



# Network Position and Throughput Performance of Seaports

César Ducruet, S.W. Lee, J.M. Song

## ► To cite this version:

César Ducruet, S.W. Lee, J.M. Song. Network Position and Throughput Performance of Seaports. T.E. Notteboom. Current Issues in Shipping, Ports, and Logistics, Academic and Scientific Publishers (ASP), pp.189-201, 2011. halshs-00561393v2

**HAL Id: halshs-00561393**

**<https://shs.hal.science/halshs-00561393v2>**

Submitted on 22 Feb 2011

**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.

## CHAPTER 10

### Network Position and Throughput Performance of Seaports



César DUCRUET, Sung-Woo LEE and Ju-Miang SONG

#### **Abstract**

The determinants of throughput volume at a given set of ports have rarely been approached from a network perspective. This paper proposes a set of novel indicators describing the relative situation of seaports in the worldwide maritime network of container shipping in 2006, which are distinguished among five categories: circulation (calls, vessels, and operators), foreland (distance to other ports and distribution of connections), connectivity itself (number of connections to other ports), centrality (betweenness and eccentricity), and neighbourhood (strength and clustering indices). Main results help to classify ports according to their location and function in the network, while they stress which parameters most influence throughput volumes. Although centrality indicators highly correlate with throughput, the latter seems to be influenced mostly by the geographic parameter of the maximum distance link to another port.

#### **1 | Introduction**

The quantitative analysis of seaports is a traditional approach of port and maritime geography aiming at understanding, among other things, the factors most influencing ports' performance and, therefore, competitiveness. A large literature on port choice and port selection insists more on the qualitative aspects of performance besides other economic factors such as cost and productivity (Ng, 2009; Notteboom, 2009). Another body of research groups together various port-related indicators describing different aspects of a port's life (e.g. traffic, infrastructure, number of calls), either for highlighting types of ports or to insist on which factor seems most relevant for comparing ports except from the classic figure of annual throughput (Joly and Martell, 2003). A wide set of possible measures have thus been proposed, some being measurable and other not (De Langen et al., 2007), all placed under the general name of "port performance indicators" or PPIs.

Because covering this large research field exhaustively would run beyond the scope of this paper, the present research aims at exploring to what extent other kinds of port performance indicators may be drawn from network analytical tools. Since ports are critical nodes in maritime (and land) networks, it seems very natural to expect that their position in such networks may be expressed by means of other variables than local characteristics but on relational measures. Such measures are already extensively used in the wider research area of “network science” that is composed of various elements from graph theory, social network analysis, and complex networks. Based on recent works introducing such approaches to the maritime world (Ducruet and Notteboom, 2010), this paper wishes to push further the statistical analysis of a varied set of indicators, notably in relation with the classic one of throughput, in order to verify which of these indicators best influence ports’ performance.

The remainder of the paper are as follows. Section 2 introduces the data on vessel movements used for building graphs of the global maritime network in which the position of ports will be measured from various perspectives. Section 3 presents the results of a factor analysis while Section 4 runs a multiple regression analysis. Finally, concluding remarks are given in Section 5 as well as some possible avenues for future research in this direction.

## **2 | Data and methodology**

Data on worldwide vessel movements was obtained from Lloyd’s Marine Intelligence Unit (LMIU) for the year 2006<sup>1</sup>. It allowed building the global graph of inter-port links with ports themselves considered as nodes. Ports are also characterized by their throughput in TEUs using Containerisation International data for 2006. Such data has the advantage of covering most of the world’s fleet of containerships (about 98% of total fleet capacity), while it offers high precision in the information: daily vessel movements are recorded as well as the capacity (TEU) of the ships.

For a full investigation of the interdependence between throughput volume and network position, we have distinguished among two types of graphs. One graph is the one of direct links between ports, taking into account successive calls between previous and next ports only. The other graph adds indirect links to the former, thus including all intermediate calls between all ports welcoming the same vessels.

Five categories of indicators were calculated for each graph in order to be compared with throughput data (see Table 1):

- Circulation indicators simply count the number of vessels, calls, and operators for each port after one year of movements;
- Foreland indicators provide the geographic dimension of distance (average and maximal) between each port and other ports connected, while adding a measure of foreland diversity reflecting upon the distribution of traffic;

---

<sup>1</sup> <http://www.seasearcher.com/>

- Connectivity indicators are local measures defined by the number of ports connected (degree); the weighted degree is the sum of all traffic links for a given port, and the nodal degree is the number of maximal flow connections with other ports. For instance, a large port is likely to attract the maximum traffic flow of many secondary ports, while the latter may have only one maximum flow directed to a larger port;
- Centrality indicators are classic measures of accessibility. They complete connectivity indicators by the fact that they illustrate the position of ports on the level of the entire graph: betweenness centrality is the number of positions on possible shortest paths and eccentricity is a normalized measure of farness from other ports;
- Neighbourhood indicators calculate the relative position of ports vis-à-vis their direct neighbours considering the configurations of their relations (strength, Strahler, clustering)<sup>2</sup>.

*Table 1. List of indicators*

| Type of measure | Variable                        | Definition                                                                                                                                                           |
|-----------------|---------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Circulation     | No. Vessels                     | Number of different vessels having called at the port through the year                                                                                               |
|                 | No. Calls                       | Number of times a vessel has called through the year                                                                                                                 |
|                 | No. Operators                   | Number of different operators having called the port through the year                                                                                                |
| Foreland        | Max. distance to another port   | Maximum orthodromic distance (km) among all connections to other ports                                                                                               |
|                 | Avg. distance to other ports    | Average orthodromic distance (km) among all connections to other ports                                                                                               |
|                 | Foreland diversity index        | Relative diversity index applied to ports' worldwide traffic distribution at country level (inverse of the sum of differences in shares compared with world average) |
| Connectivity    | Degree                          | Number of ports connected                                                                                                                                            |
|                 | Weighted degree                 | Sum of traffic on all connections                                                                                                                                    |
|                 | Nodal degree                    | Number of ports connected by dominant flows (i.e. maximum flow links)                                                                                                |
| Centrality      | Betweenness centrality          | Number of positions on possible shortest paths in the entire graph                                                                                                   |
|                 | Weighted betweenness centrality | Number of positions on possible shortest weighted paths in the entire graph                                                                                          |
|                 | Eccentricity                    | Normalized measure of remoteness from all other ports                                                                                                                |
| Neighbourhood   | Clustering coefficient          | Probability for direct neighbours to be connected with each other                                                                                                    |
|                 | Strahler index                  | Ramification level                                                                                                                                                   |
|                 | Strength index                  | Strength of adjacent connections with other ports                                                                                                                    |

*Source: own elaboration based on LMIU data*

Among the variables, the affinity with throughput volume may be appreciated by comparing linear correlations (Table 2). Results are very similar regardless of the type of graph

<sup>2</sup> Appendix 1 provides formulas for main network measures.

considered except for foreland diversity, degree centrality, weighted degree, betweenness centrality that have slightly lower correlations in the graph of all links. Unsurprisingly, highest correlations are observed with circulation indicators (vessels, calls, and operators) and with weighted degree that is equivalent to total vessel traffic. Other measures are more complex and better depict the topological dimension of ports' position. Among the latter, only betweenness centrality exceeds 0.8 in the graph of direct links. Nodal degree, degree centrality, and foreland diversity have also significant correlations (over 0.7) with throughputs. Weighted betweenness, eccentricity and distance measures all exhibit moderate correlations. Neighbourhood measures in general are negatively correlated with throughput. This may be explained by the fact that large ports multiply their connections to other ports, while the latter are not necessarily connected with each other. This is typically an expression of the hub-and-spoke configuration with one pole connecting numerous satellites. Smaller ports have higher strength and clustering indices because they are more likely to form cliques<sup>3</sup> with their neighbours.

*Table 2. Linear (Pearson) correlations with TEU throughput by type of graph*

| Type of measure | Variable                        | Direct links graph | All links graph |
|-----------------|---------------------------------|--------------------|-----------------|
| Circulation     | Vessels                         | 0.921              | 0.921           |
|                 | Calls                           | 0.906              | 0.906           |
|                 | Operators                       | 0.792              | 0.794           |
| Foreland        | Maximum distance                | 0.392              | 0.266           |
|                 | Average distance                | 0.285              | 0.293           |
|                 | Foreland diversity              | 0.702              | 0.703           |
| Connectivity    | Degree centrality               | 0.760              | 0.627           |
|                 | Nodal degree                    | 0.750              | 0.731           |
|                 | Weighted degree                 | 0.897              | 0.802           |
| Centrality      | Betweenness centrality          | 0.810              | 0.521           |
|                 | Weighted betweenness centrality | 0.280              | 0.189           |
|                 | Eccentricity                    | 0.441              | 0.436           |
| Neighbourhood   | Clustering coefficient          | -0.311             | -0.414          |
|                 | Strahler                        | -0.034             | 0.043           |
|                 | Strength                        | -0.275             | -0.335          |

*Source: own elaboration based on LMIU data*

<sup>3</sup> A clique is defined as a maximal complete subgraph of at least 3 nodes

### 3 | Factor analysis

The factor analysis of all indicators (including throughput) was applied without rotation and on each type of graph. Results for the situation of indicators on each factor (or principal component) may be summarized as follows (Table 3):

- Overall, the composition of factors is very similar between the GDL and the GAL;
- F1 groups together TEU throughput, circulation indicators (operators, vessels, calls), and degree centrality, indicating that large ports are those of intense maritime activity deploying many links to other ports. These top indicators are opposed to the neighbourhood indicators of clustering and strength. Thus, F1 has a strong hierarchical logic in which highly correlated indicators go together;
- F2 is built on an opposition between TEU throughput, circulation indicators, nodal degree on the one hand, foreland indicators (distance to other ports, foreland diversity), and eccentricity on the other. We interpret this trend as an opposition between a hierarchical logic (similar to F1) and a geographical logic: centrally located ports have more throughput than remotely located ports;
- F3 shows an interesting opposition between connectivity (nodal degree), centrality (betweenness) indicators and foreland/neighbourhood indicators. Only in the GAL, throughput is significantly included in the group of foreland/neighbourhood indicators. This means that in the GAL, ports belonging to dense neighbourhoods generally generate more traffic than centrally located ports, which role is better explained by a bridge (or intermediate) function between dense neighbourhoods. Indeed, some ports may act as strategic pivots between regions without generating large traffics, as seen in the case of Anchorage (Fleming and Hayuth, 1994);
- F4 will not be interpreted due to its different composition in the GDL and the GAL, and because port throughput does not appear as a significant contributor to the observed trends.

The extent to which such results possess a geographic dimension can be verified in Figures 1a and 1b (direct links and all links). Because F1 is too much influenced by throughput itself, we concentrate the interpretation of F2, F3 and F4. In terms of the main ports represented<sup>4</sup> on F2, there is an opposition between some major hub ports and some major gateway ports. Asian hub ports are grouped together with European gateway-hubs (Rotterdam, Hamburg) in the trend of throughput, nodal degree and number of calls. These ports are local-global pivots ensuring the redistribution of containers within and between maritime regions. They are opposed to “classic” gateway ports (e.g. US ports) which manage to connect long-distance flows: their situation in the Pacific and their importance in trans-Pacific trades (USA-Asia) clearly explain the influence of foreland (distance) indicators on their grouping on F2, as these ports generate moderate throughput volumes. The dominance of their gateway functions tends to negatively impact their centrality in the maritime network, because our measures do not take into account hinterland accessibility. In turn, the average distance of major hub ports’ connections is lowered by the intensity of their hub-and-spoke activities towards local secondary ports. We clearly observe a geographic divide between the Europe-

---

<sup>4</sup> see Appendix 2 for most representative ports.

Asia maritime corridor having high internal connectivity (cf. neighbourhood) and the rest of the world more characterized by long-distance connections.

*Table 3. Results of the factor analysis by type of graph*

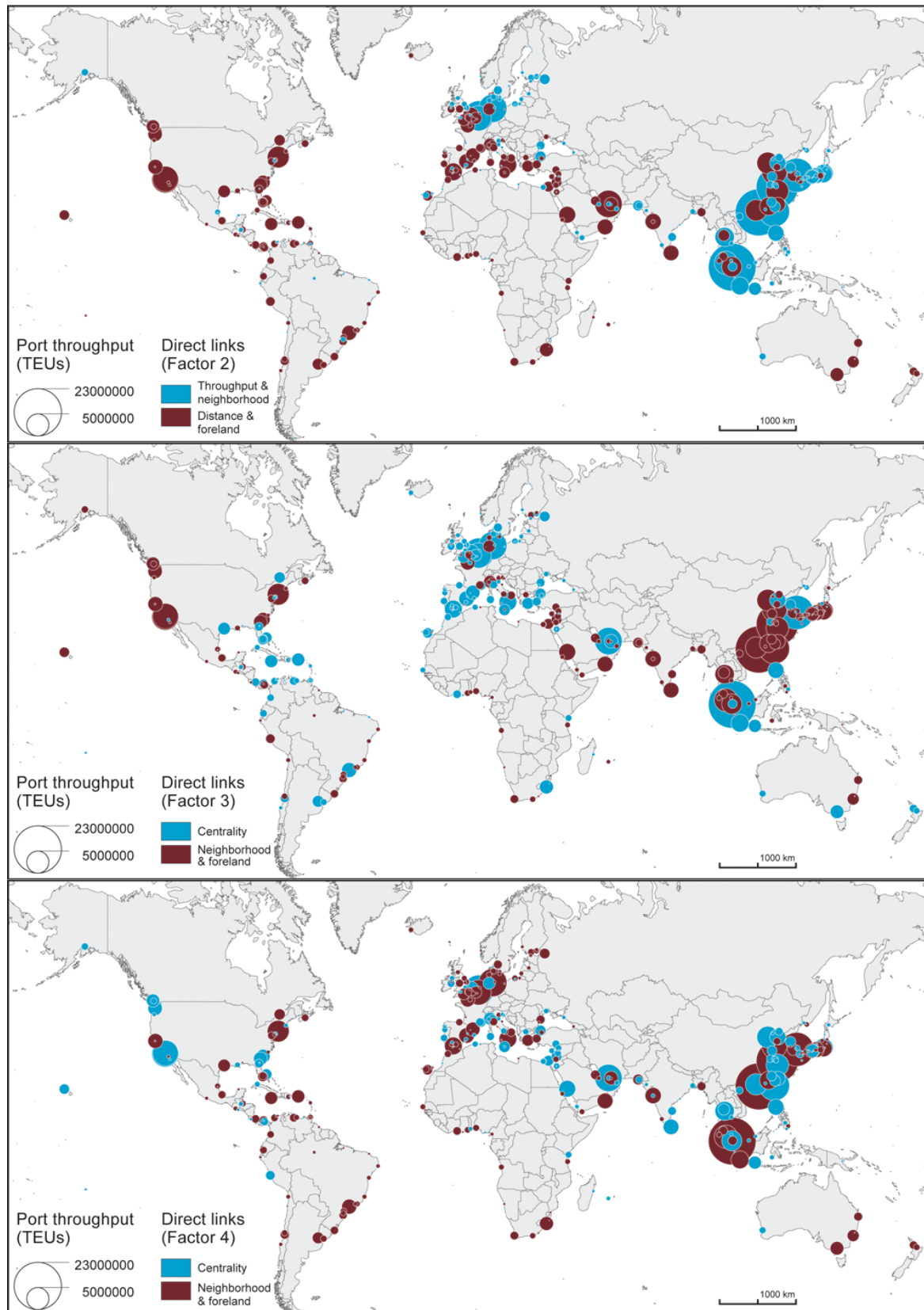
| Value                           | Graph of direct links (GDL) |          |          |          | Graph of all links (GAL) |          |          |          |
|---------------------------------|-----------------------------|----------|----------|----------|--------------------------|----------|----------|----------|
|                                 | Factor 1                    | Factor 2 | Factor 3 | Factor 4 | Factor 1                 | Factor 2 | Factor 3 | Factor 4 |
| Eigenvalue                      | 9.054                       | 1.812    | 1.407    | 1.029    | 8.437                    | 2.034    | 1.536    | 0.962    |
| % of Var.                       | 56.587                      | 11.323   | 8.791    | 6.433    | 52.730                   | 12.712   | 9.603    | 6.011    |
| Cum. %                          | 56.587                      | 67.910   | 76.701   | 83.134   | 52.730                   | 65.442   | 75.045   | 81.055   |
| Throughput (TEUs)               | 0.886                       | -0.313   | 0.076    | 0.004    | 0.856                    | -0.317   | -0.254   | 0.056    |
| Operators                       | 0.939                       | 0.002    | 0.115    | -0.008   | 0.957                    | 0.016    | -0.070   | -0.056   |
| Vessels                         | 0.956                       | -0.178   | 0.104    | 0.023    | 0.942                    | -0.188   | -0.196   | -0.003   |
| Calls                           | 0.905                       | -0.336   | 0.061    | 0.023    | 0.889                    | -0.330   | -0.218   | 0.066    |
| Nodal degree                    | 0.837                       | -0.307   | -0.156   | 0.059    | 0.705                    | -0.416   | -0.200   | 0.143    |
| Maximum distance                | 0.650                       | 0.630    | 0.237    | 0.054    | 0.544                    | 0.724    | 0.008    | 0.005    |
| Average distance                | 0.508                       | 0.634    | 0.253    | 0.145    | 0.527                    | 0.672    | -0.297   | -0.024   |
| Betweenness centrality          | 0.865                       | -0.295   | -0.129   | 0.041    | 0.644                    | -0.091   | 0.172    | -0.193   |
| Clustering coefficient          | -0.575                      | -0.322   | 0.709    | 0.040    | -0.732                   | -0.012   | -0.633   | 0.070    |
| Degree centrality               | 0.943                       | 0.025    | -0.007   | 0.007    | 0.917                    | 0.236    | 0.152    | -0.086   |
| Strahler                        | -0.080                      | -0.139   | -0.096   | 0.874    | 0.131                    | 0.303    | 0.130    | 0.839    |
| Strength                        | -0.519                      | -0.345   | 0.744    | -0.026   | -0.502                   | 0.287    | -0.750   | 0.081    |
| Weighted betweenness centrality | 0.310                       | -0.185   | -0.124   | -0.468   | 0.254                    | -0.182   | 0.306    | 0.382    |
| Eccentricity                    | 0.714                       | 0.430    | 0.212    | -0.009   | 0.791                    | 0.480    | 0.165    | -0.105   |
| Weighted degree                 | 0.884                       | -0.304   | 0.095    | 0.074    | 0.803                    | -0.329   | -0.250   | 0.096    |
| Foreland diversity              | 0.804                       | 0.187    | 0.280    | -0.092   | 0.804                    | 0.174    | -0.196   | -0.058   |

*Source: own elaboration based on LMIU data and StatiXL software*

On F3, results slightly differ according to the type of graph considered, although the overall logic is similar. In the GDL, the influence of distance indicators on the grouping of many Asia-Pacific ports together is apparent (Canada, US, Mexico, China, Japan) as opposed to a number of local hub ports having a higher centrality in the network (e.g. Surabaya for Indonesia, Bandirma for Turkey, Pietarsaari for Finland, and Aalborg for Denmark-Iceland). The latter ports are regional hubs redistributing cargoes to peripheral ports. In the GAL, the opposition has a clearer geographic distribution, with many Asia-Pacific ports and hubs opposed with several European (Scandinavia-Baltic and Atlantic) ports. In both figures, Europe appears as a rather distinct group characterized by higher centrality (and throughput) than neighbourhood. Such results may illustrate the fact that regardless of their size, European ports have a good position globally compared with other ports, which remain more embedded regionally for an equivalent throughput size. Such pattern may suggest the permanency of a centre-periphery pattern of international trade inherited from the past.

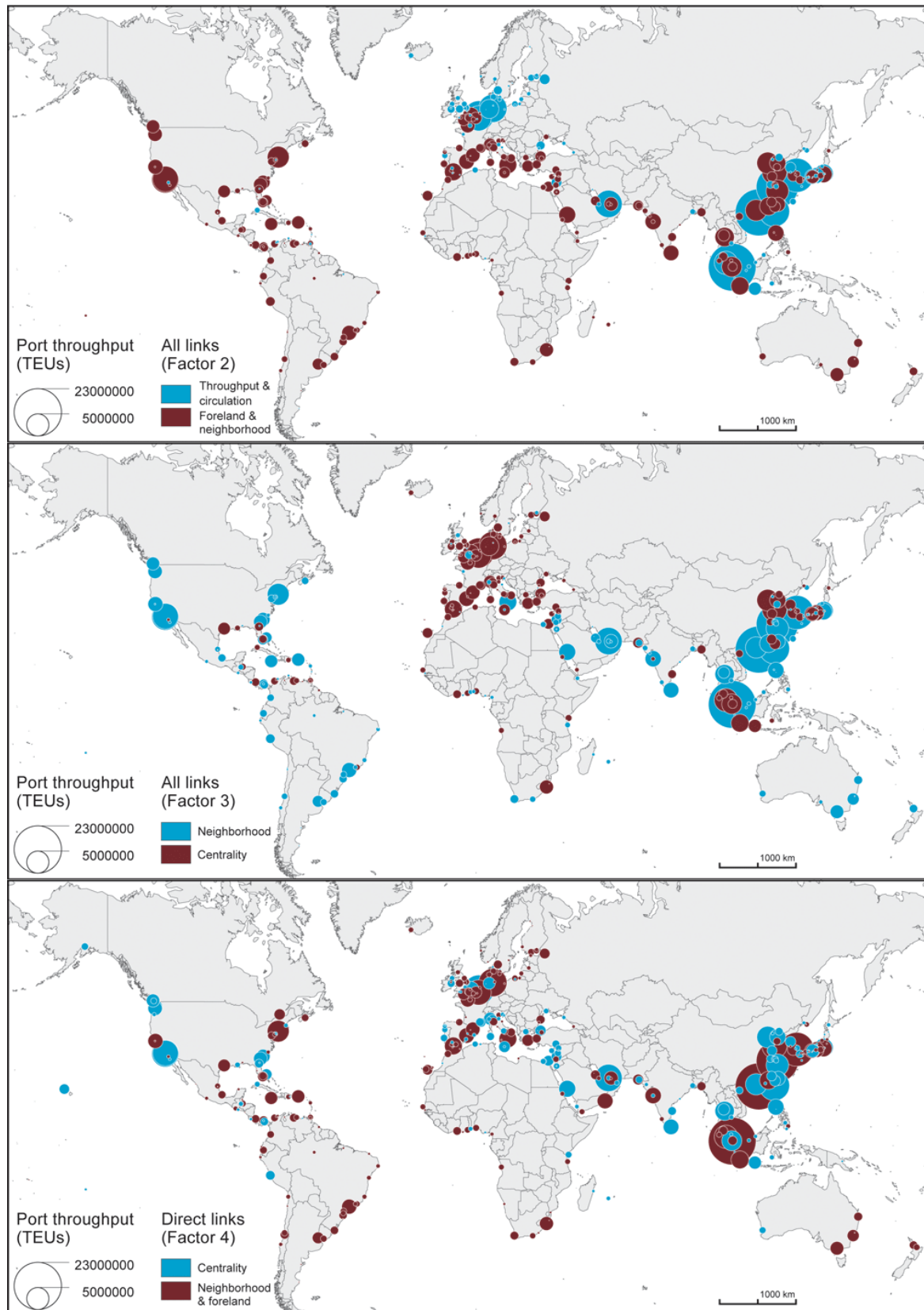
On F4, an East-West belt is made of more central ports of all throughput sizes, from Los Angeles to China. The opposite profile points at major (hub) ports with long-range connections (foreland) as well as strong local embeddedness (neighbourhood), and their distribution is comparable in the GDL and in the GAL.

Figure 1a. Position of ports on each factor (graph of direct links)



Source: own realization based on LMIU and Containerisation International data

Figure 1b. Position of ports on each factor (graph of direct links)



Source: own realization based on LMIU and Containerisation International data

## 4 | Multiple Regression Analysis

In this section we wish to assess the effects of network indicators on throughput by means of an exploratory study. We use AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) in order to select suitable combinations of network indicators in all cases by imposing penalty to explanatory variables in excess. Of course, such methods do not exclude possible shortcomings due to the variables chosen. For each type of graph, we select 2 models which AIC and BIC is near to minimum with high Adjusted R-square. Then the selection of the final model is based on the works of Anderson et al. (1972) and Allen (1974) proposing to adopt the model with minimum of the PRESS, which equals the sum of squares of predicted residual errors. Table 4 shows the best criteria of possible cases and the selected criterion of the model in each graph.

*Table 4. Criterion of Variable Selection*

|                   | Graph of direct links (GDL) |                  | Graph of all links (GAL) |                  |
|-------------------|-----------------------------|------------------|--------------------------|------------------|
|                   | Best                        | Finally Selected | Best                     | Finally Selected |
| R-Square          | Max : 84.54%                | 84.20%           | Max : 83.15%             | 82.87%           |
| Adjusted R-Square | Max : 83.76%                | 83.74%           | Max : 82.53%             | 82.46%           |
| (AIC)             | Min : 6716.1889             | 6716.1889        | Min : 5894.2948          | 5894.2948        |
| BIC               | Min : 6718.9119             | 6718.9119        | Min : 5897.0185          | 5897.0185        |

The analysis of variance (ANOVA) is presented in Table 5, where the P-value in both graphs is smaller than 0.0001, meaning that the model is significant at 5% level. In both cases, the selected model for each graph can explain about 83% of throughput's variance (dependent variable), which is a satisfactory result (Table 6).

Table 7 shows the coefficients of selected network indicators and the P-values of each coefficient. We have seven network indicators in the GDL and five in the GAL. Four indicators affect TEU throughput in both graphs: vessels, operators, maximum distance to another port, and weighted degree. The coefficients of number of vessels and number of operators are respectively positive and negative in both graphs. It means that whatever the type of graph considered, the more vessels and the lesser operators, the more throughput increases. However, the sign of the coefficients for "maximum distance to another port" and weighted degree is different both in the GDL and in the GAL. If the maximum distance to another port increases, then port throughput decreases in the GDL, but contrastingly increases in the GAL. The influence of weighted degree is the reverse to the maximum distance to another port. The role of distance is thus paradoxical: while larger ports connect longer distances on average, this is complicated by some remotely located ports (e.g. Pacific and Oceania, South Africa) generating less traffics. Thus, distance is both a measure of good performance and a measure of remoteness: it does not have the same meaning for all ports.

*Table 5. ANOVA results*

|                             | Source          | DF  | Sum of squares | Mean Square | F Value | Pr > F |
|-----------------------------|-----------------|-----|----------------|-------------|---------|--------|
| Graph of direct links (GDL) | Model           | 7   | 6.440927E14    | 9.201324E13 | 183.44  | <.0001 |
|                             | Error           | 241 | 1.208827E14    | 5.015881E11 |         |        |
|                             | Corrected Total | 248 | 7.649754E14    |             |         |        |
| Graph of all links (GAL)    | Model           | 5   | 6.217753E14    | 1.243551E14 | 204.13  | <.0001 |
|                             | Error           | 211 | 1.285381E14    | 6.091855E11 |         |        |
|                             | Corrected Total | 216 | 7.503135E14    |             |         |        |

*Table 6. Complementary results*

|                | Graph of direct links (GDL) | Graph of all links (GAL) |
|----------------|-----------------------------|--------------------------|
| Root MSE       | 708229                      | 780503                   |
| Dependent Mean | 735777                      | 824218                   |
| Coeff Var      | 96.25591                    | 94.69624                 |
| R-Square       | 0.8420                      | 0.8287                   |
| Adj R-Sq       | 0.8374                      | 0.8246                   |

Three indicators including foreland diversity index, betweenness centrality, and weighted betweenness centrality are selected in the GDL only. The average distance to other ports is selected and has a negative influence on throughput in the GAL.

In both graphs (GDL and GAL), the most influential indicator is the maximum distance to another port. Such result was not easily predictable based on simple correlations. It means that whatever the type of graph considered, ports with long-range connections are more likely to generate large throughputs than ports with short-range connections. This is a very clear illustration of the importance of geographic parameters in shaping port hierarchies. Ports with wider radiance and geographic reach, of course, are likely to connect trunk lines of high density where ocean carriers deploy their largest vessels. As seen in the factor analysis, there is also the influence of conjectural phenomena with the overwhelming importance of trans-Pacific trades in the current pattern of container shipping networks, with major US and Asian ports connecting with each other through weighty shipping lines.

Table 7. Parameter estimates

|                             | Label                           | DF | Parameter Estimate | Standard Error | t Value | Pr >  t |
|-----------------------------|---------------------------------|----|--------------------|----------------|---------|---------|
| Graph of direct links (GDL) | Intercept                       | 1  | 12274              | 69748          | 0.18    | 0.8605  |
|                             | No. vessels                     | 1  | 10756              | 1082.53567     | 9.94    | <.0001  |
|                             | No. operators                   | 1  | -12768             | 3517.18958     | -3.63   | 0.0003  |
|                             | Max. distance to another port   | 1  | -25.10418          | 15.41028       | -1.63   | 0.1046  |
|                             | Foreland diversity index        | 1  | -1850.02870        | 1011.58465     | -1.83   | 0.0687  |
|                             | Weighted degree                 | 1  | 0.01027            | 0.00363        | 2.83    | 0.0051  |
|                             | Betweenness centrality          | 1  | 11.07465           | 6.22764        | 1.78    | 0.0766  |
|                             | Weighted betweenness centrality | 1  | 2.26709            | 0.73670        | 3.08    | 0.0023  |
| Graph of all links (GAL)    | Intercept                       | 1  | -5353.56807        | 77282          | -0.07   | 0.9448  |
|                             | No. vessels                     | 1  | 12459              | 922.70239      | 13.50   | <.0001  |
|                             | No. operators                   | 1  | -12730             | 3976.01767     | -3.20   | 0.0016  |
|                             | Max. distance to another port   | 1  | 42.83943           | 24.68682       | 1.74    | 0.0841  |
|                             | Avg. distance to other ports    | 1  | -162.24881         | 64.55029       | -2.51   | 0.0127  |
|                             | Weighted degree                 | 1  | -0.00005430        | 0.00002313     | -2.35   | 0.0198  |

## 5 | Conclusion

This exploratory research has provided floor for a reflection on the influence of network position on throughput volumes. Although maritime networks are volatile with regard to land-based fixed networks (e.g. road, rail), their characteristics and configurations also express important realities of trade and port development. The main results of this paper confirm the importance of centrality in the network but, also, the influence of geographic distance between ports. Ports reaching further distance through direct or indirect shipping connections are more likely to generate container throughput than ports connecting nearest

ports. While such results would need further verification in terms of statistical goodness and policy implications, it paves the way towards complementary research in this field, perhaps by looking at how throughput growth (instead of sole volume) can be explained by network position and its possible changes. Yet, the measure of distance remains paradoxical because it does not overlap entirely the port hierarchy: some smaller ports that are remotely located also connect long distance links. This phenomenon is directly caused by the physical embedding of ports in a spatial network. Perhaps, a zoom at ports of comparable size, such as the world's largest ports facing similar issues, would provide more interesting results in terms of functional differentiation and specialization.

## References

Allen, D. M. (1974) The Relationship between Variable Selection and Data Augmentation and a Method for Prediction, *Technometric*, 16, 125-127

Anderson, R. L., D. M. Allen, F. Cady (1972) Selection of Predictor Variables in Multiple Linear Regression, in: *Statistical Papers in Honor of George W. Snedecor*, T.A. Bancroft, ED., Iowa State University Press, Ames

Akaike (1974) A New Look at the Statistical Model Identification, *IEEE Trans. Automatic Control*, 19, 716-723

De Langen, P.W., Nijdam, M., Van der Horst, M.R. (2007) New indicators to measure port performance, *Journal of Maritime Research*, 4(1), 23-36

Ducruet, C., Notteboom, T.E. (2010) The worldwide maritime network of container shipping: spatial structure and regional dynamics, *GaWC Research Bulletin* 364, <http://www.lboro.ac.uk/gawc/rb/rb364.html>

Joly, O., Martell, H. (2003) Infrastructure benchmarks for European container ports, Paper presented at the Inha-Le Havre International Conference, Incheon, South Korea

Kashyap, R. L. (1977) A Bayesian Comparison of Different Classes of Dynamic Models Using Empirical Data, *IEEE Trans. Automatic Control*, 22, 715-727

Ng, K.Y.A. (2009) *Port competition: the case of North Europe*, VDM Verlag

Notteboom, T.E. (2009) Complementarity and substitutability among adjacent gateway ports, *Environment and Planning A*, 41(3), 743-762

Tongzon, J. (1995) Determinants of port performance and efficiency, *Transportation Research Part A*, 29(3), 245-252

### Appendix 1. Illustration of main network measures

| Measure                                               | Description                                                                                           | Formula                                                                  |
|-------------------------------------------------------|-------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Degree centrality                                     | Number of adjacent nodes (k)                                                                          | $k_i = C_D(i) = \sum_j x_{ij}$                                           |
| Weighted degree                                       | Sum of link weights (s)                                                                               | $s_i = C_D^w(i) = \sum_j w_{ij}$                                         |
| Transitivity, Clustering coefficient                  | Number of observed triads divided by maximum possible number of triads                                | $C_i = \frac{2 \{e_{jk}\} }{k_i(k_i - 1)}$                               |
| Strength index                                        | Probability for adjacent links to belong to cycles of length 3 and 4                                  | $w_s(u) = \frac{\sum_{e \in adj(u)} w_s(e)}{deg(u)}$                     |
| Eccentricity (or Koenig number, associated number)    | Number of links needed to connect the farthest node                                                   | $e(x) = \max_{y \in X} d(x, y).$                                         |
| Betweenness centrality (or shortest-path betweenness) | Number of occurrences on all shortest paths                                                           | $C_B(i) = \frac{g_{jk}(i)}{g_{jk}}$                                      |
| Weighted betweenness centrality (or flow betweenness) | Traffic flow through node i between nodes j and k divided by maximum possible traffic between j and k | $C_i^F = \frac{\sum_{j < k \in G} m_{jk}(i)}{\sum_{j < k \in G} m_{jk}}$ |

Source : adapted from various sources

### Appendix 2. Statistical description of the indicators

| Indicator                 | Mean            |            | Median          |           | St. dev.         |             | Max               |              | Min        |      |
|---------------------------|-----------------|------------|-----------------|-----------|------------------|-------------|-------------------|--------------|------------|------|
|                           | direct          | all        | direct          | all       | Direct           | all         | direct            | all          | direct     | all  |
| Vessels                   | 92.4            | 92.4       | 26.5            | 26.5      | 190.1            | 190.1       | 1857.0            | 1857.0       | 1.0        | 1.0  |
| Calls                     | 623.8           | 623.8      | 181.5           | 181.5     | 1587.3           | 1587.3      | 18198.0           | 18198.0      | 1.0        | 1.0  |
| Operators                 | 38.5            | 38.5       | 19.0            | 19.0      | 51.1             | 51.1        | 336.0             | 336.0        | 1.0        | 1.0  |
| Maximum distance          | 3595.3          | 6715.2     | 2059.1          | 9073.7    | 3611.6           | 3811.9      | 10018.4           | 10018.6      | 0.0        | 0.0  |
| Average distance          | 942.7           | 2221.6     | 461.7           | 2115.2    | 1203.4           | 1711.9      | 9811.0            | 9738.4       | 0.0        | 0.0  |
| Foreland diversity        | 81.0            | 81.0       | 71.0            | 71.0      | 32.1             | 32.1        | 277.3             | 277.3        | 50.4       | 50.4 |
| Degree centrality         | 30.3            | 156.2      | 20.0            | 130.0     | 31.7             | 117.7       | 226.0             | 610.0        | 1.0        | 6.0  |
| Nodal degree              | 2.9             | 3.1        | 2.0             | 1.0       | 4.1              | 8.2         | 39.0              | 94.0         | 1.0        | 1.0  |
| Weighted degree           | 9545765.8       | 49276316.8 | 1510825.0       | 3406765.0 | 28280603.3       | 168347922.4 | 340268827.0       | 2210600634.0 | 1849.0     | 1.0  |
| Betweenness centrality    | 6123.2          | 2270.7     | 1306.5          | 387.7     | 16402.5          | 5922.4      | 174516.0          | 83246.2      | 0.3        | 0.1  |
| Weighted betw. centrality | 16149.5         | 6961.9     | 3584.0          | 879.7     | 53817.6          | 16155.9     | 840440.0          | 135225.0     | 2.0        | 0.1  |
| Eccentricity              | 0.8             | 0.7        | 0.8             | 0.8       | 0.1              | 0.1         | 1.0               | 1.0          | 0.4        | 0.3  |
| Clustering coefficient    | 0.5             | 0.6        | 0.4             | 0.6       | 0.2              | 0.2         | 1.0               | 1.0          | 0.0        | 0.2  |
| Strahler                  | 2289.3          | 12137.6    | 2303.3          | 12600.4   | 121.5            | 2540.2      | 2381.0            | 13810.5      | 1.0        | 1.0  |
| Strength                  | 0.3             | 0.4        | 0.3             | 0.4       | 0.1              | 0.1         | 0.5               | 0.7          | 0.0        | 0.2  |
| <b>Throughput (TEUs)</b>  | <b>766433.8</b> |            | <b>158851.0</b> |           | <b>2212855.8</b> |             | <b>23192200.0</b> |              | <b>5.0</b> |      |

*Appendix 3. Results of the factor analysis with most significant ports*

| Graph of direct links (GDL) |                    |                  | Graph of all links (GAL) |                |                 |
|-----------------------------|--------------------|------------------|--------------------------|----------------|-----------------|
| Factor 1                    | Factor 2           | Factor 3         | Factor 1                 | Factor 2       | Factor 3        |
| Singapore                   | Boston(USA)        | Portland(OR USA) | Singapore                | Kudamatsu      | Varna           |
| Hong Kong                   | New Orleans        | Visakhapatnam    | Hong Kong                | Altamira       | Szczecin        |
| Shanghai                    | Tampa              | Ensenada(MEX)    | Busan                    | Vancouver(CAN) | Constantza      |
| Busan                       | Wilmington(NC USA) | General Santos   | Shanghai                 | San Francisco  | Aalborg         |
| Rotterdam                   | Bremerhaven        | Guangzhou        | Rotterdam                | Port Elizabeth | Leixoes         |
| Hamburg                     | Portland(OR USA)   | Kawasaki         | Hamburg                  | Baltimore      | St. Petersburg  |
| Kaohsiung                   | Oakland            | Vancouver(CAN)   | Kaohsiung                | Sydney         | Riga            |
| Port Klang                  | Houston            | Civitavecchia    | Antwerp                  | Fremantle      | Poti            |
| Ningbo                      | Savannah           | Kolkata          | Port Klang               | Mejillones     | Ghent           |
| Antwerp                     | Tanjung Pelepas    | Trieste          | Ningbo                   | Auckland       | Bilbao          |
| Yokohama                    | Tacoma             | Progreso         | Qingdao                  | Veracruz       | Seville         |
| Qingdao                     | Manzanillo(PAN)    | Banjul           | Bremerhaven              | Mobile         | Belfast         |
| New York                    | Vancouver(CAN)     | Salaverry        | Tokyo                    | San Vicente    | Recife          |
| Tokyo                       | Rio de Janeiro     | Kandla           | Xiamen                   | Caldera(CRI)   | Gdynia          |
| Xiamen                      | Los Angeles        | Lazaro Cardenas  | Nagoya                   | Paranagua      | Tees            |
|                             | Kolkata            | Ghent            |                          | Oranjestad     | Tacoma          |
|                             | Mazatlan           | Pietarsaari      |                          | Sakai          | Lazaro Cardenas |
|                             | Ahus               | Goole            |                          | Shanghai       | San Francisco   |
|                             | Visakhapatnam      | Camden(NJ USA)   |                          | Bergen         | Santander       |
|                             | Doha(QAT)          | Halmstad         |                          | Rotterdam      | Tanga           |
|                             | Puerto Deseado     | Bandirma         |                          | Warrenpoint    | Qui Nhon        |
|                             | Kaohsiung          | Surabaya         |                          | Melilla        | Kudamatsu       |
|                             | Cagayan de Oro     | Aalborg          |                          | Kiel           | Manzanillo(CUB) |
|                             | Salaverry          | Iskenderun       |                          | Bar            | Busan           |
|                             | Shanghai           | Palma(Maj)       |                          | Hamburg        | Eilat           |
|                             | Rotterdam          | Raahe            |                          | Almeria        | Oranjestad      |
|                             | Hamburg            | Boston(GBR)      |                          | Norrkoping     | Mejillones      |
|                             | Busan              | Tekirdag         |                          | Busan          | Puerto Deseado  |
|                             | Hong Kong          | Uusikaupunki     |                          | Hong Kong      | Hong Kong       |
|                             | Singapore          | Norrkoping       |                          | Singapore      | Singapore       |

*Source: own elaboration based on LMIU data and StatiXL software*