Fostering the use of methods for geosimulation models sensitivity analysis and validation

R. Reuillon[1,2], M. Leclaire[1,2], J. Raimbault[3,1,2,*], H. Arduin[4], P. Chapron[5], G. Chérel[1], E. Delay[6], F. Lavallée[7], J. Passerat[8], P. Peigne[9], J. Perret[1,5], S. Rey-Coyrehourcq[10]

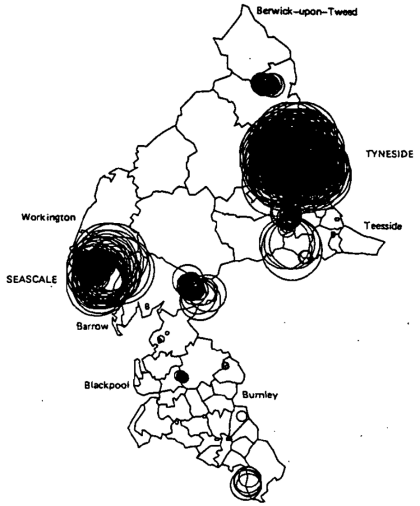
[1] UPS CNRS 3611 ISC-PIF
[2] UMR CNRS 8504 Géographie-cités
[3] CASA, UCL
[4] UMR Inserm 1037 CRCT
[5] Univ. Paris-Est, LaSTIG STRUDEL, IGN, ENSG
[6] CIRAD UPR GREEN
[7] Irstea, LISC
[8] ConsenSys
[9] École 42
[10] UMR CNRS 6266 IDEES

[*] juste.raimbault@polytechnique.edu

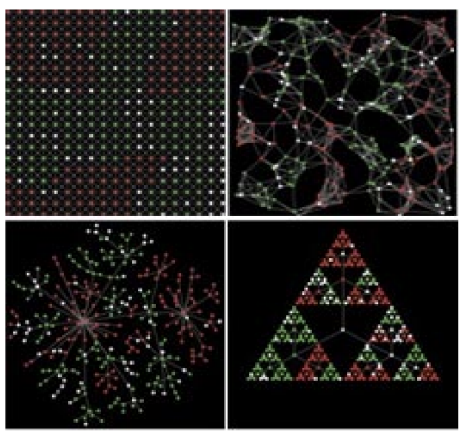
OpenMOLE

ECTQG 2019
Exploration of geosimulation models
September 6th, 2019

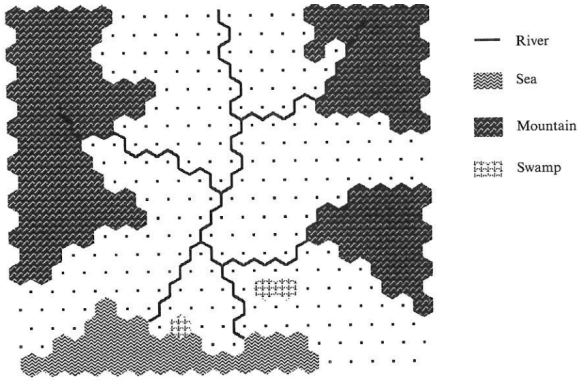
Simulation in TQG



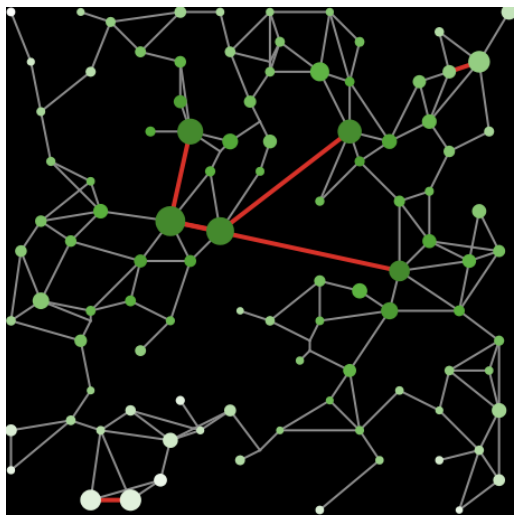
Geographical analysis machine [Openshaw et al., 1987]



Schelling on networks [Banos, 2012]



Simpop 1 model[Sanders et al., 1997]



SimpopNet model [Schmitt, 2014]

Necessity of simulation models in geography induced by complexities of these systems ?

- Ontological complexity [Pumain, 2003]
- Dynamical complexity: non-ergodicity and path-dependancy [Pumain, 2012]
- Complexity and co-evolution
- Complexity and emergence

Model exploration is basically running a model *a lot of times*, following a *design of experiments*, to gain knowledge about *model properties*.

e.g. : sensitivity analysis

Recent and significant increase in the development of methods to explore, calibrate and optimize (geo)simulation models.

$\rightarrow$ ease model validation !

Practice of model validation

Methods and tools remain underused by simulation communities, despite an easier access to HPC facilities.

e.g. mail of Bruce Edmonds (emblematic figure of social simulation) on the SIMSOC mailing list, on the 16th of May 2019.

> *Dear Colleagues,*
>
> *David Hales and I have been looking at* ***how to do massively parallel runs of NetLogo simulation models on the Cloud****. Something like (a) design your runs using NetLogo's BehaviorSpace (b) upload the model to the cloud (c) run it on the cloud (d) get the resulting table of results back.*
>
> *We are wondering how many people would be interested in something like this.[...]*

The **OpenMOLE** free and open source software provides (i) model embedding; (ii) transparent access to HPC; (iii) state-of-the-art model exploration methods.

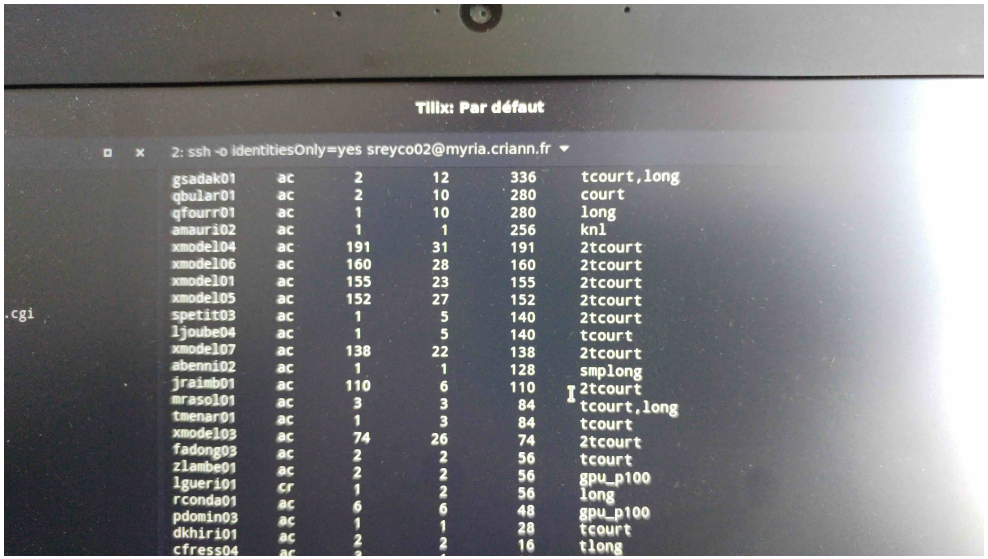
Success stories: epidemiology [Arduin, 2018], ecology [Lavallée et al., 2019], planning [Brasebin et al., 2017], urban science [Raimbault, 2018]

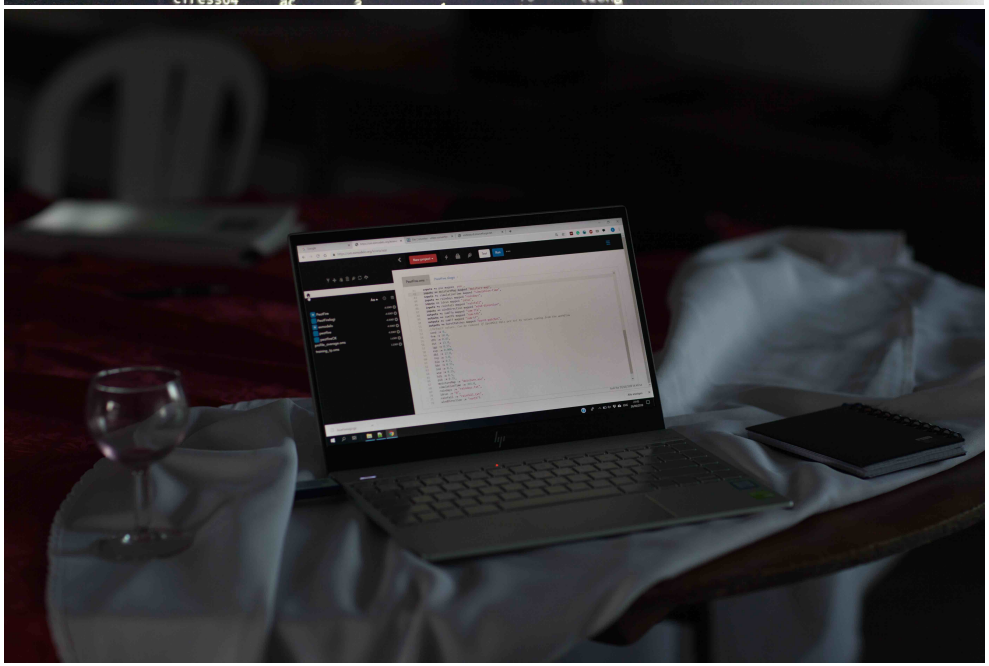
The **eXModelo school** to learn model validation methods

- OpenMOLE as the learning tool, but reproducible with other tools and methods
- between a formal summer school and a workshop (student projects)

Passionated researchers...

... and intensive computation





A common model to present the exploration and validation protocol

$\rightarrow$ modularity and complementarity of aspects

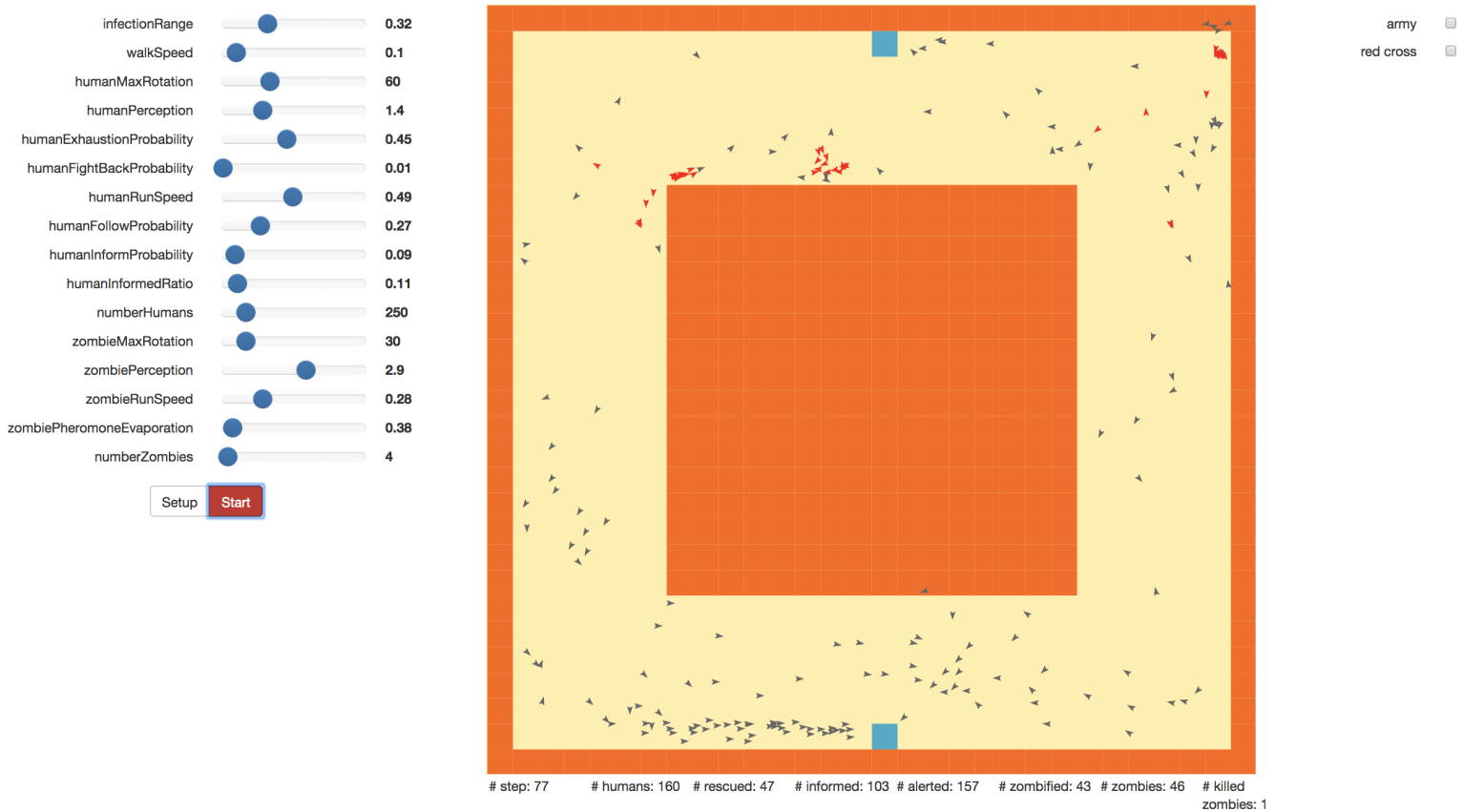
$\rightarrow$ possibility of an increased complexity and research questions left open

A discipline-agnostic model: zombie epidemiology

$\rightarrow$ difficulty of interdisciplinary dialogue

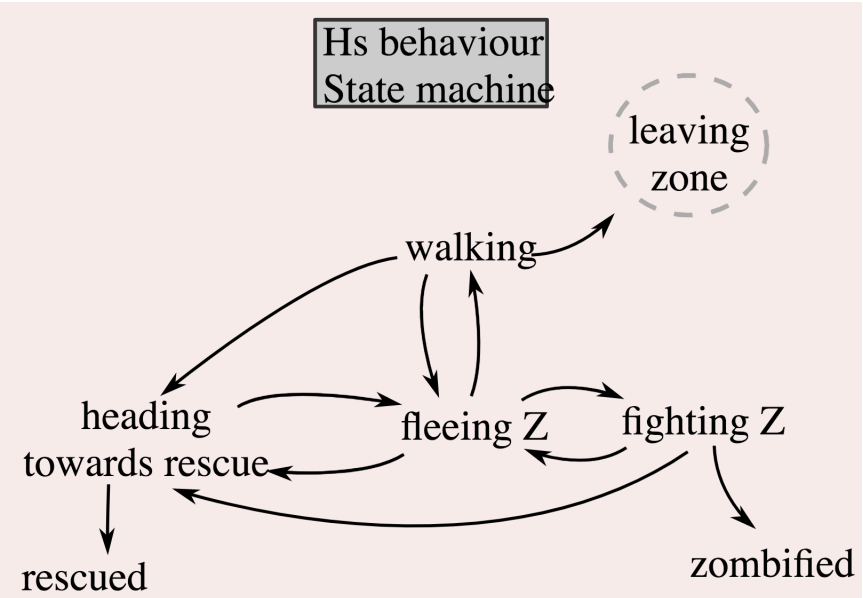
$\rightarrow$ agent-based spatial modeling as a natural way to enhance it

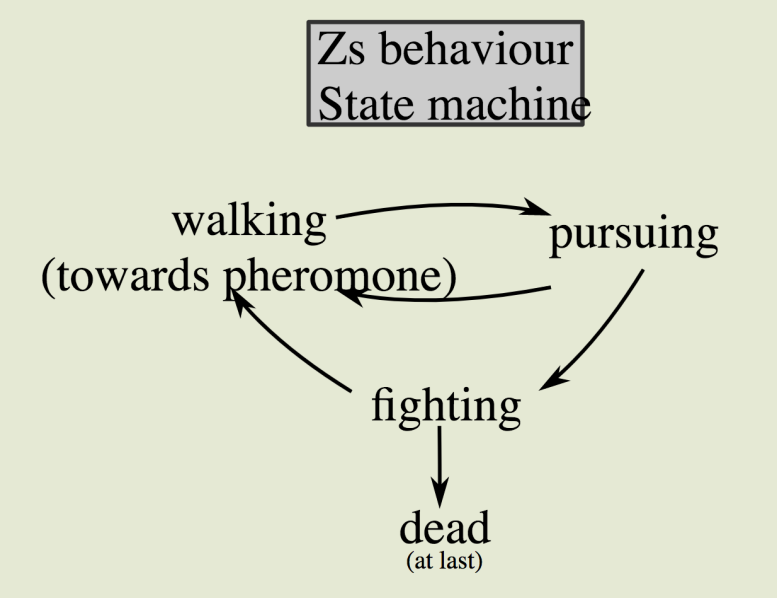
An operational model for local Zombie invasion



Local scale agent-based model

Agents state machines





Design of Experiments

	Coverage	Interpretability	Budget
One factor at a time	✗	✓	✓
Complete plan	✓	✓	✗
LHS/Sobol	✓	✗	✓

Sensitivity analysis

	Coverage	Interpretability	Budget
Morris	✗	✓	✓
Saltelli	✓	✓	✗

Syntax of a sensitivity analysis method in OpenMOLE

```
SensitivityMorris(
    evaluation = (model on env by 5000),
    inputs = List(
       humanFollowProbability in (0.0,1.0),
       humanInformedRatio in (0.0,1.0),
       humanInformProbability in (0.0,1.0)
    ),
    outputs = List(totalZombified,halfZombified),
    repetitions = 1000,
    levels = 20
) hook CSVHook(workDirectory / "morris_result.csv")
```

→ **Practice for students at eXModelo:** explore the script `morris.oms`, comment the results obtained with a large-scale experiment `morrisresults`

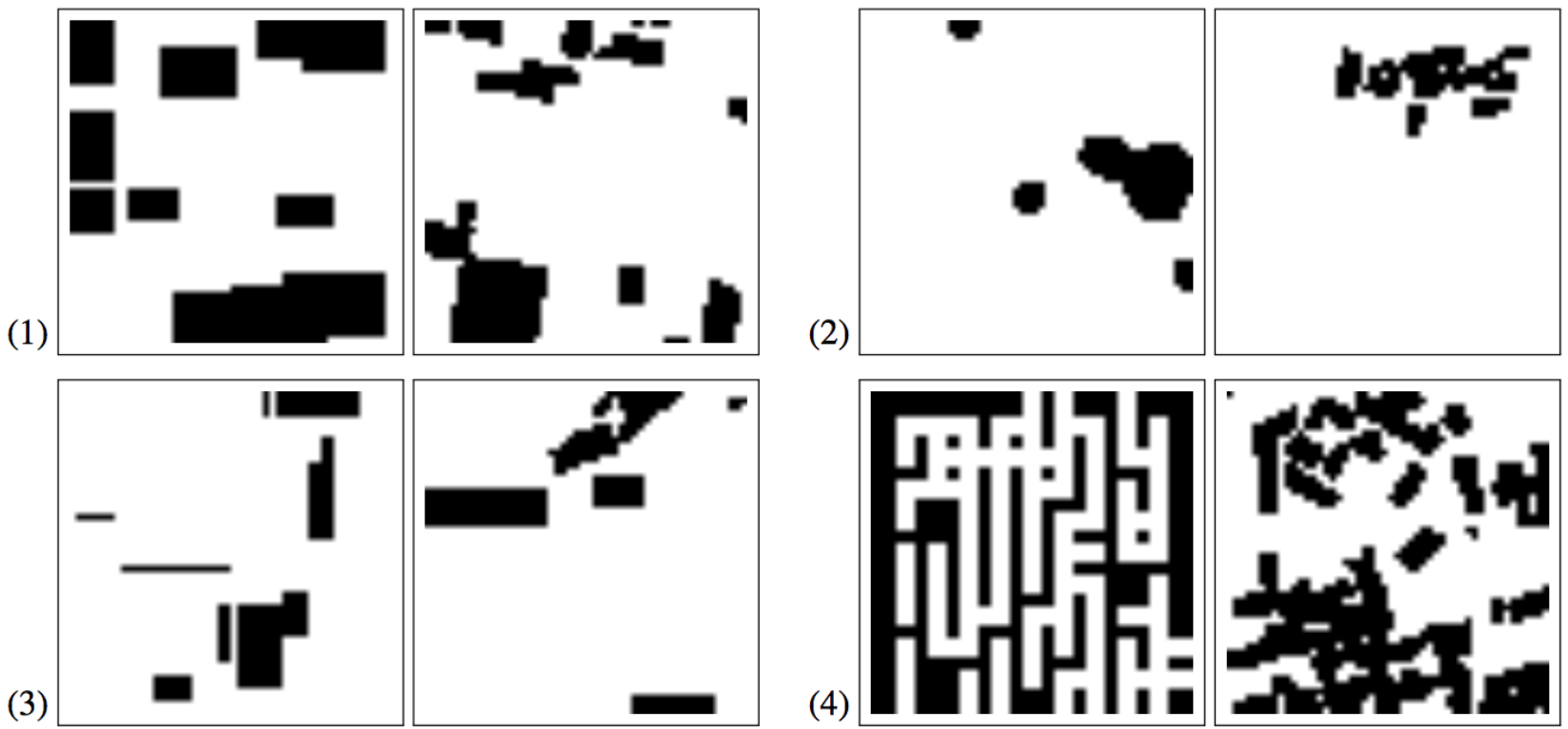
Calibration: Evolutionary (GA) and Bayesian (ABC) methods

Diversity Search: "look for the variety of obtainable patterns in output space"

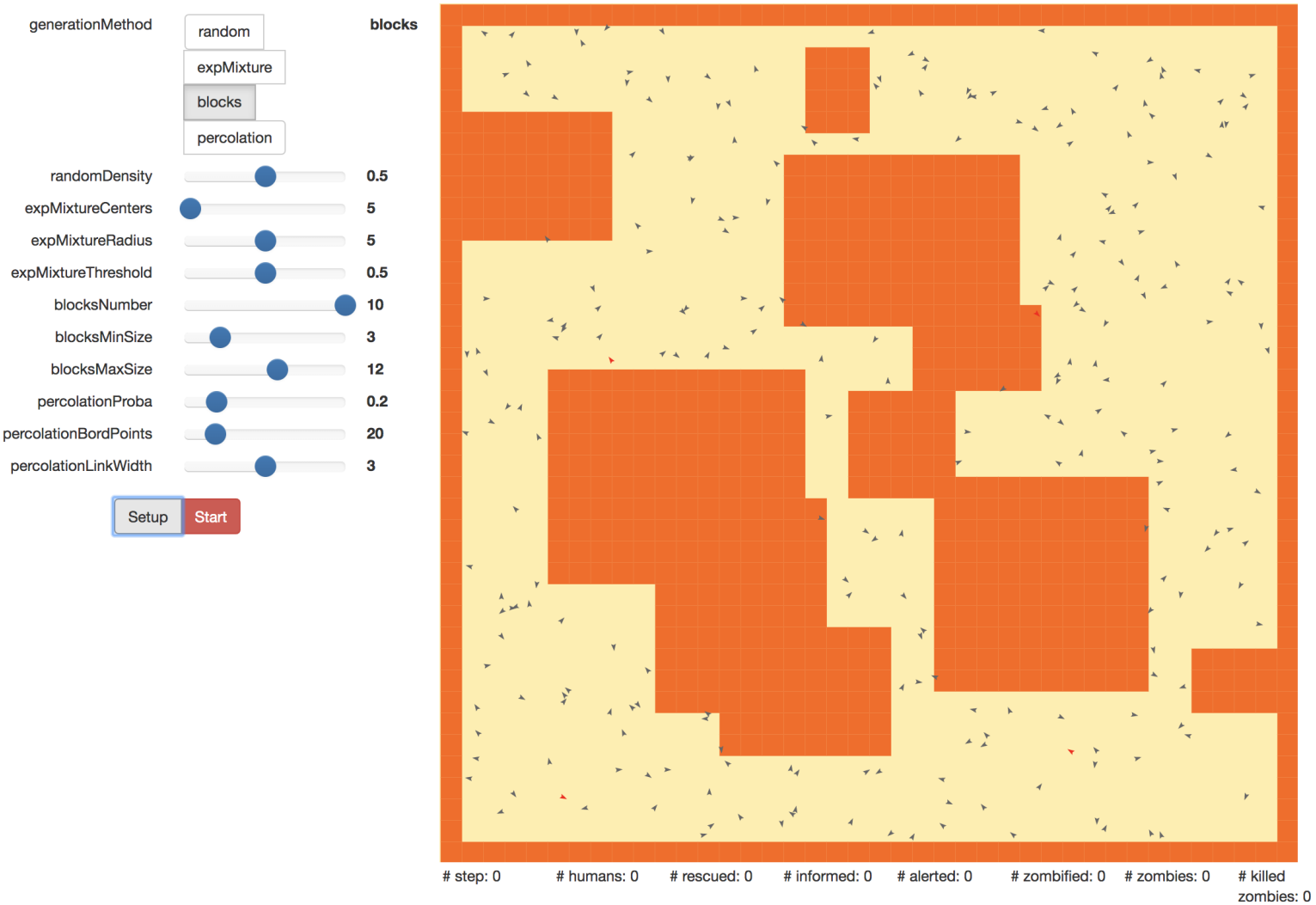
Origin Search: "look for the input that produce a given pattern in output space"

A specific focus on spatial aspects: built-in spatial sensitivity analysis library

Example: generators of synthetic urban districts
[Raimbault and Perret, 2019]



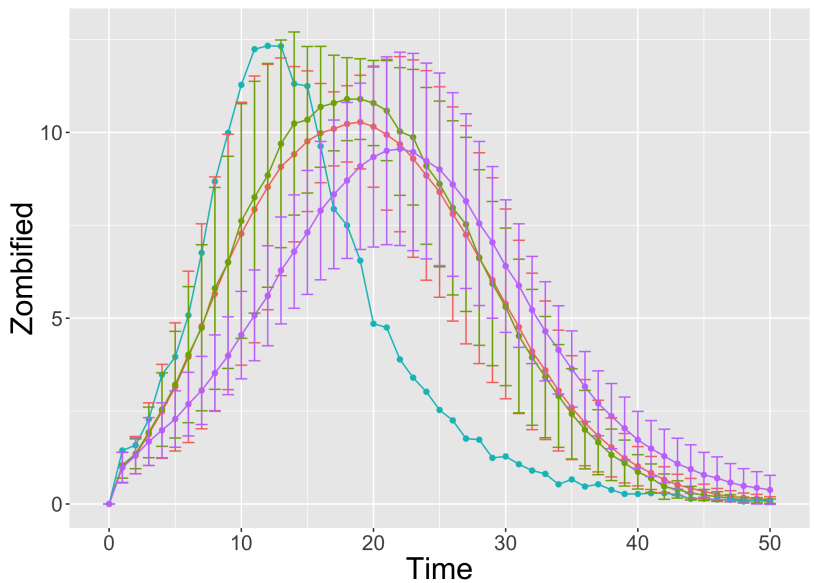
Modifying the world in the Zombie model

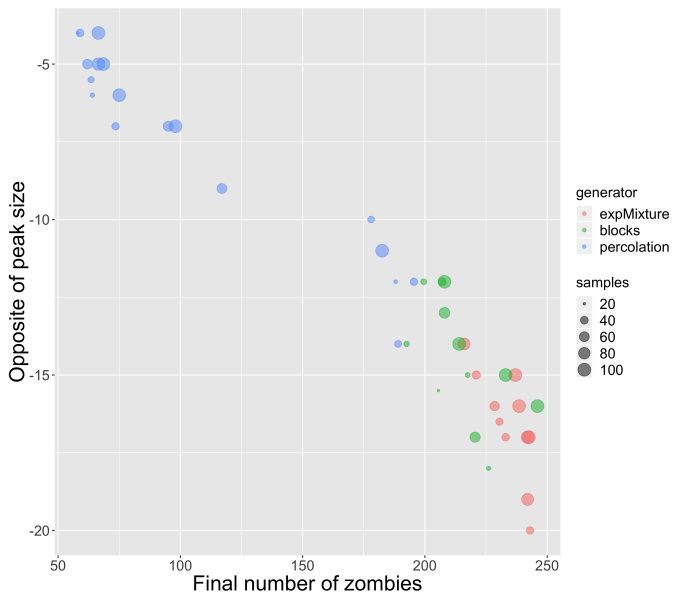


Spatial sensitivity of the Zombie model

OpenMOLE

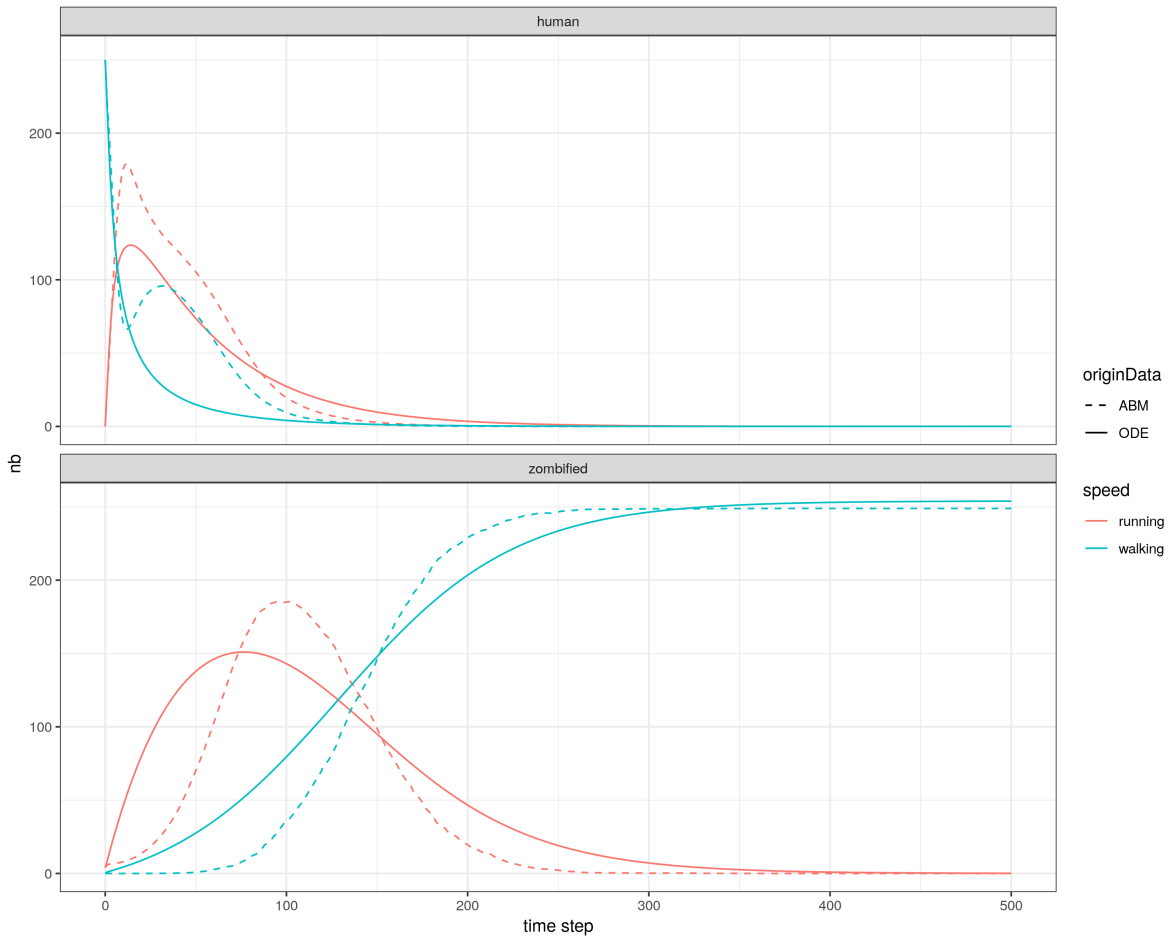
Which spatial organization impedes the epidemics?





→ *Study of submodels to foster diverse questions and approaches: cooperation between humans, army intervention, red cross and vaccination*

→ *Comparison of the ABM with alternative formalisms*

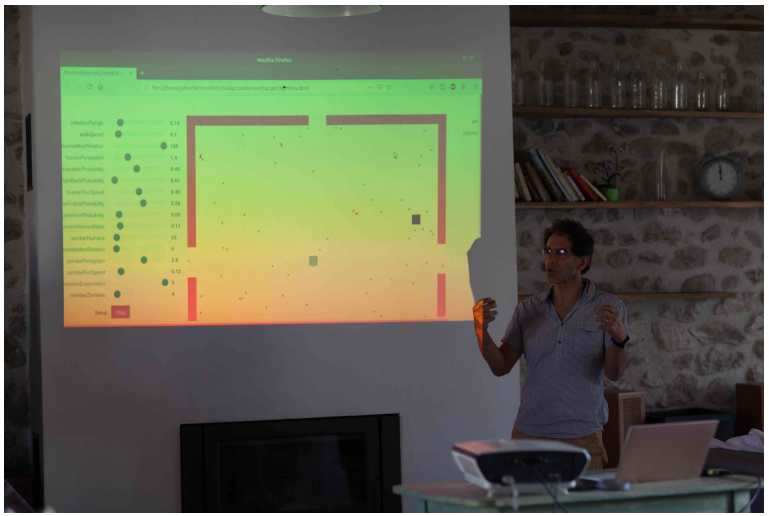


Regimes of ABM compared to ODE patterns

A **student challenge** for autonomous practice
→ they define a fresh thematic question
→ they define the adequate design of experiments, especially methods and measures.
→ possibly model code modification

Challenge results

Each group came with unforeseen questions and ideas; not much advanced methods where however used



Goals :

- Balancing theory and practice
- Flattening disciplinary heterogeneity
- Focusing on model analysis methods instead of platform/framework

Expected benefits :

- Methods and approach dissemination
- OpenMOLE visibility improvement
- For the students: more robust model studies

Agenda :

Progressive introduction of methods(2.5 day)
→ first day mostly devoted to vocabulary/concepts/framework
→ each course followed by group practice session on toy example
→ emphasis on "question ⟷ method"

Advanced methods (0.5) day
→ 3 hours focused on a specific advanced method : ODE modeling, ABC method calibration, Spatial sensitivity analysis

Challenge (1.5 day)
→ application of the method corpus on self defined question
→ 5 students group , collective restitution

$\rightarrow \approx$ 1 month of workload (ventilated during the year before) for 8-9 people, consumed in 5 days!

$\rightarrow$ successful method during the school; middle and long term impacts to be assessed

$\rightarrow$ fostered model exploration methods and practices, in an interdisciplinary environment

Stay tuned for next eXModelo: `exmodelo.org`

Use and contribute to OpenMOLE: `openmole.org`

Reproducible school:

Course contents available at
`https://github.com/openmole/exmodelo-courses`

Model available at `https://github.com/openmole/exmodelo-model`

We need you if the zombies (or something else?) come back !

$\rightarrow$ prepare you for end of May, 2020 !

References I


Arduin, H. (2018).
Modélisation mathématique des interactions entre pathogènes chez l'hôte humain: Application aux virus de la grippe et au pneumocoque.
PhD thesis, Université Paris-Saclay.


Banos, A. (2012).
Network effects in schelling's model of segregation: new evidence from agent-based simulation.
Environment and Planning B: Planning and Design, 39(2):393–405.


Brasebin, M., Chapron, P., Chérel, G., Leclaire, M., Lokhat, I., Perret, J., and Reuillon, R. (2017).
Apports des méthodes d'exploration et de distribution appliquées à la simulation des droits à bâtir.

References II


Lavallée, F., Smadi, C., Alvarez, I., Reineking, B., Martin, F.-M., Dommanget, F., and Martin, S. (2019).
A stochastic individual based model for the growth of a stand of japanese knotweed including mowing as a management technique.
arXiv preprint arXiv:1902.06971.


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.


Pumain, D. (2003).
Une approche de la complexité en géographie.
Géocarrefour, 78(1):25–31.

References III


Pumain, D. (2012).
Urban systems dynamics, urban growth and scaling laws: The question of ergodicity.
In *Complexity Theories of Cities Have Come of Age*, pages 91–103. Springer.


Raimbault, J. (2018).
Indirect evidence of network effects in a system of cities.
Environment and Planning B: Urban Analytics and City Science, page 2399808318774335.


Raimbault, J. and Perret, J. (2019).
Generating urban morphologies at large scales.
The 2019 Conference on Artificial Life, (31):179–186.

References IV


Sanders, L., Pumain, D., Mathian, H., Guérin-Pace, F., and Bura, S. (1997).
Simpop: a multiagent system for the study of urbanism.
Environment and Planning B, 24:287–306.


Schmitt, C. (2014).
Modélisation de la dynamique des systèmes de peuplement: de SimpopLocal à SimpopNet.
PhD thesis, Paris 1.