Distribution Management*, 28(8), 357-367.
- Heywood I. (1998), *Introduction to Geographical Information Systems*. New York: Addison Wesley Longman.
- Hidalgo, C. A., and Castañer, E. E. (2015). The amenity space and the evolution of neighborhoods. *arXiv preprint arXiv:1509.02868*.
- Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.
- Hillier B (1996). *Space is the machine*. Cambridge: Cambridge University Press.
- Hillier, B. (1989). *The architecture of the urban object*. *Ekistics*, 5-21.
- Hillier, B. (1999). Centrality as a process: accounting for attraction inequalities in deformed grids. *Urban Design International*, 4(3-4), 107-127.
- Hillier, B. and Iida, S., (2005), 'Network effects and psychological effects: a theory of urban movement,' In: van Nes, A. (ed.), *Proceedings of the Fifth International Space Syntax Symposium*, Delft: University of Technology, Vol. 1, p. 553-564.
- Hillier, B., and Hanson, J. (1989). *The social logic of space*. Cambridge university press.
- Hillier, B., and Vaughan, L. (2007). The city as one thing. *Progress in Planning*, 67(3), 205-230.
- Hillier, B., Hanson, J., and Graham, H. (1987). Ideas are in things: an application of the space syntax method to discovering house genotypes. *Environment and Planning B: planning and design*, 14(4), 363-385.
- Hillier, B., Penn, A., Hanson, J., Grajewski, T., and Xu, J. (1993). Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: planning and design*, 20(1), 29-66.

- Hofmeister, B. (2004). The study of urban form in Germany. *Urban Morphology*, 8(1), 3-12.
- Holgate, P. (1965). Tests of randomness based on distance methods. *Biometrika*, 52(3/4), 345-353.
- Holgate, P. (1965). Tests of randomness based on distance methods. *Biometrika*, 52(3/4), 345-353.
- Holt, D., Steel, D. G., Tranmer, M., and Wrigley, N. (1996). Aggregation and ecological effects in geographically based data. *Geographical analysis*, 28(3), 244-261.
- Horwood, E. M., and Boyce, R. R. (1959). *Studies of the central business district and urban freeway development*. Seattle: University of Washington Press.
- Hotelling H., (1929) Stability in competition, *The Economic Journal*, vol. 39, n° 153 mars, 41-57.
- Howard. E. B. and R. L. Davies. (1986), Contemporary Change in *Newcastle City Centre* and the Impact of the Metro System. CURDS. Discussion Paper No.77. Newcastle: University of Newcastle-upon-Tyne.
- Hubbard, J. (1978). Dielectric dispersion and dielectric friction in electrolyte solutions. II. *The Journal of Chemical Physics*, 68(4), 1649-1664.
- Huff, D. L. (1962). A note on the limitations of intraurban gravity models. *Land Economics*, 38(1), 64-66.
- Huff, D. L. (1963). A probabilistic analysis of shopping center trade areas. *Land economics*, 39(1), 81-90.
- Huff, D. L. (2000). Don't misuse the Huff model in GIS. *Business Geographies*, 8(8), 12.
- Hurd, R. (1903). Principles of City Land Values: New York Record and Guide. interdependencies. *Journal of Regional Science*, 36(4), 617-630.
- Jacobs, J., 1961. *The Death and Life of Great American Cities*. Vintage Books, A Division of Random House, New York, 143-222.
- Jenks, C., and Mayer, S. (1990). *The Consequences of Growing up in a Poor Neighborhood*, ed. by LL a. M. McGeary.
- Jensen, P. (2006). Network-based predictions of retail store commercial categories and optimal locations. *Physical Review E*, 74(3), 035101.
- Jensen-Butler, C. (1972). Gravity models as planning tools: a review of theoretical and operational problems. *Geografiska Annaler: Series B, Human Geography*, 54(1), 68-78.
- Jensen-Butler, C. (1972). Gravity models as planning tools: a review of theoretical and operational problems. *Geografiska Annaler: Series B, Human Geography*, 54(1), 68-78.
- Jiang, B., and Claramunt, C. (2004). A structural approach to the model generalization of an urban street network. *Geoinformatica*, 8(2), 157-171.
- Jiang, B., and Claramunt, C. (2004). Topological analysis of urban street networks. *Environment and Planning B: Planning and design*, 31(1), 151-162.
- Jiang, B., Yin, J., and Liu, Q. (2015). Zipf's law for all the natural cities around the world. *International Journal of Geographical Information Science*, 29(3), 498-522.
- Jiang, B., Yin, J., and Zhao, S. (2009). Characterizing the human mobility pattern in a large street network. *Physical Review E*, 80(2), 021136.
- Jiang, B., Zhao, S., and Yin, J. (2008). Self-organized natural roads for predicting traffic flow: a sensitivity study. *Journal of statistical mechanics: Theory and experiment*, 2008(07), P07008.
- Johnston, R.J. (1973) *Spatial Structures*, London: Methuen.
- Jones K. and Doucet M., (2000) Big-box retailing and the urban retail structure : the case of the Toronto area, *Journal of Retailing and Consumer Services*, Vol.7, 233-247.
- Jones, K., and Simmons, J. W. (1990). *Location, location, location*. Nelson Canada.

- Joosten, V., and Van Nes, A. (2005). How block types influences the natural movement economic process: Micro-spatial conditions on the dispersal of shops and Café in Berlin. In *5th International Space Syntax Symposium, Delft, The Netherlands* (Vol. 13).
- Joosten, V., and Van Nes, A. (2005). How block types influences the natural movement economic process: Micro-spatial conditions on the dispersal of shops and Café in Berlin. In *5th International Space Syntax Symposium, Delft, The Netherlands* (Vol. 13).
- Joseph, A. E., and Bantock, P. R. (1982). Measuring potential physical accessibility to general practitioners in rural areas: a method and case study. *Social science and medicine*, 16(1), 85-90.
- Judd, C. M., McClelland, G. H., and Ryan, C. S. (2011). *Data analysis: A model comparison approach*. Routledge.
- Kampschulte, A. (2002). Le commerce de détail en périphérie concurrence-t-il l'activité des centres-villes dans l'espace frontalier germano-suisse? Analyse comparée des cas de Bâle et Bad Säckingen. *Revue Géographique de l'Est*, 42(1-2).
- Kanevski M. and Pereira M.G., (2017). "Local fractality: The case of forest fires in Portugal". *Physica A: Statistical Mechanics and its Applications*, 479, 400-410.
- Karathodorou, N. (2012). The Link between Urban Form and Travel: A Focus on Public Transport Demand and Methodological Issues.
- Kärrholm, M., Barata-Salgueiro, T., Soumagne, J., Fernandes, J., and Chamusca, P. (2017). Time-space complexity and the opening hours of commerce: a study of four European cities. *Revista de Geografia e Ordenamento do Território*, (12), 153-177.
- Kaufman, J. S., and Cooper, R. S. (1999). Seeking causal explanations in social epidemiology. *American journal of epidemiology*, 150(2), 113-120.
- Kim, H. K., and Sohn, D. W. (2002). An analysis of the relationship between land use density of office buildings and urban street configuration: Case studies of two areas in Seoul by space syntax analysis. *Cities*, 19(6), 409-418.
- King R.D. (2014)" Computation of Local Fractal Dimension Values of the Human Cerebral Cortex" *Applied Mathematics* 5 (1733-1740).
- King, G. (1989). Variance specification in event count models: From restrictive assumptions to a generalized estimator. *American Journal of Political Science*, 762-784.
- Kivell, P. T., Shaw, G., and Dawson, J. A. (1980). Retail geography.
- Kleiber, C., and Zeileis, A. (2016). Visualizing count data regressions using rootograms. *The American Statistician*, 70(3), 296-303.
- Koch, D. (2007). *Structuring Fashion: Department stores as situating spatial practice*. PhD dissertation, Chalmers University of Technology, Sweden.
- Kohl J (1850) *Der Verkehr und die Ansiedelungen der Menschen in ihrer Abhsingigkeit von der Gestaltung der Erdoberflache*, 2nd edn. Leipzig
- Konishi, H. (2005). Concentration of competing retail stores. *Journal of Urban economics*, 58(3), 488-512.
- Kropf K (2017). Bridging configurational and urban tissue analysis. Proc.11th Space Syntax Symposium, Lisbon 165.1-13
- Kropf, K. (2009). Aspects of urban form. *Urban Morphology*, 13(2), 105.
- Kruger, M. (1977) *An Approach to Built Form Connectivity at an Urban Scale*, unpublished PhD dissertation, University of Cambridge, USA.

- Kulldorff, M., and Nagarwalla, N. (1995). Spatial disease clusters: detection and inference. *Statistics in medicine*, 14(8), 799-810.
- Kutner, M. H., Neter, J., Nachtsheim, C. J., and Li, W. (2004). Applied linear regression models (4th ed.). New York, NY: McGraw-Hill.
- Ladd, H. F., and Wheaton, W. (1991). Causes and consequences of the changing urban form: introduction. *Regional Science and Urban Economics*, 21(2), 157-162.
- Lagesse, C., Bordin, P., and Douady, S. (2015). A spatial multi-scale object to analyze road networks. *Network Science*, 3(1), 156-181.
- Lakshmanan, J. R., and Hansen, W. G. (1966). A retail market potential model. *Journal of the American Institute of planners*, 31(2), 134-143.
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1), 1-14.
- Lang G., Marcon E., Puech F. (2016) "Distance-based measures of spatial concentration: introducing a relative density function." <hal-01082178v3> (2016).
- Larkham, P. J., and Jones, A. N. (1991). *A glossary of urban form*. Urban Morphology Research Group, School of Geography, University of Birmingham.
- Larkham, P. J., and Morton, N. (2011). Drawing lines on maps: morphological regions and planning practices. *Urban Morphology*, 15(2), 133-151.
- Laulajainen. R.. (1987). *Spatial Strategies in Rerailing*. Dordrecht: Rcidcl.
- Law, S., Karimi, K., Penn, A., and Chiaradia, A. (2013). Measuring the influence of spatial configuration on the housing market in metropolitan London. In *Proceedings of the 2013 International Space Syntax Symposium, Seoul, Korea*(Vol. 31).
- Law, S., Karimi, K., Penn, A., and Chiaradia, A. (2013). Measuring the influence of spatial configuration on the housing market in metropolitan London. In *Proceedings of the 2013 International Space Syntax Symposium, Seoul, Korea* (Vol. 31).
- Law, S., Paige, B., and Russell, C. (2018). Take a look around: using street view and satellite images to estimate house prices. *arXiv preprint arXiv:1807.07155*.
- Law, S., Seresinhe, C. I., Shen, Y., and Gutierrez-Roig, M. (2018). Street-Frontage-Net: urban image classification using deep convolutional neural networks. *International Journal of Geographical Information Science*, 1-27.
- Lawson, A. B. (2010). Hotspot detection and clustering: ways and means. *Environmental and Ecological Statistics*, 17(2), 231-245.
- Lawson, H. (2001). Globalization, flows of culture and people, and new century frameworks for family-centered policies, practices, and development. *Family-centered policies and practices: International implications*, 338-376.
- Leamer, E. E. (1983). Let's take the con out of econometrics. *Modelling Economic Series*, 73, 31-43.
- Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, 46(1), 38.
- Lebrun, N. (2002). *Centralités urbaines et concentrations de commerces*. PhD dissertation, Université de Reims - Champagne Ardenne, France.
- Ledrut, R. 1968. *L'espace social de la ville*, Paris, Anthropos.
- Leduc, T., and Chauvat, G. (2015). Étude comparée de quatre indicateurs qualifiant le champ visuel d'un piéton en milieu urbain. *Revue Internationale de Géomatique*, 25(1), 75-98.

- Lee, C., and Moudon, A. V. (2006). The 3Ds+ R: Quantifying land use and urban form correlates of walking. *Transportation Research Part D: Transport and Environment*, 11(3), 204-215.
- Legendre P, Legendre L. (1998). Numerical Ecology, Vol.24, (Dev. in Env. Modelling).
- Legendre, P., and Legendre, L. F. (2012). *Numerical ecology*(Vol. 24). Elsevier.
- Lemarchand, N. (2008). *Géographie du commerce et de la consommation: les territoires du commerce distractif*, PhD dissertation, Université Paris-Sorbonne - Paris IV, France.
- Lemarchand, N. (2011). Nouvelles approches, nouveaux sujets en géographie du commerce. Une géographie renouvelée par l'analyse culturelle. *Géographie et cultures*, (77), 9-24.
- Lerman, Y., and Omer, I. (2016). Urban area types and spatial distribution of pedestrians: Lessons from Tel Aviv. *Computers, Environment and Urban Systems*, 55, 11-23.
- Lerman, Y., Rofè, Y., and Omer, I. (2014). Using space syntax to model pedestrian movement in urban transportation planning. *Geographical Analysis*, 46(4), 392-410.
- Levy, A. (1999). Les trois âges de l'urbanisme: Contribution au débat sur la troisième ville. *Esprit* (1940-), 46-61.
- Lévy, A. (2005). Formes urbaines et significations: revisiter la morphologie urbaine. *Espaces et sociétés*, (3), 25-48.
- Levy, M., and Weitz, B. A. (2006). *Retailing management*. New York: Irwin/McGraw-Hill.
- Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., and Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry and Urban Greening*, 14(3), 675-685.
- Li, Y., and Liu, L. (2012). Assessing the impact of retail location on store performance: A comparison of Wal-Mart and Kmart stores in Cincinnati. *Applied Geography*, 32(2), 591-600.
- Liebst, L. S. (2015). Phenomenology of the movement economy: A multilevel analysis. *The Journal of Space Syntax*, 6(1), 49-60.
- Liebst, L. S. (2019). Exploring the Sources of Collective Effervescence: A Multilevel Study. *Sociological Science*, 6, 27-42.
- Lin, G., Chen, X., and Liang, Y. (2018). The location of retail stores and street centrality in Guangzhou, China. *Applied geography*, 100, 12-20.
- Liu, J. L., Yu, Z. G., and Anh, V. (2015). Determination of multifractal dimensions of complex networks by means of the sandbox algorithm. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(2), 023103.
- Liu, L., Silva, E., Wu C., and Hui, W. (2017). A machine learning-based method for the large-scale evaluation of the urban environment. *Computers, Environment and Urban Systems*.
- Long, J. S. and J. Freese. (2003). *Regression Models for Categorical Dependent Variables Using Stata*, Revised Edition. College Station, TX: Stata Press.
- Long, N., and Kergomard, C. (2005). Classification morphologique du tissu urbain pour des applications climatologiques. *Revue Internationale de géomatique*. 15(4). DOI: 10.3166/riig.15.487-512.
- Longley, P., Batty, M., and Shepherd, J. (1991). The size, shape and dimension of urban settlements. *Transactions of the institute of British Geographers*, 75-94.
- Longley, P., Batty, M., Shepherd, J., and Sadler, G. (1992). Do green belts change the shape of urban areas? A preliminary analysis of the settlement geography of South East England. *Regional Studies*, 26(5), 437-452.
- Lopes, J. V., Paio, A., Beirão, J. N., Pinho, E. M., and Nunes, L. (2015). Multidimensional Analysis of Public Open Spaces-Urban Morphology, Parametric Modelling and Data Mining.

- Losch, A. (1954). *Economics of location*. Yale University Press.
- Louf, R., and Barthelemy, M. (2014). A typology of street patterns. *Journal of The Royal Society Interface*, 11(101), 20140924.
- Lovett, A., and Flowerdew, R. (1989). Analysis of count data using Poisson regression. *The Professional Geographer*, 41(2), 190-198. doi:10.1111/j.0033-0124.1989.00190.x
- Lu, Y., and Chen, X. (2007). On the false alarm of planar K-function when analyzing urban crime distributed along streets. *Social science research*, 36(2), 611-632. doi:10.1016/j.ssresearch.2006.05.003
- Lundevaller, E. H. (2009). The effect of travel cost on frequencies of shopping and recreational trips in Sweden. *Journal of transport geography*, 17(3), 208-215.
- Luo, W., and Wang, F., (2003). Spatial accessibility to primary care and physician shortage area designation: a case study in Illinois with GIS approaches. In: Skinner, R., Khan, O. (Eds.), *Geographic Information Systems and Health Applications*. Idea Group Publishing, Hershey, PA, 260–278.
- Lynch K (1960). *The image of the city*. Cambridge: MIT Press.
- Macintyre, S., Maciver, S., and Sooman, A. (1993). Area, class and health: should we be focusing on places or people?. *Journal of social policy*, 22(2), 213-234.
- Mackness, W. A., and Beard, K. M. (1993). Use of graph theory to support map generalization. *Cartography and Geographic Information Systems*, 20(4), 210-221.
- Mackness, W. A., and Chaudhry, O. Z. (2011). Automatic classification of retail spaces from a large scale topographic database. *Transactions in GIS*, 15(3), 291-307.
- Maddala, G. S. and Lahiri, K. (2009). *Introduction to econometrics*. 4th ed. Hoboken, NJ: Wiley
- Madry, P. (2016). *Ville et commerce à l'épreuve de la déterritorialisation* (PhD dissertation, Brest).
- Madry, P. (2018). Le nouveau rapport entre ville et commerce et l'impossible réforme des politiques d'urbanisme commercial. *Tous urbains*, (1), 42-47.
- Mahy, M., Alexander, C., Neis, H., Anninou, A., and King, I. (1987). *A new theory of urban design* (Vol. 6). Center for Environmental Struc.
- Makridakis, S., Wheelwright, S. C., and Hyndman, R. J. (2008). *Forecasting methods and applications*. John Wiley and sons.
- Mandelbrot, B. B. (1977). *Fractals: form, chance, and dimension* (Vol. 706). San Francisco: WH Freeman.
- Mandelbrot, B. B. (1983). *The fractal geometry of nature* (Vol. 173, p. 51). New York: WH freeman.
- Marcon, E., and Puech, F. (2009). Measures of the geographic concentration of industries: improving distance-based methods. *Journal of Economic Geography*, 10(5), 745-762.
- Marcon, E., and Puech, F. (2017). A typology of distance-based measures of spatial concentration. *Regional Science and Urban Economics*, 62, 56-67.
- Marcus, L. (2010). Spatial capital. *The Journal of Space Syntax*, 1(1), 30-40.
- Marshall, J.U. (1978) *On the structure of the Loschian landscape*, *Journal of Regional Science*, 18 (1): 121-5.
- Marshall, S. (2005). *Streets and Patterns*. London: Spon Press.
- Marshall, S. (2012). Science, pseudo-science and urban design. *Urban Design International*, 17(4), 257-271.

- Marshall, S., Gil, J., Kropf, K., Tomko, M., and Figueiredo, L. (2018). Street network studies: from networks to models and their representations. *Networks and Spatial Economics*, 18(3), 735-749.
- Martin, J., and Hall, D. B. (2016). R<sup>2</sup> measures for zero-inflated regression models for count data with excess zeros. *Journal of Statistical Computation and Simulation*, 86(18), 3777-3790.
- Martin, L., Leslie, M., and March, L. (1972). *Urban space and structures* (No. 1). Cambridge University Press.
- Massa, P., and Campagna, M. (2014). Social Media Geographic Information: recent findings and opportunities for smart spatial planning. *Tema. Journal of Land Use, Mobility and Environment*.
- Matheron, G. (1965). *Les variables régionalisées et leur estimation: une application de la théorie des fonctions aléatoires aux sciences de la nature*. Masson et CIE.
- Matthews, J. W., and Turnbull, G. K. (2007). Neighborhood street layout and property value: The interaction of accessibility and land use mix. *The journal of real estate finance and economics*, 35(2), 111-141.
- McCullagh, P. (2019). *Generalized linear models*. Routledge.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, ed. P. Zarembka, New York: Academic Press, 105:42.
- McKelvey, R. D., and Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology*, 4(1), 103-120.
- McMichael, A. J. (1999). Prisoners of the proximate: loosening the constraints on epidemiology in an age of change. *American journal of epidemiology*, 149(10), 887-897.
- Menard, S. (1995). *Applied logistic regression analysis*: Sage university series on quantitative applications in the social sciences. Thousand Oaks, CA: Sage.
- Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, 54(1), 17-24.
- Merenne-Schoumaker, B. (1998). Impacts des mutations du commerce sur la hiérarchie des polarités commerciales, l'exemple de la Belgique. *Hommes et Terres du Nord*, 1998(4), 217-224.
- Merenne-Schoumaker, B. (2003). *Géographie des services et des commerces*, Rennes, Presses Universitaires de Rennes, Coll. Didact Géographie.
- Merlin, P., and Choay, F. (1988). *de l'article/du chapitre Dictionnaire de l'urbanisme et de l'aménagement*. Presses Universitaires de France.
- Merry, P. R. (1955). *An inquiry into the nature and function of a string retail development: a case study of East Colfax Avenue, Denver, Colorado*, PhD dissertation, Northwestern University, USA.
- Metton, A. (2002). *Commerces et espaces fragiles*. Collection Commerce et société.
- Min, K. M., Moon, J. M., and Kim, Y. O. (2007, June). The effect of spatial configuration on land use and land value in Seoul. In *Proceedings of the Sixth Space Syntax Symposium* (080-1).
- Mittlböck, M., and Schemper, M. (1996). Explained variation for logistic regression. *Statistics in medicine*, 15(19), 1987-1997.
- Mohamed, A. A., and Van Nes, A. (2017). Spatial accessibility and commercial land use patterns planned versus unplanned areas in Cairo. *Proceedings of the 11th International Space Syntax Symposium* (SSS 2017).
- Mongin, O. (1995). *Vers la troisième ville*, Paris, Hachette, 1995.
- Monnot, M. (1973). Chiffres d'affaires et part du marché dans le commerce de détail. *Économie et Statistique*, 41(1), 29-34.



- Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17-23.
- Morganti, E., Dablanc, L., and Fortin, F. (2014). Final deliveries for online shopping: The deployment of pickup point networks in urban and suburban areas. *Research in Transportation Business and Management*, 11, 23-31.
- Morganti, E., Seidel, S., Blanquart, C., Dablanc, L., and Lenz, B. (2014). The impact of e-commerce on final deliveries: alternative parcel delivery services in France and Germany. *Transportation Research Procedia*, 4, 178-190.
- Morris, R. L., and Zisman, S. B. (1962). The pedestrian, downtown, and the planner. *Journal of the American Institute of Planners*, 28(3), 152-158.
- Moudon, A. V. (1997). Urban morphology as an emerging interdisciplinary field. *Urban morphology*, 1(1), 3-10.
- Moudon, A. V., Hubner, M., and Hubner, M. H. (2000). *Monitoring land supply with geographic information systems: theory, practice, and parcel-based approaches*. John Wiley and Sons, Wiley, New York, USA.
- Muldoon-Smith, K., Greenhalgh, P., Conroy-Dalton, R., Alvanides, S., King, H., and Sparkes, B. (2015). Urban transactions. *Proceedings of the 10th International Space Syntax Symposium*, 94:1.
- Mulíček, O., and Osman, R. (2018). Rhythm of urban retail landscapes: Shopping hours and the urban chronotopes. *Moravian Geographical Reports*, 26(1), 2-13.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of econometrics*, 33(3), 341-365.
- Muratori, S. (1959). Studi per una operante storia urbana di Venezia. *Palladio*, 1959, 1-113.
- Nagelkerke, N. J. (1992). Maximum likelihood estimation of functional relationships. In *Maximum Likelihood Estimation of Functional Relationships* (11-61). Springer, New York, USA.
- Naik, N., Philipoom, J., Raskar, R., and Hidalgo, C. (2014). Streetscore-predicting the perceived safety of one million streetscapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (779-785).
- Narvaez, L., Penn, A., and Griffiths, S. (2012). Space syntax economics: decoding accessibility using property value and housing price in Cardiff, Wales. In *Proceedings of the Eighth International Space Syntax Symposium, Santiago de Chile* (1-19).
- Neelon, B., Chang, H. H., Ling, Q., and Hastings, N. S. (2016). Spatiotemporal hurdle models for zero-inflated count data: exploring trends in emergency department visits. *Statistical methods in medical research*, 25(6), 2558-2576.
- Neelon, B., O'Malley, A. J., and Smith, V. A. (2016). Modeling zero-modified count and semicontinuous data in health services research Part 1: Background and overview. *Statistics in medicine*, 35(27).
- Nelder, J. A., and Baker, R. J. (1972). *Generalized linear models*. Wiley Online Library.
- Nelder, J. A., and Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, 135(3), 370-384.
- Neter, J., Wasserman, W., and Kutner, M. H. (1989). *Applied linear regression models*. Homewood, IL Richard D. Irwin
- Netto, V. M. (2016). 'What is space syntax not?' Reflections on space syntax as sociospatial theory. *Urban Design International*, 21(1), 25-40.
- Netto, V. M., Saboya, R., Vargas, J. C., Figueiredo, L., Freitas, C., and Pinheiro, M. (2012). The convergence of patterns in the city:(isolating) the effects of architectural morphology on

- movement and activity. In *Proceedings of the 8th International Space Syntax Symposium* (1-32). Santiago de Chile: Pontificia Universidad Católica de Chile.
- Netzell, O. (2013). The effect of accessibility on retail rents: testing integration value as a measure of geographic location. *Journal of Property Research*, 30(1), 1-23.
- Newman, P., and Kenworthy, J. (1999). Costs of automobile dependence: global survey of cities. *Transportation Research Record*, 1670(1), 17-26.
- Nezlek, J. B. (2011). *Multilevel modeling for social and personality psychology*. SAGE Publications Ltd.
- Nguyen, L., and Teller, J. (2017). Color in the urban environment: A user-oriented protocol for chromatic characterization and the development of a parametric typology. *Color Research and Application*, 42(1), 131-142.
- Nottale, L. (1993). *Fractal Space–Time and Microphysics: Towards a Theory of Scale Relativity* World Scientific, Singapore.
- O’hara, R. B., and Kotze, D. J. (2010). Do not log-transform count data. *Methods in ecology and Evolution*, 1(2), 118-122.
- Oakes, J. M. (2004). The (mis) estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Social science and medicine*, 58(10), 1929-1952.
- Okabe A., Funamoto S. (2000). An exploratory method for detecting multi-level clumps in the distribution of points—a computational tool, VCM (variable clumping method). *Journal of Geographical Systems*, 2(2):111-120.
- Okabe, A., and Satoh, T. (2006). Uniform network transformation for points pattern analysis on a non-uniform network. *Journal of Geographical Systems*, 8(1), 25-37.
- Okabe, A., and Sugihara, K. (2012). *Spatial analysis along networks: statistical and computational methods*. John Wiley and Sons.
- Okabe, A., and Yamada, I. (2001). The K-function method on a network and its computational implementation. *Geographical Analysis*, 33(3), 271-290.
- Okabe, A., Asami, Y., and Miki, F. (1985). Statistical analysis of the spatial association of convenience goods stores by use of a random clumping model. *Journal of Regional Science*, 25(1), 11-28.
- Okabe, A., Boots, B., Sugihara, K., and Chiu, S. N. (2009). *Spatial tessellations: concepts and applications of Voronoi diagrams* (Vol. 501). John Wiley and Sons.
- Okabe, A., Yomono, H., and Kitamura, M. (1995). Statistical analysis of the distribution of points on a network. *Geographical Analysis*, 27(2), 152-175.
- Okabe, A., Yomono, H., and Kitamura, M. (1995). Statistical analysis of the distribution of points on a network. *Geographical Analysis*, 27(2), 152-175.
- Oliveira, E. A., Furtado, V., Andrade, J. S., and Makse, H. A. (2018). A worldwide model for boundaries of urban settlements. *Royal Society open science*, 5(5), 180468.
- Oliveira, V. (2013). Morpho: a methodology for assessing urban form. *Urban Morphology*, 17(1), 21-33.
- Oliveira, V., and Medeiros, V. (2016). Morpho: Combining morphological measures. *Environment and Planning B: Planning and Design*, 43(5), 805-825.
- Oliveira, V., and Silva, M. (2013). Morpho. *Revista de Morfologia Urbana*, 1(1), 31-44.
- Omer, I., and Jiang, B. (2015). Can cognitive inferences be made from aggregate traffic flow data?. *Computers, Environment and Urban Systems*, 54, 219-229.
- Omer, I., and Kaplan, N. (2017). Using space syntax and agent-based approaches for modeling pedestrian volume at the urban scale. *Computers, Environment and Urban Systems*, 64, 57-67.

- Omer, I., and Kaplan, N. (2019). Structural properties of the angular and metric street network's centralities and their implications for movement flows. *Environment and Planning B: Urban Analytics and City Science*, 46(6), 1182-1200.
- Omer, I., and Zafir-Reuven, O. (2015). The development of street patterns in Israeli cities. *Journal of Urban and Regional Analysis*, 7(2), 113.
- Omer, I., Rofè, Y., and Lerman, Y. (2015). The impact of planning on pedestrian movement: contrasting pedestrian movement models in pre-modern and modern neighborhoods in Israel. *International Journal of Geographical Information Science*, 29(12), 2121-2142.
- Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data. *Environment and planning A*, 16(1), 17-31.
- Openshaw, S., and Albanides, S. (1999). Applying geocomputation to the analysis of spatial distributions. *Geographical Information Systems: Principles and Technical issues*, 1, 267-282.
- Openshaw, S., Charlton, M., Wymer, C., and Craft, A. (1987). A mark 1 geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information System*, 1(4), 335-358.
- Ortiz-Chao, C. G. (2008). Land use patterns and access in Mexico City. *ACSP-AESOP Fourth Joint Congress*, Chicago, Illinois.
- Panerai, P., Castex, J., and Depaule, J. C. (1997). *Formes urbaines : de l'îlot à la barre*. Editions Parentheses, Marseille.
- Panerai, P., Castex, J., Depaule, J. C., and Samuels, I. (2004). *Urban forms: the death and life of the urban block*. Routledge.
- Parkins, A. E. (1930). Profiles of the Retail Business Section of Nashville, Tenn., and Their Interpretation. *Annals of the Association of American Geographers*, 20(3), 164-175.
- Parr, J. B. (2007). Spatial definitions of the city: four perspectives. *Urban Studies*, 44(2), 381-392.
- Pascal, A. H., and McCall, J. J. (1980). Agglomeration economies, search costs, and industrial location. *Journal of Urban Economics*, 8(3), 383-388.
- Pavlis, M., Dolega, L., and Singleton, A. (2018). A Modified DBSCAN Clustering Method to Estimate Retail Center Extent. *Geographical Analysis*, 50(2), 141-161.
- Pavlis, M., Dolega, L., and Singleton, A. (2018). A Modified DBSCAN Clustering Method to Estimate Retail Center Extent. *Geographical Analysis*, 50(2), 141-161.
- Penn, A., and Turner, A. (2002). *Space syntax based agent simulation*. Springer-Verlag.
- Peponis, J., Allen, D., Haynie, D., Scoppa, M., and Zhang, Z. (2007). Measuring the configuration of street networks: the spatial profiles of 118 urban areas in the 12 most populated metropolitan regions in the US. *6th International Space Syntax Symposium Proceedings*, Istanbul Technical University-Turkey 002:1-002:16.
- Peponis, J., Bafna, S., and Zhang, Z. (2008). The connectivity of streets: reach and directional distance. *Environment and Planning B: Planning and Design*, 35(5), 881-901.
- Perez J., Fusco G., Araldi A. and Fuse T. (2019). Identifying building Typologies and their spatial patterns in the metropolitan areas of Marseille and Osaka, *Asia-Pacific Journal of Regional Science*, 1-25.
- Perez, J., Fusco, G., Araldi, A., and Fuse, T. (2018). Building typologies for urban fabric classification: Osaka and Marseille case studies. In *International Conference on Spatial Analysis and Modeling, 8th–9th September* (p. 13).
- Péron, R. (2004). *Les boîtes. Les grandes surfaces dans la ville*, L'Atalante, Paris.

- Perry, G. L. (2004). SpPack: spatial point pattern analysis in Excel using Visual Basic for Applications (VBA). *Environmental Modelling and Software*, 19(6), 559-569.
- Perry, G. L., Miller, B. P., and Enright, N. J. (2006). A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant ecology*, 187(1), 59-82.
- Pierre Frankhauser, Hélène Houot, Cécile Tannier, Gilles Vuidel. Vers des déplacements périurbains plus durables : proposition de modèles fractals opérationnels d'urbanisation. *Colloque du PREDIT Mobilité des personnes : Comportements et Services, les comportements de mobilité*, Apr 2007, Lyon.
- Pinon, P. (1991). *Lire et composer l'espace public*, STU, Paris.
- Pinon, P. (1994). *Composition urbaine I, II*, Paris, dau-stu.
- Pinon, P., and Henry, D. D. (1991). *Lire et composer l'espace public*. Les Éditions du STU.
- Pipkin, J. S. (1993). A Partitioning Model of Urban Retail Structure. *Geographical Analysis*, 25(3), 179-198.
- Pollak, M. D. (1988). From Castrum to Capital: Autograph Plans and Planning Studies of Turin, 1615-1673. *Journal of the Society of Architectural Historians*, 47(3), 263-280.
- Pont, M. B. (2007). *Urbanism Laboratory for Cities and Regions: Progress of Research Issues in Urbanism 2007*. IOS Press.
- Pont, M. B., and Haupt, P. (2007). The relation between urban form and density. *Urban Morphology*, 11(1), 62.
- Pont, M. B., and Marcus, L. (2014). Innovations in measuring density: From area and location density to accessible and perceived density. *NA*, 26(2).
- Pont, M. B., and Marcus, L. (2015). What can typology explain that configuration cannot. In *SSS10 Proceedings of the 10th International Space Syntax Symposium*, 41:1-43:16.
- Porta, S., Crucitti, P., and Latora, V. (2006). The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications*, 369(2), 853-866.
- Porta, S., Crucitti, P., and Latora, V. (2006). The network analysis of urban streets: a primal approach. *Environment and Planning B: planning and design*, 33(5), 705-725.
- Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., and Latora, L. (2012). Street centrality and the location of economic activities in Barcelona. *Urban Studies*, 49(7), 1471-1488.
- Porta, S., Strano, E., Iacoviello, V., Messori, R., Latora, V., Cardillo, A., and Scellato, S. (2009). Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and design*, 36(3), 450-465.
- Portzamparc, C. (1995). « L'âge III », Projet urbain, n° 3, *La ville hors la ville*, Paris, Ministère de l'Équipement, 1995, p. 4-6.
- Potter, R. B. (1979). Perception of Urban Retailing Facilities: An Analysis of Consumer Information Fields, *Geografiska Annaler. Series B, Human Geography*, Vol.61, No.1, 19-29.
- Potter, R. B. (1980). Spatial and Structural Variations in the Quality Characteristics of Intra-Urban Retailing Centres, *Transactions of the Institute of British Geographers*, 5(2), 207-228.
- Potter, R. B. (1981). The multivariate functional structure of the urban retailing system: A British case study. *Transactions of the Institute of British Geographers*, 188-213.
- Potter, R. B., and Beaumont, C. D. (1984). Parsimonious multivariate classifications of urban retailing systems. *The Professional Geographer*, 36(3), 353-358.
- Potter, R.B. (1981). The multivariate functional structure of the urban retailing system: a British case study, *Transactions, Institute of British Geographers*, 6: 188-213.

- Potter, R.B. (1982). *The Urban Retailing System: Location, Cognition and Behaviour*, Aldershot: Gower.
- Proudfoot, M. J. (1936). The Major Outlying Business Centers of Chicago, Phd dissertation, University of Chicago.
- Proudfoot, M. J. (1937). City retail structure. *Economic Geography*, 13(4), 425-428.
- Puissant, A., Skupinski, G., Lachiche, N., Braud, A., and Perret, J. (2010). Classification des tissus urbains à partir de bases de données vectorielles-application à Strasbourg. In *Conférence Internationale de Géomatique et Analyse Spatiale SAGEO'10* ( 198-211).
- Punakivi, Mikko, Hannu Yrjölä, and Jan Holmström (2001). Solving the last mile issue: reception box or delivery box? *International Journal of Physical Distribution and Logistics Management* 31.6 (2001): 427-439.
- Purciel, M., Neckerman, K. M., Lovasi, G. S., Quinn, J. W., Weiss, C., Bader, M. D., and Rundle, A. (2009). Creating and validating GIS measures of urban design for health research. *Journal of environmental psychology*, 29(4), 457-466.
- Rao, P. V., Rao, P. S. C., Davidson, J. M., and Hammond, L. C. (1979). Use of Goodness-of-fit Tests for Characterizing the Spatial Variability of Soil Properties 1. *Soil Science Society of America Journal*, 43(2), 274-278.
- Ratti, C. (2004). Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design*, 31(4), 487-499.
- Raudenbush, S. W., and Willms, J. (1995). The estimation of school effects. *Journal of educational and behavioral statistics*, 20(4), 307-335.
- Raudenbush, S., and Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of education*, 1-17.
- Raynaud, D. (1999). Forme urbaine: une notion exemplaire du point de vue de l'épistémologie des sciences sociales. In Ph. Boudon, (éd.) *Langages singuliers et partagés de l'urbain* Proceeding papers LOUEST, CNRS UMR 7544-L'Harmattan, 93-120, 1999.
- Reilly W. J. (1931). *The law of retail gravitation*, New York, Knickerbocker Press, 183 p.
- Remali, A. M., Porta, S., and Romice, O. (2014). Correlating street quality, street life and street centrality in Tripoli, Libya. *The past, present and future of high streets*, 104-129.
- Renard-Grandmontagne, C. (2016). Commerce et espaces urbains. Bulletin de la Société Géographique de Liège, Société Géographique de Liège, 2016, 66, 19-24.
- Reynolds, J., and Schiller, R. (1992). A new classification of shopping centres in Great Britain using multiple branch numbers. *Journal of Property Research*, 9(2), 122-160.
- Reynolds, J., and Wood, S. (2010). Location decision making in retail firms: evolution and challenge. *International Journal of Retail and Distribution Management*, 38(11/12), 828-845.
- Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of applied probability*, 13(2), 255-266.
- Ripley, B.D., (1981). *Spatial Statistics*, New York: John Wiley.
- Robinson W S (1950). Ecological correlations and the behavior of individuals. *American Sociological Review* 15, 351-57.
- Rodríguez, D. A., Brisson, E. M., and Estupiñán, N. (2009). The relationship between segment-level built environment attributes and pedestrian activity around Bogota's BRT stations. *Transportation research part D: transport and environment*, 14(7), 470-478.
- Rogers, A. (1965). A stochastic analysis of the spatial clustering of retail establishments. *Journal of the American Statistical Association*, 60(312), 1094-1103.

- Rogers, A. (1969). Quadrat analysis of urban dispersion: 2. case studies of urban retail systems, *Environment and Planning*, 1 (2): 155-71.
- Rolph, I.K. (1935). *The Locational Structure of Retail Trade* (Washington: U. S. Bureau of Foreign and Domestic Commerce, Domestic Commerce Series, 1929).
- Roncayolo, M. (1996). *Les grammaires d'une ville : essai sur la genèse des structures urbaines à Marseille*, Paris, ehess.
- Roncayolo, M. (2002). *Lectures de ville, Formes et temps*, Paris, Parenthèses.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1), 34-55.
- Rosenberg D. (2001). *Cloning Silicon Valley*, New York, Prentice Hall.
- Rosvall, M., and Bergstrom, C. T. (2007). An information-theoretic framework for resolving community structure in complex networks. *Proceedings of the National Academy of Sciences*, 104(18), 7327-7331.
- Roth Tran, B. (2016). Blame It On the Rain: Weather Shocks and Retail Sales. Available at SSRN 3381302.
- Roux, A. V. D., Merkin, S. S., Arnett, D., Chambless, L., Massing, M., Nieto, F. J., and Watson, R. L. (2001). Neighborhood of residence and incidence of coronary heart disease. *New England Journal of Medicine*, 345(2), 99-106.
- Rowley, A. (1994). Definitions of urban design: The nature and concerns of urban design. *Planning Practice and Research*, 9(3), 179-197.
- Rowlingson, B. S., and Diggle, P. J. (1993). Splancs: spatial point pattern analysis code in S-Plus. *Computers and Geosciences*, 19(5), 627-655.
- Roy, J. R., and Thill, J. C. (2004). Spatial interaction modelling. *Papers in Regional Science*, 83(1), 339-361.
- Rozenfeld, H. D., Rybski, D., Andrade, J. S., Batty, M., Stanley, H. E., and Makse, H. A. (2008). Laws of population growth. *Proceedings of the National Academy of Sciences*, 105(48), 18702-18707.
- Rui, Y., and Ban, Y. (2014). Exploring the relationship between street centrality and land use in Stockholm. *International Journal of Geographical Information Science*, 28(7), 1425-1438.
- Rui, Y., Yang, Z., Qian, T., Khalid, S., Xia, N., and Wang, J. (2016). Network-constrained and category-based point pattern analysis for Suguo retail stores in Nanjing, China. *International Journal of Geographical Information Science*, 30(2), 186-199.
- Ruilan, D., and Xinqi, Z. (2004). The relation of the city road structure and the land price based on the space syntax. *Science of surveying and mapping*, 29(5), 76-79.
- Rushton, G. (1969). Analysis of spatial behaviour by revealed space preference, *Annals of the Association of American Geographers*, 59: 391-400.
- Sadahiro, Y. (2000). A PDF-based analysis of the spatial structure of retailing. *GeoJournal*, 52(3), 237-252.
- Salesses, P., Schechtner, K., and Hidalgo, C. A. (2013). The collaborative image of the city: mapping the inequality of urban perception. *PloS one*, 8(7), e68400.
- Salgueiro, T. B., and Cachinho, H. A. (2011). *Retail planning for the resilient city: consumption and urban regeneration*. Centro de Estudos Geográficos, Universidade de Lisboa.
- Salgueiro, T. B., and Erkip, F. (2014). Retail planning and urban resilience - an introduction to the special issue. *Cities*, 36, 107-111.
- Salingaros, N.A. (2003). "The Sensory Value of Ornament", *Communication Cognition*, Volume 36, No. 3-4, pages 331-351.

- Salovey, P., and Mayer, J. D. (1990). Emotional intelligence. *Imagination, cognition and personality*, 9(3), 185-211.
- Samuels, I., Panerai, P., Castec, J. and Depaule, J. (2004). *Urban forms. The death and life of the urban block*. London: Architectural Press.
- Saraiva, M. M. (2013). *The morphological sense of commerce: Symbioses between commercial activity and the form and structure of Portuguese medium-sized cities*. PhD dissertation, Universidade do Porto.
- Saraiva, M., and Pinho, P. (2017). Spatial modelling of commercial spaces in medium-sized cities. *GeoJournal*, 82(3), 433-454.
- Saraiva, M., Marques T.S., and Pinho, P. (2019). Vacant Shops in a Crisis Period—A Morphological Analysis in Portuguese Medium-Sized Cities. *Planning Practice and Research* 34(3): 255-287.
- Sarma, A. K. (2006). *The social logic of shopping: a syntactic approach to the analysis of spatial and positional trends of community centre markets in New Delhi*, PhD dissertation, University College London, UK.
- Scheer, B. C. (2001). The anatomy of sprawl. *Places*, 14(2).
- Scheer, B. C. (2015). The epistemology of urban morphology. *Urban Morphology*, 19(2), 117-134.
- Schonlau, M. (2002). The clustergram: a graph for visualizing hierarchical and nonhierarchical cluster analyses. *The Stata Journal*, 2(4), 391-402.
- Scoppa, M. D. (2013). *Towards a theory of distributed attraction: the effects of street network configuration upon the distribution of retail in the City of Buenos Aires*. PhD dissertation, Georgia Institute of Technology, USA.
- Scoppa, M. D., and Peponis, J. (2015). Distributed attraction: the effects of street network connectivity upon the distribution of retail frontage in the City of Buenos Aires. *Environment and Planning B: Planning and Design*, 42(2), 354-378.
- Scott, P. (1970). *Geography and retailing* (Hutchinson, London).
- SDEC (2005). Schéma Départemental d'Équipement Commercial des Alpes-Maritimes <http://www.alpes-maritimes.gouv.fr/content/download/14684/148842/file/SDEC06.pdf>.
- Seamon, D. (2015). The life of the place: A phenomenological commentary on Bill Hillier's theory of space syntax. *NA*, 7(1).
- Seresinhe, C. I., Moat, H. S., and Preis, T. (2018). Quantifying scenic areas using crowdsourced data. *Environment and Planning B: Urban Analytics and City Science*, 45(3), 567-582.
- Seresinhe, C. I., Preis, T., and Moat, H. S. (2017). Using deep learning to quantify the beauty of outdoor places. *Royal Society open science*, 4(7), 170170.
- Serra, C. P., and Karas, E. W. (1997). *Fractais gerados por sistemas dinâmicos complexos*. Universitaria Champagnat.
- Serra, M., and Hillier, B. (2019). Angular and Metric Distance in Road Network Analysis: A nationwide correlation study. *Computers, Environment and Urban Systems*, 74, 194-207.
- Serra, M., Gil, J., and Pinho, P. (2012). Unsupervised classification of evolving metropolitan street patterns. In *Proceedings of the EAAE/ISUF International Conference "New Urban Configurations"*.
- Serra, M., Gil, J., and Pinho, P. (2017). Towards an understanding of morphogenesis in metropolitan street-networks. *Environment and Planning B: Urban Analytics and City Science*, 44(2), 272-293.

- Serra, M., Psarra, S., and O'Brien, J. (2018). Social and physical characterization of urban contexts: Techniques and methods for quantification, classification and purposive sampling. *Urban Planning*, 3(1), 58-75.
- Sevenet, M. (2013). *La forme en 3D dans l'analyse spatiale des territoires urbanisés et de la durabilité urbaine*. PhD dissertation, Université Nice Sophia-Antipolis, France.
- Sevtsuk, A. (2010). *Path and place: a study of urban geometry and retail activity in Cambridge and Somerville*, PhD dissertation, Massachusetts Institute of Technology, USA.
- Sevtsuk, A. (2014). Location and agglomeration: The distribution of retail and food businesses in dense urban environments. *Journal of Planning Education and Research*, 34(4), 374-393.
- Sevtsuk, A., and Mekonnen, M. (2012). Urban network analysis. *Revue internationale de géomatique—n*, 287, 305.
- Shaw, M. E. (1954). Group structure and the behavior of individuals in small groups. *The Journal of psychology*, 38(1), 139-149.
- Shimbel, A. (1953). Structural parameters of communication networks. *The bulletin of mathematical biophysics*, 15(4), 501-507.
- Shiode S., Shiode N. (2009). "Detection of multi-scale clusters in network space." *International Journal of Geographical Information Science*, 23(1):75-92.
- Shonkwiler, J. S., and Harris, T. R. (1993). A Non-Gaussian Time Series Analysis of Rural Retail Business Counts. *Journal of Regional Science*, 33(1), 37-48.
- Shonkwiler, J. S., and Harris, T. R. (1996). Rural retail business thresholds and interdependencies. *Journal of Regional Science*, 36(4), 617-630.
- Siksna, A. (1997). The effects of block size and form in North American and Australian city centres. *Urban morphology*, 1(1), 19-33.
- Sileshi, G., Hailu, G., and Nyadzi, G. I. (2009). Traditional occupancy–abundance models are inadequate for zero-inflated ecological count data. *Ecological Modelling*, 220(15), 1764-1775.
- Sitte, C. (1889). *City Planning According to Artistic Principles*, reprint 1986. Rizzoli, New York, USA.
- Soesanti, S., and Nakai, N. (2010). Street network in Bandung City, Indonesia: Comparison between City Center and new commercial area. *World Academy of Science, Engineering and Technology*, 42, 1104-1114.
- Song, Y., and Knaap, G. J. (2007). Quantitative classification of neighbourhoods: The neighbourhoods of new single-family homes in the Portland Metropolitan Area. *Journal of Urban Design*, 12(1), 1-24.
- Song, Y., and Knaap, G. J. (2007). Quantitative classification of neighbourhoods: The neighbourhoods of new single-family homes in the Portland Metropolitan Area. *Journal of Urban Design*, 12(1), 1-24. doi: 10.1080/13574800601072640.
- Song, Y., Popkin, B., and Gordon-Larsen, P. (2013). A national-level analysis of neighborhood form metrics. *Landscape and urban planning*, 116, 73-85. doi: 10.1016/j.landurbplan.2013.04.002.
- Stahl, K. (1982). Differentiated products, consumer search, and locational oligopoly. *The Journal of Industrial Economics*, 97-113.
- Stähle A. (2005). *Compact sprawl*. PhD dissertation, Chalmers tekniska högskola, Sweden.
- Stahle, A., Marcus, L., and Karlström, A. (2007). Place Syntax Tool—GIS Software for Analysing Geographic Accessibility with Axial Lines. *New developments in space syntax software*, 35-42.
- Steadman, P. (1973). Graph theoretic representation of architectural arrangement. *Architectural Research and Teaching*, 161-172.



- Steadman, P. (1973). Graph theoretic representation of architectural arrangement. *Architectural Research and Teaching*, 161-172.
- Steadman, P. (1983). *Architectural morphology: an introduction to the geometry of building plans*. London: Pion. Hillier, B., and Hanson, J. (1984). *The Social Logic of Space*, Cambridge Univ.
- Steadman, P. (1983). *Architectural morphology: an introduction to the geometry of building plans*. London: Pion.
- Steadman, P. (2013). Density and built form: integrating «Spacemate» with the work of Martin and March. *Environment and Planning B: Planning and Design*, (40), 341–358.
- Steadman, R., Myers, R. P., Leggett, L., Lorenzetti, D., Noseworthy, T., Rose, S., and Clement, F. (2013). A health technology assessment of transient elastography in adult liver disease. *Canadian Journal of Gastroenterology and Hepatology*, 27(3), 149-158.
- Steenberghen, T., Aerts, K., and Thomas, I. (2010). Spatial clustering of events on a network. *Journal of Transport Geography*, 18(3), 411-418.
- Stephen, M. (2005). *StreetsandPatterns*. New York: Spon Press.
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347), 730-737.
- Stevens, B.H. (1985). Location of economic activities: the JRS contribution to the research literature, *Journal of Regional Science*, 25 (4): 663-85.
- Stoyan, D. (1984). On correlations of marked point processes. *Mathematische Nachrichten*, 116(1), 197-207.
- Stoyan, D. (1987). Statistical analysis of spatial point processes: a soft-core model and cross-correlations of marks. *Biometrical Journal*, 29(8), 971-980.
- Stoyan, D., and Penttinen, A. (2000). Recent applications of point process methods in forestry statistics. *Statistical Science*, 15(1), 61-78.
- Stoyan, D., and Penttinen, A. (2000). Recent applications of point process methods in forestry statistics. *Statistical Science*, 15(1), 61-78.
- St-Pierre, A. P., Shikon, V., and Schneider, D. C. (2018). Count data in biology—Data transformation or model reformation?. *Ecology and evolution*, 8(6), 3077-3085.
- Takahashi, T. (2013). Agglomeration in a city with choosy consumers under imperfect information. *Journal of Urban Economics*, 76, 28-42.
- Taleb, N. N. (2001). *Foiled by randomness: The hidden role of choice in the markets and in life*. New York: Texere.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable* (Vol. 2). Random house.
- Talen, E. (2005). *New urbanism and American planning: the conflict of cultures*. Routledge.
- Talen, E. (2011). The geospatial dimension in urban design. *Journal of Urban Design*, 16(1), 127-149.
- Tannier, C., and Pumain, D. (2005). Fractals in urban geography: a theoretical outline and an empirical example. *Cybergeo: European Journal of Geography*.
- Tannier, C., Thomas, I., Vuidel, G., and Frankhauser, P. (2011). A Fractal Approach to Identifying Urban Boundaries. **城市边界识别的分形方法**. *Geographical Analysis*, 43(2), 211-227.
- Taud, H. and Parrot J.-F. (2005). “Measurement of DEM roughness using the local fractal dimension” *Géomorphologie: relief, processus, environnement*, 4(327-338).
- Tél, T., Fülöp Á., Vicsek T. (1989). Determination of fractal dimensions for geometrical multifractals, *Physica A: Statistical Mechanics and its Applications* 159:155-166.

- Thiessen, A. H. (1911). Precipitation averages for large areas. *Monthly weather review*, 39(7), 1082-1089.
- Thill, J. C. (1992). Spatial competition and market interdependence. *Papers in Regional Science*, 71(3), 259-275.
- Thill, J. C., and Thomas, I. (1987). Toward conceptualizing trip-chaining behavior: A review. *Geographical Analysis*, 19(1), 1-17.
- Thill, J. C., and Timmermans, H. (1992). Analyse des décisions spatiales et du processus de choix des consommateurs: théorie, méthodes et exemples d'applications. *L'Espace géographique*, 143-166.
- Thomas, E. N. (1960). *Maps of residuals from regression: their characteristics and uses in geographic research* (No. 2). Department of Geography, State University of Iowa.
- Thomas, G.B. (1960). *Calculus and analytical geometry*. Boston, MA: Addison-Wesley.
- Thünen, V. J., and Heinrich, J. (1826). Der Isolierte Staat in Beziehung auf Landschaft und Nationalökonomie. *Trans. By CM Wartenberg (1966) Von Thünen's Isolated State*. Oxford: Pergamon Press.
- Thurstain-Goodwin, M., and Gong, Y. (2005). Mapping long-term retail trends in London. *Journal of Targeting, Measurement and Analysis for Marketing*, 13(3), 220-233.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
- Timothée, P., Nicolas, L. B., Emanuele, S., Sergio, P., and Stéphane, J. (2010, March). A network based kernel density estimator applied to Barcelona economic activities. In *International conference on computational science and its applications* ( 32-45). Springer, Berlin, Heidelberg.
- Tiwari, P., Doi, M., and Kawakami, T. (2006). Analysis of household leisure and shopping behavior in Ibaraki Prefecture, Japan. In *Review of Urban and Regional Development Studies: Journal of the Applied Regional Science Conference 18(2)*, 165-178. Melbourne, Australia: Blackwell Publishing Asia.
- Tjur, T. (2009). Coefficients of determination in logistic regression models—A new proposal: The coefficient of discrimination. *The American Statistician*, 63(4), 366-372.
- Todem, D., Hsu, W. W., and Kim, K. (2012). On the efficiency of score tests for homogeneity in two-component parametric models for discrete data. *Biometrics*, 68(3), 975-982.
- Topçu, M., and Kubat, A. S. (2009). The analysis of urban features that affect land values in residential areas. In *Proceedings of the 7th International Space Syntax Symposium, Stockholm, Sweden* (8-11).
- Torrens, P. M. (2008). A toolkit for measuring sprawl. *Applied Spatial Analysis and Policy*, 1(1), 5-36.
- Tsai, Y. H. (2005). Quantifying urban form: compactness versus 'sprawl'. *Urban studies*, 42(1), 141-161.
- Tsou, K. W., and Cheng, H. T. (2013). The effect of multiple urban network structures on retail patterns—A case study in Taipei, Taiwan. *Cities*, 32, 13-23.
- Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
- Turner, A. (2001a). Depthmap: a program to perform visibility graph analysis. In *Proceedings of the 3rd International Symposium on Space Syntax* 31, 31-12.
- Turner, A. (2001b). Angular analysis. In *Proceedings of the 3rd international symposium on space syntax* (30-1). Atlanta, GA: Georgia Institute of Technology.

- Turner, A. (2007). From axial to road-centre lines: a new representation for space syntax and a new model of route choice for transport network analysis. *Environment and Planning B: Planning and Design*, 34(3), 539-555.
- Vakhutinsky, A., Mihic, K., and Wu, S. M. (2019). A Prescriptive Analytics Approach to Markdown Pricing for a in E-Commerce Retailer. *Journal of Pattern Recognition Research*, 1, 1-21.
- Van Nes, A. (2005). Typology of shopping areas in Amsterdam. In *proceedings space syntax. 5th international symposium*. Delft.
- Van Nes, A. (2014). Indicating street vitality in excavated towns: Spatial configurative analyses applied to Pompeii. *Spatial analysis and social spaces: Interdisciplinary approaches to the interpretation of prehistoric and historic built environments*.
- Van Nes, A., and Nguyen, T. M. (2009). Gender Differences in the Urban Environment: The flâneur and flâneuse of the 21st Century. In *Proceedings of the 7th International Space Syntax Symposium* 122, 1-7.
- Van Nes, A., and Yamu, C. (2017). Space Syntax: A method to measure urban space related to social, economic and cognitive factors. In *The Virtual and the Real in Planning and Urban Design* (136-150). Routledge.
- Van Nes, A., Berghauer Pont, M., and Mashhoodi, B. (2012). Combination of Space syntax with spacematrix and the mixed use index: The Rotterdam South test case. In *8th International Space Syntax Symposium, Santiago de Chile, Jan. 3-6, 2012*. PUC, Santiago, Chili.
- Vazquez Parras, J. (2011). Classification géométrique des espaces commerçants wallons-Vers une objectivation des caractéristiques morphologiques des nodules? *Bulletin de la Société Géographique de Liège*, 56(thématique).
- Velicer, W. F., and Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate behavioral research*, 25(1), 1-28.
- Venerandi, A., Quattrone, G., Capra, L., Quercia, D., and Saez-Trumper, D. (2015, February). Measuring urban deprivation from user generated content. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing* ( 254-264). ACM.
- Verstraete, G., Aghezzaf, E. H., and Desmet, B. (2019). A data-driven framework for predicting weather impact on high-volume low-margin retail products. *Journal of Retailing and Consumer Services*, 48, 169-177.
- Verstraete, G., Aghezzaf, E. H., and Desmet, B. (2019). A data-driven framework for predicting weather impact on high-volume low-margin retail products. *Journal of Retailing and Consumer Services*, 48, 169-177.
- Vialard, A. (2013). *A Typology of Block-faces*, PhD dissertation, Georgia Institute of Technology, USA.
- Vialard, A., and Carpenter, A. (2015). Building density of parcels and block-faces from a syntactical, morphological and planning perspective. In *Proceedings of the 10th international space syntax symposium* (p. 100).
- Vialard, A., Ingleby, T., and Ring, P. (2017). *Fight for Frontage: A Shopping Mall On a High Street*. In: *Proceedings of the 11th International Space Syntax Symposium*. Instituto Superior Tecnico - Universidade de Lisboa, Lisbon.
- Villain, J. (2011). *The Impact of the Urban Form on the Spatial Distribution of Commercial Activities in Montréal*, MSc dissertation, Concordia University, Canada.
- Voiron-Canicio, C. (1995). *Analyse spatiale et analyse d'image*, Montpellier, Reclus, Espace mode d'emploi.

- Von Bertalanffy, L. (1968). General system theory. *New York*, 41973(1968), 40.
- Voronoi, G. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *Journal für die Reine und Angewandte Mathematik*.
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, 307-333.
- Wang, C. L., Chen, Z. X., Chan, A. K., and Zheng, Z. C. (2000). The influence of hedonic values on consumer behaviors: an empirical investigation in China. *Journal of Global Marketing*, 14(1-2), 169-186.
- Wang, E., and Song, J. (2008). The political economy of retail change in Chinese cities. *Environment and Planning C: Government and Policy*, 26(6), 1197-1226.
- Wang, F., Antipova, A., and Porta, S. (2011). Street centrality and land use intensity in Baton Rouge, Louisiana. *Journal of Transport Geography*, 19(2), 285-293.
- Wang, F., Chen, C., Xiu, C., and Zhang, P. (2014). Location analysis of retail stores in Changchun, China: A street centrality perspective. *Cities*, 41, 54-63.
- Wang, S., and Jones, K. (2002). Retail Structure of Beijing. *Environment and Planning A: Economy and Space*, 34(10), 1785-1808.
- Wang, S., and Zhang, Y. (2005). The new retail economy of Shanghai. *Growth and Change*, 36(1), 41-73.
- Wang, S., Xu, G., and Guo, Q. (2018). Street centralities and land use intensities based on points of interest (POI) in Shenzhen, China. *ISPRS International Journal of Geo-Information*, 7(11), 425.
- Wang, Z., and Harris, T. R. (2012). Retail Business Pattern Changes in Rural Development. Available at SSRN 2698160.
- Wang, Z., Ma, S., and Wang, C. Y. (2015). Variable selection for zero-inflated and overdispersed data with application to health care demand in Germany. *Biometrical Journal*, 57(5), 867-884.
- Wang, Z., and Harris, T. R. (2012). Retail Business Pattern Changes in Rural Development, Available at SSRN 2698160.
- Warton, D. I. (2005). Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics: The official journal of the International Environmetrics Society*, 16(3), 275-289.
- Wayens, B. (2006). *Structure et dynamique du commerce de détail bruxellois. Une approche sous l'angle des facteurs de localisation différentielle*. PhD Dissertation Université Libre de Bruxelles, Belgium.
- Wayens, B., and Grimmeau, J. P. (2004). Hiérarchie des noyaux commerciaux de la Région de Bruxelles-capitale. *Cahier bruxellois de l'urbanisme*, 2004, 24.
- Weber, M. (1958). *Essays in sociology*. A Galaxy Book.
- Wei, H. L., and Billings, S. A. (2007). Feature subset selection and ranking for data dimensionality reduction. *IEEE transactions on pattern analysis and machine intelligence*, 29(1), 162-166.
- Westhuizen van der, D. (2009). Contributions of accessibility and visibility characteristics to neighborhood typologies and their predictions of physical activity and health. 7th International Space Syntax Symposium.
- White, P.G. (1975). *The Composite Structure of the Central Area*, Newcastle: University of Newcastle upon-Tyne, Department of Geography, 30.
- White, R., and Engelen, G. (1993). Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment and planning A*, 25(8), 1175-1199.

- Whitehand, J. W. (2001). British urban morphology: the Conzenian tradition. *Urban morphology*, 5(2), 103-109.
- Wilkinson, L., and Dallal, G. E. (1981). Tests of significance in forward selection regression with an F-to-enter stopping rule. *Technometrics*, 23(4), 377-380.
- Wilson, P. (2015). The misuse of the Vuong test for non-nested models to test for zero-inflation. *Economics Letters*, 127, 51-53.
- Wrigley N. and Lowe M. (2002). *Reading Retail*, Londres, Arnold.
- Wrigley, N., and Lowe, M. (1996). *Retailing, consumption and capital: towards the new retail geography*. Longman.
- Wrigley, N., and Lowe, M. (2014). *Reading retail: A geographical perspective on retailing and consumption spaces*. Routledge.
- Xie, Z., and Yan, J. (2008). Kernel density estimation of traffic accidents in a network space. *Computers, environment and urban systems*, 32(5), 396-406.
- Xie, Z., and Yan, J. (2013). Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach. *Journal of transport geography*, 31, 64-71.
- Yamada, I., and Thill, J. C. (2004). Comparison of planar and network K-functions in traffic accident analysis. *Journal of Transport Geography*, 12(2), 149-158.
- Yamada, I., and Thill, J. C. (2007). Local indicators of network-constrained clusters in spatial point patterns. *Geographical Analysis*, 39(3), 268-292.
- Yamada, I., and Thill, J. C. (2010). Local indicators of network-constrained clusters in spatial patterns represented by a link attribute. *Annals of the Association of American Geographers*, 100(2), 269-285.
- Ye, Y., and Van Nes, A. (2013). Measuring urban maturation processes in Dutch and Chinese new towns: Combining street network configuration with building density and degree of land use diversification through GIS. *The Journal of Space Syntax*, 4(1), 18-37.
- Ye, Y., and Van Nes, A. (2014). Quantitative tools in urban morphology: Combining space syntax, spacematrix and mixed-use index in a GIS framework. *Urban morphology*, 18(2), 97-118.
- Ye, Y., Li, D., and Liu, X. (2018). How block density and typology affect urban vitality: An exploratory analysis in Shenzhen, China. *Urban Geography*, 39(4), 631-652.
- Yoon, H. K. (1997). The Origin and Decline of a Japanese Temple Town: Saidaiji Monzenmachi. *Urban Geography*, 18(5), 434-450.
- Yoshida, H., and Omae, M. (2005). An approach for analysis of urban morphology: methods to derive morphological properties of city blocks by using an urban landscape model and their interpretations. *Computers, Environment and Urban Systems*, 29(2), 223-247.
- Zhang, H. (2004). The optimality of naive Bayes. *AA*, 1(2), 3.
- Zhang, L., Ye, Y., Zeng, W., and Chiaradia, A. (2019). A Systematic Measurement of Street Quality through Multi-Sourced Urban Data: A Human-Oriented Analysis. *International journal of environmental research and public health*, 16(10), 1782.
- Zhang, M., and Kukadia, N. (2005). Metrics of urban form and the modifiable areal unit problem. *Transportation Research Record*, 1902(1), 71-79.
- Zhao, X., Lim, A., Guo, H., Ding, C., and Song, J. S. (2018). Retail Clusters in Developing Economies. *Manufacturing and Service Operations Management*, 21(2), 452-467.
- Zipf, G. K. (1949). *Human behaviour and the principle of least-effort*. Cambridge MA edn. Reading: Addison-Wesley.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., and Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer Science and Business Media.



# LIST of ACRONYMS and ABBREVIATIONS

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BP	Breaking Point
BR	Backward Regression
BRT	Bid Rent Theory
CBD	Central Business District
CCI	Chamber of Commerce and Industry
CDF	Cumulative Distribution Function
C-REC	Count-Regression Error Characteristic
EDA	Exploratory Data Analysis
Enet	Elastic Net Penalised Regression
GLM	Generalized Linear Model
HAC	Hierarchical Agglomerative clustering
IDW	Inverse Distance Weight
KDE	Kernel Density Estimation
LASSO	Least Absolute Shrinkage Selection Operator
LDA	Linear Discriminant Analysis
LR	Likelihood Ratio test
MAUP	Modifiable Areal Unit Problem
MCA	Multiple Centrality Assessment
MCP	Minimax Concave Penalty
MFA	Multiple Fabric Assessment
MLE	Maximum Likelihood Estimation
MLR	Multiple Linear Regression
NetKDE	Network Constrained KDE
Nnh	Nearest Neighbour
OKM	Ordinary Kriging Method
OLS	Ordinary Least Square
P	Poisson
PCA	Principal Component Analysis
PDF	Probability Density Function
PMD	Principal of Minimum Differentiation



QT	Quality Threshold
REC	Regression Error Characteristic
ROC	Receiver Operating Characteristic
SCAD	Smoothly Clipped Absolute Deviation
SIT	Spatial Interaction Theory
SSx	Space Syntax
VIF	Variance Inflated Factor
ZA	Zero Altered
ZANB	Zero Altered Negative Binomial
ZAP	Zero Altered Poisson
ZI	Zero Inflated
ZINB	Zero Inflated Negative Binomial
ZIP	Zero Inflated Poisson

# LIST of TABLES

Table 1 Literature review-part I .....	55
Table 2 Literature review-part II .....	56
Table 3 Repartition of Network Buildings and Retail in the three main geographical sectors of the French Riviera .....	99
Table 4 Methodological Literature Review of Micro-Retail Studies in Economic Geography and Point pattern analysis.....	121
Table 5 Form and Function measures implemented in Retail Fabric Assessment methodology...	132
Table 6 RFA Indicator mutual information with cluster variable (left) and frequency of observation among the 10 collectively most informative variables (right) .....	148
Table 7 RFA Mutual Information of each Retail Fabric.....	151
Table 8 Street Network Configurational Indicators .....	222
Table 9 Directional Centrality Indicators (Squares, Coastline and Anchor Stores).....	224
Table 10 Streetscape indicators from Street Centerline Sightlines; (* corrected measures with lane width).....	230
Table 11 The statistical dispersion of different spatial unit dimensions at the origin of the MAUP	232
Table 12 Streetscape indicators implemented through Proximity Band procedure.....	236
Table 13 MFA Indicator mutual information with cluster variable (left) and frequency of observation among the 10 collectively most informative variables (right.) .....	249
Table 14 MFA Mutual Information of each Urban Fabric Class.....	250
Table 15 The distribution of streets, stores and the relative store density for each morphological regions.....	297
Table 16 Repartition of retail agglomeration distribution (RFA typology, Section 5.3) within urban morphological regions (Urban Fabrics and Morphological Macro Regions from MFA Section 9.3). Upper part: proportion of Urban Fabric j with Retail Fabric i; Lower part: proportion of Retail Fabric i within Urban Fabric j. ....	299
Table 17 Model selection: AIC-Akaike Information Criteria; c-number of parameters in the model; n-feature number; G-Gaussian; P-Poisson; NB-Negative Binomial; ZI-Zero Inflated; * and ** indicate that AIC have been obtained with, respectively, logit and probit functions. ....	338
Table 18 Results of the Vuong test between our seven models for the overall study area and for each morphological region. ....	339
Table 19 Comparison of results of NB and ZINB models when global and sub-region are evaluated on the same subgroup of features. For each goodness of fit measure the raw values and the percent chance between the models implemented on the overall space study each specific subregion. ....	342
Table 20 Confusion matrix displaying the results of the multinomial model for the prediction of the retail fabrics (a), its sensitivity (b) and precision (c). ....	376
Table 21 Odd Ratios of the ElasticNet penalised regression procedure applied to the multinomial model describing the relationship between retail agglomeration types (10-levels RFs) and urban morphological indicators. A total of 36 regressors are selected.....	378

# LIST of FIGURES

Figure 1 The illustration of nested hexagons in Christaller's central place theory (Christaller 1933, p. 71). .....	25
Figure 2 Illustration of the Bid Rent Theory: A) von Thünen 1927, B) Alonso 1964, C) Scott 1970 ...	27
Figure 3 Berry's hierarchical model for retail system (A) and its application to the city of Chicago (B) (reproduction from the original cover of Berry's 1963 "Commercial structure and commercial blight"). .....	36
Figure 4 Davies' Structural models of central are core retailing facilities (A), and its application to the Coventry central area (B) (Davies 1972).....	37
Figure 5 Brown's post-hierarchical model (1991).....	38
Figure 6 The chronological and thematic evolution of the retail geography approaches and the relative socio-technological context .....	44
Figure 7 Schematic illustration of the different approaches of retail geography (left) and the specific approach chosen in this dissertation: the physical city (right) .....	45
Figure 8: A) The retail center of Nashville, Tennessee (Parkins 1930); B) The phases of the urban development and retail distribution (Beaujeu -Garnier 1977) .....	46
Figure 9 Some morphological fractures identified by Lebrun 2002 in Reims(left) and Nancy (right)	47
Figure 10 Graphical representation of Hillier's Movement Economy Theory .....	54
Figure 11 Urban morphological context, might outline different relationship between retail distribution and network configurational properties, traditionally hidden when the overall study area values are considered .....	63
Figure 12 Seven different urban landscapes extracted from the central area of Nice .....	65
Figure 13 A graphical representation of the spatial interactions between retail and flows elaborated from Lebrun's description (2002). .....	69
Figure 14 Graphical representation of the retail and physical systems, their elementary components .....	72
Figure 15 Possible causes at the origin of the Survivorship Bias .....	80
Figure 16 The study of the relationship between the retail spatial organisation and urban morphology: while the Movement Economy Theory (Hillier 1999) focuses on the relation between network configuration and retail density (dark grey), in this dissertation we also consider additional descriptors of both retail and urban morphological system (light grey)	82
Figure 17 The administrative and natural context surrounding the Alpes Maritimes Department and the repartition of the population in the littoral, sublittoral and inland sectors.....	84
Figure 18 The French Riviera along the history: a region of borders and passageway.....	85
Figure 19 Repartition of the retail surface in our study area; in red, the national highway and its accesses (elaboration from CCINCA 2016).....	92
Figure 20 Repartition of the retail activities categories of the French Riviera: retail number (left) and retail large surfaces (>300m <sup>2</sup> , right). .....	98
Figure 21 Spatial distribution of small and average-sized retail. On the left: street segment retail count; On the right: street segment retail density. ....	101
Figure 22 Retail agglomeration classification outcomes in the three seminal works of: a) Davies (1976), Coventry; b) Potter(1981), Stockport; c) Beaujeu-Garnier (1977), Paris.....	110
Figure 23 The three-step procedure traditionally implemented in Retail Geography studies for the identification of retail structure at the intra-urban scale.....	122

Figure 24 The network-based catchment area around three stores (A, B, C) and their relative number of store families at three different scales (street, neighbourhood and district) ...	128
Figure 25 Example of Function Pattern applied to clothing store distribution .....	130
Figure 26 Example of Form Pattern: distribution variable at the district level .....	131
Figure 27 NC-LFD (a), Global K (b) and Local K (c-e) for three different theoretical layouts of retailers disposed over a regular meshed grid.(source: Araldi and Fusco 2019).....	140
Figure 28 Retail activity fractal dimension, coefficient of determination and local relative concentration, calculated in a 300 m radius. (Source: Araldi and Fusco 2019).....	141
Figure 29 Classification of the retail fabric at different scales. (Source: Araldi and Fusco 2019). ...	142
Figure 30 Retail fabric classification in Menton (left) and Antibes (right). (Source: Araldi and Fusco 2019) .....	143
Figure 31 Retail fabric classification through NC-LFD (a) and G-ILINCS (b) (Source: Araldi and Fusco 2019) .....	143
Figure 32 Projection in mutual information space of the variables (form-green, function-blue)....	148
Figure 33 Projection in mutual information space of the clusters .....	150
Figure 34 The core-centers of Nice, Cannes, Antibes (point-pattern representation of RFA results) .....	152
Figure 35 RF1 Touristic-oriented agglomerations in Antibes, Cannes and Juan-les-Pins (above); street view of RF1 (Source: Google Street View 2019).....	154
Figure 36 RF2 Clothing and Leisure-oriented agglomerations in Cannes, Menton and Nice (above); street view of RF2 (Source: Google Street View 2019).....	157
Figure 37 RF3 Mixed and Service-oriented agglomerations in Nice, Menton and Antibes.....	158
Figure 38 Street view of RF3 (Source: Google Street View 2019) .....	159
Figure 39 RF4 Residential pervasive agglomeration in Nice and Cannes .....	160
Figure 40 Almost-360° street view of RF4 (Source: Google Street View 2019).....	161
Figure 41 RFA outcomes in peripheral areas.....	162
Figure 42 RF5.a Pedestrian Ribbons in Nice metropolitan area(above); street view of RF5.a (Source: Google Street View 2019) .....	164
Figure 43 Pedestrian Ribbon (RF5.a) and scattered out-of-ribbon (RF5.b) in Nice (above). Some street view pictures of RF5.b (below) (Source: Google Street View 2019) .....	165
Figure 44 RF6.a: Scattered Vehicle Ribbons in the French Riviera and some local examples in Cagnes-sur-Mer(a), Menton(b) and Antibes(c). (Source: Google Map-Google Street View 2019) .....	166
Figure 45 Specialised Nucleated Peripheral agglomerations (RF6.b).....	167
Figure 46 The concentration of Car Seller along the Avenue du Campon, Le Cannet .....	168
Figure 47 RF7 The identification of the 12 average-sized cities and neighborhood centralities (above); some Street View of RF7 (Source: Google Street View 2019) .....	169
Figure 48 The identification of the isolated touristic/leisure oriented retail agglomerations, associated with small historical centres (above); Street View of RF8 (Source: Google Street View 2019) .....	171
Figure 49 RF9.a Fragmented pedestrian ribbons in a residential local-neighbourhood community. Some examples in Le Cannet (a), Villefranche (b) and Nice (c); Street View of RF9 (Source: Google Street View 2019). .....	173
Figure 50 Retail districts identified by local authorities. Source: Interscot Commerce (2012). .....	174
Figure 51 The spatial organisation of retail fabrics: a model for the French Riviera Conurbation ..	175
Figure 52 Example of VGA analysis (on the left) and segment analysis (on the right) (source: UCL Space Syntax: <a href="http://otp.spacesyntax.net/">http://otp.spacesyntax.net/</a> ). .....	186

Figure 53 Overview scheme of configurational approaches in urban morphology studies. (Readaptation from: Fusco and Tirico 2016) .....	187
Figure 54 Marshall's Street pattern taxonomy and ABCD diagram (Source: Marshall 2005) .....	194
Figure 55 Spacematrix (source: Berghauser Pont and Haupt 2007).....	195
Figure 56 Traditional classification of blocks and buildings do not consider their relative spatial distribution. Shuffling street-block position or building position within blocks, the outcomes of the traditional three-step classification would result in the same typologies of building/block morphological type. ....	198
Figure 57 Location-based measures: while Song and Knaap (2007) and Berghauser Pont and Marcus (2014, 2017) focus on specific radii for the identification of local morphological regions, Fractal and geostatistical studies show the evolution of density at different scales and specific behaviours (such as, respectively, fractal transition or catastrophes in Badariotti). .....	200
Figure 58 Skeletal streetscape (Source: Harvey 2017) .....	204
Figure 59 Methodological literature review of typo-morphological and urban design studies developing quantitative approaches for the analysis of the urban form. ....	207
Figure 60 Reach and Betweenness configurational centralities implemented at different scales (pedestrian-oriented on the left and vehicle oriented on the right).....	216
Figure 61 Projection in the geographic space (Nice) of Reach-R and Betweenness-B indicators (left) and their normalized version through the 2SFCA method (right) .....	220
Figure 62 Projection in the geographic space (Cannes and Antibes) of network configurational indicators (left) and their normalised version (right). ....	221
Figure 63 Two different approaches for the compositional analysis of the interface between streets and blocks. On the left, Vialard (2013) block-face approaches (radial sightlines and 20-metre buffer); on the right, the skeletal streetscape of Harvey et Aultman Hall (2015). ....	225
Figure 64 Some example of semi-peripheral suburbs in Nice, France (left) and Baltimore, USA (right). ....	227
Figure 65 Streetscape measures implemented with sightlines. Graphical representation of the procedure (left), and combination given by Façades alignment and Vertical Alignment. ...	228
Figure 66 Projection in the geographical space of streetscape indicators. ....	230
Figure 67 Streetscape measures implemented on the Proximity Band definition.....	232
Figure 68 Proximity Band implementation from street segments (above) and some projection on real case studies (below). ....	233
Figure 70 Example of combinations of Street Corridor Effect and Façade Misalignment. ....	237
Figure 69 Skeletal streetscape morphological indicators. ....	237
Figure 71 Projection in the geographical space of some skeletal streetscape morphological indicators. ....	238
Figure 72 Moran's I Scatterplot. ....	242
Figure 73 Outcomes from the implementation of LISA (on Generalized Thiessen Partition) and ILINCS (on Proximity Bands partitions) to morphological streetscape indicators (Corridor Effect). Zoom on the urban area between Cagnes-sur-Mer and Saint Laurent-du-Var. Source: Araldi and Fusco 2016.....	243
Figure 74 Patterns of morphological streetscape indicators, resulting from the application of network constrained local Moran's I statistics (ILINCS). Source: Araldi and Fusco 2017. ...	246
Figure 75 Projection in mutual information space of the MFA variables (left) and outcomes (right). .....	249
Figure 76 Aerial and Street View images of streets belonging to UF1, UF2 and UF3. (Source: Google Map and Google Street view 2017).....	251

Figure 77 Aerial and Street View images of streets belonging to UF4, UF5 and UF6. (Source: Google Map and Google Street view 2017).....	252
Figure 78 Aerial and Street View images of streets belonging to UF7, UF8 and UF9. (Source: Google Map and Google Street view 2017).....	253
Figure 79 Characterising urban fabrics in the French Riviera through probabilities of morphological patterns; synthetic descriptions of MFA outcomes (Source: Araldi and Fusco 2019).....	254
Figure 80 Projection in the geographical space of the probabilistic values associated with each Urban Fabric class. Darker-lighter colours correspond to higher-lower probabilities; grey segments correspond to a null probability. ....	256
Figure 81 Archetypical spatial organisation of morphological regions in French Riviera cities. (Source: Fusco and Araldi 2017) .....	258
Figure 82 Projection in the geographical space of the nine Urban Fabric: Grasse, Cannes and Antibes (Source: Araldi and Fusco 2019). ....	259
Figure 83 Projection in the geographical space of the probabilistic MFA outcomes: Cannes and Antibes. ....	259
Figure 84 Projection in the geographical space of the nine Urban Fabric: Nice metropolitan area (Source: Araldi and Fusco 2019). ....	260
Figure 85 Projection in the geographical space of the probabilistic MFA outcomes: Nice. ....	260
Figure 86 Plan de Nice, anonyme,1790. Archives municipales de Nice. Source (Graff 2013).....	262
Figure 87 The city of Nice (grey) and the new plan of the Consiglio d’Ornato (orange) using squares with a key role for the street grid organisation: A) Place Massena; B) Place Garibaldi; C) Port Lympia.....	263
Figure 88 "Aublet" Zoning Plan of Nice 1950 (Source: Graff 2013).....	265
Figure 89 The city of Nice. On the top: the historic development of the urban system and its main transport infrastructures. On the bottom: MFA outcomes.....	267
Figure 90 The Var Valley in 1780. DR/ Archives départementales-Conseil général des Alpes-Maritimes.....	268
Figure 91 On the left: MFA outcomes in the southern section of the Var Valley; on the right: the current(2016) landscape of the neighbourhood of Meridia and its master plan within the OIN Eco-Valley program.....	270
Figure 92 Antibes and its urban evolution: a) the Greek (IV b.C.) and Roman (I a.C) grids of Antipolis; b) the plan for the reinforcement of Antibes fortification walls and the construction of the Fort Carré (XVI century); c) Ernest Mace urban plan for the extension of Antibes in the region previously occupied by the military walls (1895); Today's Antibes as perceived from aerial image from Google Earth (d) and the urban fabric classification of the local Authorities (e); in (f) MFA outcomes. ....	272
Figure 93 Antibes and its urban evolution: a) the Greek (IV b.C.) and Roman (I a.C) grids of Antipolis; b) the plan for the fortification of Antibes and the construction of the Fort Carré (XVI century); c) Ernest Mace urban plan for the extension of Antibes in the region previously occupied by the military walls (1895); Today's Antibes as perceived from aerial image from Google Earth (d) and the urban fabric classification of the local Authorities (e); in (f) MFA outcomes. ....	273
Figure 94 a) Yves Bayard original plan for Sophia Antipolis; b) the four phases of the planning evolution of Sophia Antipolis in the last 5 decades; c) Aerial image of the Sophia Antipolis region from Google Earth 2019; d) MFA outcomes.....	277
Figure 95 The urban morphological context at the meso- and macro-scales: projection of the nine Urban Fabrics and First/Second-Age Cities regions.....	280

Figure 96 The urban morphological context at the meso- and macro-scales: the distribution of the street within the nine Urban Fabrics and First/Second-Age Cities regions and natural spaces. ....	280
Figure 97 The street as a keystone: bridging urban form (Configurational, Contextual, Compositional) and retail measures (magnitude, presence/absence and agglomeration typology) at the street level.....	290
Figure 98 Schematic visualization of store presence/absence repartition in the French Riviera. Vertical axes: absence/presence of stores. Horizontal axes: Urban Fabric classes (the width of each column corresponds to the number of streets with built form on their relative 50-metres proximity bands).....	298
Figure 99 Histograms of the retail count (upper part) and density (lower part) for the streets with buildings in the overall study area.....	300
Figure 100 Network-constrained topological variogram (blue) and correlogram (orange) describing the spatial structure and autocorrelation of the retail distribution on the French Riviera.	303
Figure 101 Rootgram : a graphical approach to assess the fit of count regression models.....	330
Figure 102 A simple schematic representation for the implementation of Performance Parameters using the Confusion Matrix representation of a regression model outcomes.....	332
Figure 103 Schematic example of the REC-CR curves implement for the evaluation of the goodness of fit of the models. ....	334
Figure 104 Schematic illustration of model performance parameters derived from a confusion matrix representation of its outcomes. ....	335
Figure 105 Rootgram of the Global model (overall space study and variables) for: Linear model (top), Poissonian Regressions P-ZIP-ZAP (bottom-left) and Negative Binomial regressions NB-ZINB-ZANB (bottom-right). The last group of models shows a better fit between expected and observed values. ....	340
Figure 106 Projection in the geographical space of the observed, predicted and residual values for the Global model, First/Second-Age and Urban Fabric composite models.....	345
Figure 107 Projection in the geographical space of the model prediction differences. From the top to the bottom: (First/Second-age-City – Global); (Urban Fabric –Global);( First/Second-age-City –Urban Fabrics).....	346
Figure 108 Correlogram of model errors and of store counts. The spatial structure of the urban morphological system allows explaining the spatial organisation of retail at the mesoscale while residual autocorrelation is detected at the local-scale (about 3 topological steps)..	347
Figure 109 REC-Curves for the Global ZINB model.....	347
Figure 110 Selected variables from the Backward Regression and Enet Penalised Regression to the Global model.....	352
Figure 111 Selected variables from the Backward Regression and Enet Penalised Regression to the First-Age City Model.....	355
Figure 112 Selected variables from the Backward Regression and Enet Penalised Regression to the Second-Age City Model.....	357
Figure 113 Selected variables from the Backward Regression and Enet Penalised Regression to the UF1 Model.....	361
Figure 114 Selected variables from the Backward Regression and Enet Penalised Regression to the UF3 Model.....	364
Figure 115 Selected variables from the Backward Regression and Enet Penalised Regression to the UF4 Model.....	366
Figure 116 Selected variables from the Backward Regression and Enet Penalised Regression to the UF5 Model.....	367

Figure 117 Selected variables from the Backward Regression and Enet Penalised Regression to the UF6 Model..... 369

Figure 118 A-B: example of streetscapes in historical centres with strong street corridor effect and high building façade values, positively related to retail presence. C-D: examples of streetscapes with streetscape conditions negatively correlated to retail activities: lower corridor effect (interruptions observed with small empty plot -C- and presence of specialized building –D-), higher acclivities, and lower building heights. E-F: examples of streetscape where despite the low corridor effect (buildings along only one side of the street), low building height (in E) and high acclivities (in F) they still have high retail densities due to their proximity to the coastline..... 371

Figure 119 Image A-B) High Built-up-unit fragmentation associated with high number of stores. Image C) In the same street, with the same values of configurational and contextual descriptors, we might observe: on the left hand-side, high building fragmentation associated with eight small-sized stores at the ground floor. On the right hand-side, a lower built-up-unit fragmentation (only two contiguous building) with only 3 average-sized stores. Image D) Example of street with presence of exceptional large surface buildings with specialized function, the National Theatre of Nice. Image E) Example of small built-up-unit fragmentation associated with lower retail concentrations: individual buildings (Historical Palaces and Hotels) along the Promenade des Anglais. .... 372

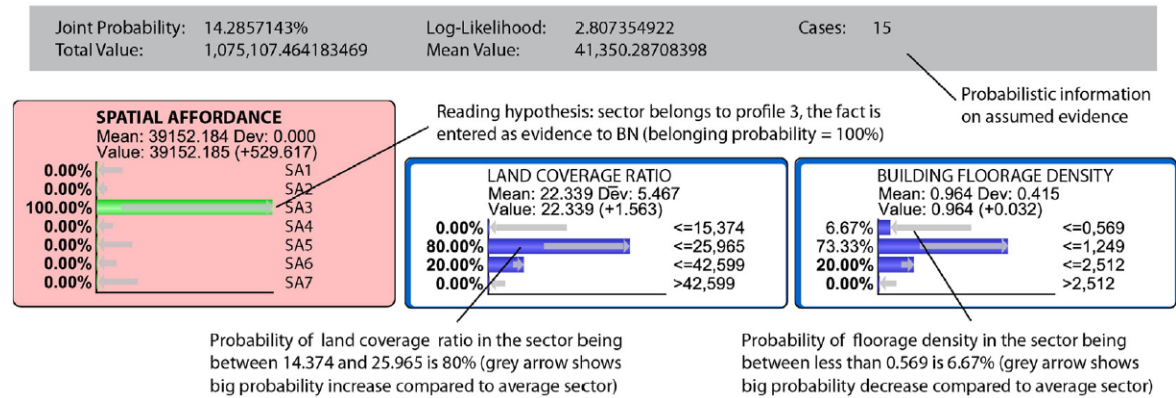
Figure 120 Projection in the geographical space of the observed values (left) and of the predictions of the multinomial model for the retail fabrics typology ..... 377



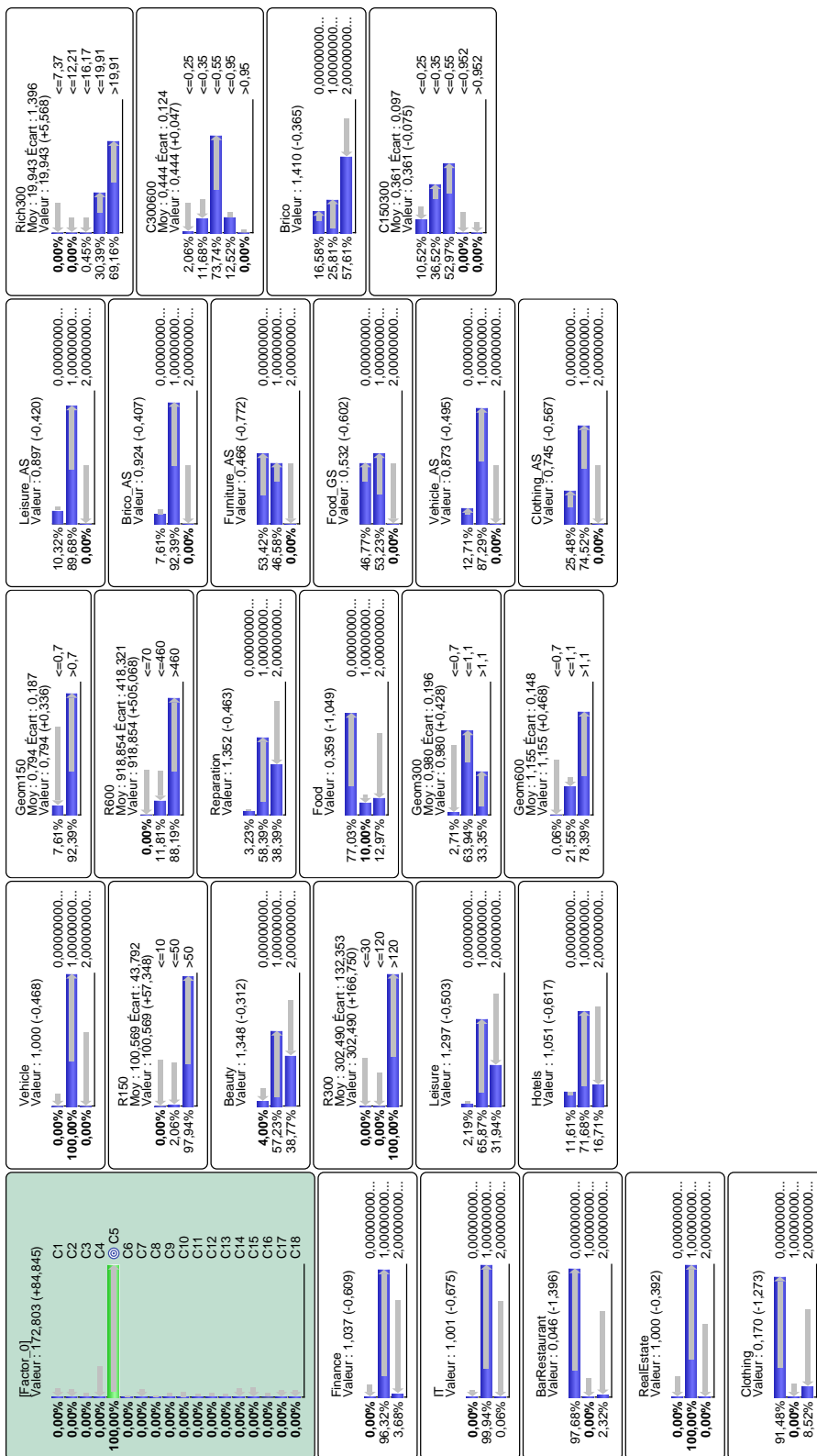
# APPENDIX 1

## The probabilistic profiles of RFA classes. (Chapter 3).

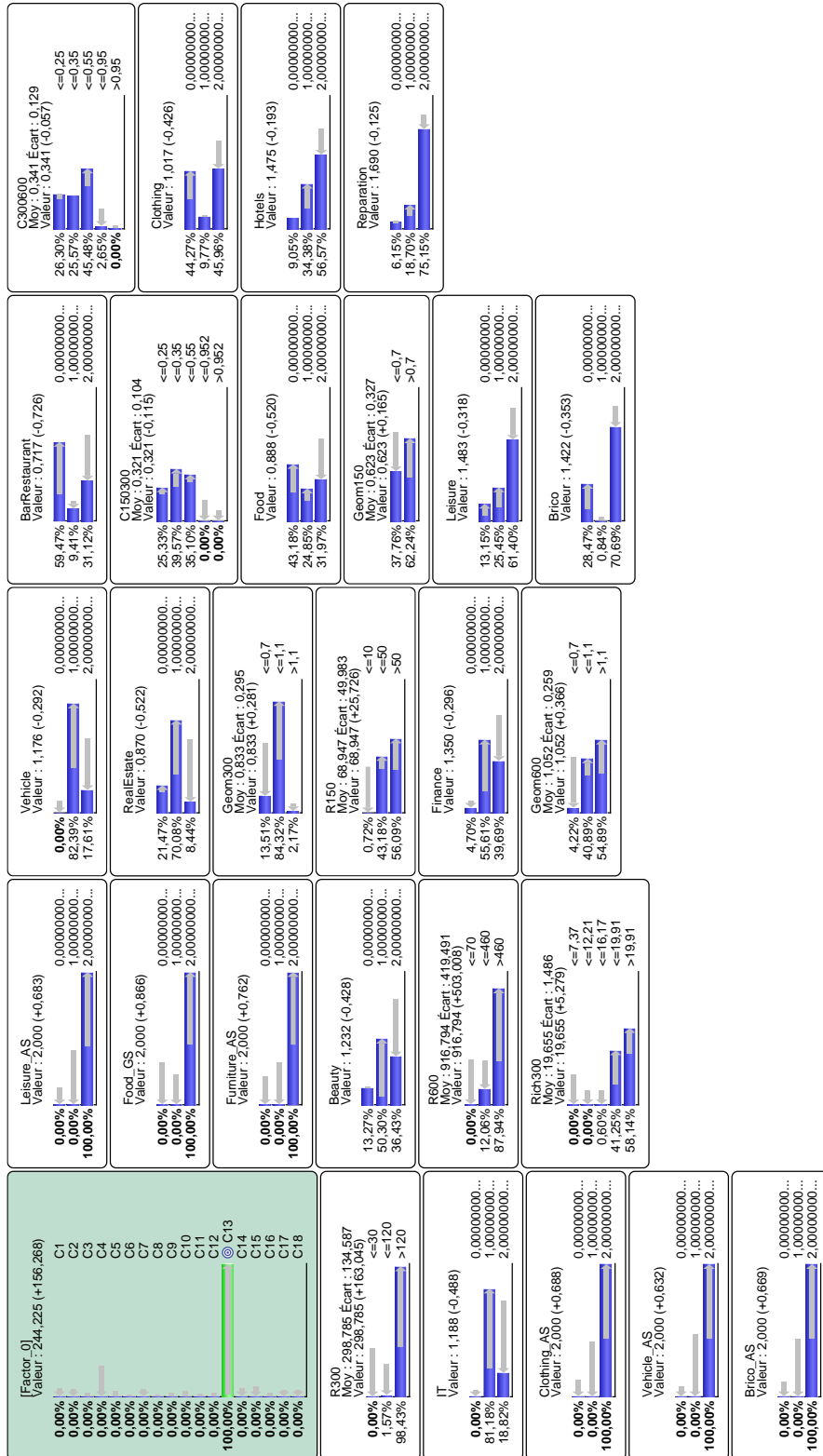
In the following image: how to read a Bayesian probabilistic profile. (from Fusco 2016)



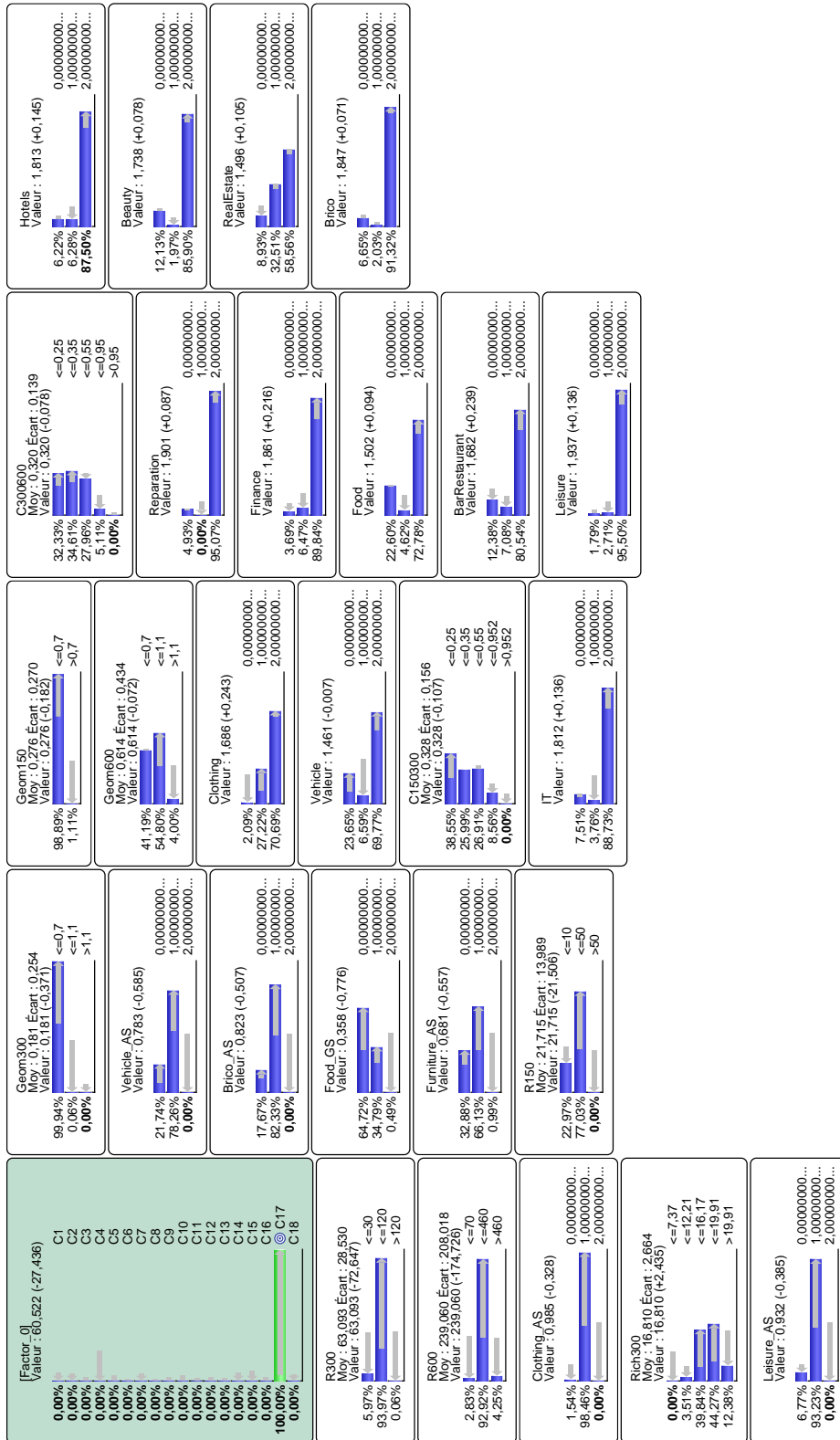
# RF1.a



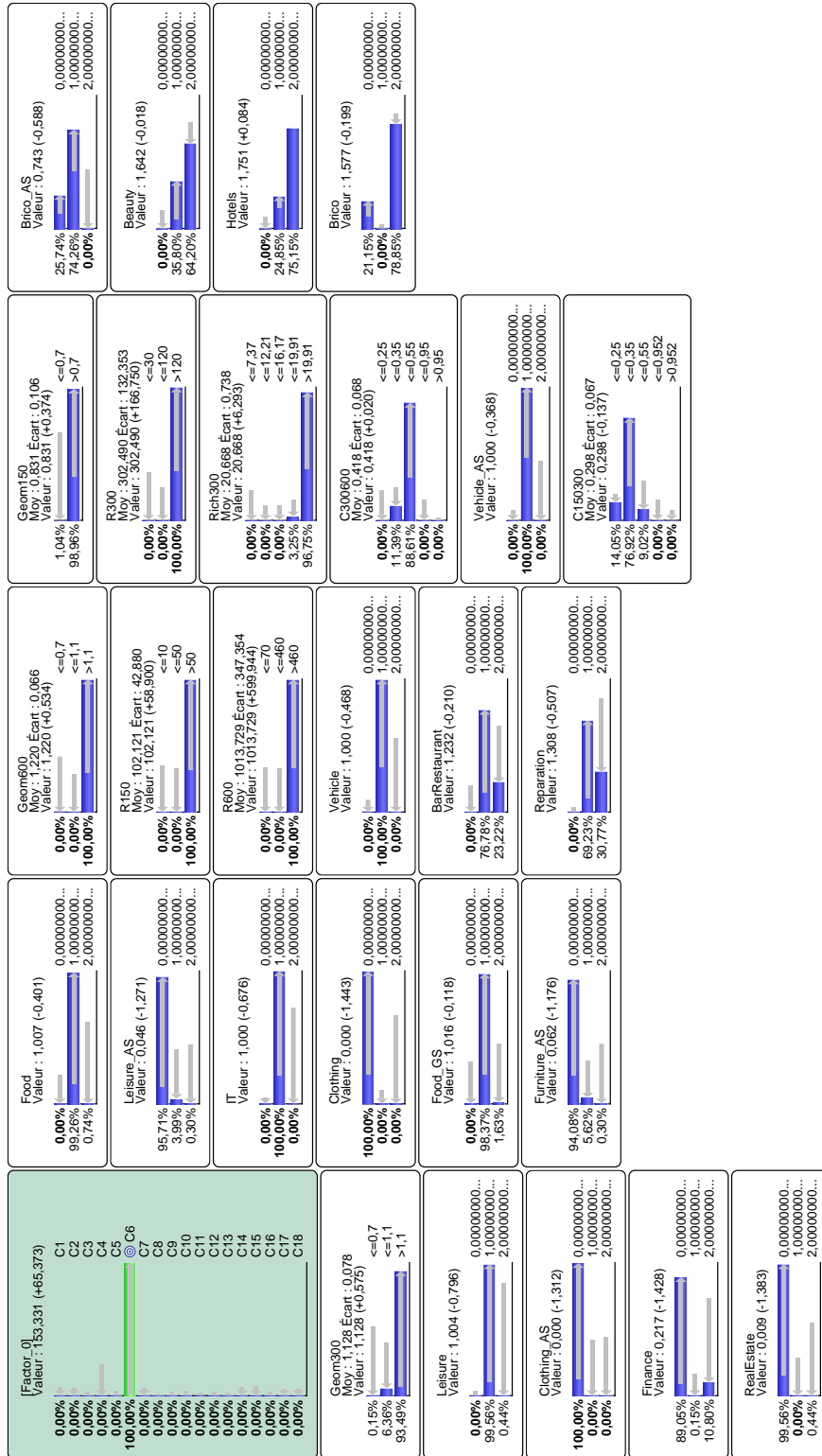
# RF1b



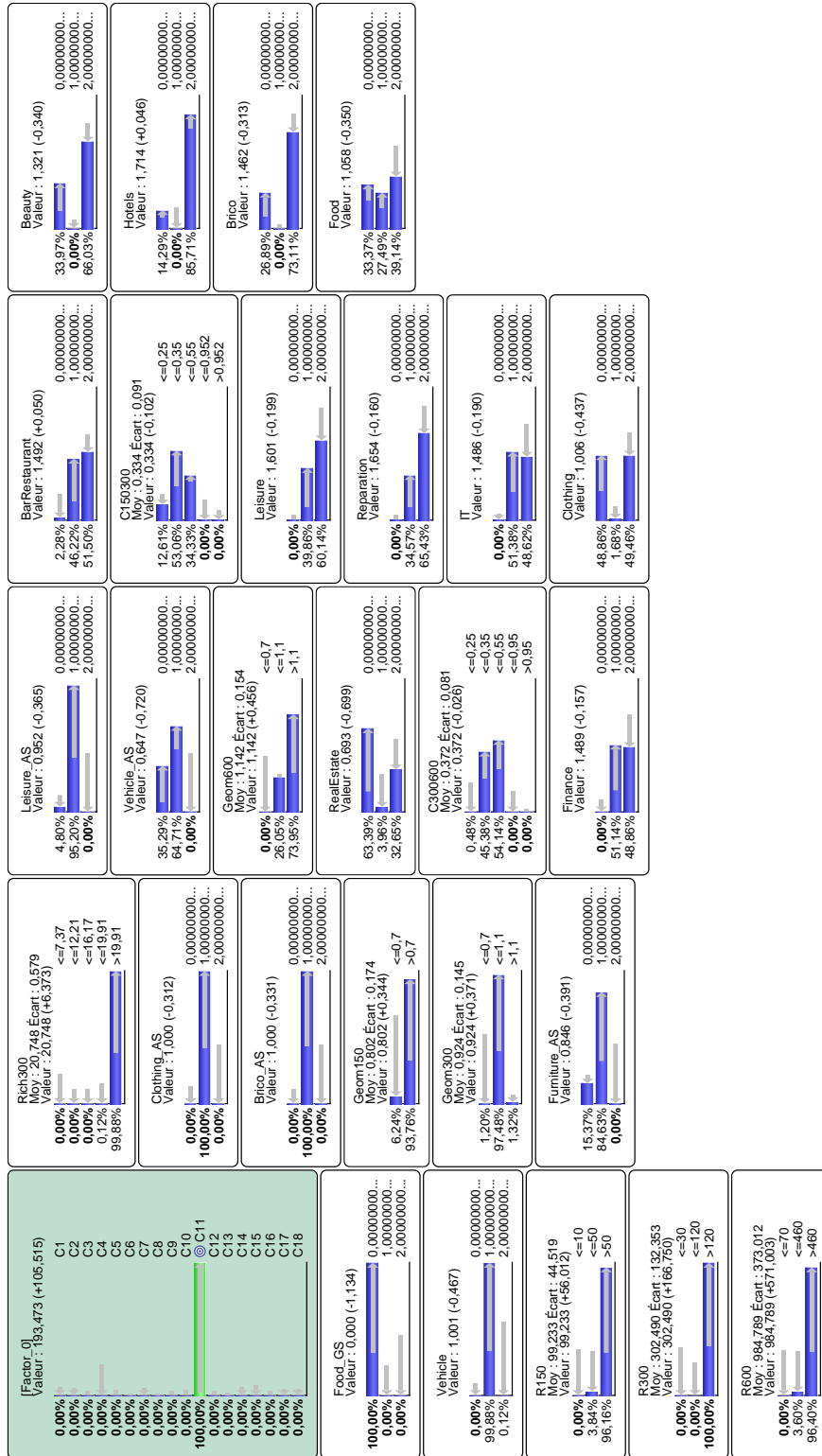
# RF2a



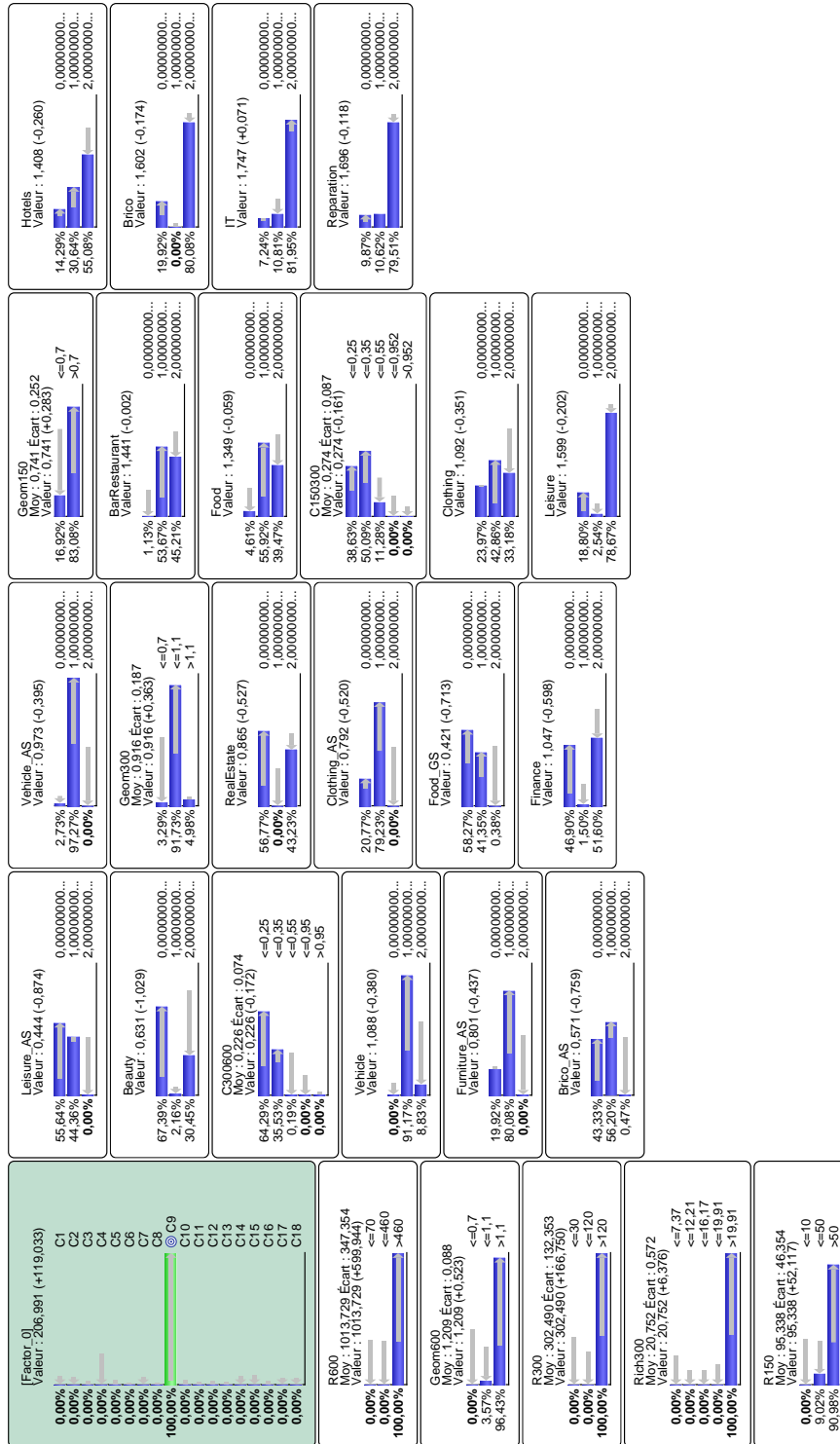
# RF2b



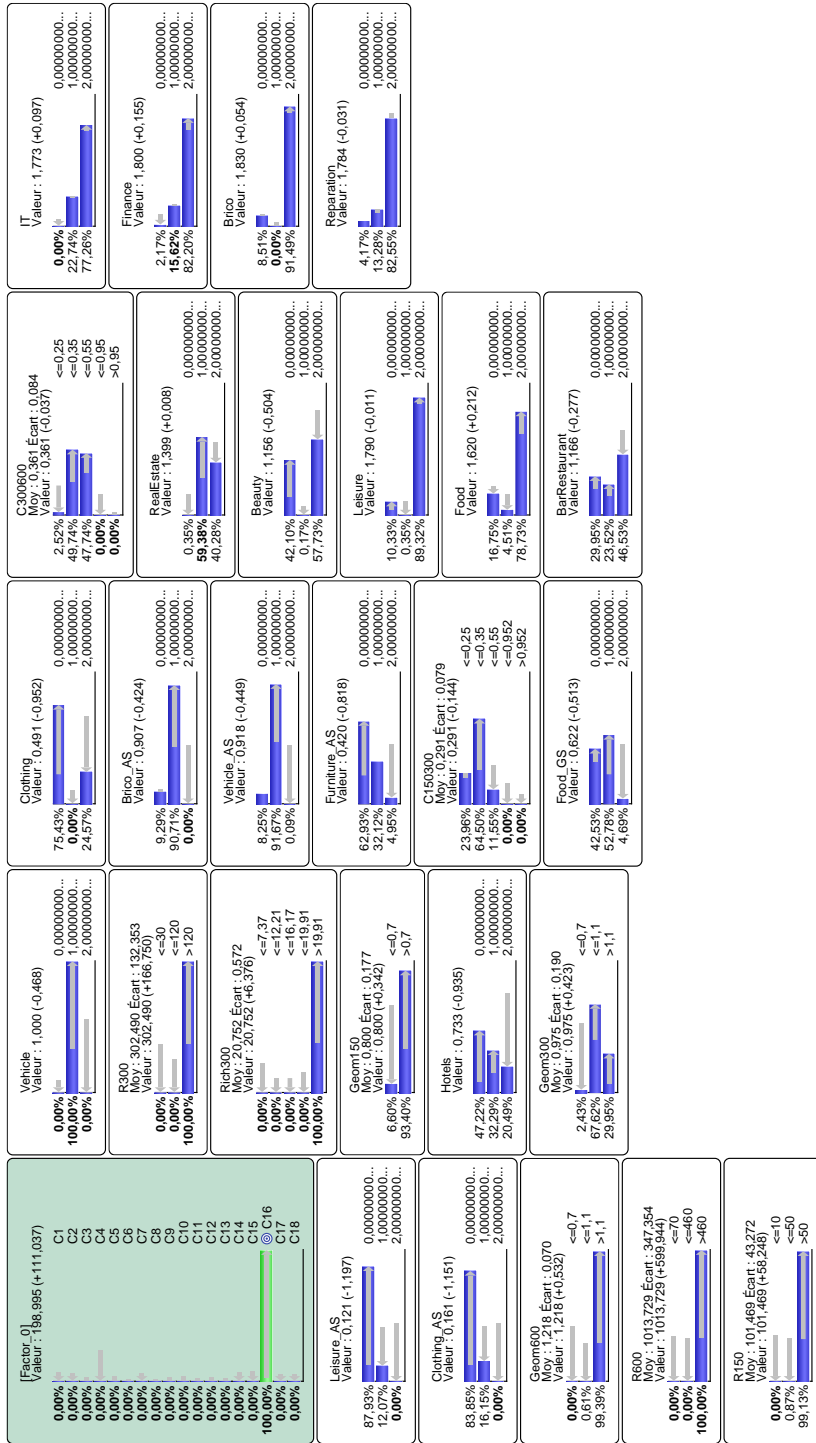
# RF2c



# RF3a

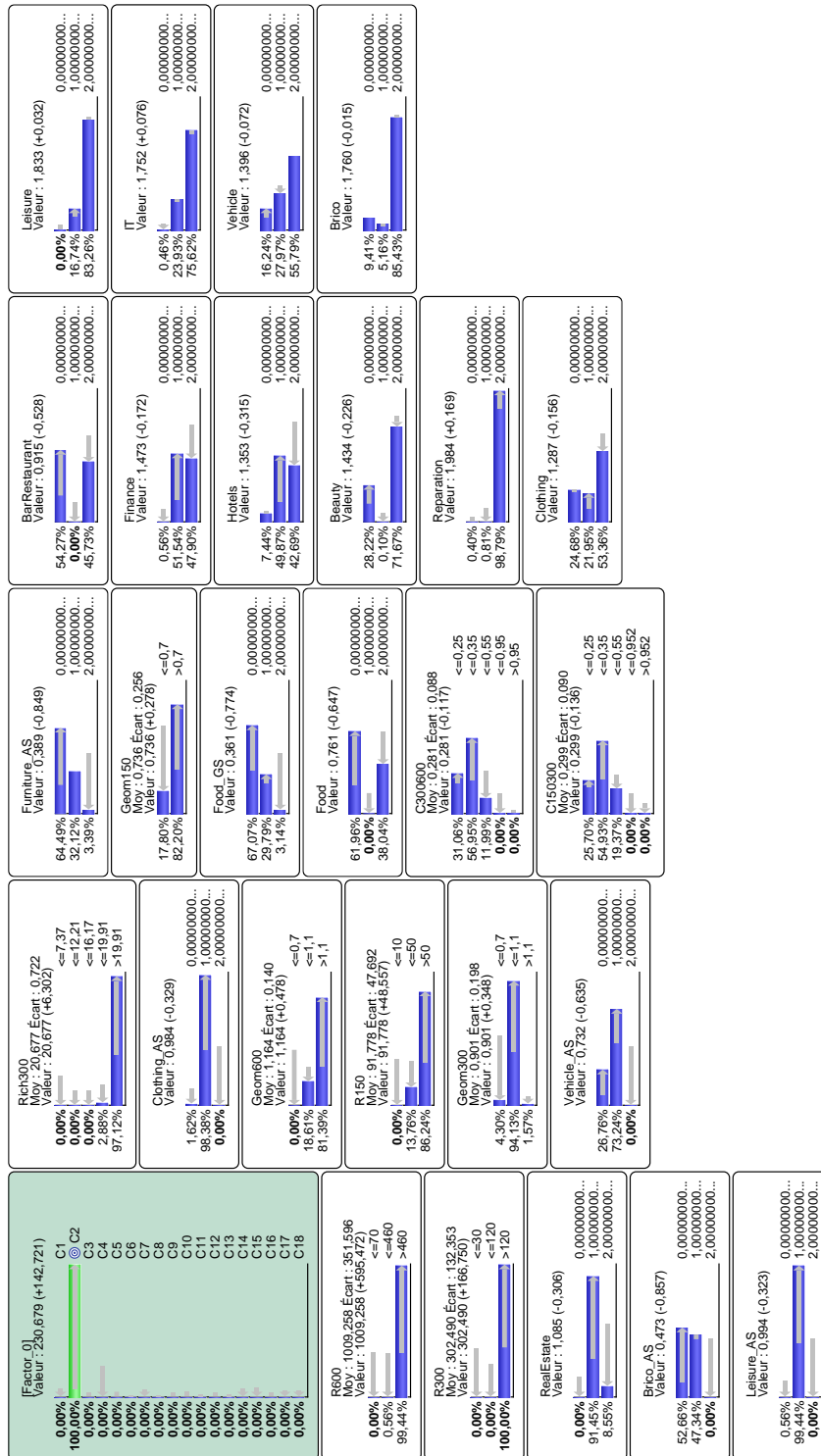


# RF3b

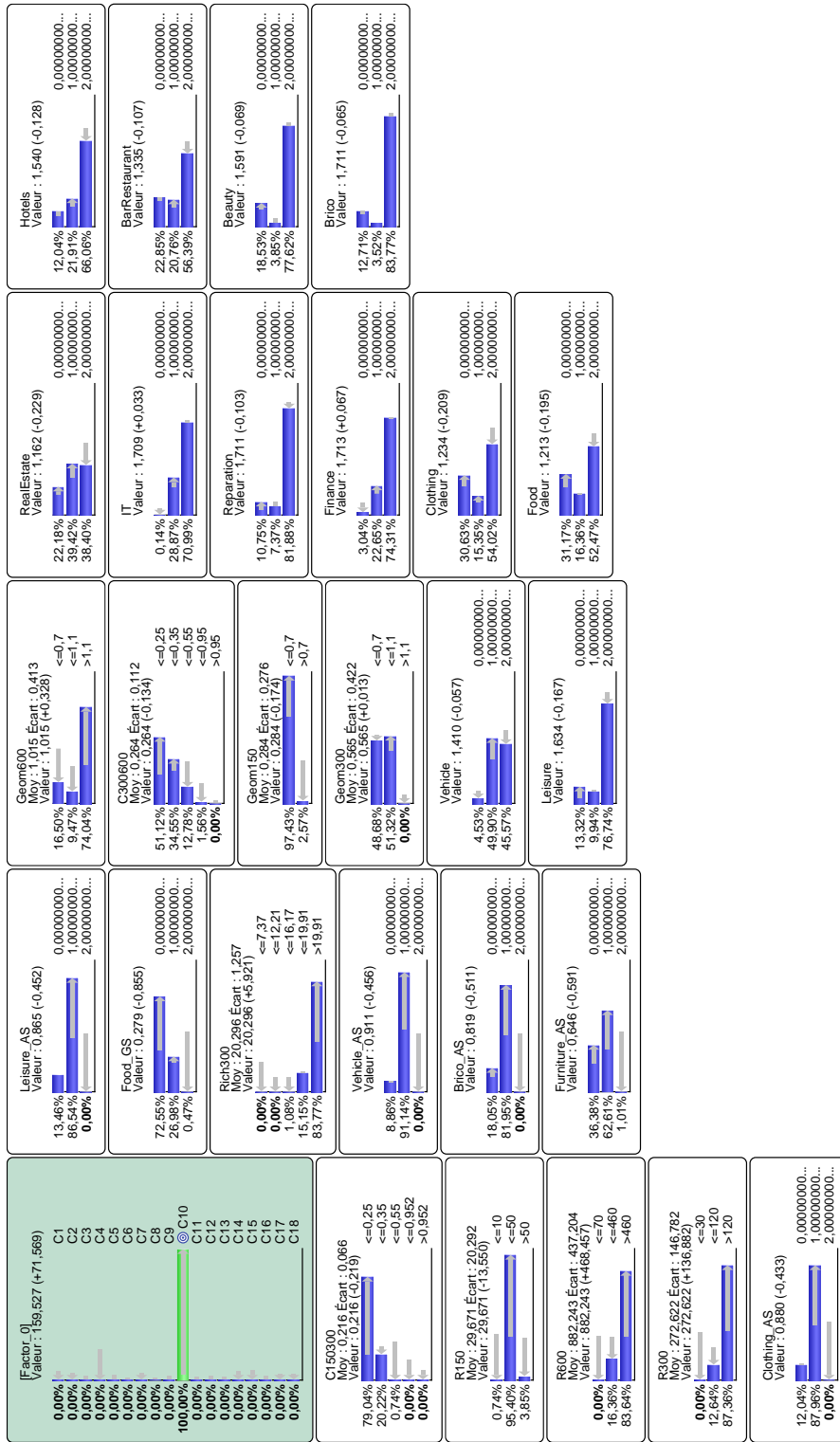




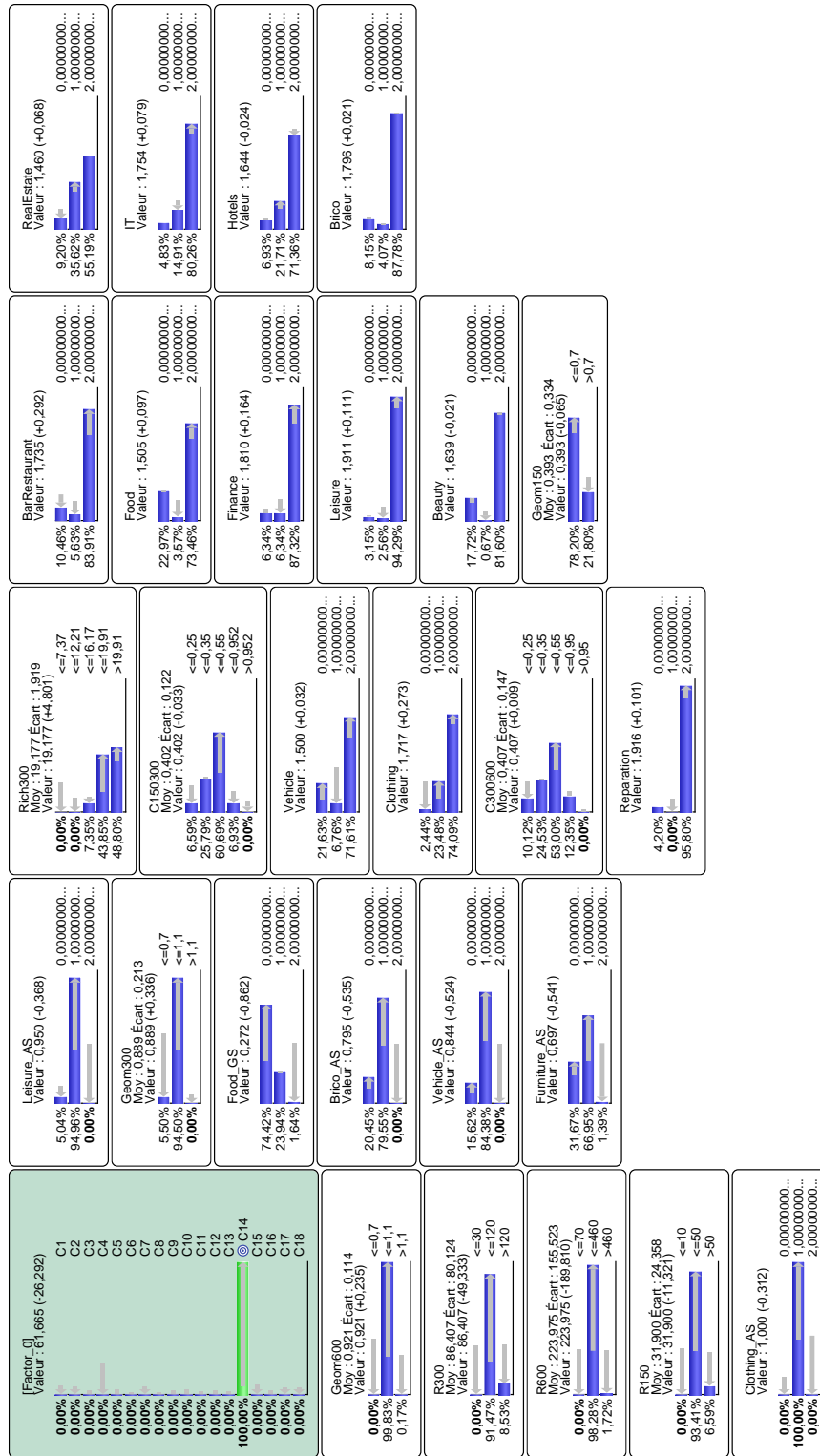
# RF4a



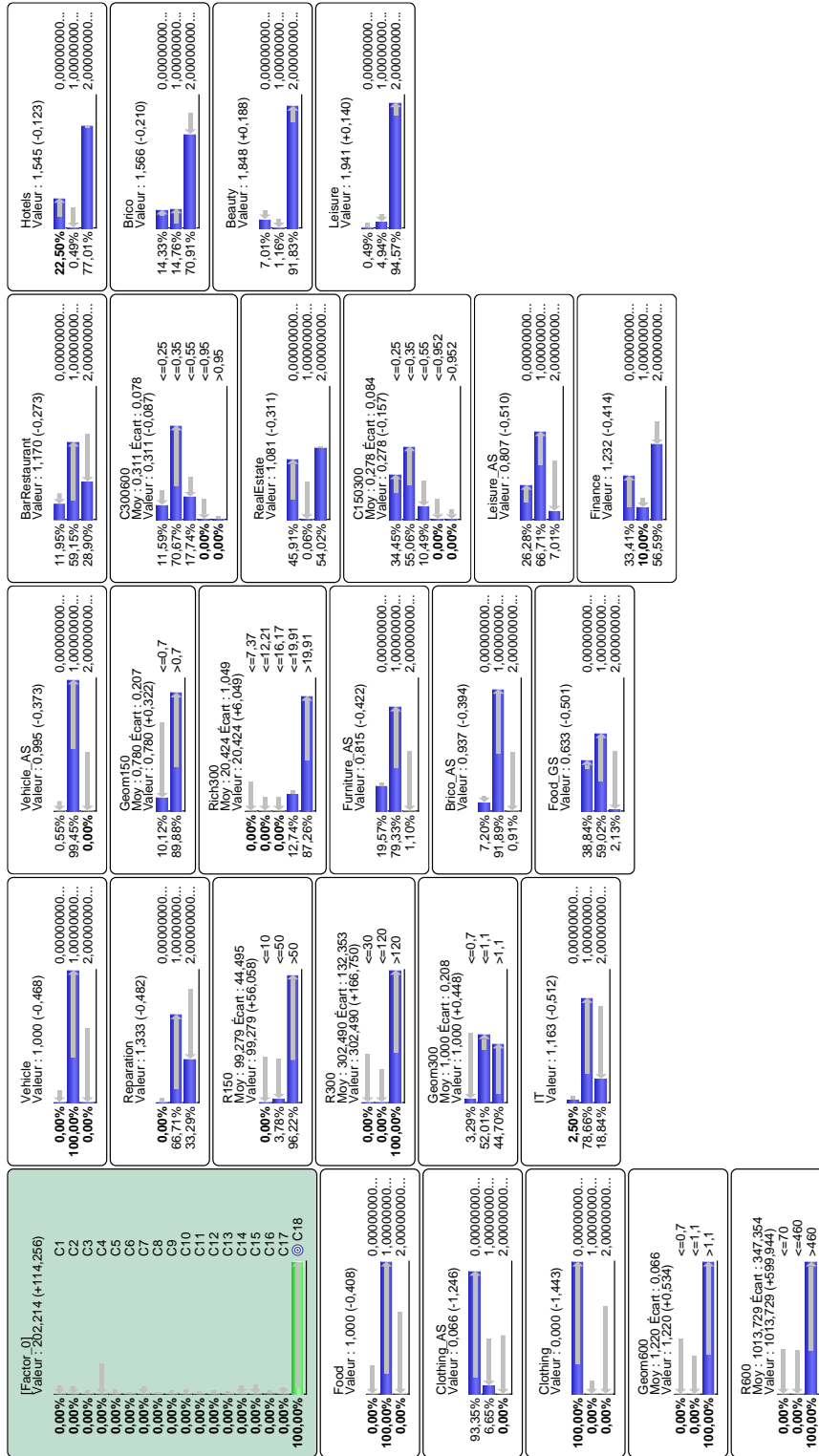
# RF4b



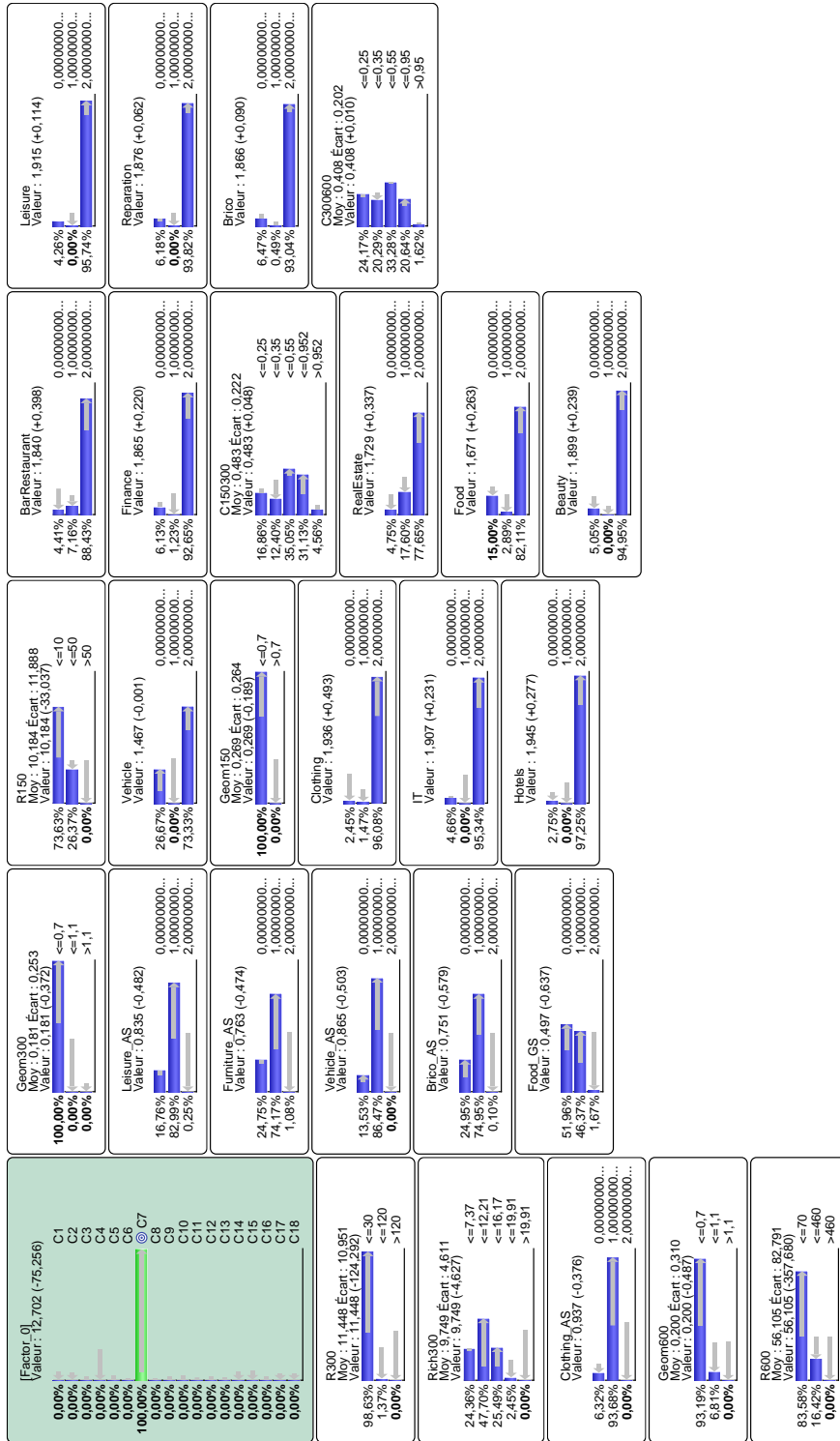
# RF5a



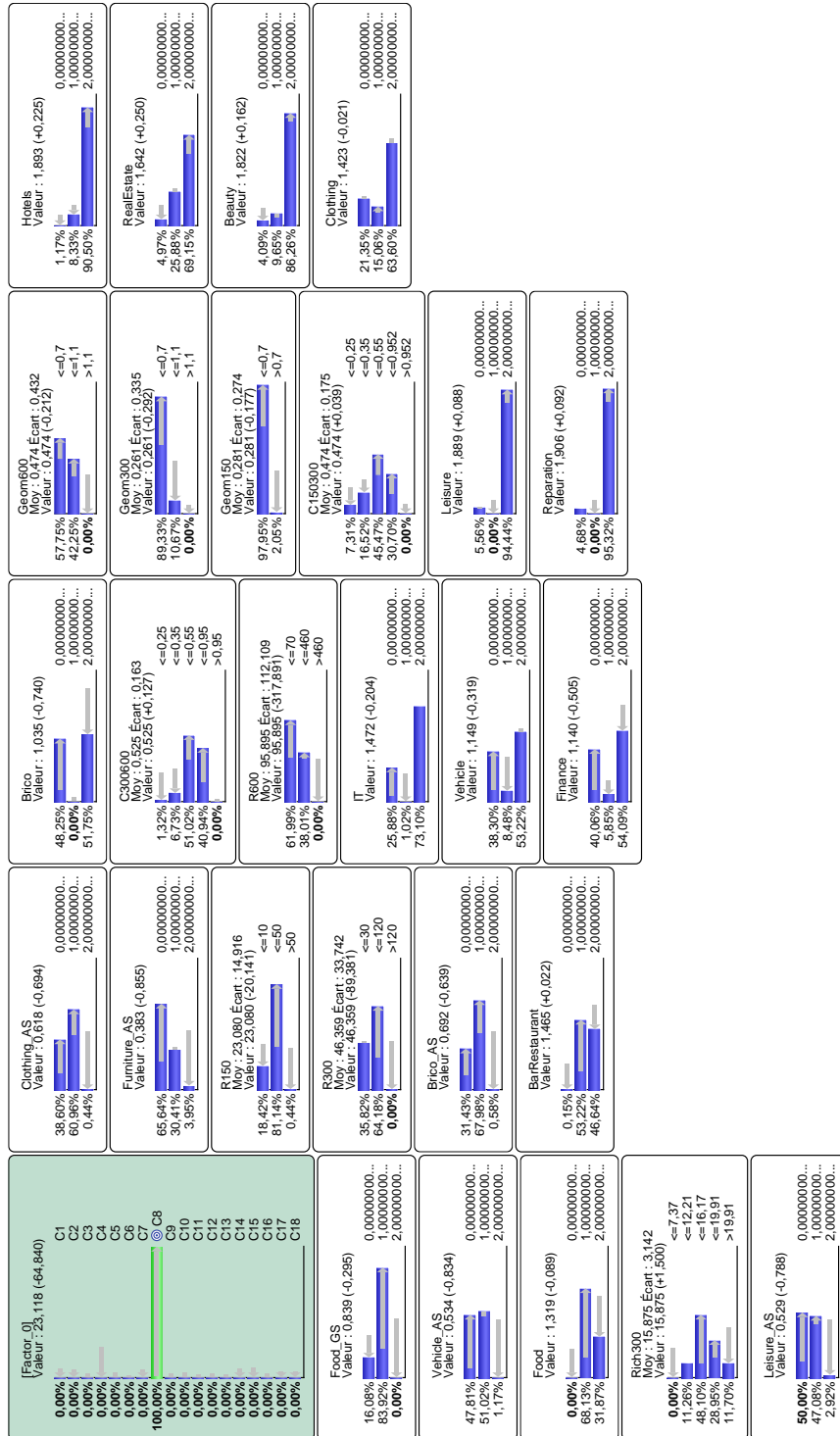
# RF5b



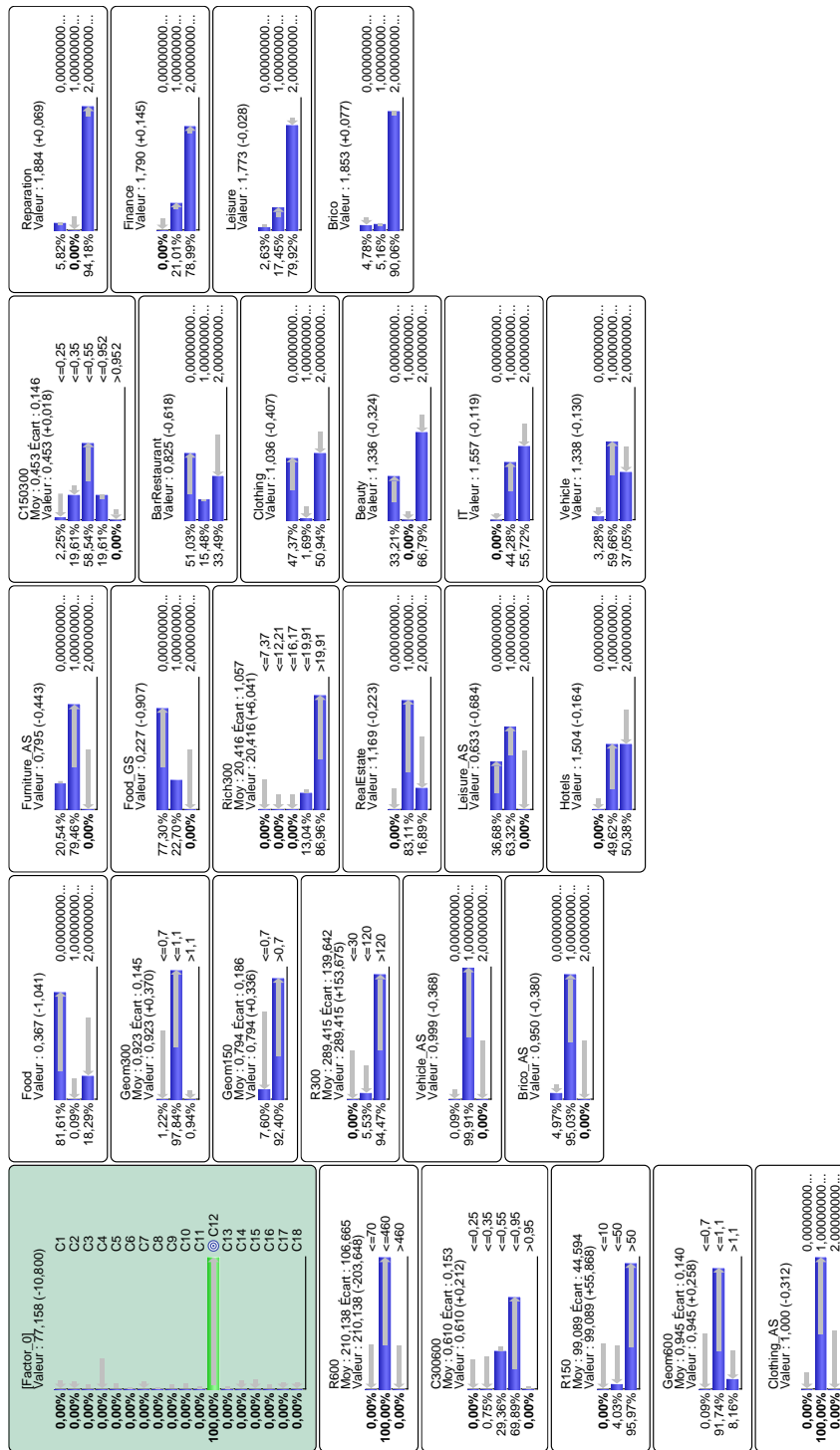
# RF6a



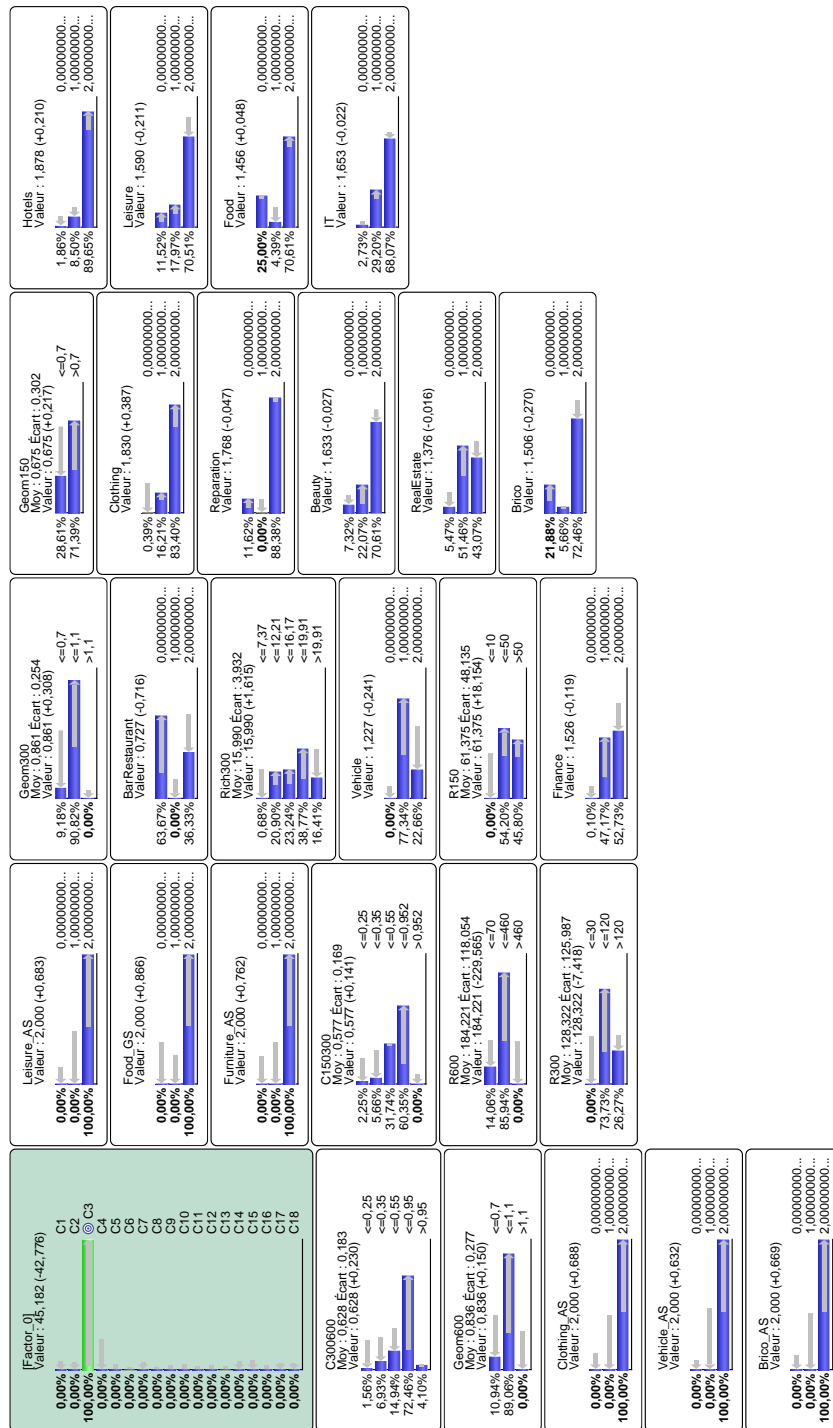
# RF6b



# RF7

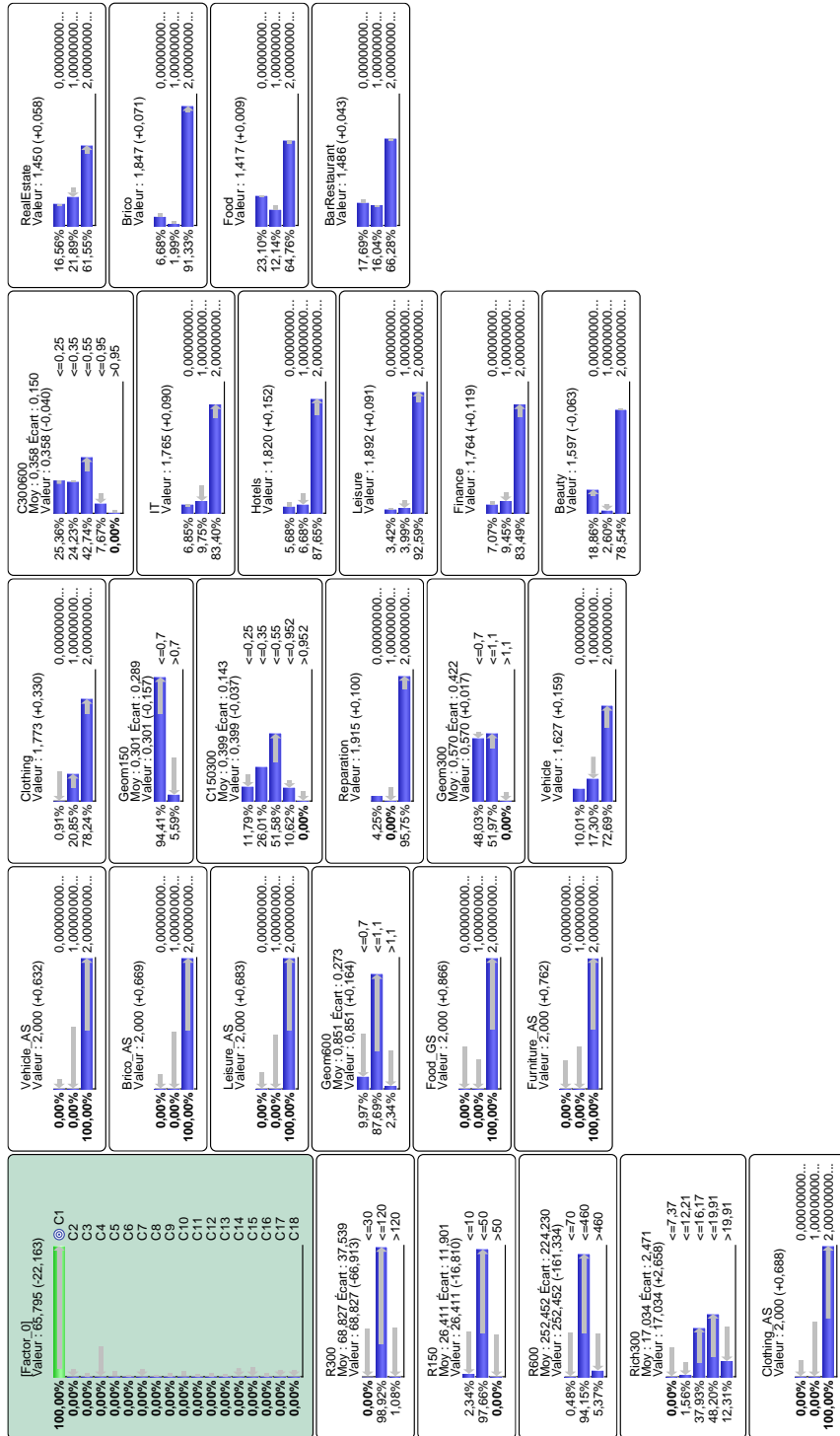


# RF8

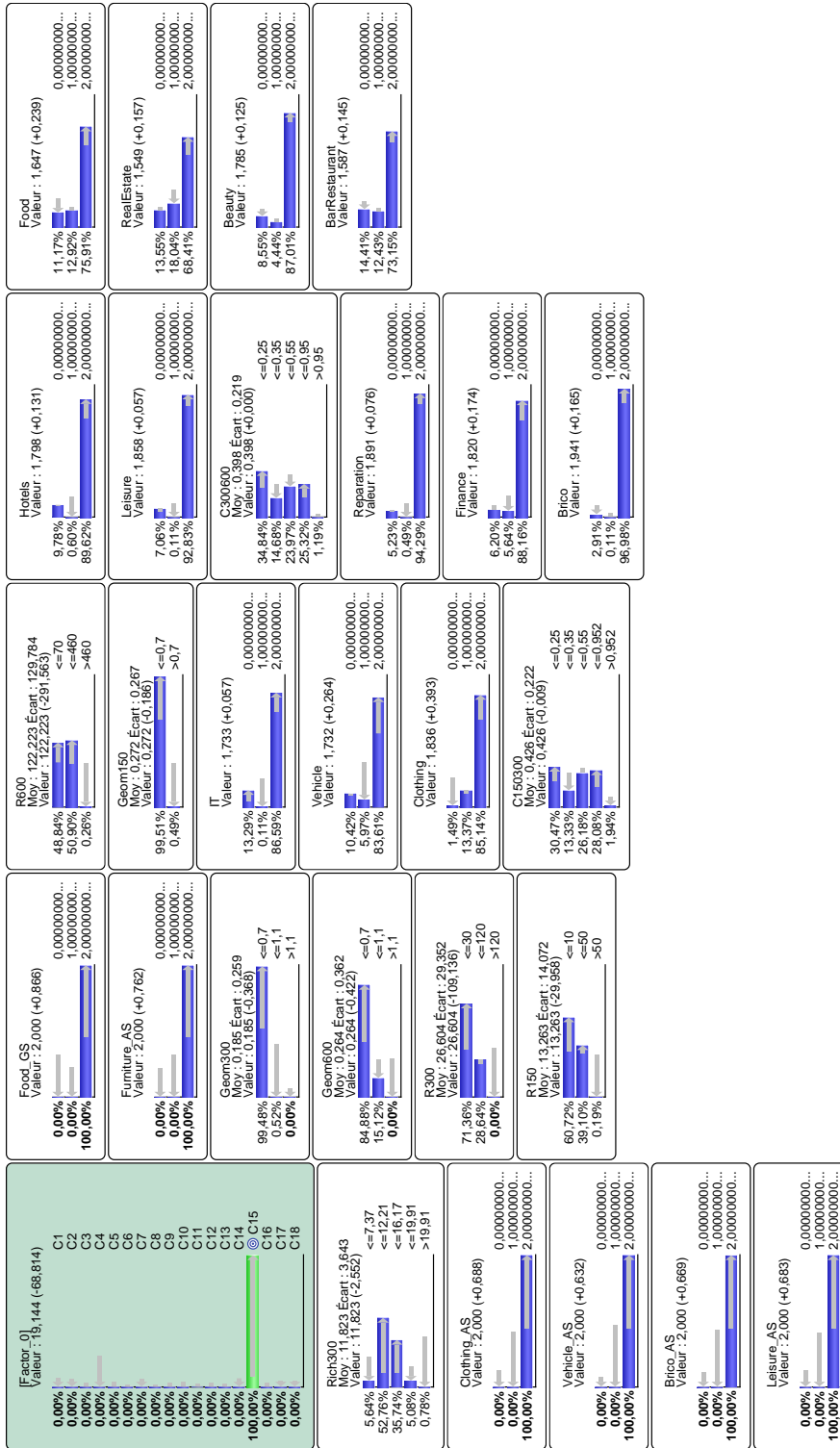




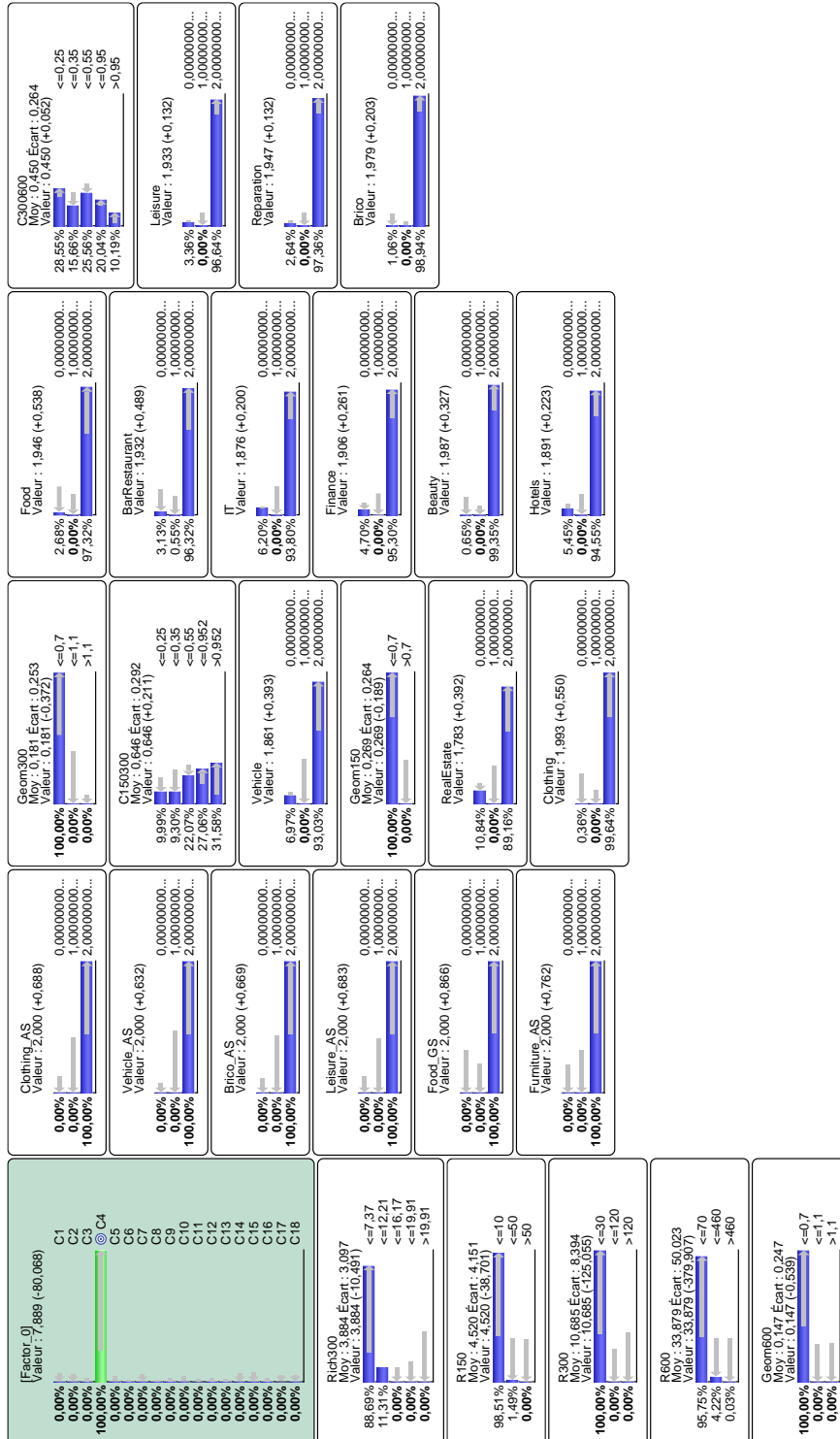
# RF9a



# RF9b

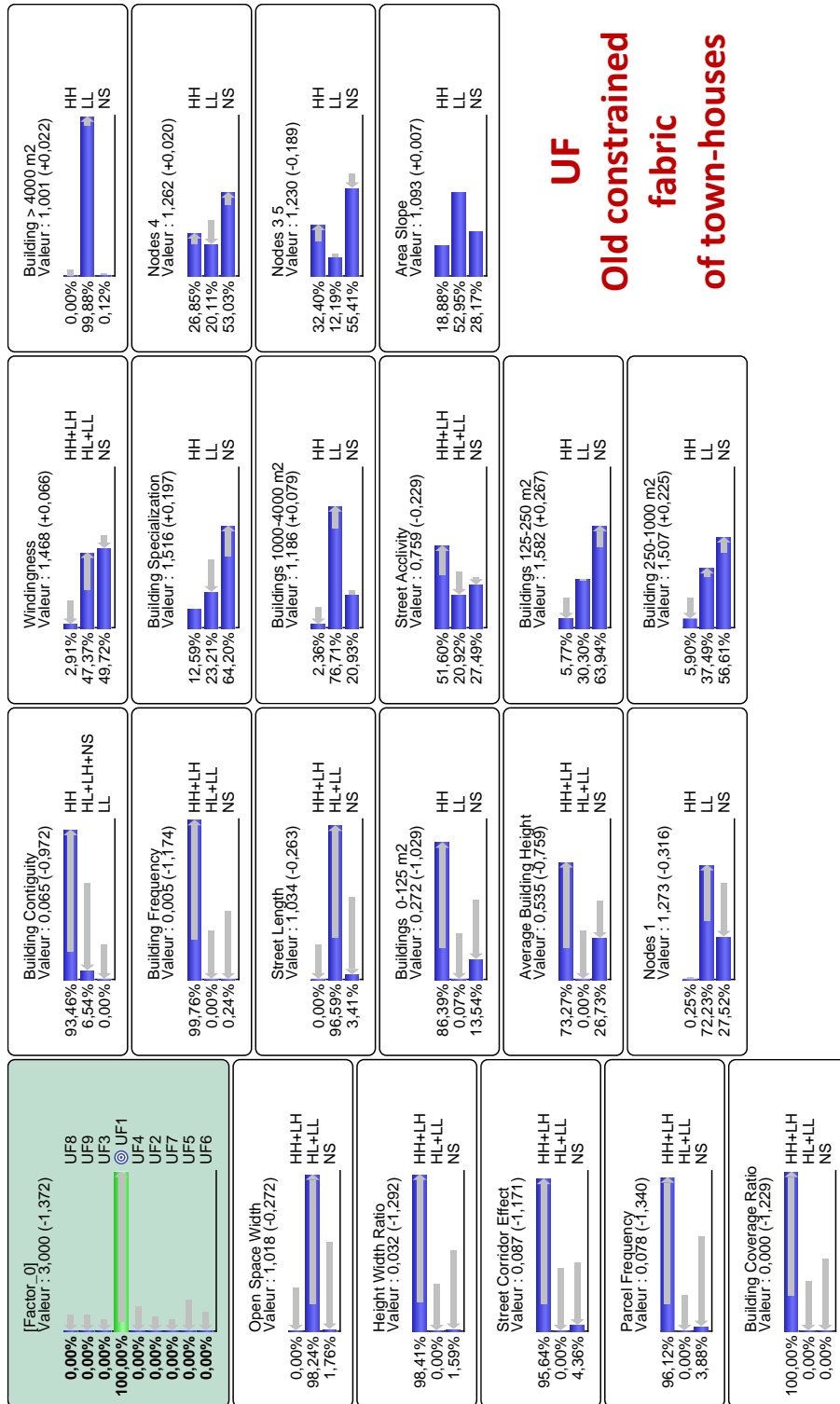


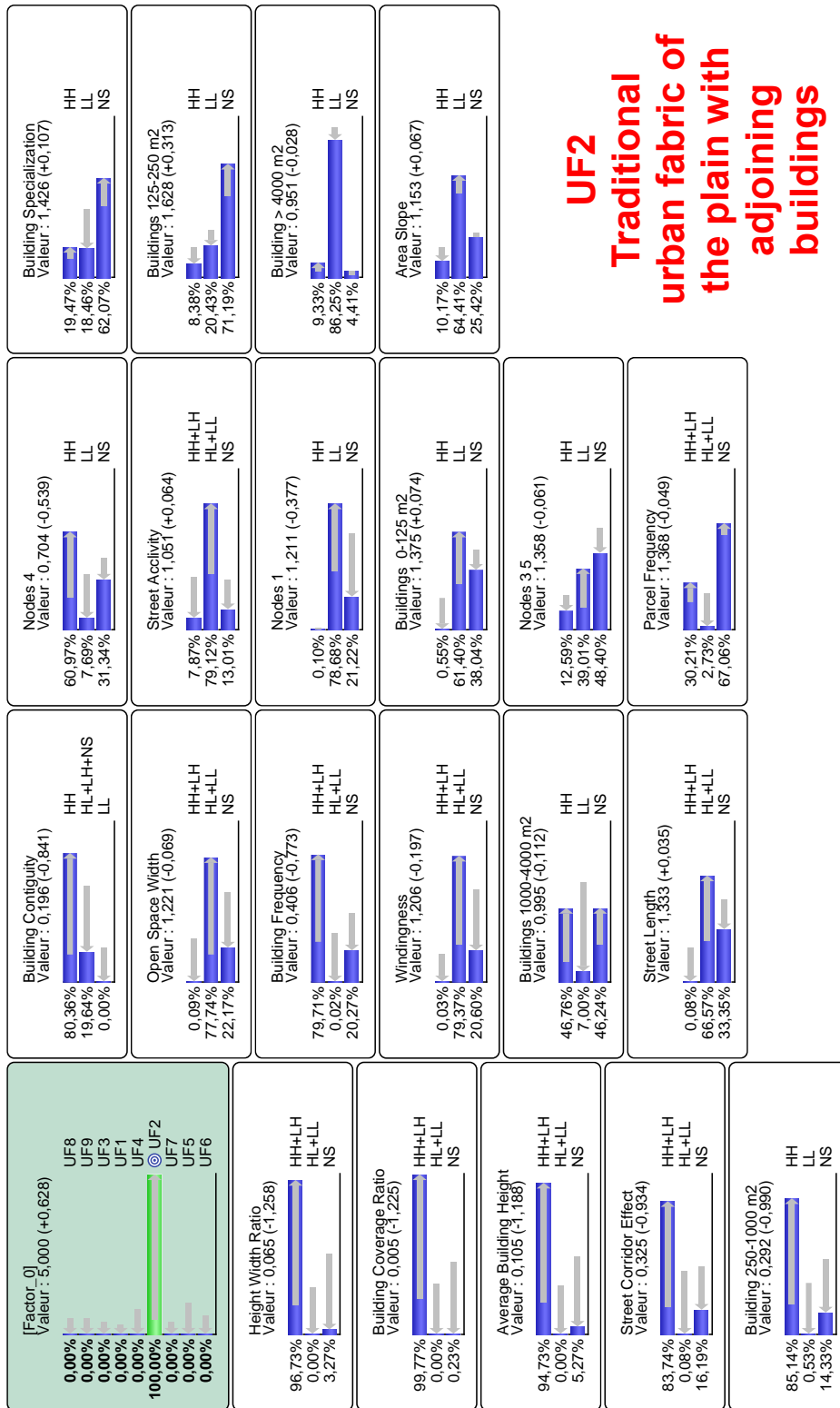
# RF10



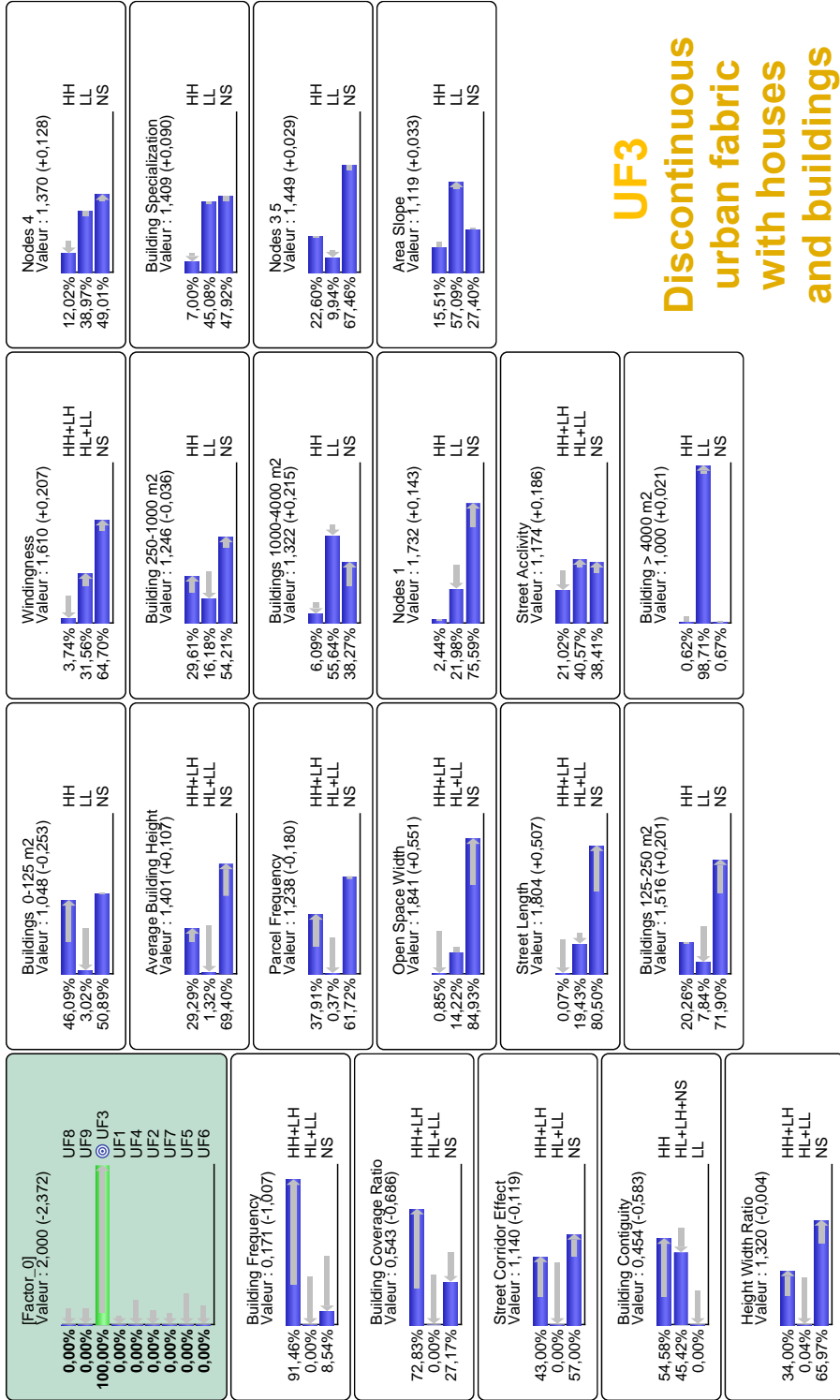
# APPENDIX 2

## The probabilistic profiles of MFA classes. (Chapter 9)

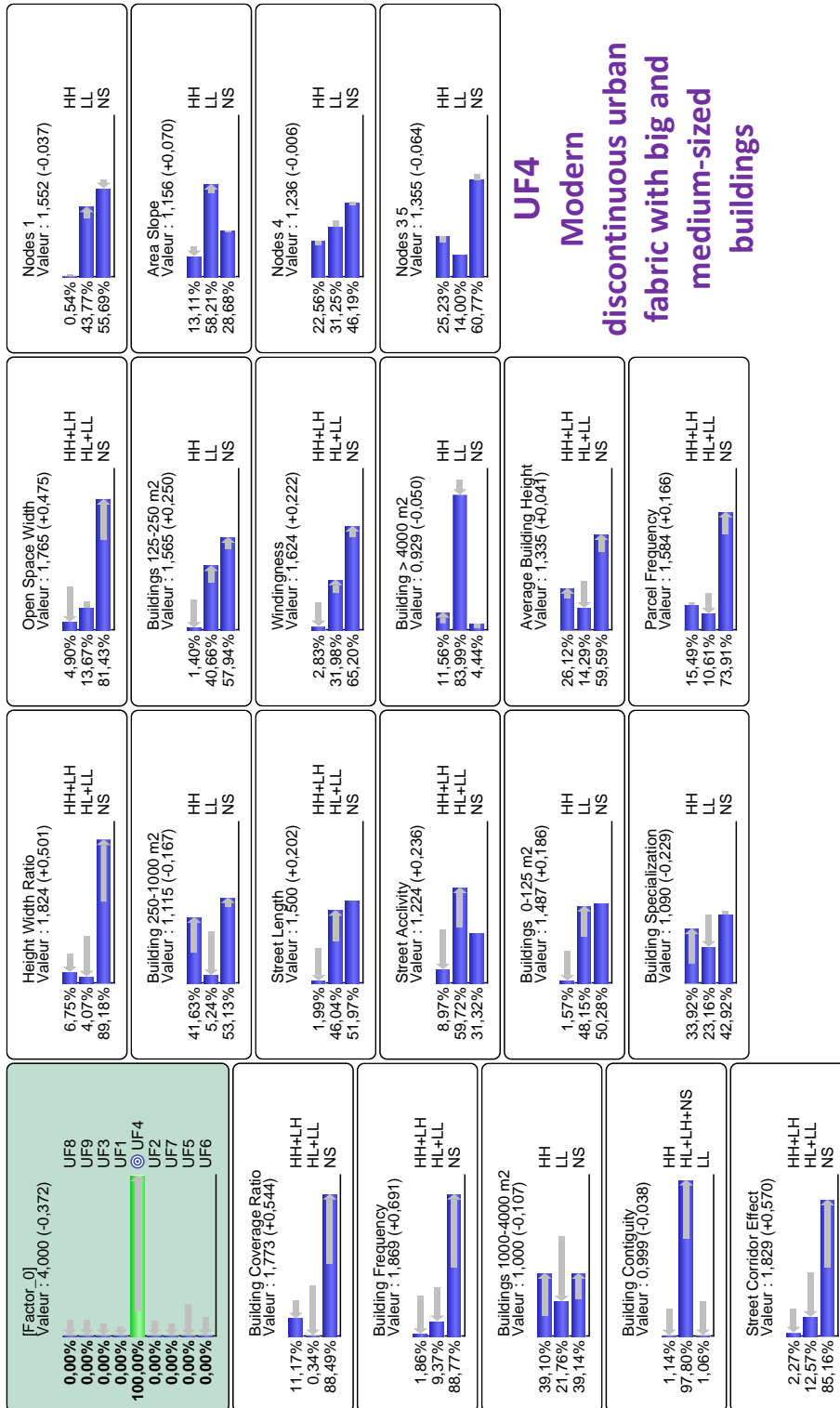


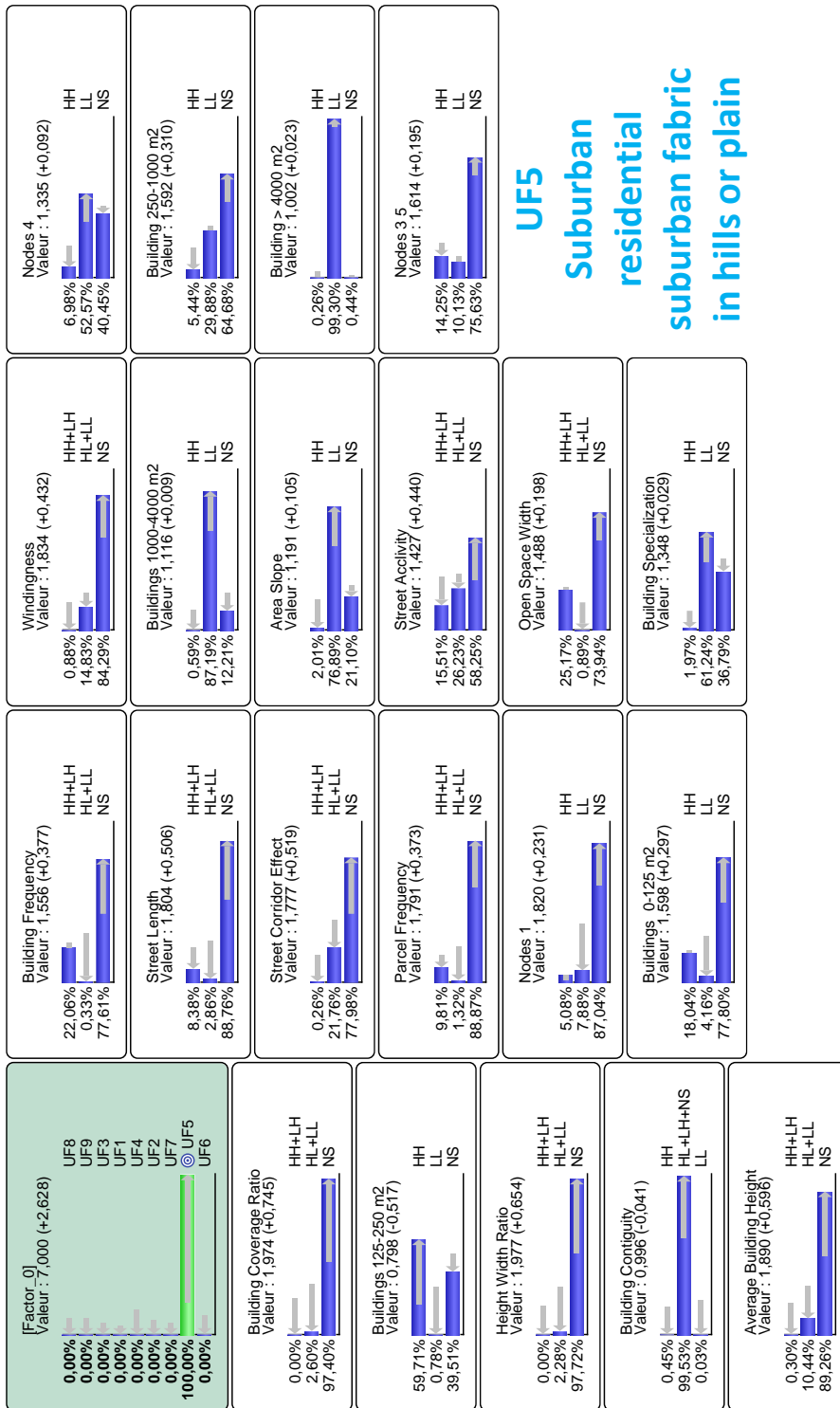


**UF2**  
**Traditional urban fabric of the plain with adjoining buildings**



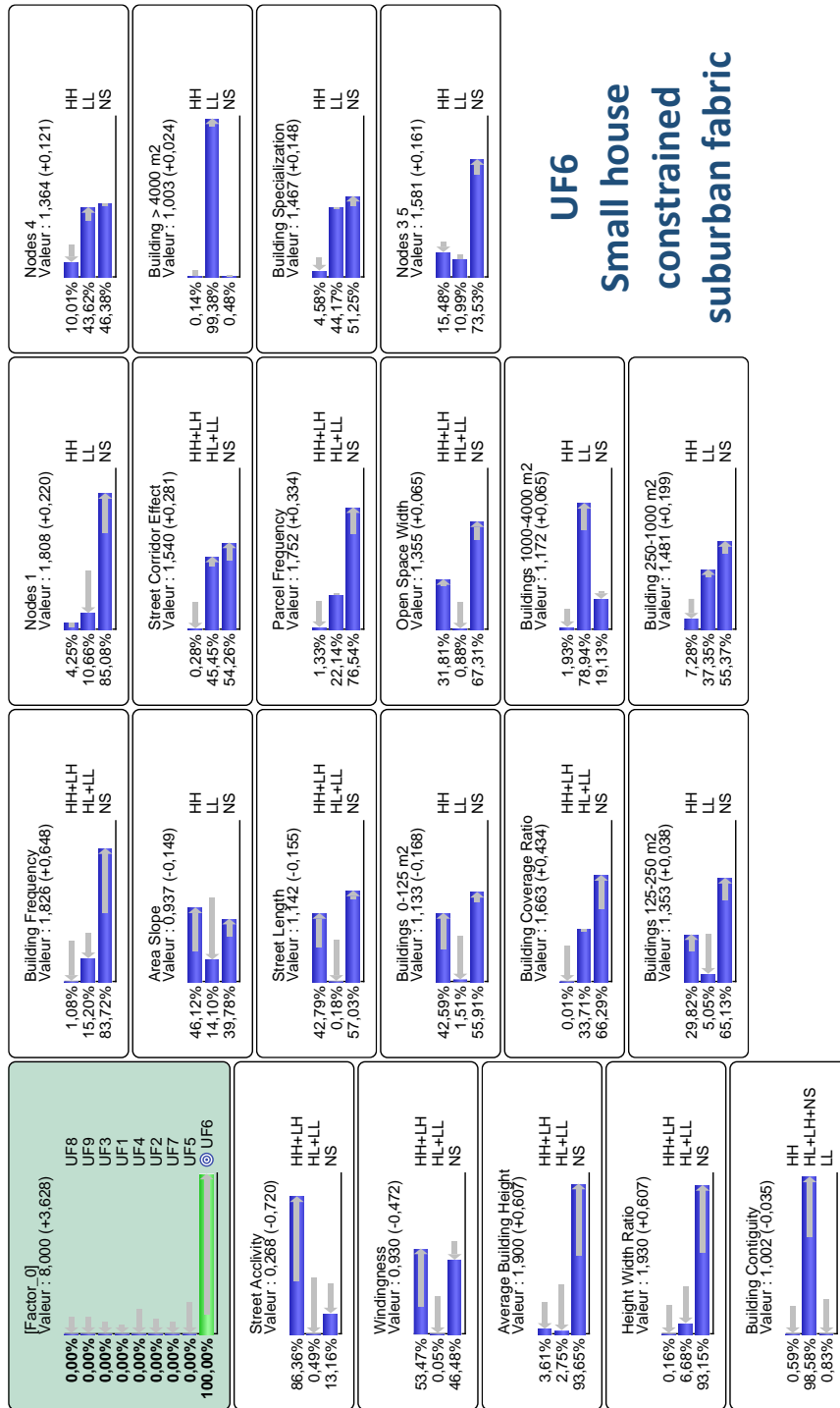
## UF3 Discontinuous urban fabric with houses and buildings

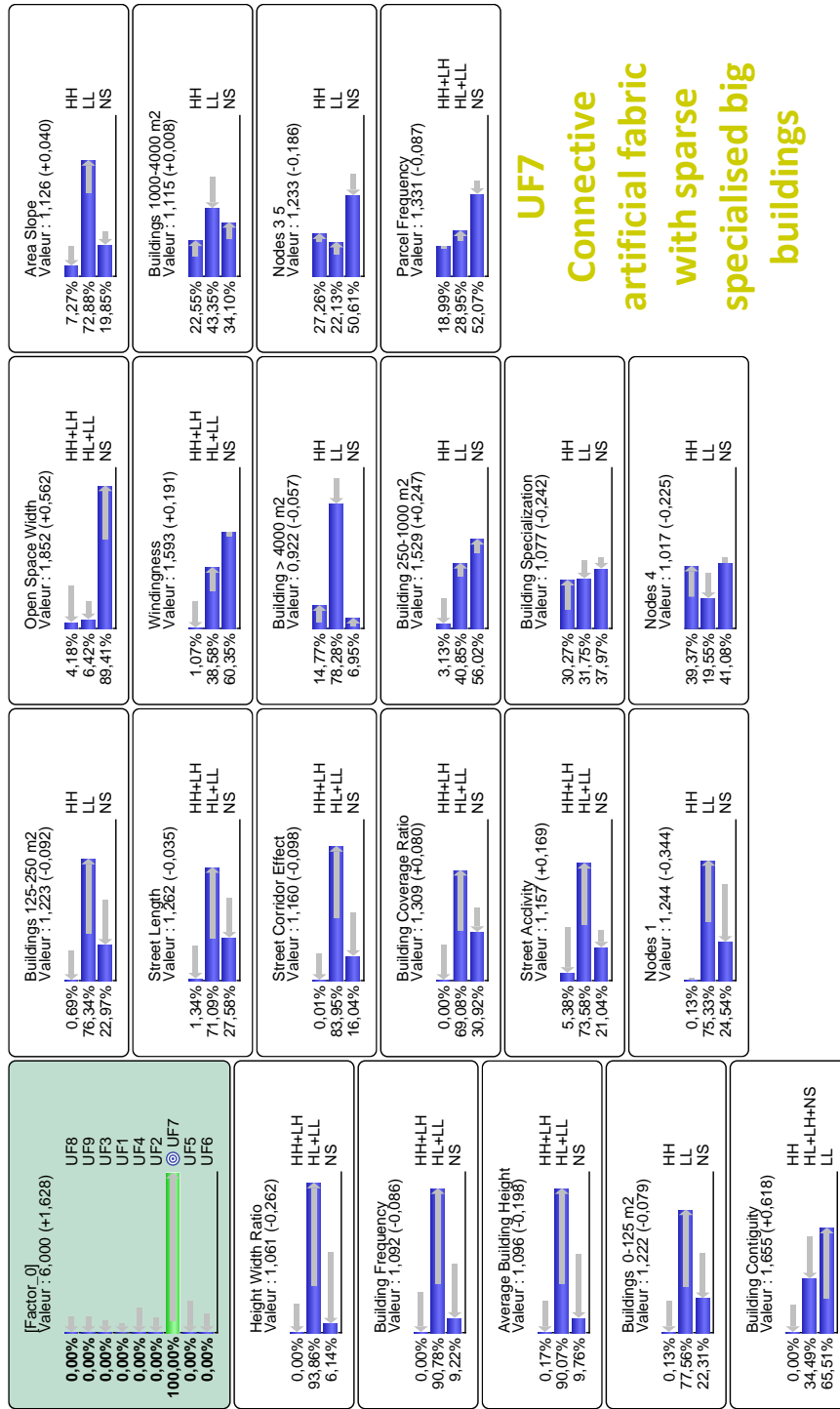




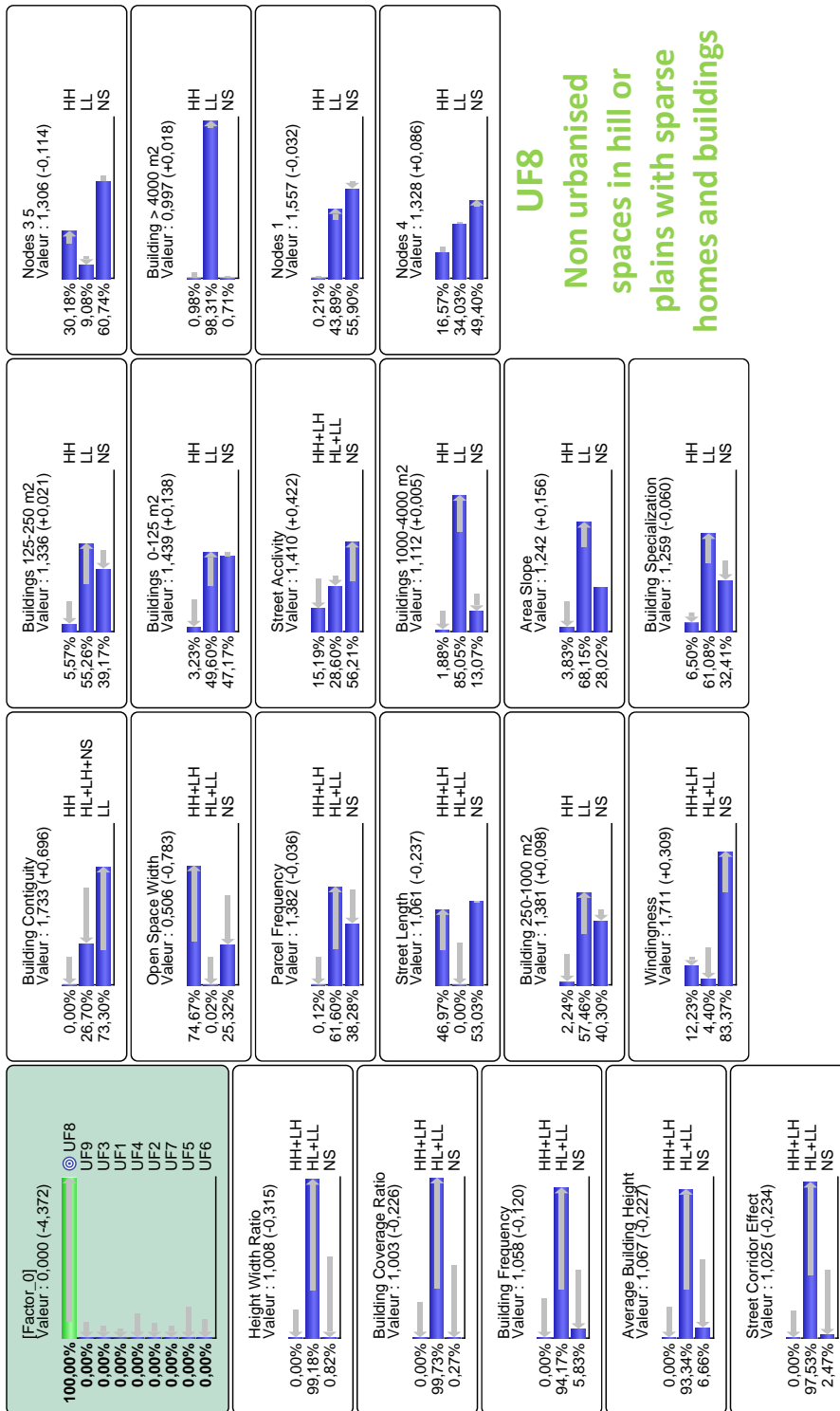
**UF5**  
**Suburban residential**  
**in hills or plain**

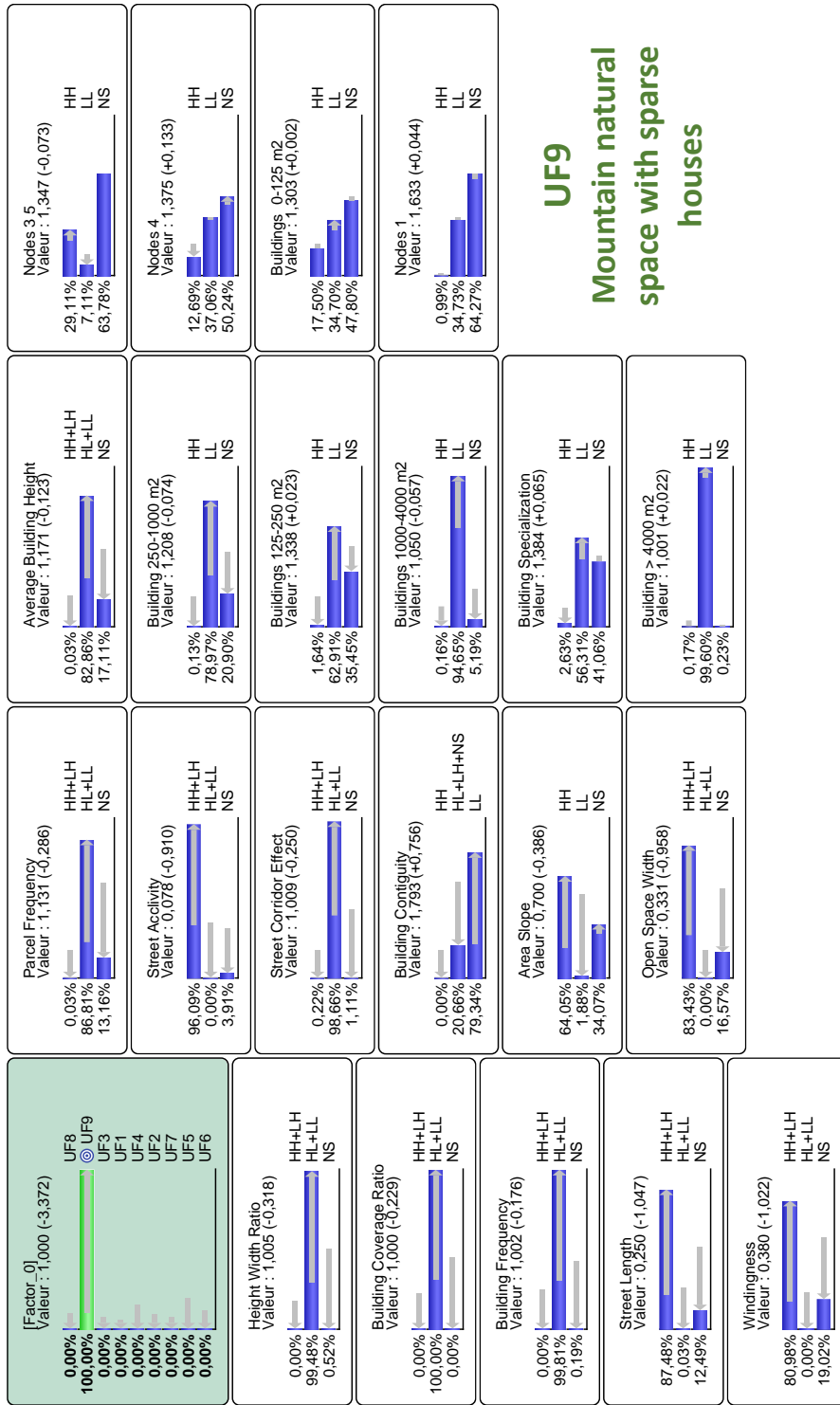






**UF7**  
Connective  
artificial fabric  
with sparse  
specialised big  
buildings





# APPENDIX 3

## Outcomes of the model and variable selection procedures. (Part 4)

		C	Sc	Sz	Pc	Pz	F1	E(T)	E(Tz)	E(Tc)	Sd(T)	Sd(Tz)	Sd(Tc)
Global	G	0,5760	0,6545	0,2337	0,9196	0,1035	0,4606	0,8264	0,4793	2,3415	2,1421	0,8343	4,3308
	P	0,7109	0,8330	0,1780	0,9049	0,1322	0,5267	0,6190	0,2257	2,3357	1,9602	0,6490	3,8925
	NB	0,7175	0,8426	0,1714	0,9040	0,1322	0,5309	0,6295	0,2206	2,4141	2,0395	0,7225	4,0152
	ZIP	0,7144	0,8371	0,1789	0,9058	0,1344	0,5319	0,6193	0,2278	2,3276	1,9720	0,6841	3,9027
	ZINB	0,7155	0,8387	0,1782	0,9068	0,1342	0,5361	0,6238	0,2283	2,3501	1,9969	0,7073	3,9427
	ZAP	0,7201	0,8450	0,1750	0,9039	0,1363	0,5323	0,6144	0,2190	2,3400	1,9731	0,6860	3,8949
	ZANB	0,7155	0,8389	0,1769	0,9051	0,1341	0,5312	0,6262	0,2268	2,3689	2,0202	0,7020	4,0012
	Global*	G	0,5561	0,6353	0,2347	0,9150	0,1047	0,4692	0,8779	0,5111	2,3667	2,2170	0,8566
P		0,6914	0,8171	0,1814	0,8998	0,1320	0,5316	0,6639	0,2474	2,3545	2,0308	0,6761	3,9306
NB		0,6984	0,8275	0,1744	0,8989	0,1319	0,5358	0,6746	0,2410	2,4350	2,1042	0,7195	4,0545
ZIP		0,6956	0,8222	0,1819	0,9007	0,1343	0,5371	0,6634	0,2487	2,3470	2,0402	0,7014	3,9409
ZINB		0,6971	0,8242	0,1811	0,9019	0,1342	0,5414	0,6681	0,2488	2,3703	2,0666	0,7274	3,9812
ZAP		0,7017	0,8308	0,1779	0,8987	0,1361	0,5374	0,6582	0,2390	2,3597	2,0419	0,7054	3,9327
ZANB		0,6968	0,8242	0,1799	0,9000	0,1341	0,5364	0,6710	0,2478	2,3892	2,0924	0,7274	4,0404
Global(F+S)		G	0,5825	0,6696	0,2288	0,9128	0,1099	0,4805	0,8322	0,4555	2,3610	2,1628	0,8384
	P	0,6965	0,8250	0,1749	0,8987	0,1312	0,5336	0,6529	0,2396	2,3308	1,9832	0,6694	3,8174
	NB	0,7011	0,8308	0,1748	0,8991	0,1336	0,5385	0,6700	0,2374	2,4263	2,1178	0,7238	4,0893
	ZIP	0,7004	0,8295	0,1766	0,8998	0,1341	0,5394	0,6523	0,2403	2,3249	2,0040	0,7134	3,8428
	ZINB	0,7040	0,8338	0,1772	0,9017	0,1357	0,5472	0,6622	0,2392	2,3795	2,0906	0,7511	4,0184
	ZAP	0,7065	0,8380	0,1727	0,8979	0,1360	0,5400	0,6492	0,2348	2,3314	2,0134	0,7755	3,8135
	ZANB	0,7005	0,8301	0,1745	0,8987	0,1332	0,5369	0,6652	0,2449	2,3711	2,0830	0,7788	3,9831
	First(Global)	G	0,3334	0,4203	0,1847	0,8859	0,0973	0,6261	1,8225	1,0775	3,0964	3,2960	1,3243
P		0,4544	0,6081	0,1914	0,8719	0,1262	0,6732	1,4529	0,6178	2,8810	2,9116	1,1386	4,1865
NB		0,4597	0,6201	0,1853	0,8693	0,1243	0,6751	1,5063	0,6165	3,0276	3,0849	1,2295	4,4203
ZIP		0,4524	0,6045	0,1922	0,8748	0,1258	0,6740	1,4633	0,6465	2,8599	2,9230	1,1943	4,1982
ZINB		0,4556	0,6098	0,1920	0,8783	0,1261	0,6783	1,4764	0,6392	2,9080	2,9954	1,2314	4,2982
ZAP		0,4549	0,6069	0,1949	0,8706	0,1284	0,6718	1,4614	0,6401	2,8657	2,9291	1,2137	4,1962
ZANB		0,4479	0,5961	0,1943	0,8734	0,1259	0,6703	1,4969	0,6568	2,9333	3,0486	1,2477	4,3881
First		G	0,3722	0,4920	0,1673	0,8771	0,0955	0,6414	1,7841	1,0394	3,0573	3,1483	1,4312
	P	0,4601	0,6177	0,1906	0,8715	0,1273	0,6759	1,4195	0,5983	2,8237	2,8322	1,0945	4,0701
	NB	0,4701	0,6333	0,1910	0,8741	0,1299	0,6832	1,4710	0,5812	2,9923	3,1016	1,1654	4,4812
	ZIP	0,4741	0,6438	0,1839	0,8731	0,1269	0,6860	1,4225	0,6065	2,8178	2,8506	1,2377	4,0396
	ZINB	0,4836	0,6542	0,1918	0,8816	0,1331	0,6962	1,4344	0,5741	2,9055	2,9977	1,2478	4,2734
	ZAP	0,4773	0,6476	0,1861	0,8694	0,1295	0,6844	1,4189	0,6054	2,8097	2,8404	1,2687	4,0058
	ZANB	0,4695	0,6354	0,1859	0,8731	0,1268	0,6831	1,4501	0,6200	2,8695	2,9675	1,3270	4,2025
	Second(Global)	G	0,6277	0,6861	0,2764	0,9194	0,1094	0,3618	0,5740	0,3770	1,7591	1,6195	0,6313
P		0,7677	0,8665	0,1731	0,9047	0,1381	0,3999	0,4101	0,1598	1,9161	1,5683	0,4673	3,6467
NB		0,7753	0,8766	0,1655	0,9040	0,1399	0,4035	0,4071	0,1521	1,9414	1,5802	0,4912	3,6502
ZIP		0,7739	0,8737	0,1733	0,9051	0,1434	0,4065	0,4061	0,1545	1,9199	1,5752	0,4738	3,6589
ZINB		0,7748	0,8750	0,1719	0,9059	0,1426	0,4111	0,4081	0,1564	1,9225	1,5751	0,5009	3,6366
ZAP		0,7811	0,8838	0,1636	0,9034	0,1446	0,4052	0,3998	0,1441	1,9382	1,5737	0,4688	3,6461
ZANB		0,7769	0,8782	0,1679	0,9044	0,1430	0,4060	0,4054	0,1510	1,9360	1,5834	0,4868	3,6654
Second		G	0,6501	0,7116	0,2800	0,9189	0,1188	0,3707	0,5260	0,3173	1,7811	1,6152	0,5336
	P	0,7726	0,8741	0,1618	0,9034	0,1354	0,3987	0,4064	0,1547	1,9202	1,5390	0,4827	3,5418
	NB	0,7754	0,8775	0,1612	0,9035	0,1373	0,4016	0,4124	0,1560	1,9548	1,6003	0,5402	3,6660
	ZIP	0,7734	0,8740	0,1685	0,9044	0,1402	0,4032	0,4051	0,1537	1,9177	1,5641	0,4786	3,6223
	ZINB	0,7748	0,8761	0,1655	0,9053	0,1386	0,4088	0,4140	0,1597	1,9438	1,6244	0,5353	3,7498
	ZAP	0,7803	0,8831	0,1620	0,9030	0,1430	0,4026	0,4017	0,1471	1,9331	1,5847	0,5681	3,5978
	ZANB	0,7748	0,8762	0,1650	0,9032	0,1398	0,3992	0,4127	0,1562	1,9561	1,6250	0,5405	3,7407

Set of thirteen descriptors for the models implemented on the overall study area and for the First/ Second-Age morphological regions; in light grey, the ZINB models are highlighted.

(G-Gaussian; P- Poisson; NB-Negative Binomial; ZIP-Zero Inflated Poisson; ZINB-Zero Inflated Negative Binomial; ZAP-Zero Alternated Poisson; ZANB-Zero Alternated Negative Binomial; C-Accuracy; Sc-Sensitivity count part; Sz- Sensitivity zero part; Pc-Precision count part; Pz-Precision zero part; F1 score; E(T)-Average Tolerance; E(Tz)-Average Tolerance Zero Part; E(Tc)-Average Tolerance count part; Sd(T)-Standard Deviation Tolerance, Sd(Tz)-Standard Deviation Tolerance Zero Part; Sd(Tc)-Standard Deviation Tolerance Count Part)

		C	Sc	Sz	Pc	Pz	F1	E(T)	E(Tz)	E(Tc)	Sd(T)	Sd(Tz)	Sd(Tc)
Global*	G	0,562	0,642	0,234	0,916	0,106	0,474	0,863	0,501	2,332	2,152	0,852	4,217
	P	0,699	0,826	0,181	0,900	0,135	0,539	0,648	0,236	2,318	1,952	0,667	3,739
	NB	0,705	0,836	0,173	0,899	0,135	0,543	0,658	0,229	2,399	2,026	0,704	3,869
	ZIP	0,701	0,829	0,181	0,901	0,137	0,543	0,648	0,239	2,309	1,960	0,689	3,749
	ZINB	0,703	0,831	0,180	0,902	0,137	0,547	0,652	0,237	2,332	1,985	0,706	3,792
	ZAP	0,707	0,837	0,178	0,899	0,139	0,543	0,643	0,229	2,321	1,961	0,691	3,741
	ZANB	0,702	0,831	0,180	0,900	0,137	0,543	0,655	0,237	2,350	2,012	0,708	3,855
	G	0,632	0,735	0,212	0,909	0,119	0,505	0,769	0,381	2,338	2,059	0,836	3,937
MFA(6)	P	0,707	0,837	0,177	0,901	0,138	0,548	0,626	0,223	2,259	1,863	0,648	3,536
	NB	0,711	0,843	0,178	0,902	0,141	0,555	0,652	0,222	2,394	2,083	0,740	3,987
	ZIP	0,713	0,845	0,182	0,903	0,144	0,560	0,624	0,224	2,246	1,887	0,756	3,520
	ZINB	0,714	0,845	0,186	0,907	0,146	0,569	0,634	0,224	2,293	1,967	0,725	3,739
	ZAP	0,715	0,848	0,176	0,901	0,142	0,556	0,626	0,225	2,253	1,904	0,831	3,494
	ZANB	0,711	0,842	0,180	0,902	0,142	0,556	0,643	0,235	2,293	2,004	0,885	3,703
	G	0,481	0,561	0,241	0,889	0,115	0,509	0,989	0,695	1,867	1,650	0,972	2,647
	UF1(Global)	P	0,619	0,751	0,224	0,882	0,155	0,573	0,707	0,325	1,851	1,531	0,688
NB		0,619	0,756	0,211	0,879	0,148	0,570	0,720	0,331	1,883	1,559	0,730	2,510
ZIP		0,617	0,748	0,225	0,879	0,156	0,567	0,711	0,327	1,861	1,515	0,671	2,459
ZINB		0,628	0,756	0,246	0,888	0,171	0,584	0,685	0,301	1,833	1,512	0,615	2,496
ZAP		0,624	0,761	0,214	0,875	0,154	0,565	0,701	0,304	1,891	1,523	0,639	2,480
ZANB		0,618	0,749	0,226	0,879	0,157	0,566	0,711	0,316	1,891	1,551	0,640	2,552
G		0,516	0,587	0,304	0,911	0,147	0,541	0,836	0,532	1,745	1,580	0,759	2,673
UF1		P	0,645	0,780	0,240	0,888	0,176	0,597	0,638	0,277	1,715	1,390	0,624
	NB	0,647	0,785	0,236	0,888	0,175	0,599	0,663	0,273	1,829	1,550	0,634	2,563
	ZIP	0,644	0,772	0,260	0,893	0,185	0,601	0,640	0,292	1,681	1,393	0,649	2,244
	ZINB	0,659	0,793	0,259	0,903	0,190	0,631	0,650	0,261	1,811	1,649	0,621	2,812
	ZAP	0,647	0,780	0,250	0,888	0,183	0,597	0,638	0,283	1,699	1,396	0,651	2,237
	ZANB	0,645	0,775	0,257	0,892	0,184	0,601	0,654	0,284	1,760	1,493	0,634	2,463
	G	0,162	0,175	0,147	0,914	0,079	0,685	2,685	1,765	3,670	4,115	1,486	5,555
	UF2(Global)	P	0,281	0,375	0,180	0,870	0,112	0,721	2,171	1,070	3,351	3,638	1,429
NB		0,288	0,392	0,176	0,872	0,111	0,725	2,263	1,079	3,532	3,867	1,593	5,016
ZIP		0,284	0,389	0,172	0,870	0,108	0,723	2,180	1,111	3,325	3,658	1,508	4,770
ZINB		0,288	0,400	0,167	0,876	0,106	0,728	2,209	1,102	3,395	3,747	1,563	4,874
ZAP		0,287	0,385	0,181	0,868	0,113	0,722	2,178	1,112	3,322	3,660	1,542	4,763
ZANB		0,275	0,366	0,179	0,879	0,110	0,721	2,236	1,144	3,406	3,811	1,579	4,977
G		0,286	0,437	0,124	0,847	0,082	0,727	2,509	1,484	3,608	3,833	1,858	4,941
UF2		P	0,284	0,388	0,173	0,860	0,109	0,721	2,121	1,042	3,277	3,500	1,365
	NB	0,304	0,425	0,175	0,875	0,113	0,733	2,224	1,092	3,555	4,063	1,468	5,336
	ZIP	0,325	0,477	0,163	0,870	0,110	0,743	2,105	1,019	3,269	3,560	1,852	4,468
	ZINB	0,328	0,479	0,166	0,885	0,111	0,749	2,113	0,995	3,312	3,651	1,652	4,681
	ZAP	0,327	0,480	0,162	0,867	0,110	0,743	2,117	1,047	3,264	3,574	1,991	4,436
	ZANB	0,320	0,462	0,167	0,878	0,111	0,742	2,154	1,075	3,310	3,697	1,906	4,675
	G	0,588	0,648	0,316	0,903	0,139	0,419	0,569	0,403	1,317	1,032	0,594	1,897
	UF3(Global)	P	0,734	0,847	0,227	0,886	0,189	0,463	0,402	0,165	1,476	1,004	0,408
NB		0,753	0,876	0,199	0,881	0,193	0,461	0,385	0,132	1,526	1,003	0,368	1,831
ZIP		0,735	0,847	0,232	0,889	0,191	0,472	0,401	0,165	1,467	1,007	0,411	1,858
ZINB		0,743	0,858	0,225	0,888	0,195	0,475	0,394	0,151	1,490	1,013	0,389	1,873
ZAP		0,741	0,857	0,217	0,885	0,190	0,464	0,396	0,153	1,498	1,011	0,395	1,859
ZANB		0,747	0,866	0,212	0,884	0,194	0,465	0,391	0,141	1,522	1,024	0,370	1,899
G		0,625	0,689	0,332	0,909	0,159	0,466	0,510	0,331	1,320	1,016	0,516	1,923
UF3		P	0,759	0,881	0,208	0,887	0,202	0,484	0,370	0,136	1,425	0,910	0,423
	NB	0,761	0,884	0,205	0,886	0,203	0,485	0,387	0,145	1,477	1,106	0,708	1,742
	ZIP	0,759	0,877	0,222	0,891	0,208	0,497	0,373	0,145	1,401	0,922	0,467	1,553
	ZINB	0,763	0,881	0,228	0,892	0,216	0,503	0,380	0,145	1,443	1,000	0,510	1,722
	ZAP	0,760	0,881	0,214	0,887	0,209	0,484	0,372	0,142	1,411	0,918	0,465	1,534
	ZANB	0,759	0,879	0,213	0,886	0,206	0,480	0,382	0,147	1,446	0,967	0,496	1,634
	G	0,482	0,536	0,259	0,898	0,097	0,405	0,940	0,594	2,384	2,385	0,763	4,940
	UF4(Global)	P	0,597	0,683	0,235	0,898	0,118	0,452	0,783	0,388	2,431	2,315	0,697
NB		0,610	0,702	0,228	0,897	0,120	0,456	0,779	0,371	2,481	2,319	0,712	4,701
ZIP		0,622	0,718	0,222	0,895	0,122	0,460	0,766	0,358	2,466	2,330	0,709	4,732
ZINB		0,611	0,702	0,230	0,900	0,120	0,463	0,787	0,389	2,448	2,319	0,778	4,677
ZAP		0,633	0,732	0,220	0,893	0,125	0,461	0,758	0,345	2,478	2,329	0,717	4,714
ZANB		0,615	0,708	0,228	0,896	0,122	0,458	0,779	0,374	2,466	2,337	0,746	4,732
G		0,478	0,521	0,301	0,912	0,108	0,417	0,916	0,580	2,316	2,356	0,699	4,923
UF4		P	0,607	0,700	0,220	0,893	0,115	0,449	0,760	0,370	2,387	2,154	0,691
	NB	0,619	0,715	0,220	0,896	0,119	0,461	0,779	0,373	2,471	2,268	0,834	4,489
	ZIP	0,618	0,711	0,233	0,898	0,125	0,463	0,753	0,362	2,384	2,198	0,705	4,429
	ZINB	0,615	0,707	0,236	0,902	0,124	0,469	0,777	0,379	2,439	2,319	0,756	4,690
	ZAP	0,623	0,718	0,225	0,895	0,124	0,460	0,749	0,355	2,395	2,194	0,728	4,396
	ZANB	0,616	0,709	0,231	0,897	0,123	0,460	0,770	0,366	2,453	2,331	0,725	4,731
	G	0,721	0,772	0,293	0,925	0,123	0,282	0,337	0,239	1,148	0,719	0,456	1,537
	UF5(Global)	P	0,872	0,965	0,098	0,906	0,214	0,223	0,187	0,036	1,447	0,687	0,192
NB		0,873	0,967	0,096	0,906	0,217	0,230	0,185	0,035	1,444	0,686	0,190	1,523
ZIP		0,865	0,955	0,119	0,908	0,207	0,250	0,193	0,046	1,414	0,687	0,216	1,532
ZINB		0,867	0,958	0,108	0,907	0,204	0,239	0,191	0,043	1,427	0,686	0,206	1,527
ZAP		0,872	0,964	0,102	0,907	0,213	0,243	0,186	0,036	1,431	0,683	0,191	1,521
ZANB		0,872	0,964	0,103	0,907	0,214	0,242	0,186	0,037	1,430	0,684	0,194	1,524
G		0,827	0,908	0,153	0,909	0,152	0,241	0,229	0,093	1,364	0,695	0,295	1,531
UF5		P	0,874	0,967	0,095	0,907	0,212	0,237	0,185	0,035	1,431	0,672	0,199
	NB	0,869	0,962	0,099	0,908	0,197	0,244	0,194	0,045	1,439	0,737	0,356	1,510
	ZIP	0,871	0,963	0,104	0,908	0,209	0,251	0,187	0,040	1,411	0,670	0,213	1,462
	ZINB	0,868	0,959	0,112	0,910	0,204	0,266	0,192	0,045	1,417	0,690	0,241	1,510
	ZAP	0,872	0,965	0,101	0,908	0,213	0,246	0,185	0,038	1,416	0,669	0,210	1,453
	ZANB	0,869	0,962	0,091	0,907	0,185	0,239	0,191	0,042	1,433	0,684	0,239	1,473
	G	0,742	0,790	0,297	0,931	0,124	0,272	0,297	0,215	1,053	0,599	0,423	1,173
	UF6(Global)	P	0,885	0,972	0,088	0,915	0,208	0,237	0,157	0,028	1,333	0,545	0,171
NB		0,888	0,975	0,092	0,915	0,230	0,241	0,154	0,026	1,336	0,547	0,163	1,118
ZIP		0,881	0,966	0,105	0,916	0,210	0,250	0,161	0,035	1,317	0,550	0,187	1,128
ZINB		0,886	0,971	0,100	0,916	0,225	0,248	0,156	0,029	1,326	0,546	0,169	1,119
ZAP		0,884	0,969	0,101	0,916	0,218	0,247	0,158	0,031	1,324	0,548	0,177	1,123
ZANB		0,884	0,968	0,107	0,916	0,223	0,253	0,159	0,032	1,320	0,549	0,180	1,130
G		0,832	0,902	0,186	0,921	0,156	0,267	0,208	0,098	1,216	0,568	0,299	1,154
UF6		P	0,887	0,970	0,12								

	Change																									
	C	Sc	Sz	Pc	Pz	F1	E(T)	E(Tz)	E(Tc)	Sd(T)	Sd(Tz)	Sd(Tc)	C	Sc	Sz	Pc	Pz	F1	E(T)	E(Tz)	E(Tc)	Sd(T)	Sd(Tz)	Sd(Tc)		
Global	ZINB_allvar	0.7155	0.8387	0.1782	0.9068	0.1342	0.5361	0.6238	0.2283	2.3501	1.9969	0.7073	3.9427	-2%	-2%	0%	0%	-5%	-3%	8%	16%	4%	9%	26%	7%	
	BR	0.7015	0.8215	0.1780	0.9053	0.1268	0.5200	0.6723	0.2638	2.4550	2.1670	0.8937	4.2188	9%	12%	-69%	-8%	3%	-59%	9%	-64%	42%	30%	-29%	32%	
	ENET	0.7800	0.9432	0.0560	0.8319	0.1377	0.2198	0.6783	0.0813	3.3272	2.5962	0.5024	5.1894	-20%	-24%	2%	0%	-16%	-7%	29%	92%	8%	28%	118%	16%	
First-age	ZINB_allvar	0.4836	0.6542	0.1918	0.8816	0.1331	0.6962	1.4344	0.5741	2.9055	2.9977	1.2478	4.2734	19%	33%	-65%	-23%	-5%	-43%	25%	-54%	52%	43%	2%	41%	
	BR	0.3849	0.4954	0.1958	0.8803	0.1120	0.6441	1.8541	1.1048	3.1353	3.8367	2.7160	4.9678	-19%	-21%	39%	1%	-32%	-16%	53%	169%	-5%	9%	91%	0%	
	ENET	0.5735	0.8700	0.0665	0.6803	0.1271	0.3952	1.7966	0.2622	4.4204	4.2863	1.7718	6.0096	6%	8%	-62%	-4%	-8%	-43%	5%	-31%	22%	23%	108%	7%	
Second-age	ZINB	0.7748	0.8761	0.1655	0.9053	0.1386	0.4088	0.4140	0.1597	1.9438	1.6244	0.5353	3.7498	-9%	-9%	-12%	-2%	-22%	-11%	18%	44%	7%	0%	23%	-4%	
	BR	0.6264	0.6922	0.2306	0.9099	0.0945	0.3418	0.6318	0.4289	1.8526	1.7745	1.0230	3.7496	7%	16%	-74%	-14%	-24%	-52%	11%	-64%	43%	15%	-43%	9%	
	ENET	0.8208	0.9468	0.0632	0.8733	0.1280	0.2321	0.4336	0.1108	2.3758	2.0021	1.1156	4.0297	-22%	-30%	0%	1%	-11%	-4%	19%	32%	15%	22%	26%	23%	
UF1	ZINB_allvar	0.6589	0.7925	0.2594	0.9033	0.1898	0.6307	0.6497	0.2613	1.8109	1.6494	0.6213	2.8117	36%	66%	-59%	-34%	-2%	-34%	28%	-58%	56%	85%	-30%	92%	
	BR	0.5984	0.7222	0.2283	0.8830	0.1478	0.5611	0.7675	0.3766	1.9362	1.6498	0.7661	2.6994	-3%	-3%	-7%	-1%	-15%	-11%	11%	23%	6%	10%	20%	8%	
	ENET	0.7054	0.9190	0.0667	0.7793	0.1438	0.3030	0.7219	0.0950	2.5964	1.9035	0.3553	3.0649	3%	6%	-49%	-5%	0%	-41%	4%	-46%	27%	17%	-37%	24%	
UF2	ZINB_allvar	0.3279	0.4792	0.1657	0.8851	0.1111	0.7492	2.1129	0.9947	3.3115	3.6508	1.6522	4.6813	-7%	-7%	7%	0%	-5%	-5%	9%	17%	4%	13%	36%	11%	
	BR	0.2542	0.3375	0.1649	0.8904	0.0990	0.7169	2.5155	1.3086	3.8091	4.4567	2.0865	5.7668	19%	26%	-69%	-7%	-18%	-34%	-2%	-45%	25%	22%	86%	10%	
	ENET	0.4443	0.7961	0.0672	0.5878	0.1084	0.4958	2.7052	0.4168	5.1581	6.7426	1.1550	9.0079	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%	
UF3	ZINB_allvar	0.7625	0.8808	0.2283	0.8918	0.2162	0.5030	0.3804	0.1454	1.4425	0.9999	0.5099	1.7223	0%	0%	1%	-29%	-1%	-5%	-29%	-1%	-23%	4%	2%	-20%	3%
	BR	0.7365	0.8527	0.2112	0.8806	0.1848	0.4457	0.4226	0.1792	1.5225	1.0979	0.6121	1.8664	0%	0%	1%	-29%	-1%	-5%	-29%	-1%	-23%	4%	2%	-20%	3%
	ENET	0.7825	0.9297	0.1172	0.8447	0.2151	0.2945	0.3970	0.0789	1.8346	1.1687	0.3235	2.1316	0%	0%	1%	-29%	-1%	-5%	-29%	-1%	-23%	4%	2%	-20%	3%
UF4	ZINB_allvar	0.6155	0.7066	0.2358	0.9024	0.1238	0.4693	0.7774	0.3787	2.4393	2.3189	0.7560	4.6897	-7%	-7%	7%	0%	-5%	-5%	9%	17%	4%	13%	36%	11%	
	BR	0.5744	0.6514	0.2534	0.9004	0.1177	0.4439	0.8492	0.4448	2.5350	2.6114	1.0279	5.2276	19%	26%	-69%	-7%	-18%	-34%	-2%	-45%	25%	22%	86%	10%	
	ENET	0.7338	0.8924	0.0725	0.8349	0.1018	0.3088	0.7587	0.2066	3.0600	2.8325	1.4028	5.1677	0%	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%
UF5	ZINB_allvar	0.8682	0.9589	0.1118	0.9097	0.2037	0.2660	0.1918	0.0449	1.4166	0.6897	0.2407	1.5098	0%	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%
	BR	0.8724	0.9669	0.0846	0.9046	0.1984	0.2101	0.1906	0.0392	1.4529	0.8008	0.4832	1.5027	0%	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%
	ENET	0.8718	0.9669	0.0789	0.9028	0.1931	0.1878	0.1893	0.0345	1.4801	0.7008	0.1923	1.5531	0%	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%
UF6	ZINB_allvar	0.8855	0.9667	0.1397	0.9217	0.2531	0.3173	0.1561	0.0348	1.2703	0.5828	0.1926	1.3204	0%	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%
	BW	0.8885	0.9798	0.0504	0.9092	0.1753	0.1584	0.1599	0.0210	1.4353	0.5959	0.1492	1.2683	0%	0%	1%	-24%	-1%	-3%	-21%	-1%	-13%	3%	16%	101%	0%
	ENET	0.9004	0.9984	0.0011	0.9021	0.0588	0.0090	0.1500	0.0020	1.5086	0.5752	0.0547	1.1385	2%	3%	-99%	-2%	-77%	-97%	-4%	-94%	19%	-1%	-72%	-14%	

Set of thirteen goodness-of-fit measures describing the complete (ZINB-allvar) and reduced models (BR-Backward Regression and Enet Penalised Regression) for each morphological region.

(G-Gaussian; P- Poisson; NB-Negative Binomial; ZIP-Zero Inflated Poisson; ZINB-Zero Inflated Negative Binomial; ZAP-Zero Alternated Poisson; ZANB-Zero Alternated Negative Binomial; R2- Coefficient of determination; C-Accuracy; Sc-Sensitivity count part; Sz- Sensitivity zero part; Pc-Precision count part; Pz-Precision zero part; F1 score; E(T)-Average Tolerance; E(Tz)-Average Tolerance Zero Part; E(Tc)-Average Tolerance count part; Sd(T)-Standard Deviation Tolerance, Sd(Tz)-Standard Deviation Tolerance Zero Part; Sd(Tc)-Standard Deviation Tolerance Count Part).