Validation of geosimulation models: a systematic review

J. Raimbault[1,2,3,4,*]

* juste.raimbault@ign.fr

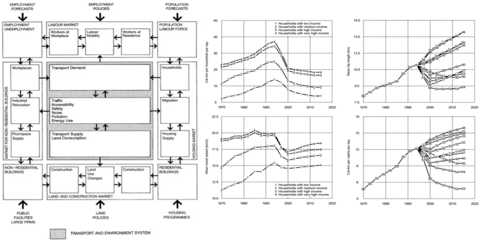
[1] LASTIG, IGN-ENSG
[2] CASA, UCL
[3] UPS CNRS 3611 ISC-PIF
[4] UMR CNRS 8504 Géographie-cités

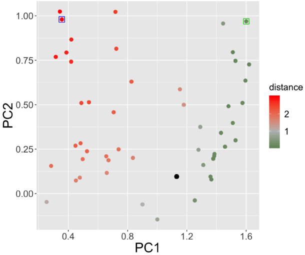
ECTQG 2023

Session Theoretical Geography 2

September 15th 2023

When is a simulation model "validated"?





(Left) Operational large microsimulation model for land-use transport interactions [Wegener, 2011]; (Right) Systematic exploration and comparison of phase diagrams across meta-parameter space for a toy model (sugarscape) [Raimbault et al., 2019]

Specificities for simulation models

Advantages of geosimulation models: spatial complexity, heterogeneity of agents and processes, multiple scales, come at the cost of a more difficult validation:

- High-dimensional parameter space and output space
- Strongly non-linear, non-ergodic, non-stationary dynamics
- Stronger role of stochasticity
- Need to be fully simulated (weak emergence)

Disciplinary definitions of validation

Definition of "model validation" strongly depends on disciplines, methods, type of models, for example:

- metrics linked to prediction performance for machine learning
- resembling patterns in Pattern Oriented Modelling [Grimm et al., 2005]
- "face validity" in agent-based models
- advanced validation toolset with the OpenMOLE platform [Reuillon et al., 2013]
- stakeholders' feedback in participatory approaches
- ...

A proposal based on model functions

Proposal for a typology of model validation techniques and standards by [Raimbault, 2020b], based on [Varenne, 2018]'s typology of model functions and modeller purpose [Giere, 2019]:

- **Perception and observation**: how much information is extracted
- **Description**: how much information is contained within
- **Prediction**: predictive power (quantitative indicators or qualitative behavior)
- **Explication and comprehension**: how much of the causal structure of the system is grasped
- **Theory construction**: how does the model contributes to the theory, to coupling of its components (e.g. medium for interdisciplinarity)
- **Communication**: how much information is conveyed and to which agents
- **Decision-making**: how are decision supported, which benefits and for what dimension (societal, environmental, etc.)?

Contribution

→ geosimulation/spatial simulation models are used in a wide range of disciplines

→ what are effective "validation" practices of these diverse scientific communities?

This contribution: *propose a Systematic Review of the concept of validation for geosimulation models.*

Systematic review

→ using data collection tools developed by [Raimbault, 2019], construct a paper corpus through a keyword request to google scholar

→ systematic screening of paper titles (and abstracts/full texts if needed) to keep relevant papers

→ extraction of method and models characteristics: discipline (from journal/affiliation), type of method (ad hoc typology), type of model (idem), generic methodological contibution

→ "meta-analysis" of corpus characteristics

PRISMA systematic review description

Relevant PRISMA reporting points:

5 - **Eligibility criteria:** explicit reference to "validation" (and synonyms) in the title/abstract

6 - **Information sources:** google scholar, accessed through the `BiblioData` API [Raimbault, 2019]

7 - **Search strategy:** direct query: ("spatial simulation"OR"geosimulation")AND("validation"OR"calibration"OR" sensitivity analysis"OR"exploration"OR"evaluation"OR" assessment")

8 - **Selection process:** manual screening of titles (/abstracts/full text) by one expert

9 - **Data collection process:** manual extraction by one expert during screening (if included)

10 - **Data items:** discipline (journal/affiliation), method (ad hoc typology), type of model, methodological work

13 - **Synthesis methods:** statistical analysis of categorical variables

27 - **Availability of data and code:** open repository at `https://github.com/JusteRaimbault/SimulationModels`

Systematic review flowchart

Direct query to google scholar

852 papers (out of 1840 from front-end search)

Title/abstract manual screening

132 papers

Abstract/full text manual screening
when title was not explicit ($\sim 50\%$ of papers)

130 papers with characteristics

Statistical analysis of categorical properties

Typology of validation methods depending on disciplines

Results: disciplines

Disciplines (on 130 papers): ecology (28), geosimulation (26), land-use change (18), environmental science (15), urban science (11), hydrology (11), computer science (8), sustainability (6), climate science (4), archeology (1), biology (1), social science (1)

→ sampling rate of the corpus unknown (need to cross with other databases), but proportions seem reasonable

→ some "hard" disciplines missing or under-represented (physics, climate science): validation is intrinsic to their modelling enterprise and not explicitly mentioned?

→ validation methodologies from ecology to social sciences? (cf ABM and POM)

Results: typology of validation methods

Ad hoc typology of validation methods:

- prediction (24)
- SA (19)
- uncertainty (18)
- multiple (13)
- benchmark (12)
- calibration (11)
- optimisation (9)
- visualisation (8)
- POM (6)
- participatory (4)
- exploration (3)
- mixed (2)
- surrogate (1)

→ A relatively exhaustive list of method types?

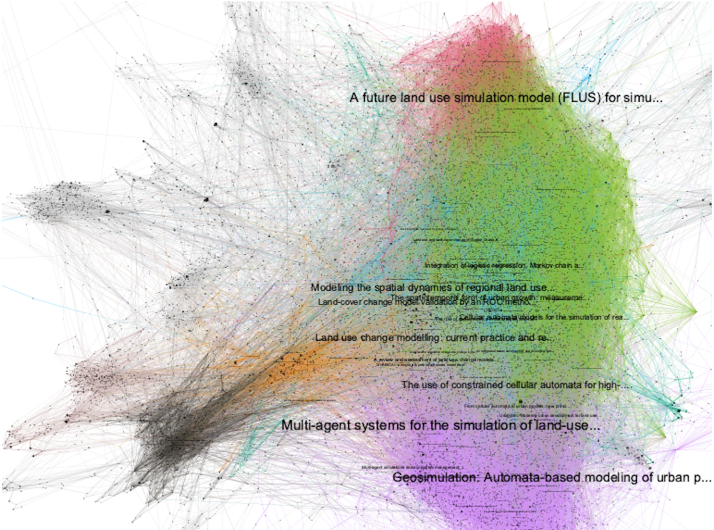
→ proportions not representative: implicit validation in many disciplines and methods

Results: disciplinary context

	benchmark	calibration	multiple	optimisation	POM	prediction	SA	uncertainty	visualisation
computer science	0	3	3	1	0	0	0	0	1
ecology	3	2	2	3	2	6	4	4	0
environmental science	2	2	0	0	0	0	2	8	1
geosimulation	0	2	2	0	0	4	5	5	3
hydrology	1	1	0	2	0	4	2	0	0
land-use change	3	0	3	1	3	7	1	0	0
sustainability	0	1	0	1	0	0	0	0	3
urban science	0	0	3	1	1	1	4	0	0

Broader literature mapping: citation network

Reconstruct scientific neighborhood using backward citation propagation at level 2 from the SR corpus (see [Raimbault, 2019]): preliminary citation network with 30434 nodes and 96806 edges



Discussion: SR/MA methodology

Issues with systematic reviews and meta-analysis on fuzzy/interdisciplinary/social science concepts (confirms observations made in [Raimbault, 2020a]:

→ difficult to control corpus and disciplinary bias, translation between disciplines

→ need new methodologies to build endogenous typologies of models and methodologies, for example based on expert model decomposition (see WIP with F. Le Néchet), or using new NLP techniques to extract concepts from full texts (but still issue of various languages)

→ which level of generalisation/detail for methods and models

→ need much more robustness check / multiple information sources cross-validation, than "classical" systematic reviews

→ impossibility to proceed to meta-analysis when outcome measures are not uniformised (link with model equivalence, integration and interoperability)

Conclusion: towards interdisciplinary validation standards?

→ although with quite (unquantifiable) bias, we find a wide diversity of validation methodologies across a variety of disciplines

→ importance of reflexivity and quantitative epistemology when working in such interdisciplinary settings (cf the CybergeoNetworks tool [Raimbault et al., 2021])

→ towards "unified" interdisciplinary validation standards and methods? Importance of model coupling and integration. The **OpenMOLE platform** provides already most of validation methods found in the typology (participatory under development by P. Chapron et al., visualisation being improved, surrogate and uncertainty being tested, novel methods such as spatial sensitivity analysis being developed [Raimbault et al., 2020]): `https://openmole.org`.

OpenMOLE

ExModelo workshop in November: still time to apply, directly contact me if interested!

https://workshop.exmodelo.org/

References I

Giere, R. N. (2019).
Scientific perspectivism.
University of Chicago press.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., and DeAngelis, D. L. (2005).
Pattern-oriented modeling of agent-based complex systems: lessons from ecology.
science, 310(5750):987–991.

Raimbault, J. (2019).
Exploration of an interdisciplinary scientific landscape.
Scientometrics, 119(2):617–641.

References II

Raimbault, J. (2020a).
A systematic review and meta-analysis of interaction models between transportation networks and territories.
arXiv preprint arXiv:2012.13367.

Raimbault, J. (2020b).
Validation levels and standards depending on models types and functions.
In *Conference on Complex Systems 2019-SIMEXPLO II Satellite*.

Raimbault, J., Chasset, P.-O., Cottineau, C., Commenges, H., Pumain, D., Kosmopoulos, C., and Banos, A. (2021).
Empowering open science with reflexive and spatialised indicators.
Environment and Planning B: Urban Analytics and City Science, 48(2):298–313.

References III

Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., and Reuillon, R. (2019).
Space matters: Extending sensitivity analysis to initial spatial conditions in geosimulation models.
Journal of Artificial Societies and Social Simulation, 22(4).

Raimbault, J., Perret, J., and Reuillon, R. (2020).
A scala library for spatial sensitivity analysis.
GISRUK.

Reuillon, R., Leclaire, M., and Rey-Coyrehourcq, S. (2013).
Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models.
Future Generation Computer Systems, 29(8):1981–1990.

Varenne, F. (2018).
Théories et modèles en sciences humaines: le cas de la géographie.
Éditions matériologiques.

References IV

Wegener, M. (2011).
The irpud model.
Spiekermann & Wegener in Dortmund. Available online: http://www.spiekermann-wegener. com/mod/pdf/AP_1101_IRPUD_Model. pdf (accessed on 1 December 2011).