

Le calcul intensif en géographie : une tradition bien ancrée

Le cas de la simulation

Sébastien Rey-Coyrehourcq (UMR IDEES)

Arnaud Banos (UMR IDEES)



@reyman64



<https://slides.com/sebastienreycoyrehourcq/calcul-shs-paris/>

ideas

16 avril 2019 - Paris

Aperçu des pratiques historiques ...

simulation & centre de calculs



chapitre de thèse [Rey-Coyrehourcq 2015]
stage 4 mois 2019 : "Pérénisation & valorisation code sources des pionniers de la quantitative en géographie"

Simulation en géographie

Les pionniers Américains et Suédois

1950's - 1960's

Univ. Washington
(Marble, Pitts, Tobler, etc.)

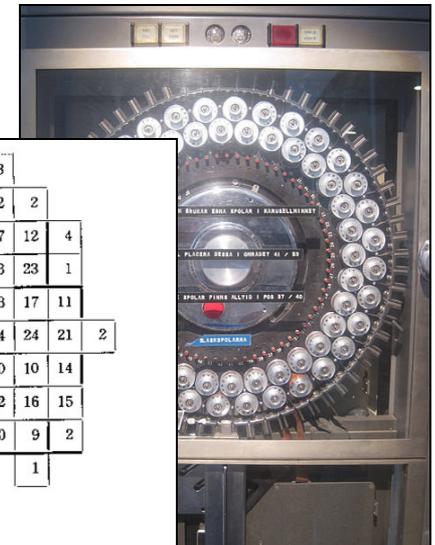
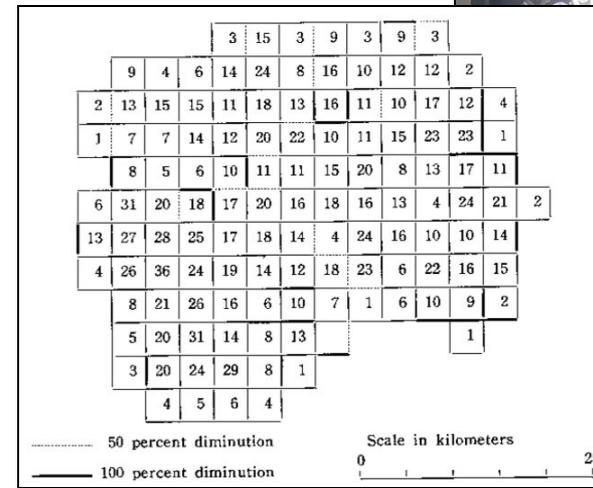


T. Hagerstrand

1956 / SMIL at Lund



Washington University **IBM 650**
& Northwestern's Vogelback computing
center **CDC 3400/ 6400**



A monte carlo simulation of diffusion

[ex **Hager III** and **Hager IV** Fortran program]

Simulation en géographie

Les pionniers au Royaume-Uni
1970's & 1980's

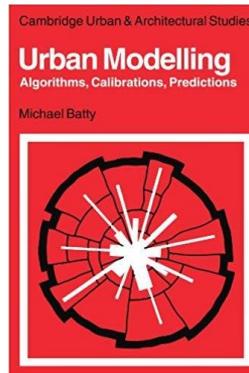
Urban Modelling

H. Couclelis
M. Batty
A. Wilson

...



pratiques peu documentées



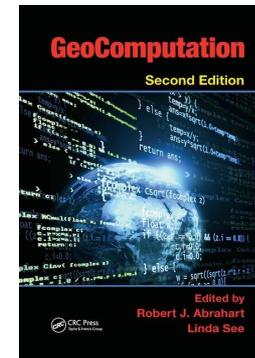
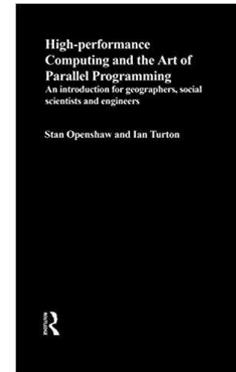
M. Batty 1976

Equipe de Leeds

" GeoComputation is [...] meant to imply the adoption of a large-scale computationally intensive scientific paradigm as a tool for doing all manner of geographical research."

S. Openshaw
R. Abrahart
I. Turton

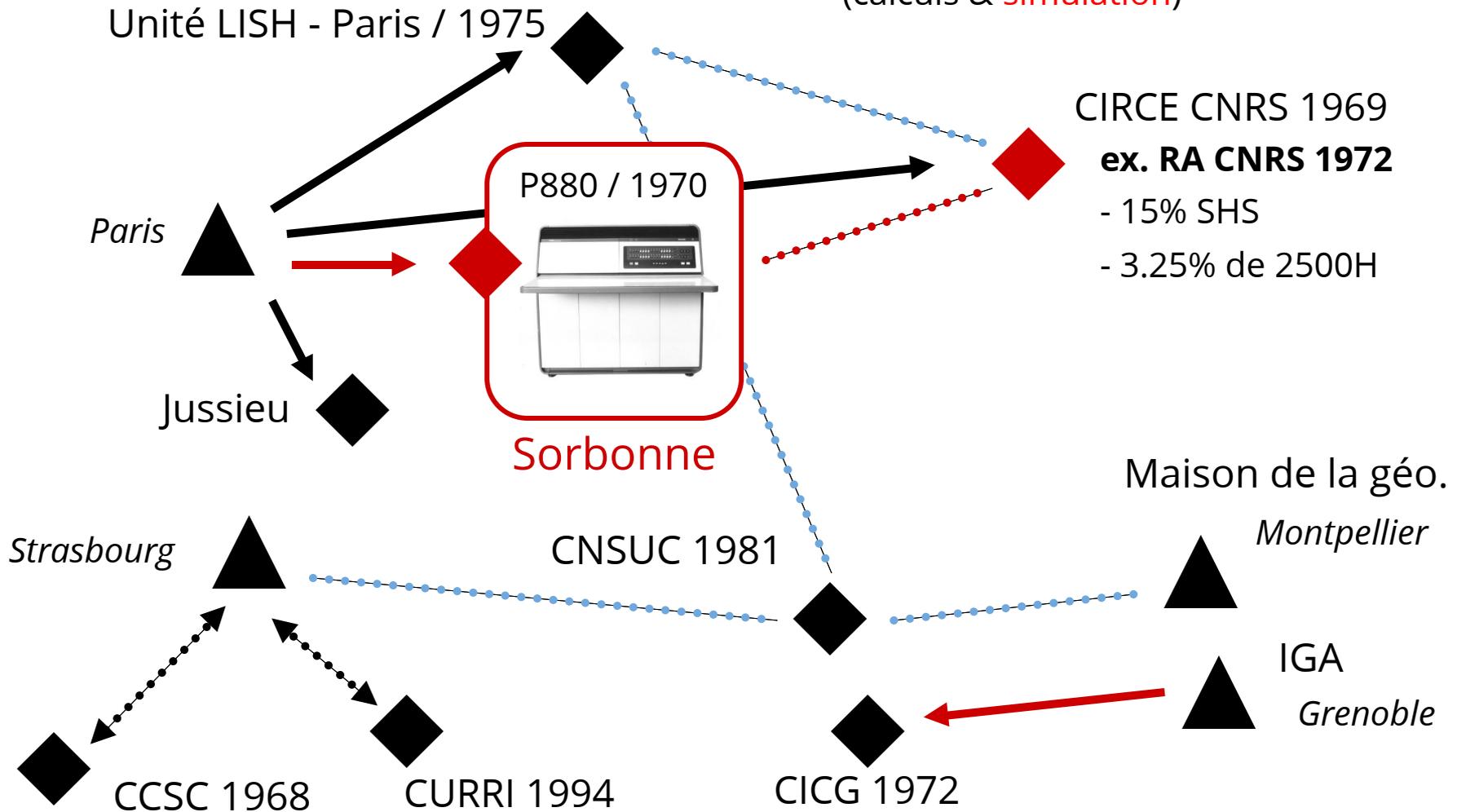
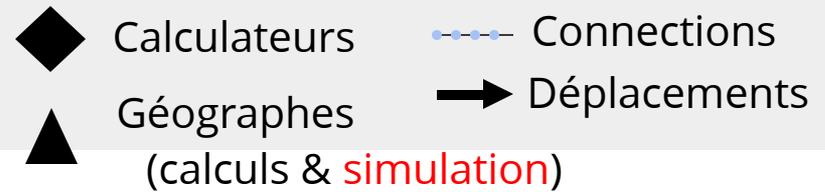
...



HPC & Geocomputation - 2000

Simulation en géographie

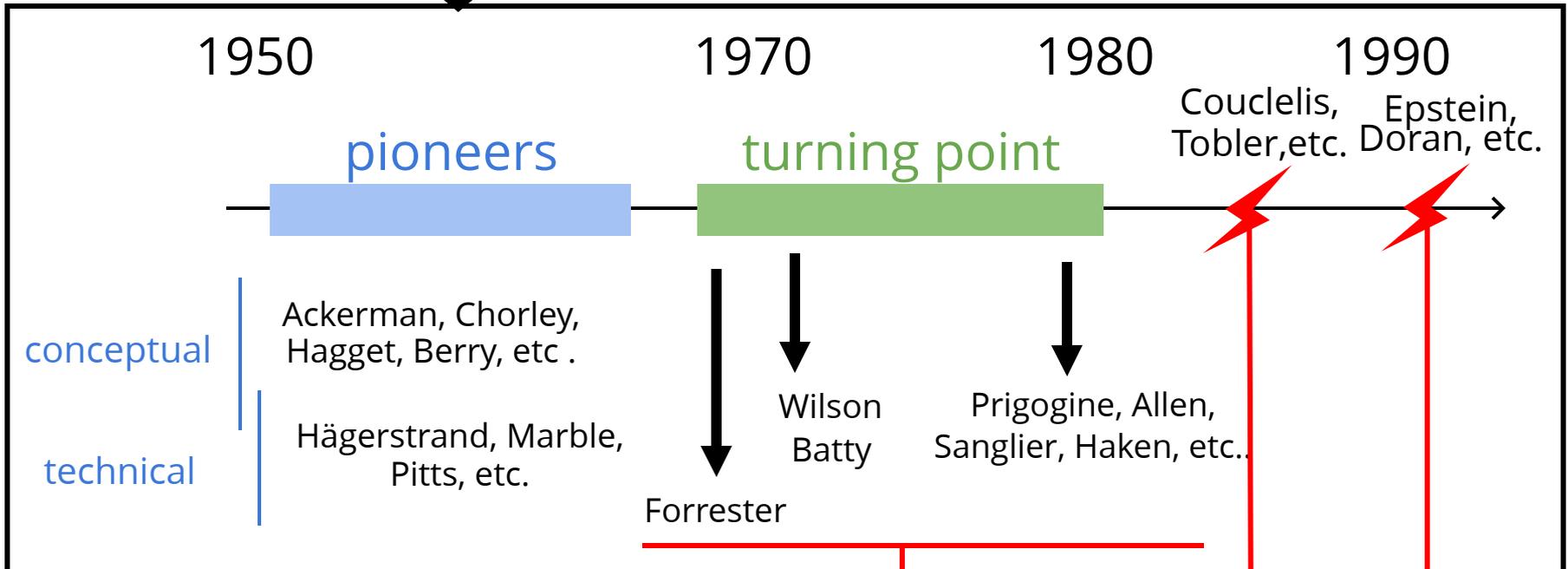
Les pionniers en France
1970's - 1990's



Simulation en géographie

[Rey-Coyrehourcq 2015]

Influences répétées du **projet systémique** (Von Bertalanfy et cie.)
open systems, retro-action, self-organization, equifinality, bifurcation, etc. [Pouvreau 2013]



~ 1980 in France



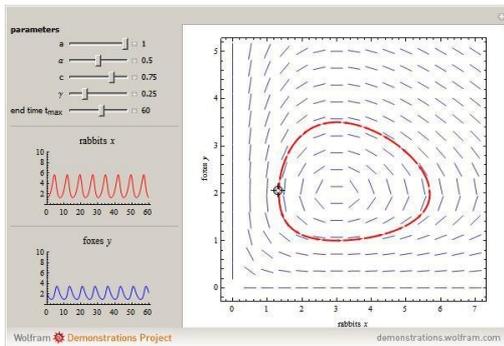
Pumain, Sanders, etc.
 Guermond, Langlois, etc.
 Chamussy, Le Berre, Uvietta, etc.

Non Linear Dynamical Systems

Cellular Automata

Agent Based Modelling

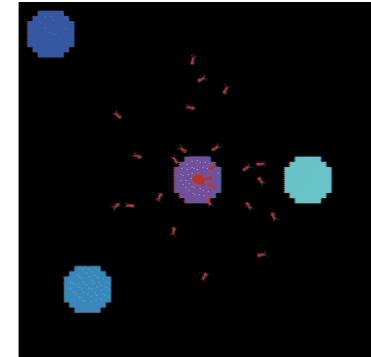
Modèles dynamiques



proie prédateur



jeu de la vie



fourmis

Systemes dynamiques

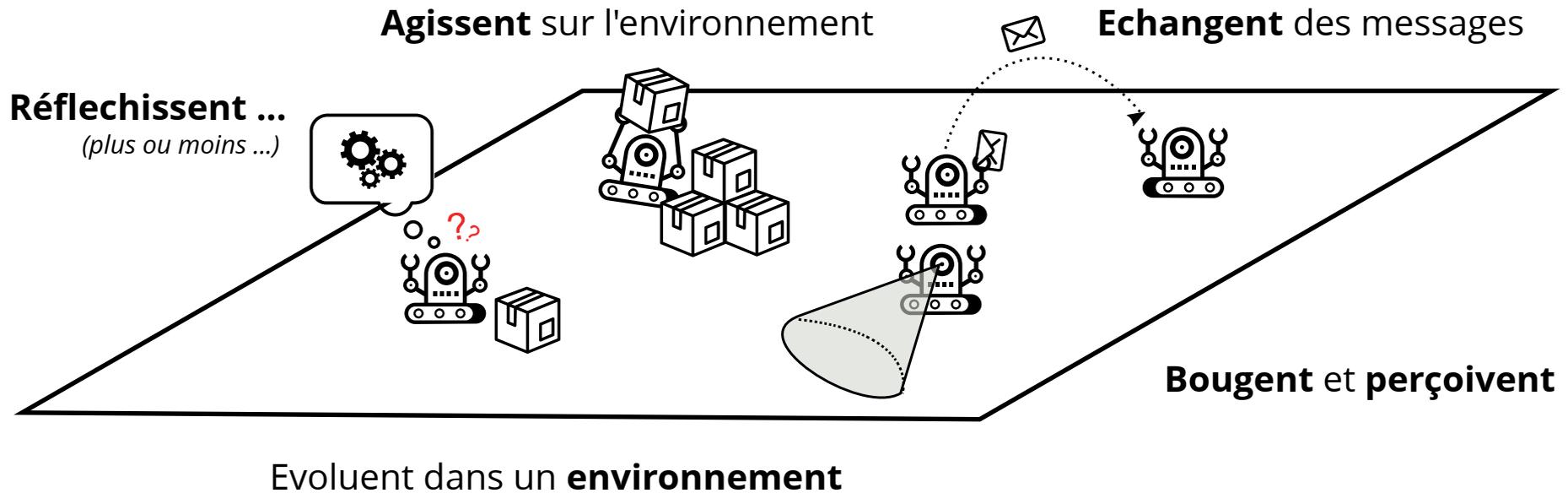
Automates
Cellulaires

Modèles multi-agent

Couplage possible de ces trois approches

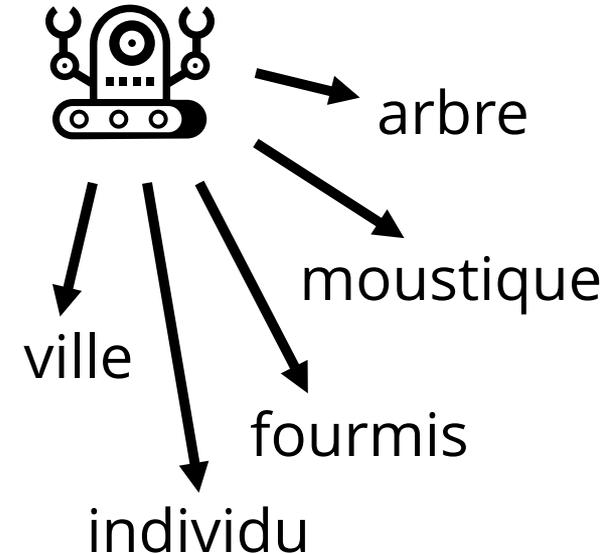
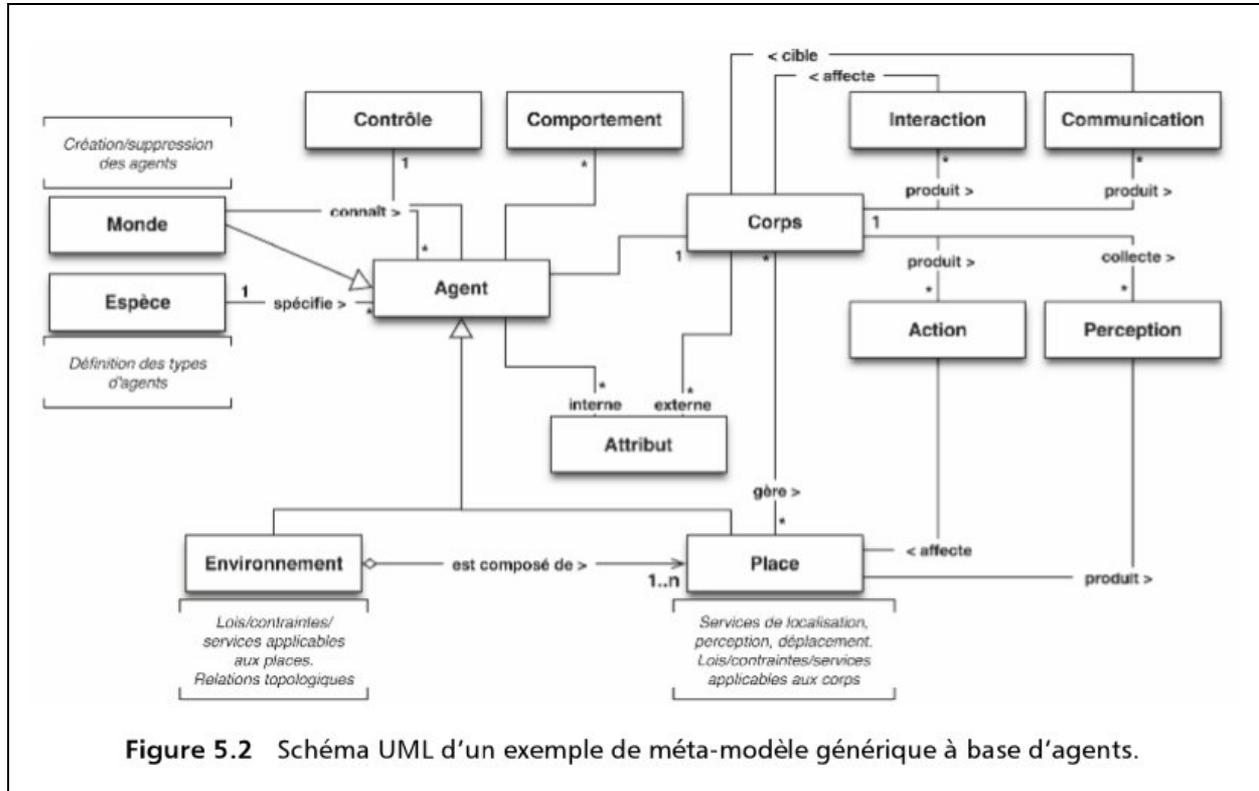
Modèle multi-agents

[Ferber 1999, Drogoul 2009]



Modèle multi-agents

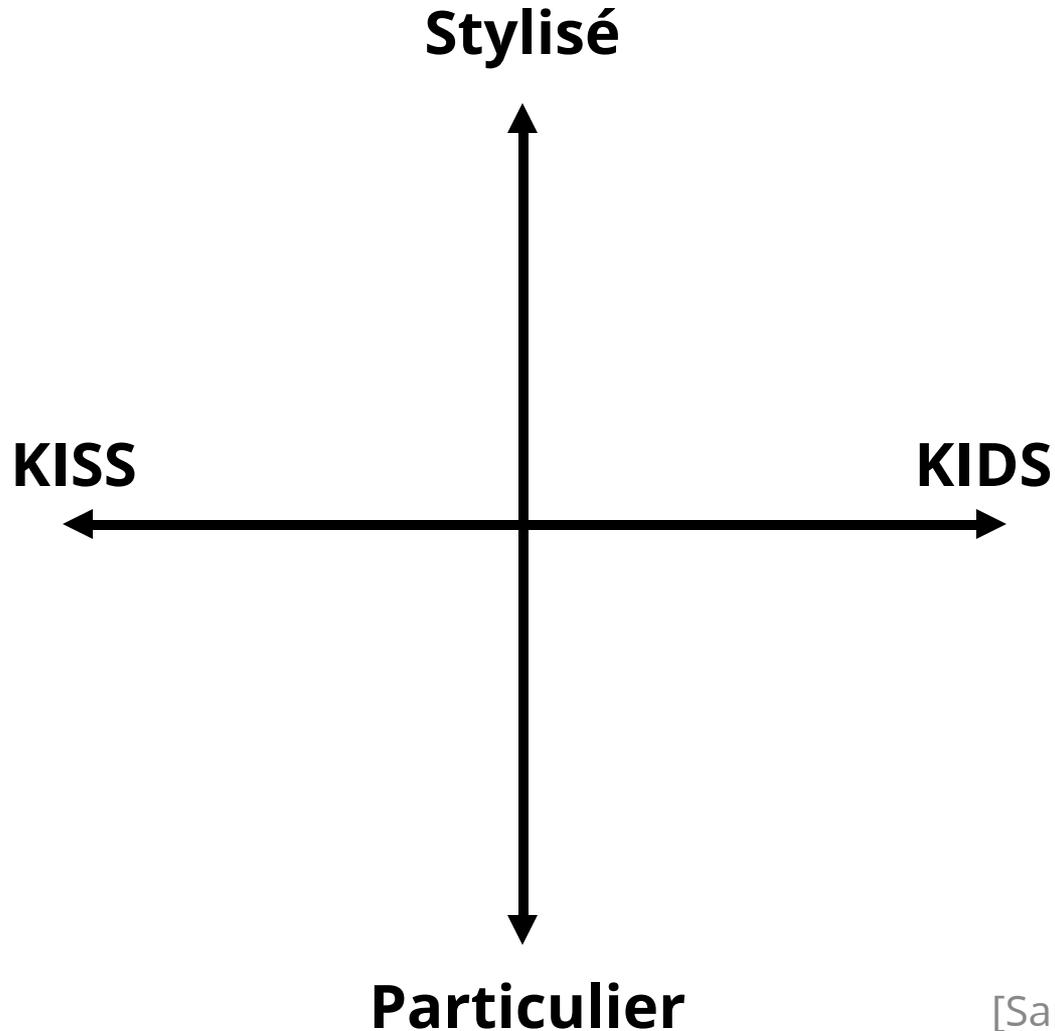
Un **méta-modèle** plus qu'un modèle



[Drogoul Treuil 2008]

flexibilité de représentation &
emboîtement d'**échelle** !

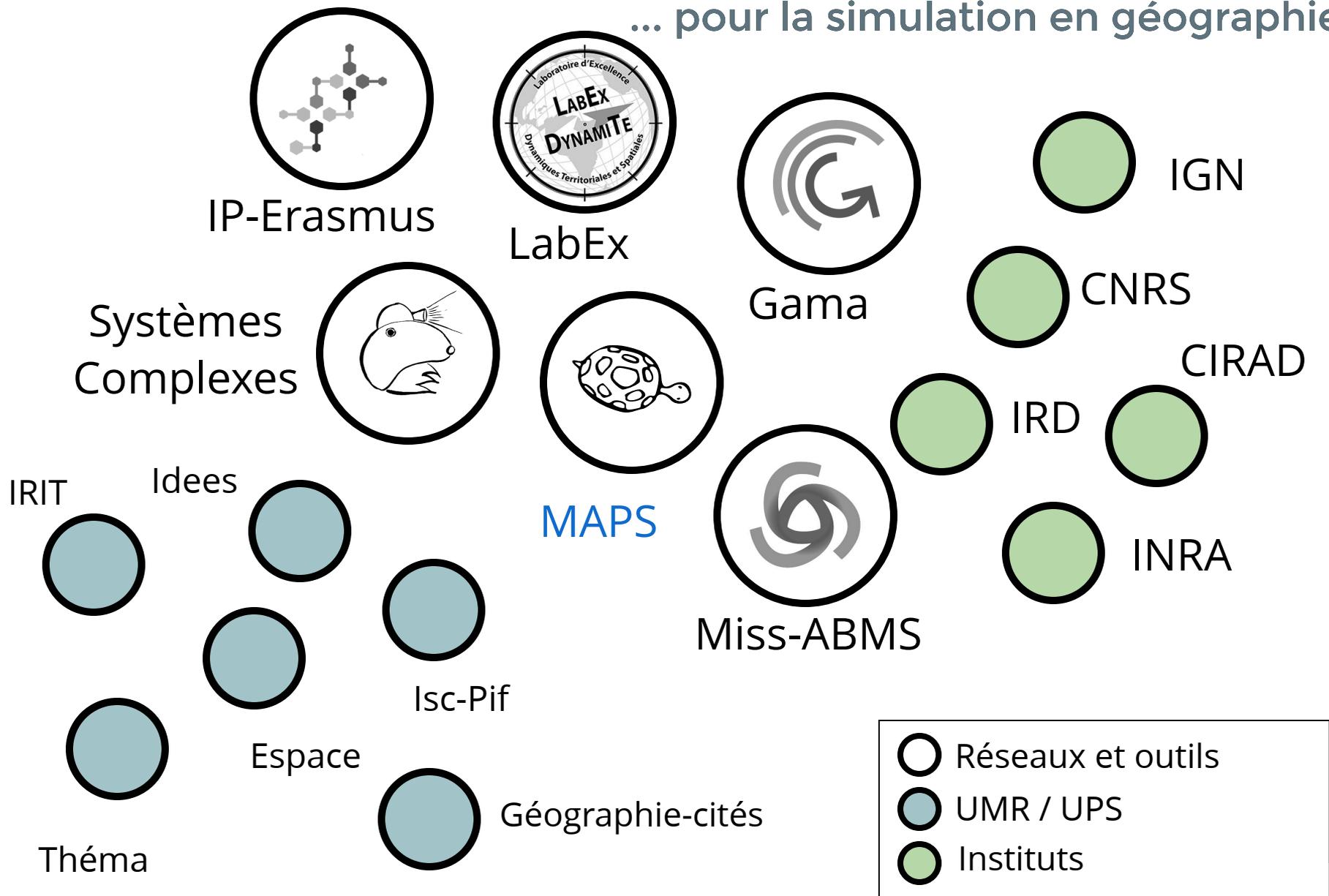
Une variété de pratiques



[Sanders & Banos 2011,
2013, Banos 2013]

Une diversité d'acteurs ...

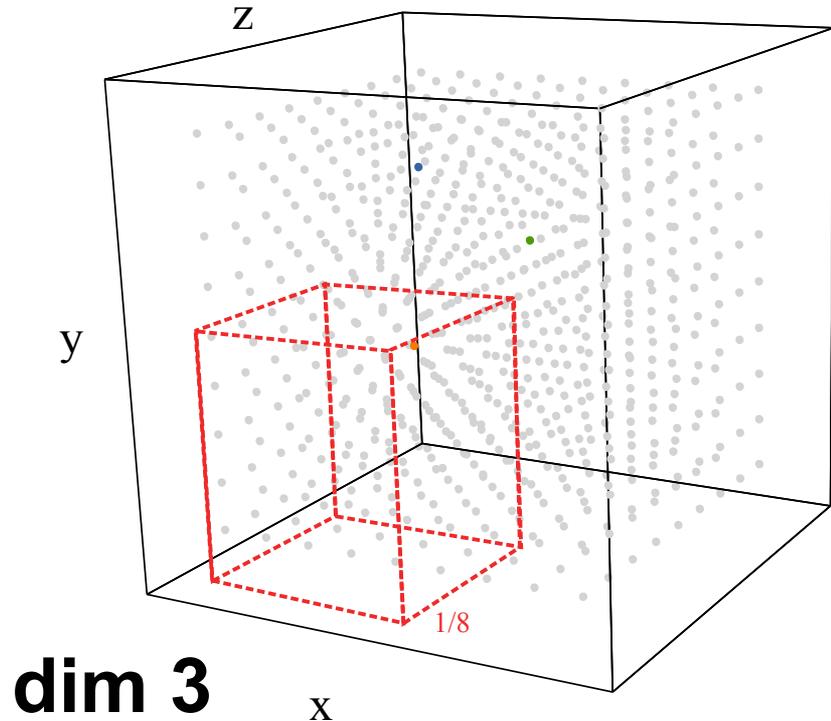
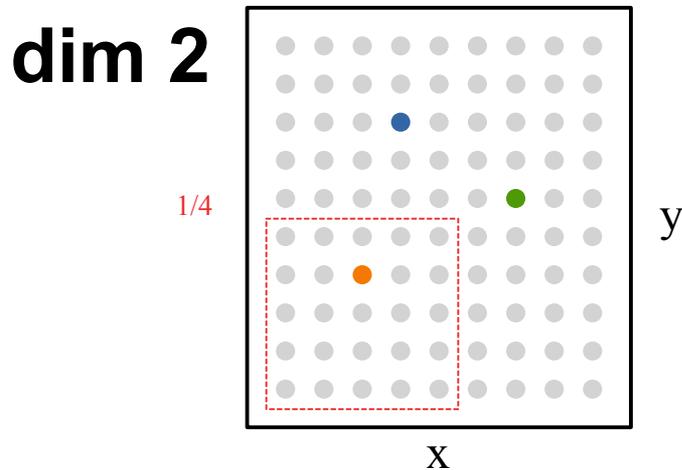
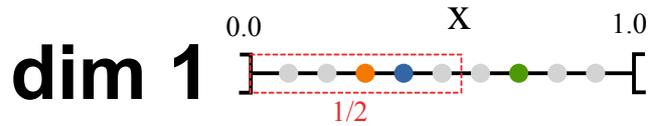
... pour la simulation en géographie



**Besoins en "calcul" ?
Quelles solutions pour la simulation**

Espace des paramètres

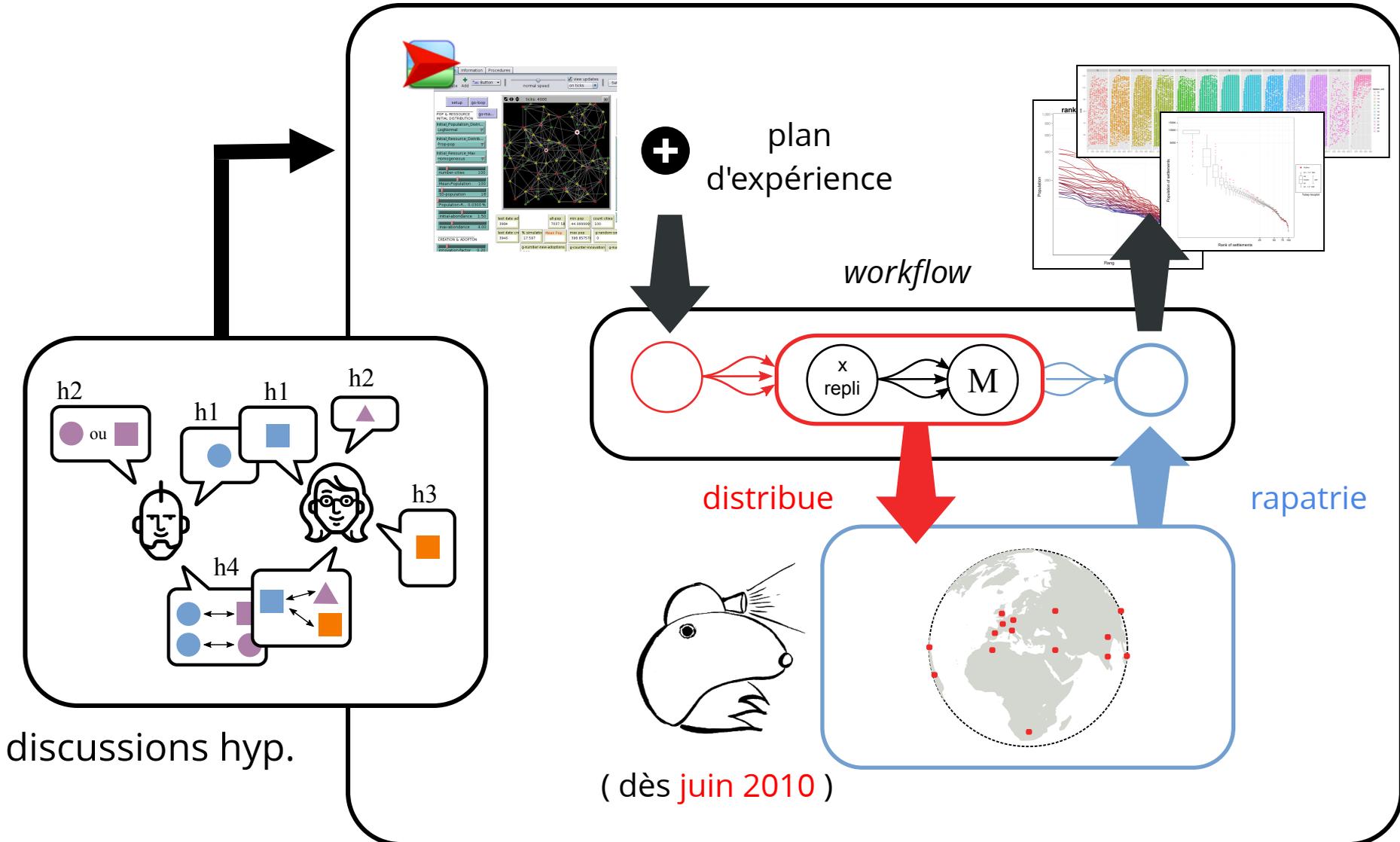
... curse dimensionality ...



+ dimensions = **-** couverture à nb points égal

Accès aux ressources

... de nombreux aller retours entre modèles et résultats ...





OpenMOLE

next.openmole.org

iscipif.fr

Une plateforme et une communauté **en expansion**

Créer des "chaînes de traitements"
parallélisables, reproductibles, simples (ou complexes)
sans **aucune connaissance** des technologies **HPC**

```
val i = Val[Double]
val res = Val[Double]

val exploration =
  ExplorationTask(i in (0.0 to 100.0 by 1.0))

val model =
  ScalaTask("val res = i * 2") set (
    inputs += i,
    outputs += (i, res)
  )

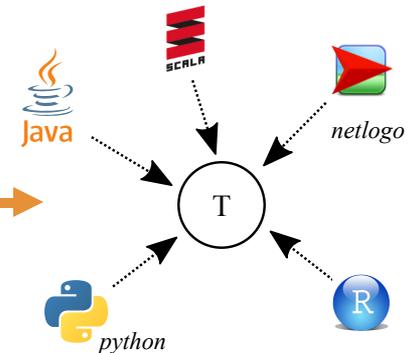
val env = LocalEnvironment(4)

exploration -< (model on env hook ToStringHook())
```

Langage de scripts (DSL) **simple**
et flexible pour composer les
workflow

Plusieurs types de plan
d'expériences : **optimisation &**
exploration.

Ecosystème riche:
simulation, etc.



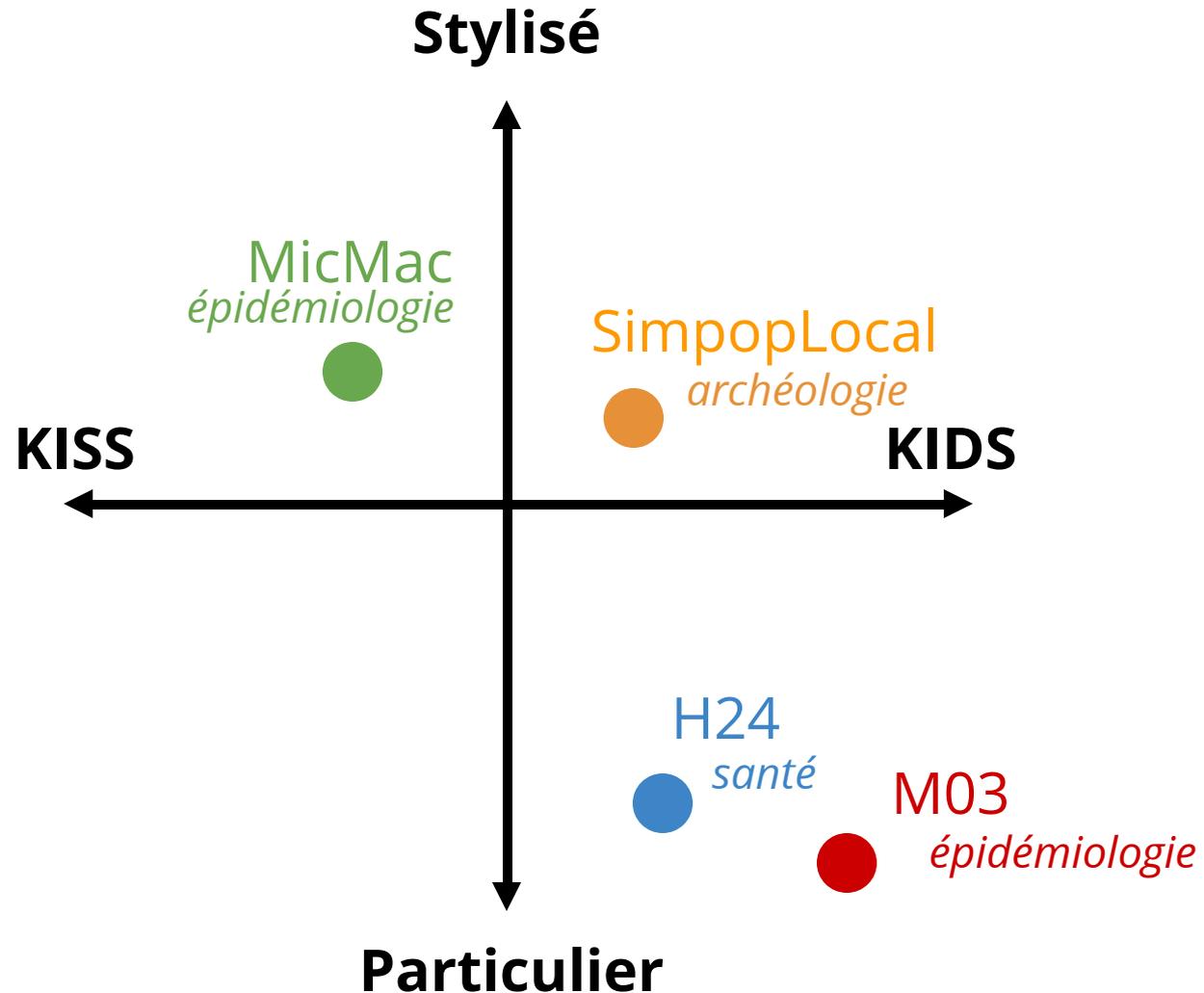
Support de multiples
technologies HPC & Grilles

Examples

4 exemples



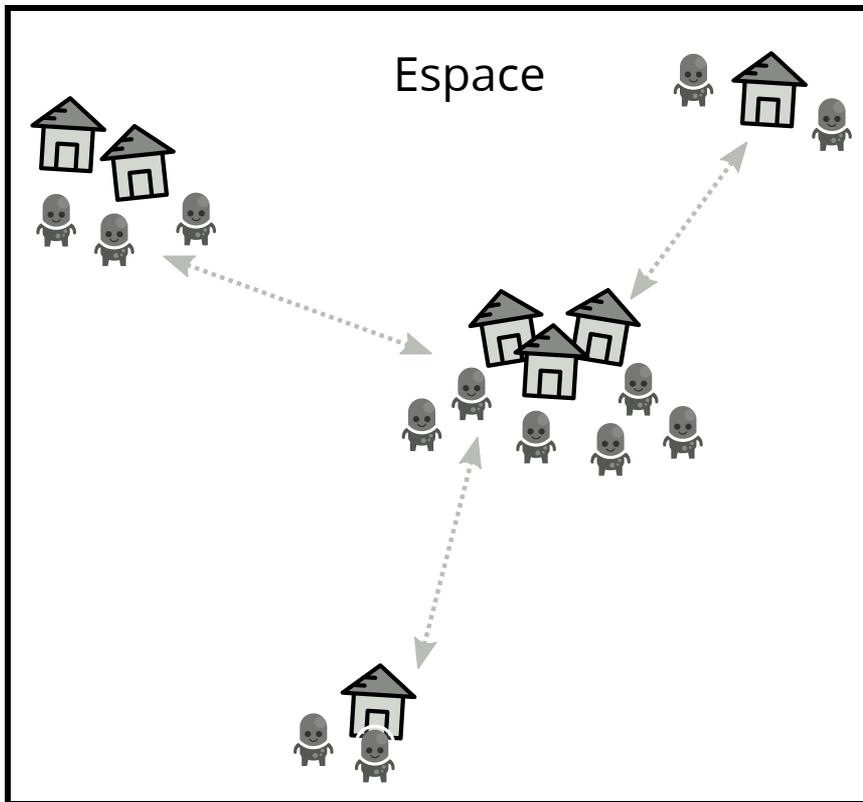
& une grande diversité de besoins en calculs : **Grille & HPC !**



SimpopLocal

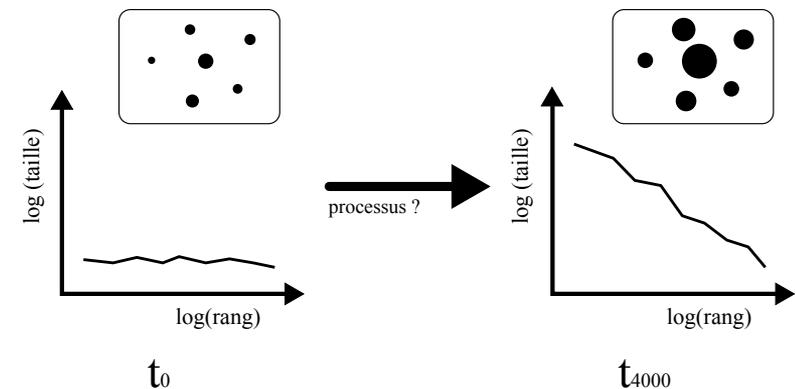
[Schmitt 2014,
Schmitt et al. 2015]

Question : La production et la diffusion de l'innovation est il un mécanisme nécessaire pour produire une **hiérarchisation** crédible du système de peuplement en 4000 ans au **néolithique** ?



système de peuplements

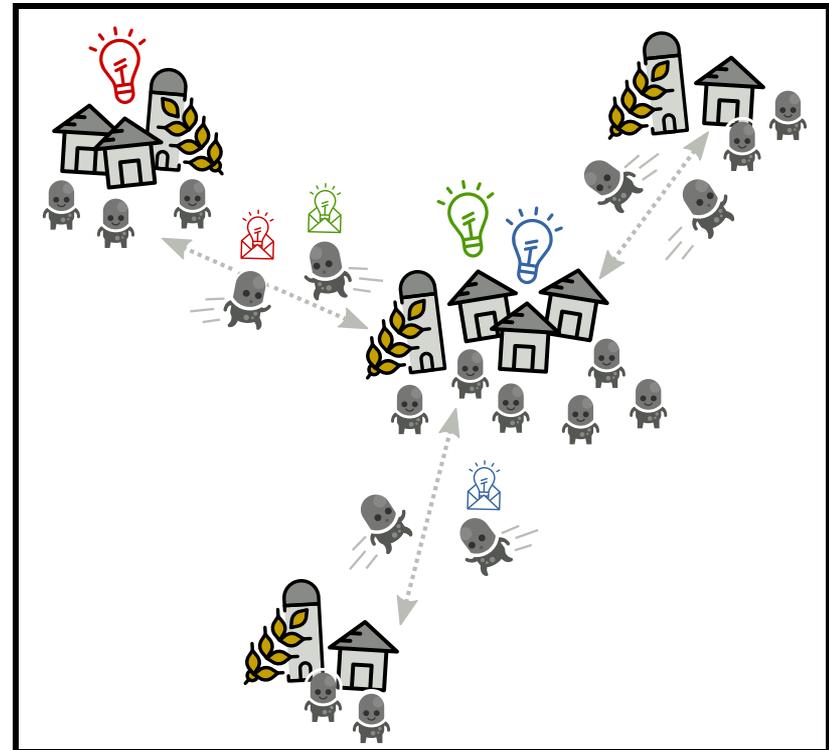
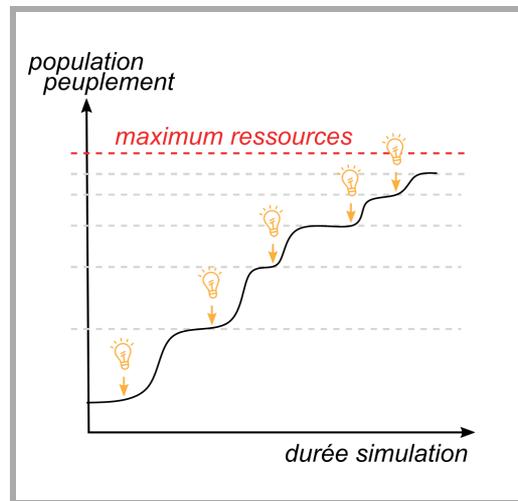
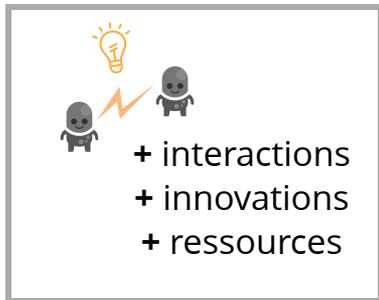
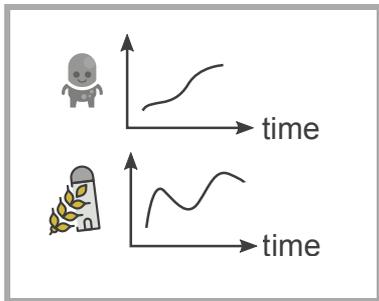
agents: villes & innovations
7 paramètres libres
4 mécanismes
très peu de données en entrée/sorties
1 loi empirique à reproduire



SimpopLocal

Question : La production et la diffusion de l'innovation est il un mécanisme nécessaire pour produire une **hiérarchisation** crédible du système de peuplement en 4000 ans au néolithique ?

Mécanismes du modèles

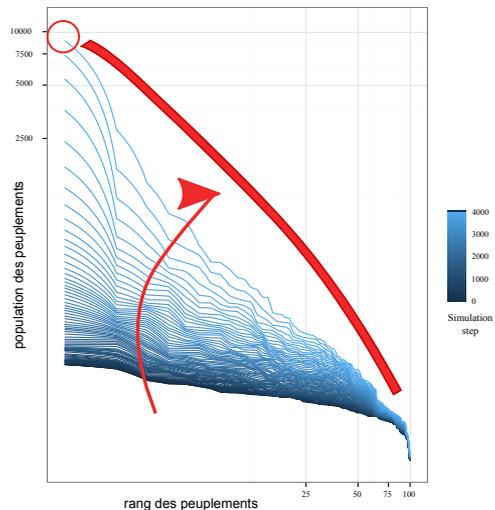


SimpopLocal

Calibration

NSGA 2 - Ilôts

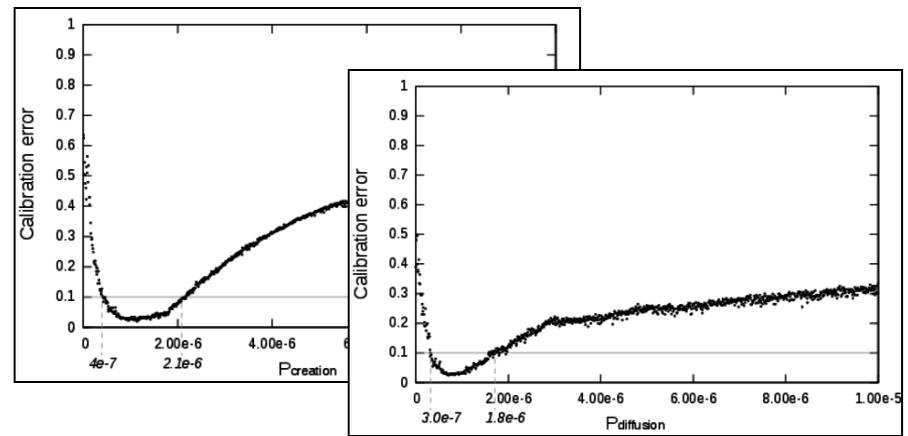
3 objectifs : (temps, forme, taille)



Analyse de sensibilité

CP-Profile

1 objectif : $f = \max(01,02,03)$



obj. f est **impossible** si 0 innovations

1 simulation = 1 coeur
1 Go Ram
~ **5/10** minutes d'exécution
500 millions de simulation
eq. 20 ans de calculs
4000 cpu EGI Biomed

→
réécriture modèle

1 simulation = 1 coeur
1 Go Ram
~ **1/2** minutes d'exécution
eq. 12 ans calcul **par profil**
15 jours de calculs (profil x 6)
4000 cpu EGI Biomed

MicMac

[Banos et al. 2015]

~ **5/10** minutes d'exécution

1 simulation par CPU

agents = personnes & avion & villes ;
couplage avec ODE

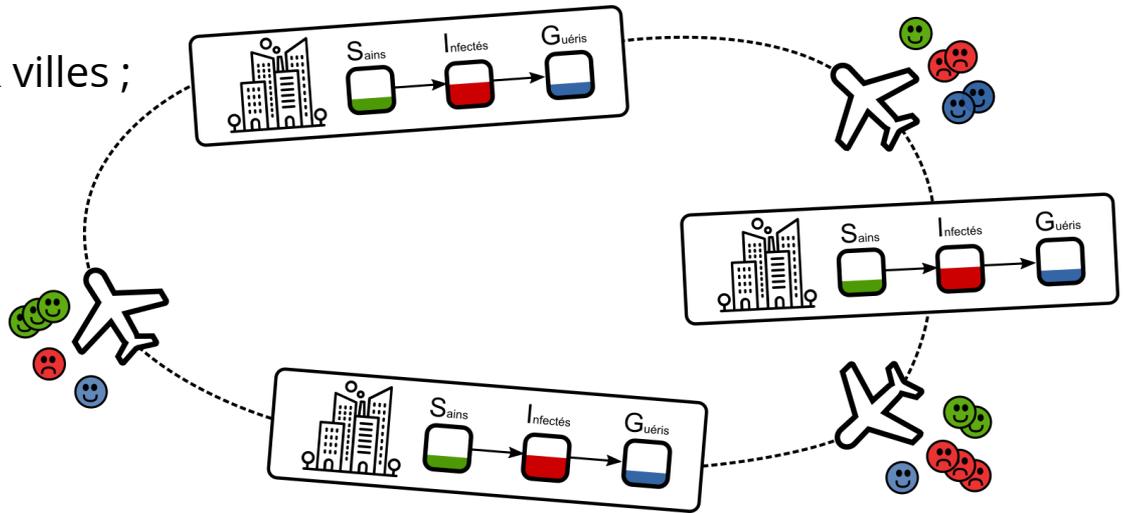
stochastique

4 stratégies à explorer

2 objectifs à minimiser

pas de données

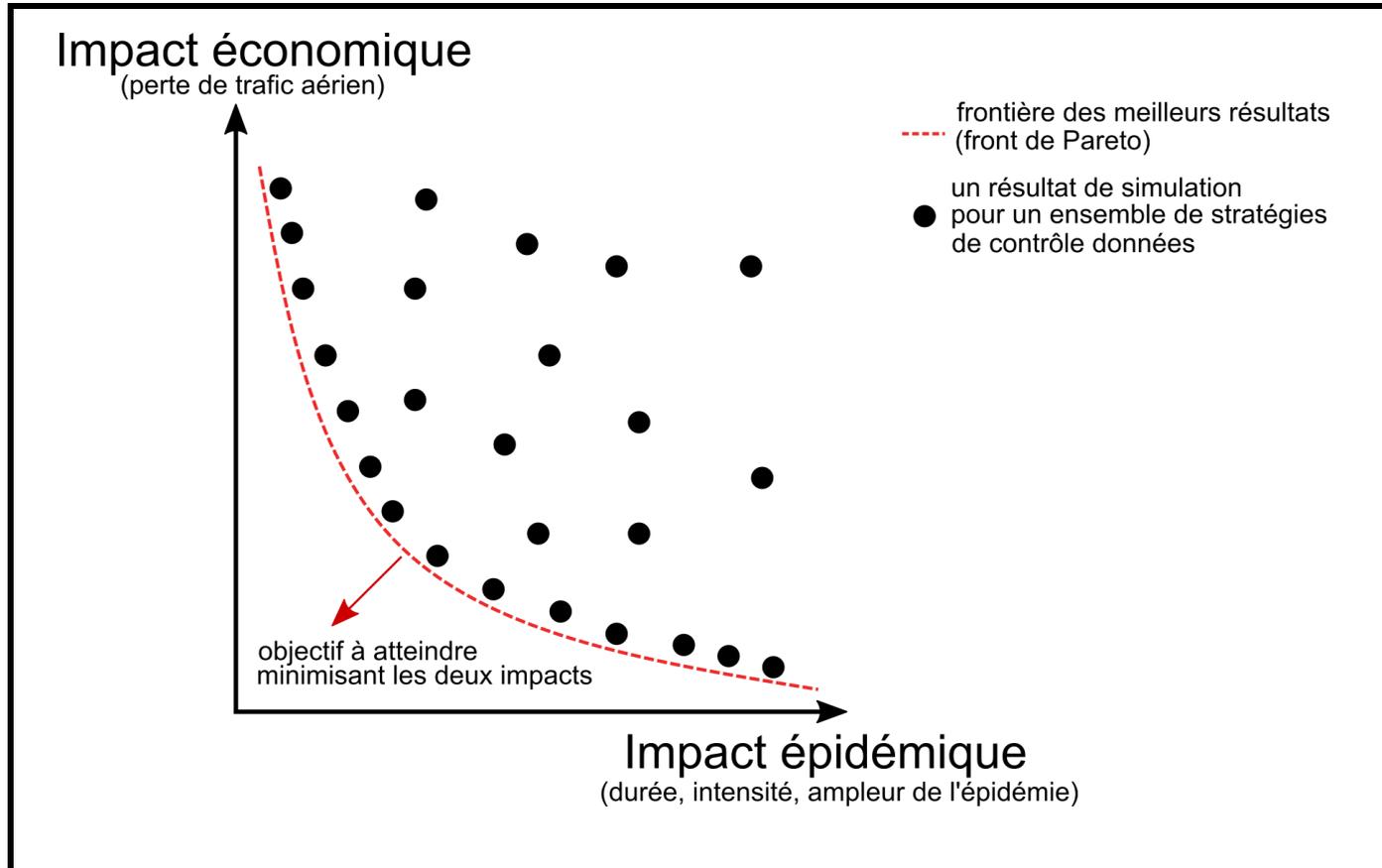
$$\begin{cases} \frac{dS}{dt} = -\beta IS \\ \frac{dI}{dt} = \beta IS - \alpha I \\ \frac{dR}{dt} = \alpha I \end{cases}$$



Question : Quel est la combinaison de **stratégies (quarantaine, évitement, responsabilisation ind., protection ind.)** qui permet de limiter à la fois les pertes économiques, et la diffusion de l'épidémie dans un tel système ?

MicMac

[Banos et al. 2015]



h24



Etudier l'effet de la segregation sociale dans l'évolution des comportements de santé sur une journée type de 24h.

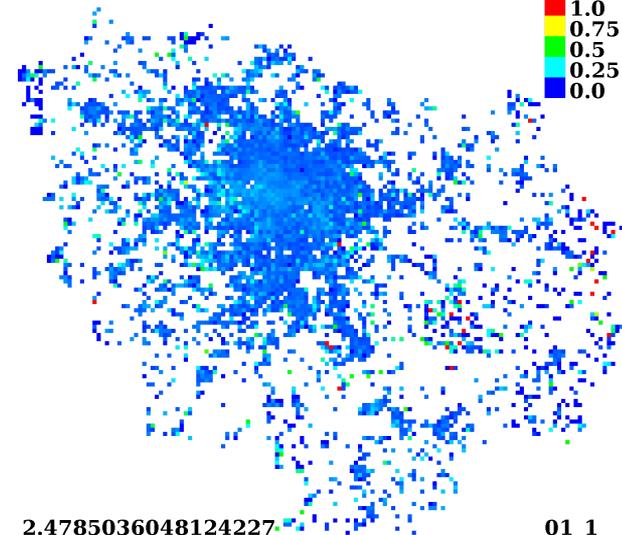
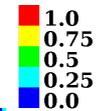
[Cottineau et al. 2017]

~ **25** minutes d'exécution
3 Go ram par simulation
1 core, mais **multithreading possible**
test : 100000 execution
calibrage: 200000 **execution**

agents: personnes / 8.16 M agents mobiles, en interaction
résolution: IDF à 1km x 1km
données: nombreuses, et à différentes dates (NPPS, CSP 18 catégories, OD, baromètre santé, etc.)

1 modèle changement d'opinion
5 paramètres
3 objectifs à minimiser

% personnes en bonne santé



2.4785036048124227

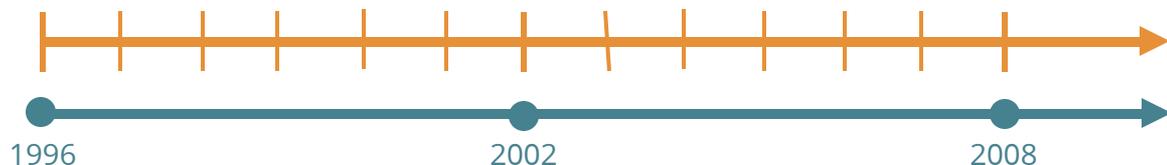
01_1

mesure d'inégalité entre les extrême des groupe d'éducation à sexe et âge égal

3 déplacements / jour (breakfast, lunch, diner)



simulated
data



steps

years

h24

2 scénarii à comparer :

	Attribution résidence aléatoire	Attribution résidence & mobilité via les données
pas de mobilité (résidence)	Scenario 1	Scenario 2
mobilité journalière (nuit,matin,soir)	Scenario 3	Scenario 4

via 3 objectifs à minimiser :

Obj 1. : MSE sum **Scenario 1**

$$\sum_{i \in (2002, 2008)} MSE_i$$

Obj. 2 : MSE sum **Scenario 4**

Obj. 3 : Le delta sur la mesure d'**inégalité (SI)** entre S1 et S4

$$\Delta(SI_{random}, SI_{mobility})$$

Question : Arrive-t-on à **reproduire** de façon cohérente les données et l'inégalité de comportement de santé fonction des groupes sociaux avec un scénario **qui ne tient pas compte de la mobilité ?!**

=> progression par falsification

Objectif : Comprendre et lutter contre l'émergence et la diffusion de la **Dengue** dans les milieux urbains

Couplage de 3 modèles en 1; *travail en cours*

**Moma**

[Maneerat et al. 2017]

**Mode**

[Misslin 2017]

**Momos**

[Cebeillac et al. 2017]

agents = moustiques & nids
résolution inf. 1m
exécution : + 12 heures
+ 50 paramètres
1 simulation = 1 cpu
16 Go Ram
DOE : LHS sur grille EGI
500000 sim, ~20 ans calculs

exploration

Micro

paysage synthétique**résolution** inf. 1m,
quartier 10km²*exploration**modèle surrogate*

Macro

indice env.**résolution** 30m,
ville entière

agent: individu mobile avec agenda
résolution 30m, ville entière
données: twitter

Questions ?



<https://slides.com/sebastienreycoyrehourcq/calcul-shs-paris>



Lien vers la bibliographie Zotero

OpenMOLE

Coding Camp OpenMOLE
27/05/2018 - 01/06/2018



- cap des 10.000 *commits*
- x 2 participants en 2018
- Amélioration de la documentation
- Ouverture à de nouvelles méthodes :
 - méthodes d'exploration *OSE*
 - générateur de grilles spatiales
 - 2 nvx cas d'utilisation (santé, éco)
- passage à *uDocker*
- gestion et déploiement multi-utilisateurs !!
- etc.

- création d'une SCIC "*Trampoline*" adossée à l'isc-pif
- attraction de nouvelles communautés scientifique
- 2 session coding-camp en 2019 : dev & utilisateurs