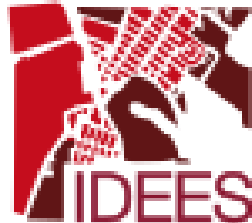


Le **Calcul** pour la **simulation** des **systems complexes** en **géographie**

histoire, enjeux, exemples

Sébastien Rey-Coyrehourcq (UMR IDEES)

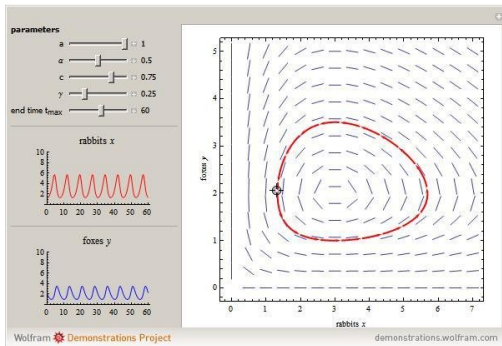
Arnaud Banos (UMR Géographie-cités)



4 juin 2018 - Caen

Historique des pratiques

Modèles dynamiques



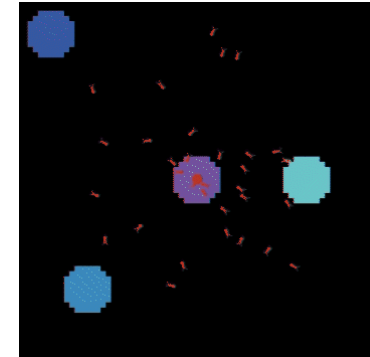
proie prédateur

Systemes dynamiques



jeu de la vie

Automates
Cellulaires



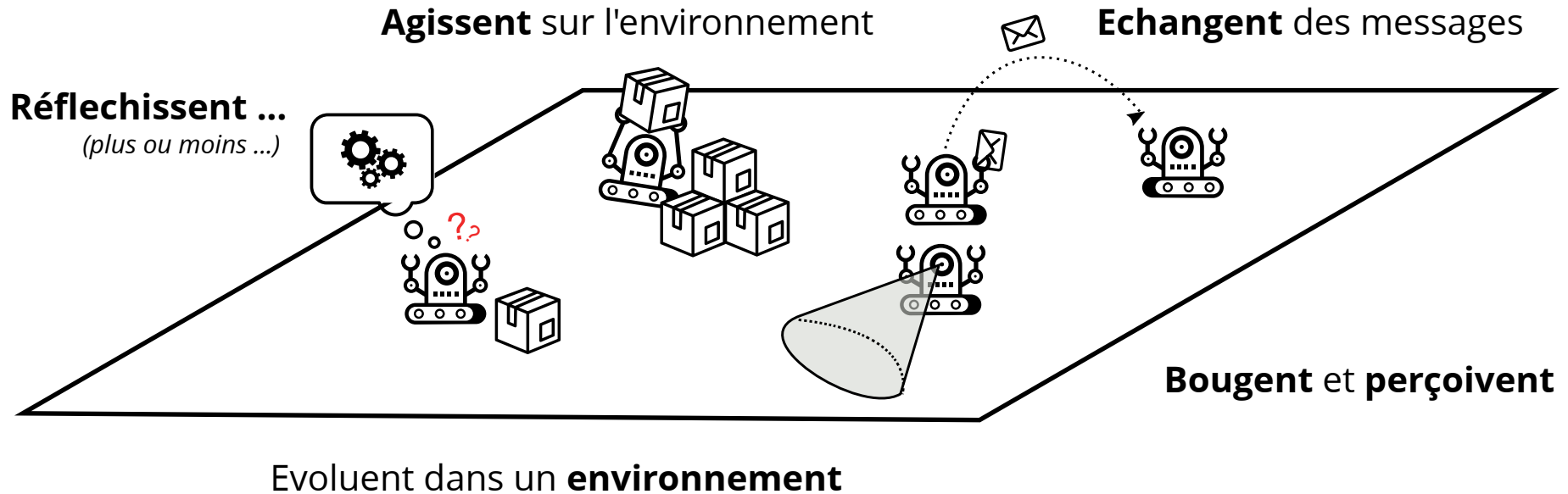
fourmis

Modèles multi-agent

Couplage possible de ces trois approches

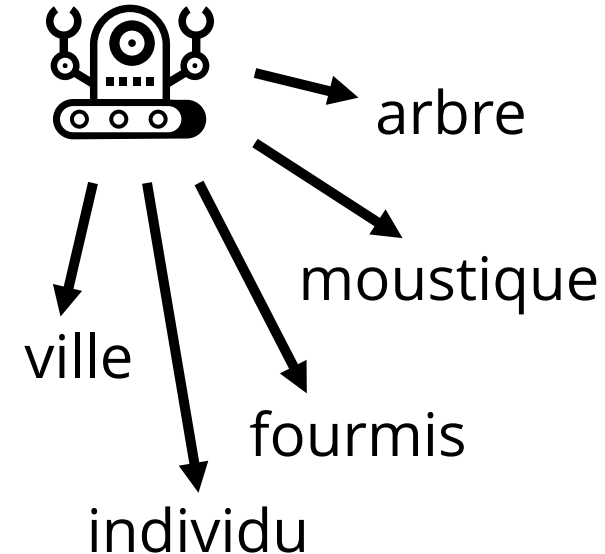
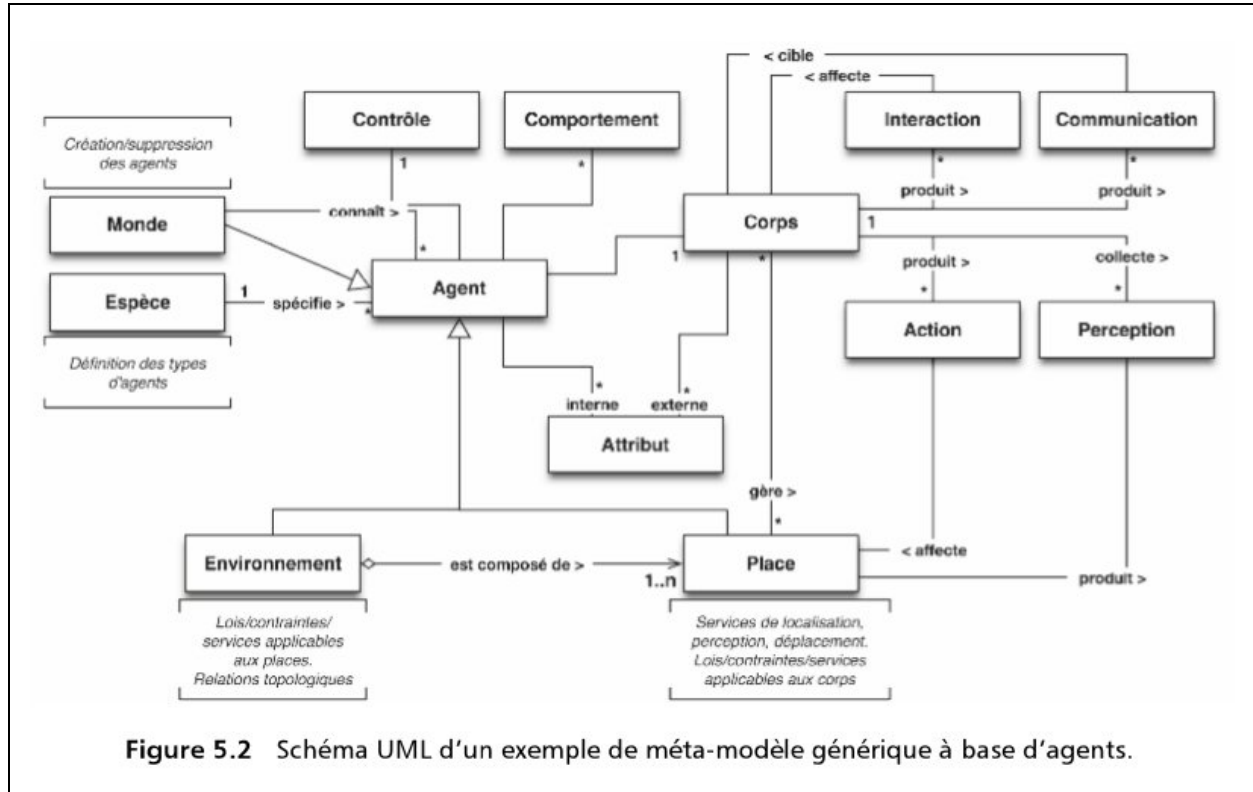
Modèle multi-agents

[Ferber 1999, Drogoul 2009]



Modèle multi-agents

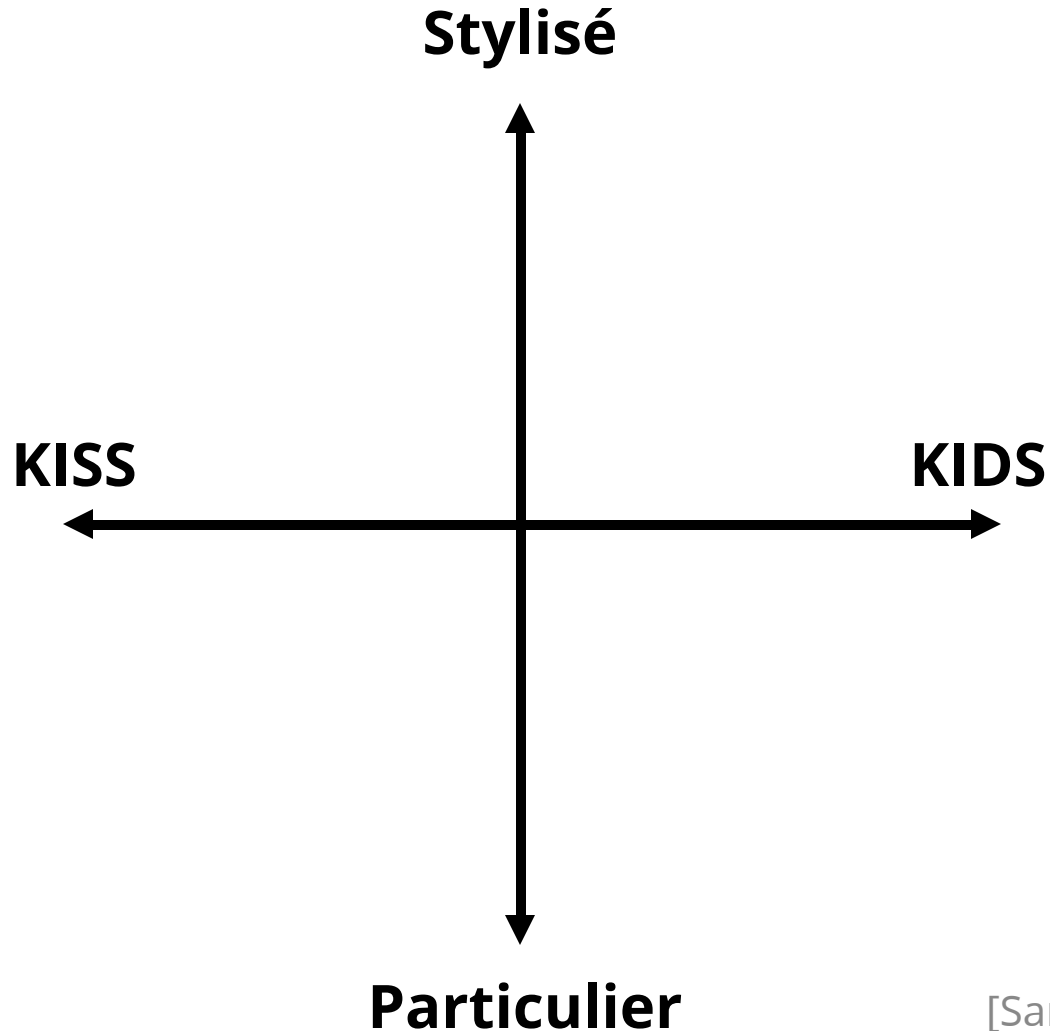
Un **méta-modèle** plus qu'un modèle



[Drogoul Treuil 2008]

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

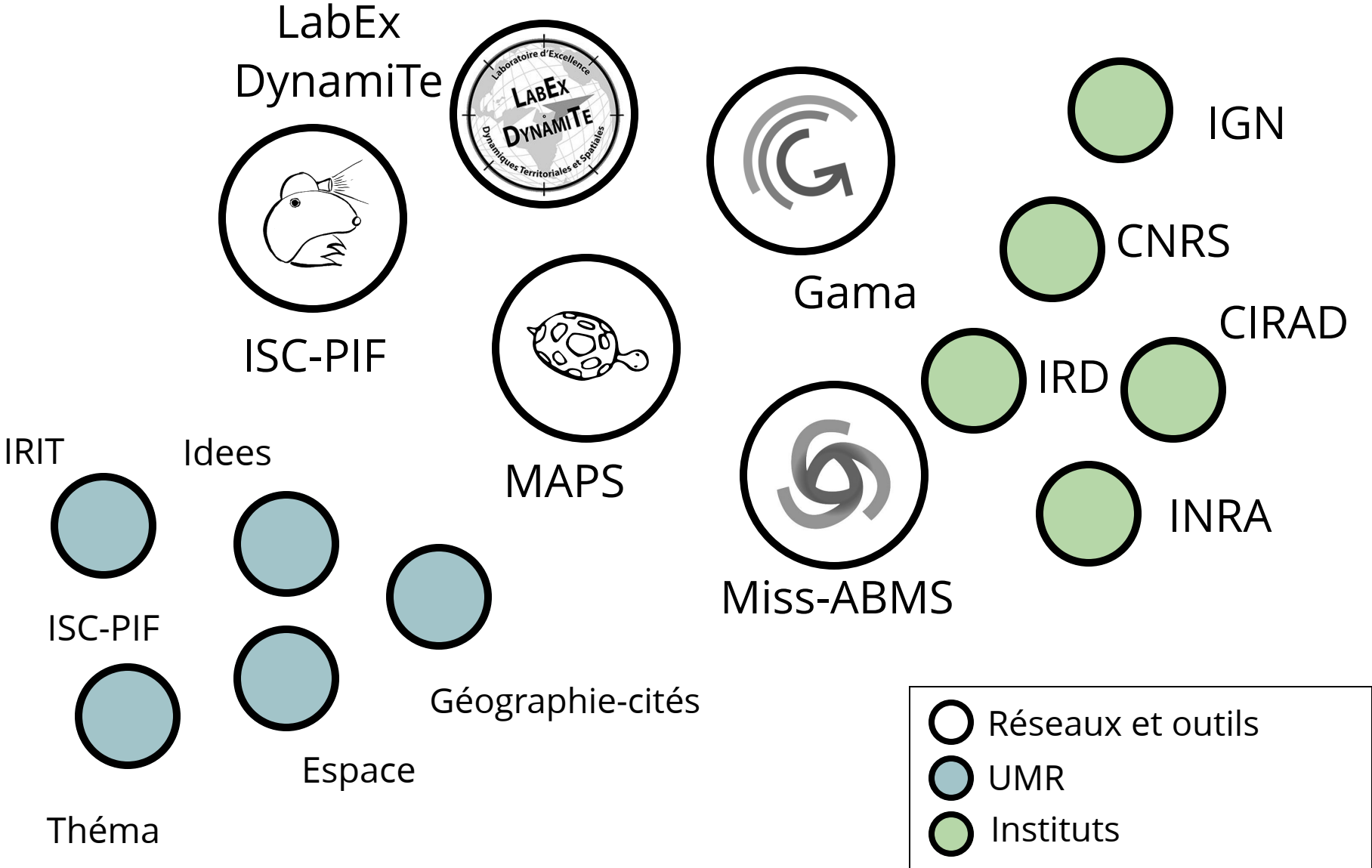
Une variété de pratiques



[Sanders & Banos 2011,
2013, Banos 2013]

Une diversité d'acteurs ...

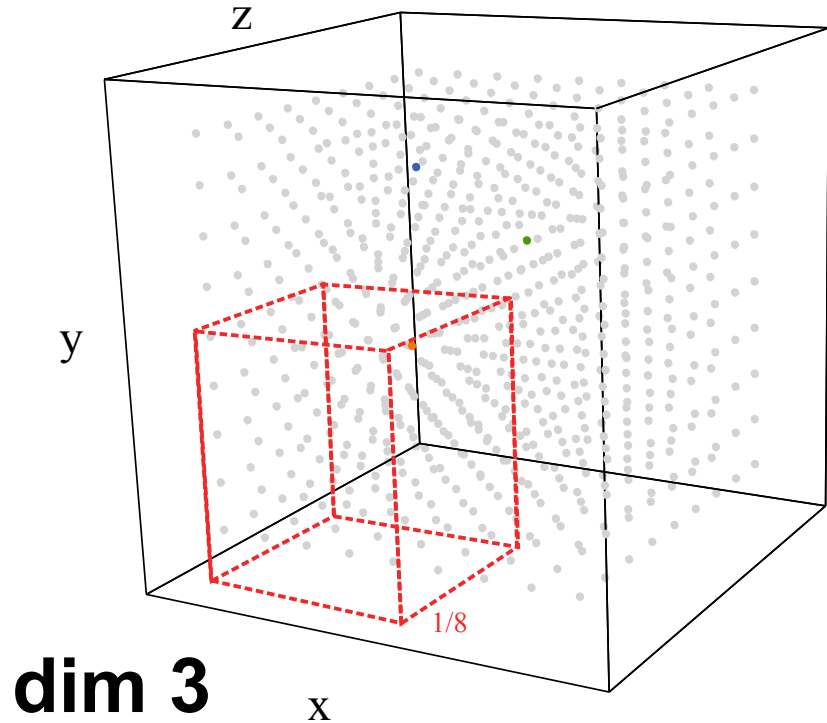
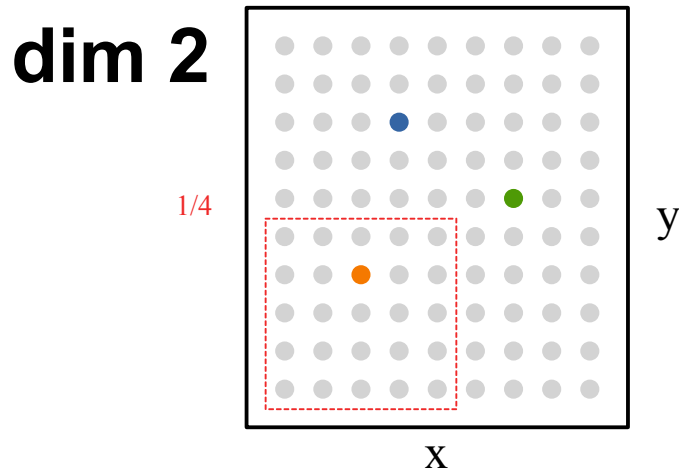
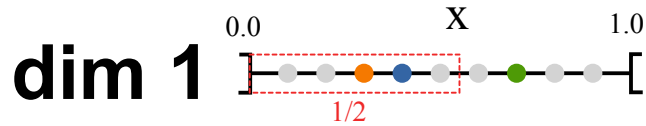
... pour la simulation en géographie



	Réseaux et outils
	UMR
	Instituts

Besoins en "calcul"

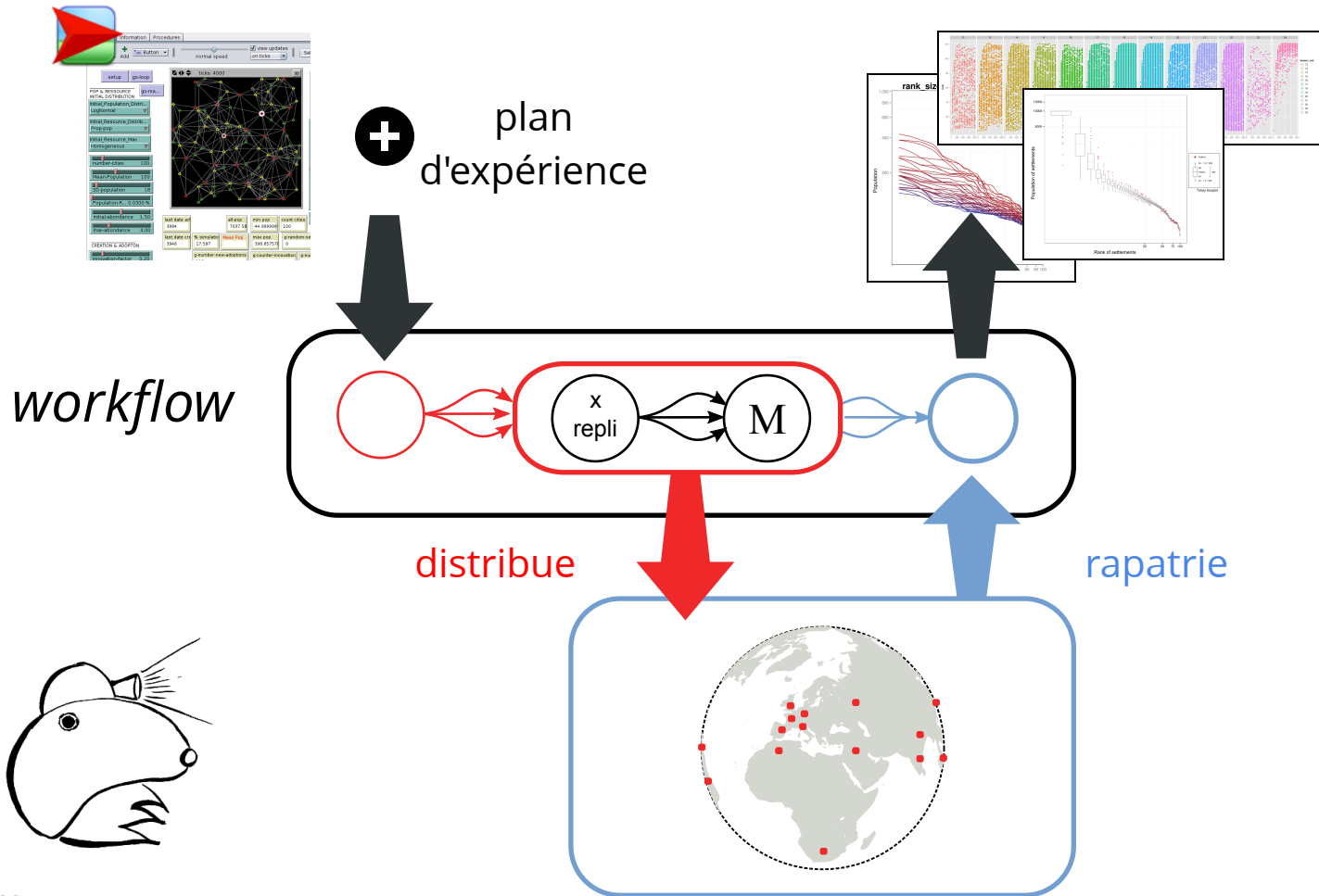
Espace des paramètres



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

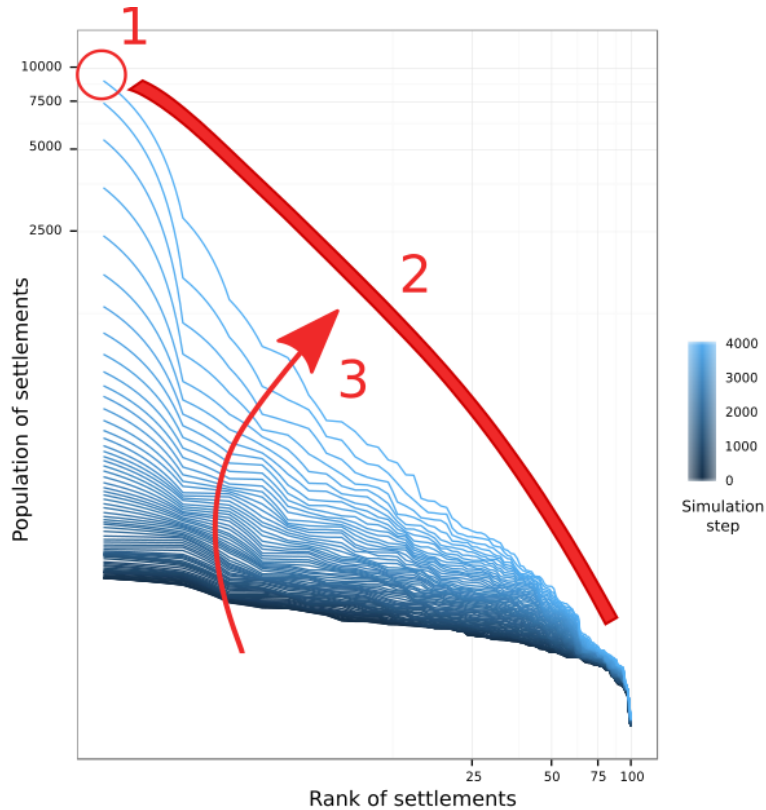
Accès aux ressources

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



(dès juin 2010)

Recherche de *patterns*



3 objectifs simultanés

obj 1 - population max
= 10000 hab.

obj 2 - distribution villes
= lognormalité

obj 3 - temps
= 4000 ans



OpenMOLE

next.openmole.org

iscpif.fr

... et l'institut des systèmes complexes Paris-IDF !

- Accès **transparent** aux ressources Grille (EGI 4000 coeurs) & HPC
- Utilisation **au fil de l'eau** des ressources
- Méthodes d'explorations **clef en main** dédiés **Grille**
- **Tâches spécifiques** aux plateformes Agents (Netlogo, Gama, Cormas) mais pas que ...
- **Outils & méthodes génériques** à la **communauté** des systèmes complexes



OpenMOLE

next.openmole.org

iscpif.fr

... et l'institut des systèmes complex Paris-IDF !

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)
  )

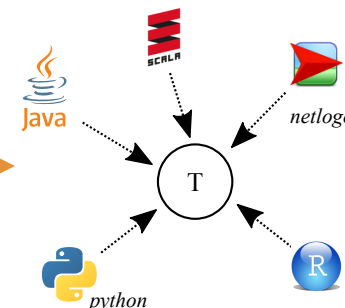
val env = LocalEnvironment(4)

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

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

Plusieurs types de plan d'expériences : Plan complet, LSH, Analyse de sensibilités, Algorithmes Evolutionnaires, etc.

Plusieurs tâches différentes !

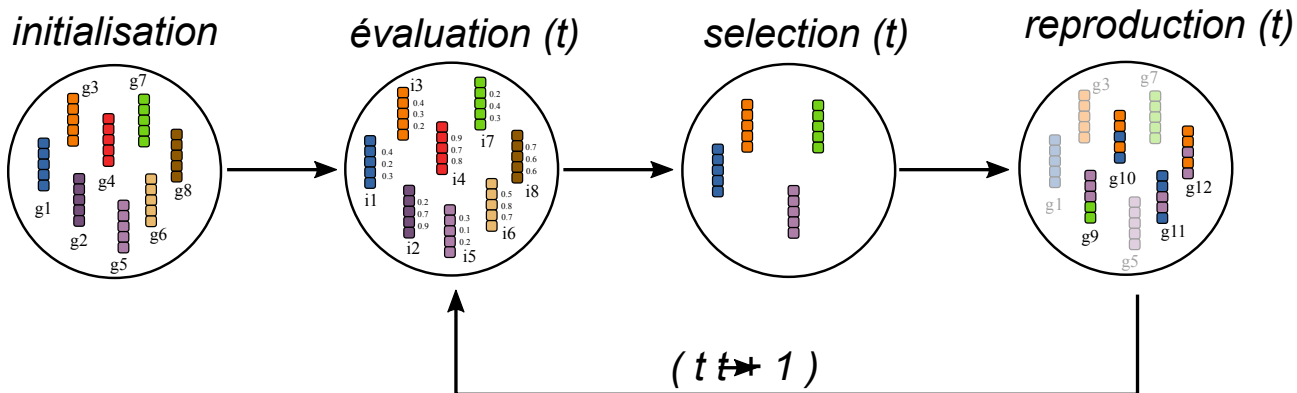
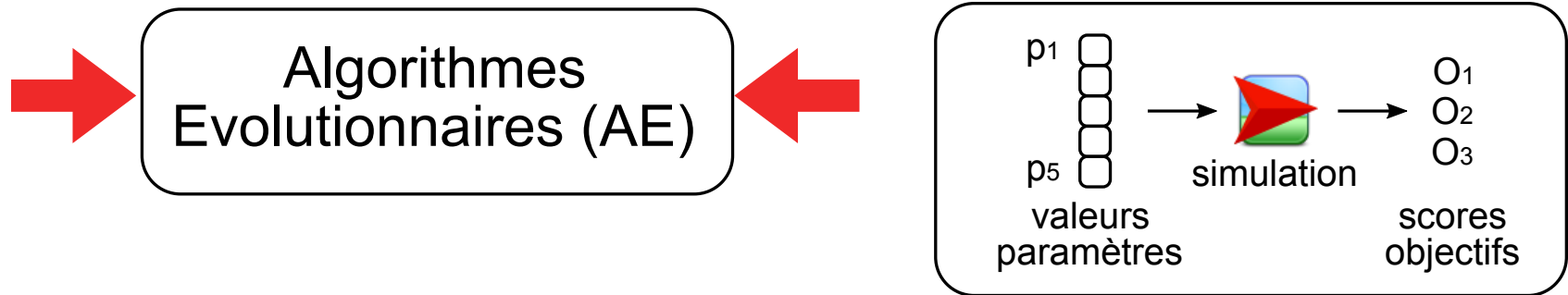


Support de multiples technologies HPC : Cluster, grille de calcul, etc.

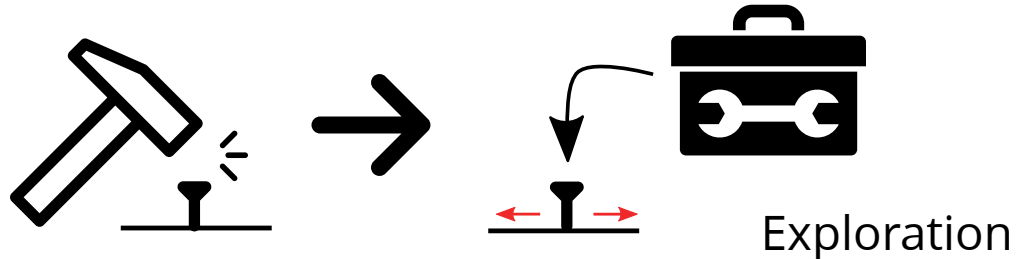
Méthodes EA intégrées

Très bien adapté à une utilisation grille de calcul :

EA 3 modes : generationnel / steady - state / Ilôts



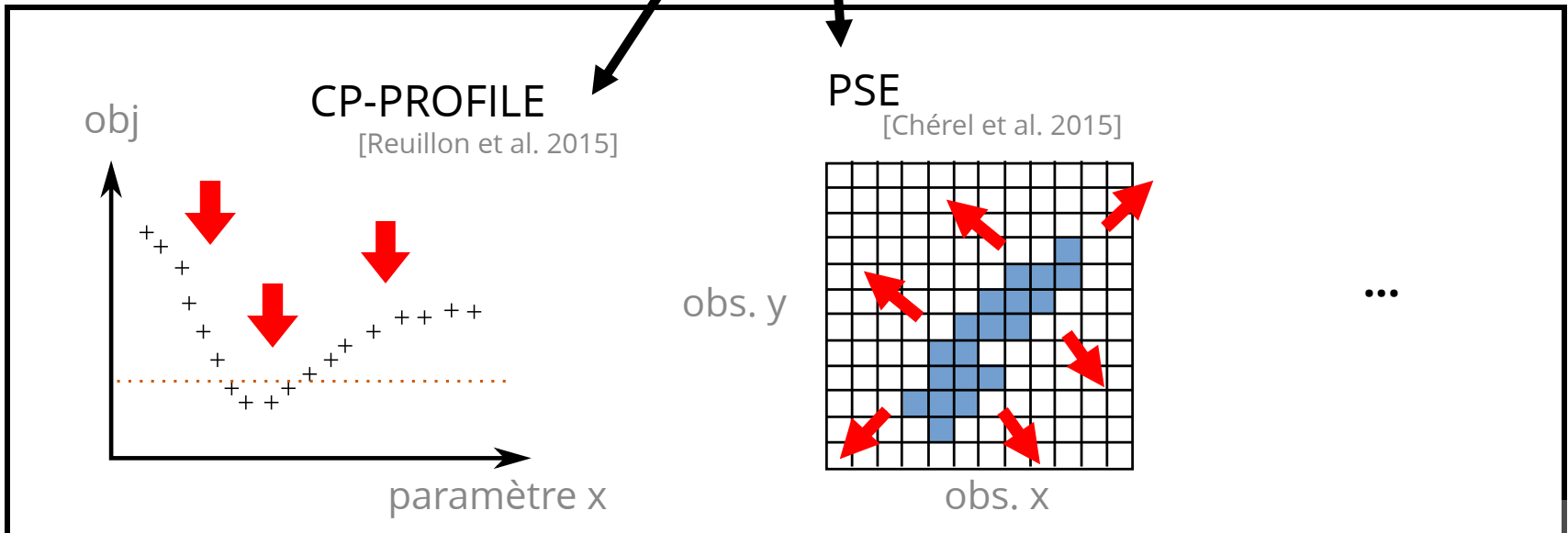
Méthodes : 2 pistes



Calibrage
pattern or not patterns?

Exploration

*diversité patterns (+ ou -)
guidée*



Perspectives

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

... un logiciel et une communauté en expansion



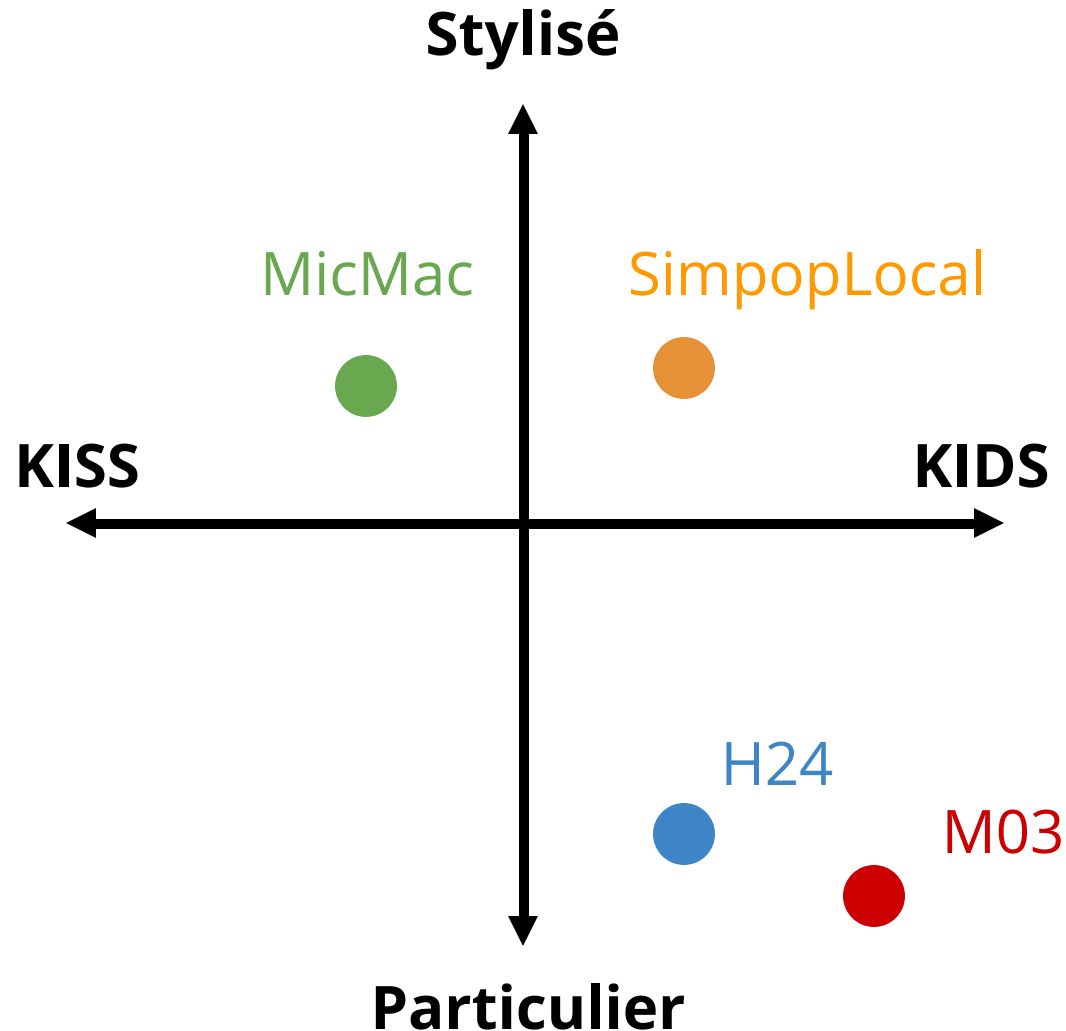
- cap des 10.000 *commits*
- x 2 participants en 2018
- Amélioration de la documentation
- Ouverture à de nouvelles méthodes :
 - méthodes d'exploration *OSE*
 - générateur de grilles spatiales
 - 2 nvx cas d'utilisation (santé, éco)
- passage à *uDocker*
- gestion et déploiement multi-utilisateurs !!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

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

Examples

4 exemples

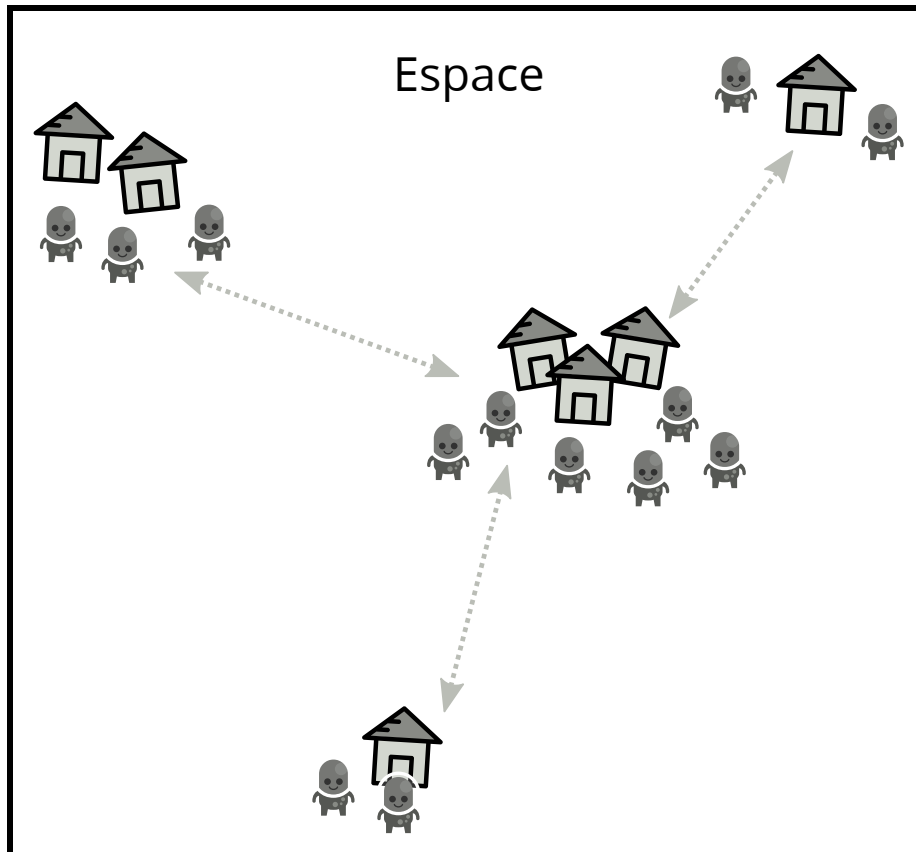
& une grande diversité de besoins en calculs : Grille & HPC !



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** ?



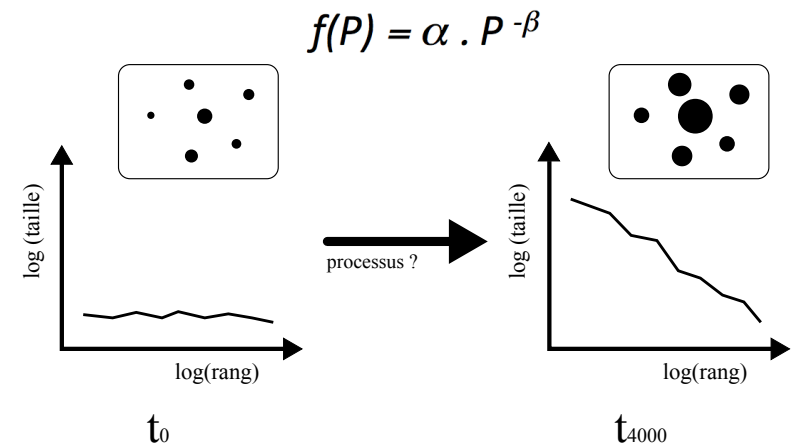
agents: villes & innovations

7 paramètres libres

4 mécanismes

très peu de données en entrée/sorties

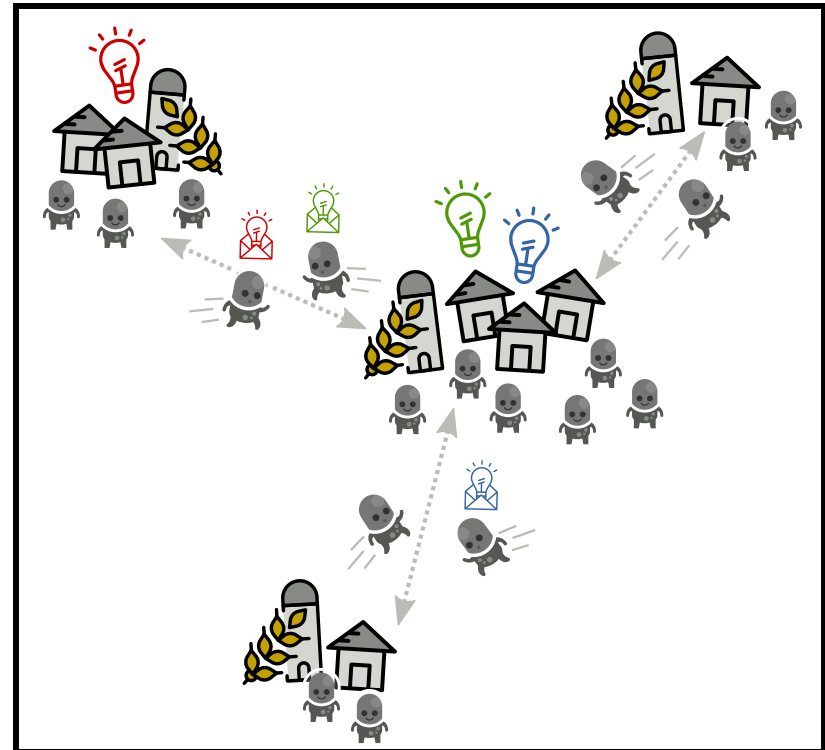
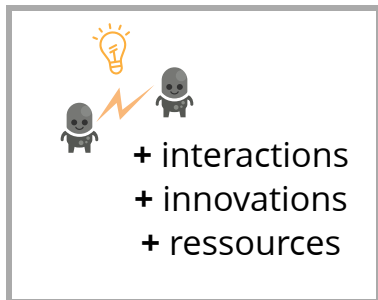
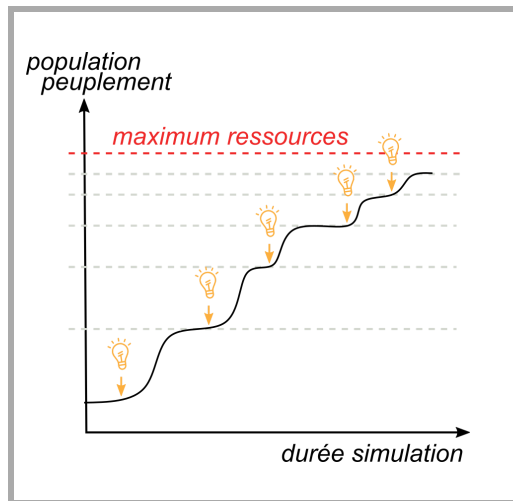
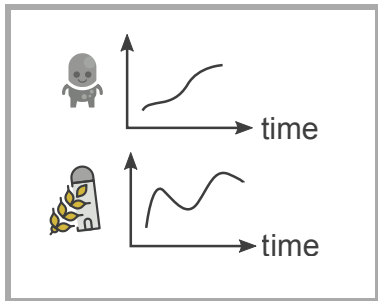
1 loi empirique à reproduire



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

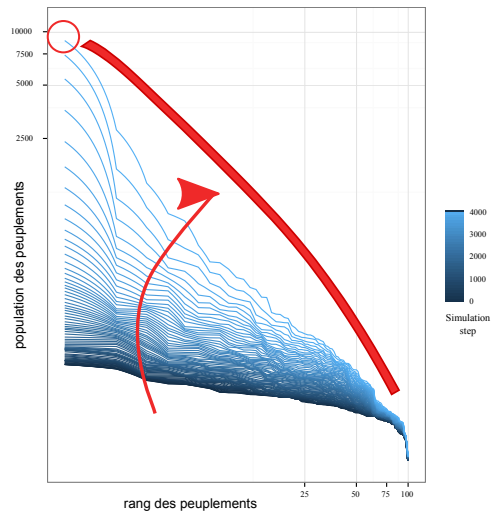


SimpopLocal

Calibration

NSGA 2 - Ilôts

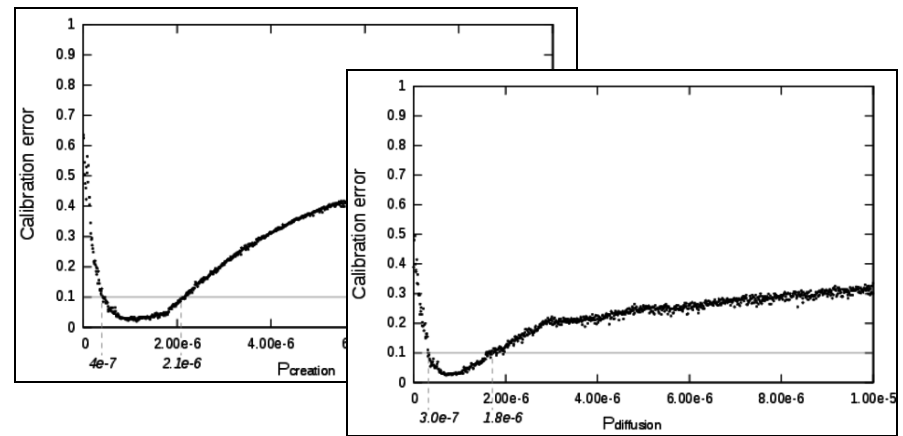
3 objectifs : (temps, forme, taille)



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

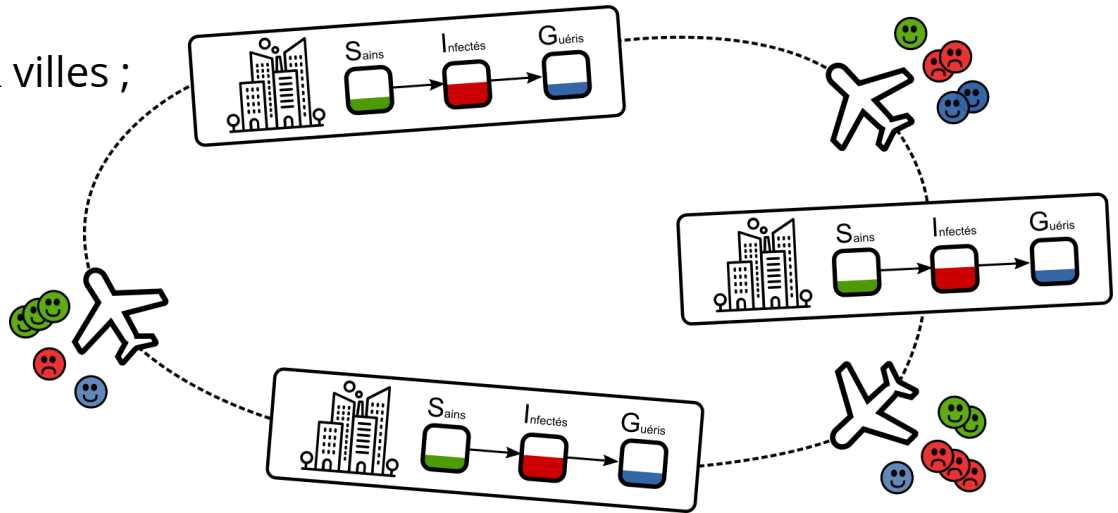
stochastique

4 stratégies à explorer

2 objectifs à minimiser

pas de données

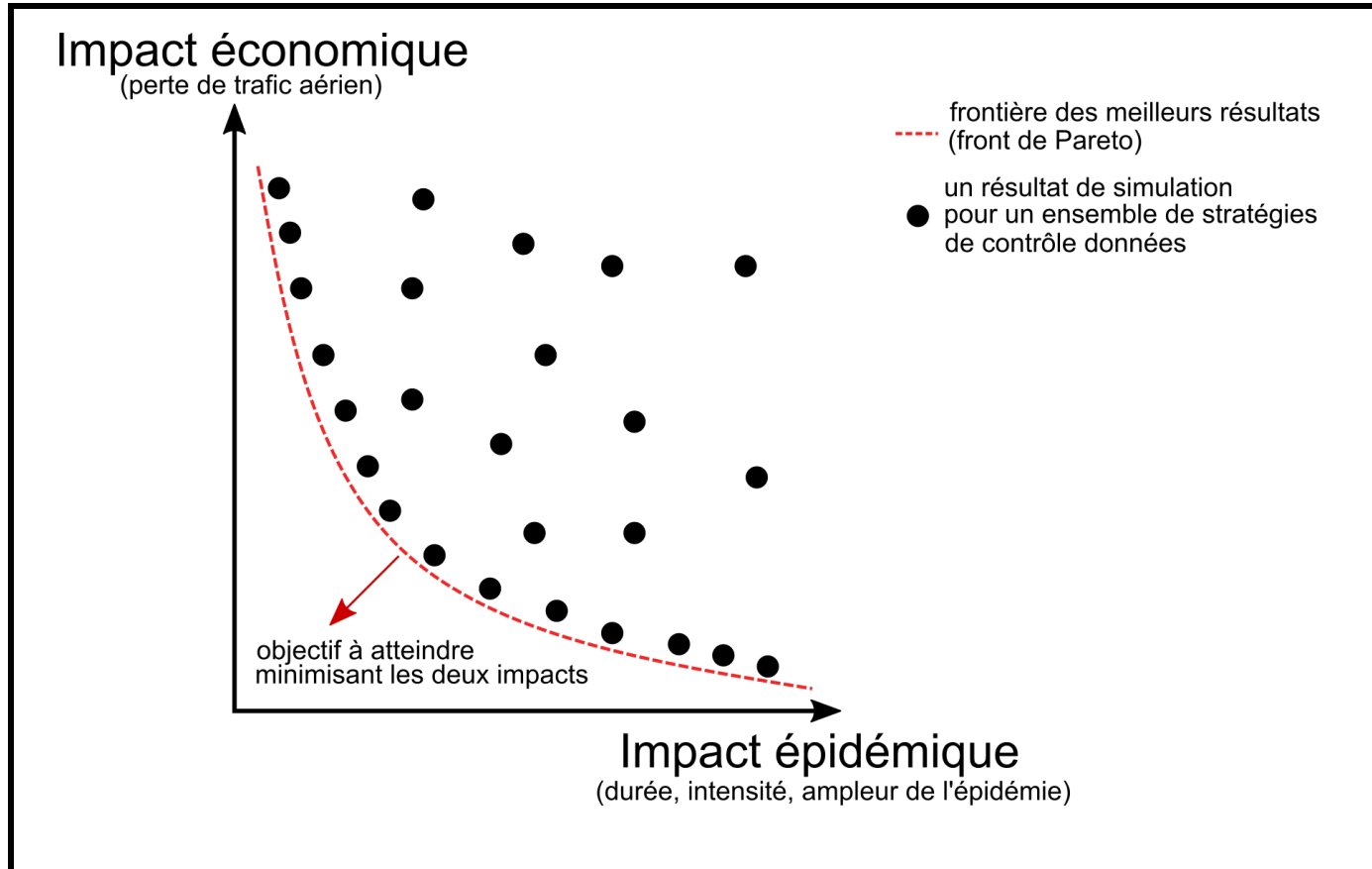
$$\begin{cases} \frac{dS}{dt} = -\beta IS \\ \frac{dI}{dt} = \beta IS - \alpha I \\ \frac{dR}{dt} = \alpha I \end{cases}$$



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.

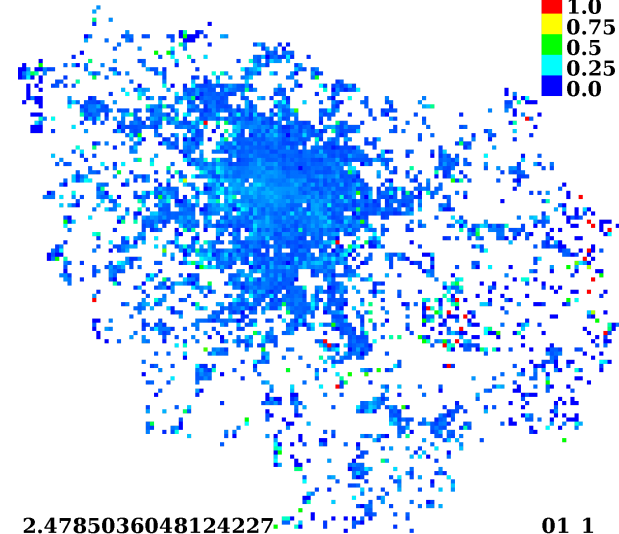
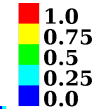
[Cottineau et al. 2017]

~ **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



01_1

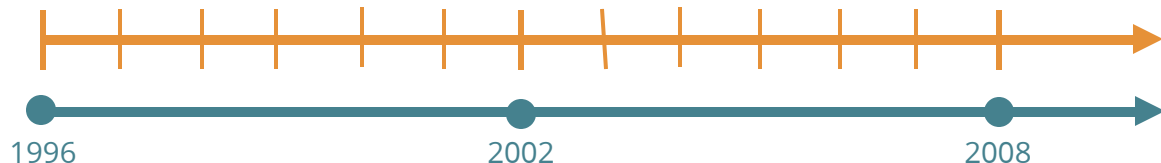
mesure d'inégalité entre les extrême des groupe d'éducation à sexe et âge égal

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



[Daudé et al. 2017]

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/sebastienreycyhourcq/hpc-criann-caen>



Lien vers la bibliographie Zotero