Validation levels and standards depending on models types and functions

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

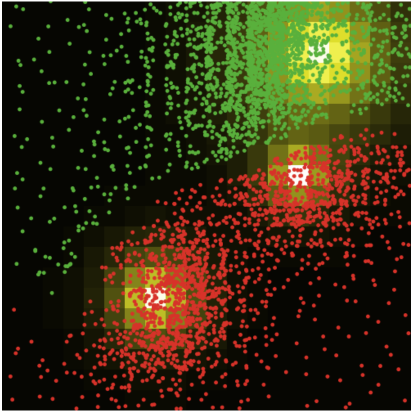
[1] CASA, UCL
[2] UPS CNRS 3611 ISC-PIF
[3] UMR CNRS 8504 Géographie-cités

* juste.raimbault@polytechnique.edu

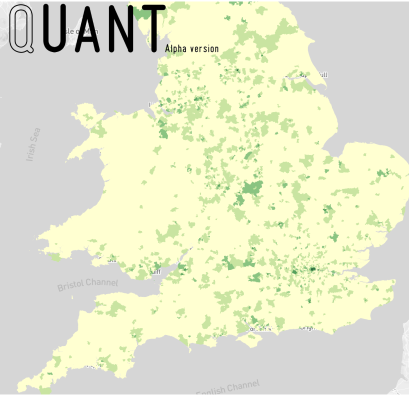
OpenMOLE

CCS 2019
Satellite SIMEXPLO
October 2nd, 2019

(Un-)validated models?



Schelling model (toy model)



Quant model (operational models)

Model validation

Proposed definition: *increasing the confidence in a model to fit its purpose*

Depends on:

- model nature/type
- model purpose
- discipline
- particular problem or application case
- expected standards
- background or mood (!) of the reviewer/reader/listener
- ...

Model validation

→ Validation has very different implications depending on epistemological positioning: from an objective procedure (reductionism) to a more conversational and reflexive process (holistic) [Barlas and Carpenter, 1990]

→ How disciplines are positioned, political relations, effective citation practices, etc. are all aspects of implicit "social" model validation

Predictive models

In geosciences (hydrology e.g. [Legates and McCabe Jr, 1999]), quantitative agreement between model and data

$\rightarrow$ choice among numerous indicators to quantify the agreement

$\rightarrow$ robust indicators? choice can be validated itself

In practice, not systematically done, as for example for land-use change models [van Vliet et al., 2016]

Microsimulation models enter a similar context (e.g. [Park and Schneeberger, 2003] for the Vissim traffic model), in a slightly different way than agent-based models

Statistical models exhibit different measures of “model quality”:

- predictive power (explained variance)
- p-value (alpha errors) and beta power (false positives)

Following [Saltelli, 2019], mathematical modeling may benefit similar standards as in statistics

$\rightarrow$ to what extent of analytical resolution is a model "validated"? (limit theorem, restricting assumptions, unfeasible ranges in practice, ...)

$\rightarrow$ finally most of the time coupled with numerical simulation? see coupling of machine learning and mathematical modeling [Butler et al., 2018] or statistical inference [Bzdok et al., 2018]

$\rightarrow$ computational turn of science [Arthur, 2013]?

On the link with simulation models:

- Formal proof systems remain limited
- Undecidability of the Turing machine Halting problem

Overview of simulation model validation methods and processes by [Sargent, 2010]

1. independent validation and verification (modelers as cognitive agents [Giere, 1990]
2. iterative process between conceptual, computerized models, and the system itself
3. Numerous validation techniques: comparison, extreme conditions, historical data, internal validity, sensitivity analysis, predictive performance, Turing test
4. Specific techniques for operational validity
5. Documentation of the validation process is crucial
6. Accreditation: science as a social process

[Landry et al., 1983] similar in operations research

Simulating the evolution of a system in a generative way: [Epstein and Axtell, 1996]: "*if you did not grow it, you did not explained it*"

→ similar to *Pattern Oriented Modeling* [Grimm et al., 2005]: reconstruct (macro) patterns from the bottom-up

Implications for validation:

- Crucial role of indicator choice (see e.g. link prediction vs. network structure reconstruction)
- fine understanding of model behavior
- role of processes and parameters
- controlled experiments (*virtual laboratories*)

→ typical example of **explication/comprehension** models (but which can also be statistical, analytical)

Sensitivity analysis

→ Sensitivity analysis is part of a model validation process [Saltelli et al., 2010]: how does a model behave in response to variations in its parameters/variables/input data?

→ Articulation of complementary methods [Cariboni et al., 2007] (validation is then the full cascade of successive methods applied)

Sensitivity analysis: examples of properties

Design of Experiments

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

Sensitivity analysis

	Coverage	Interpretability	Budget
Morris	✗	✓	✓
Saltelli	✓	✓	✗

Exploration of simulation models

OpenMOLE

Model exploration is running a simulation model, 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.

$\rightarrow$ part of model validation also

Explicative / comprehensive models are mostly made useful by their exploration

Advanced exploration methods

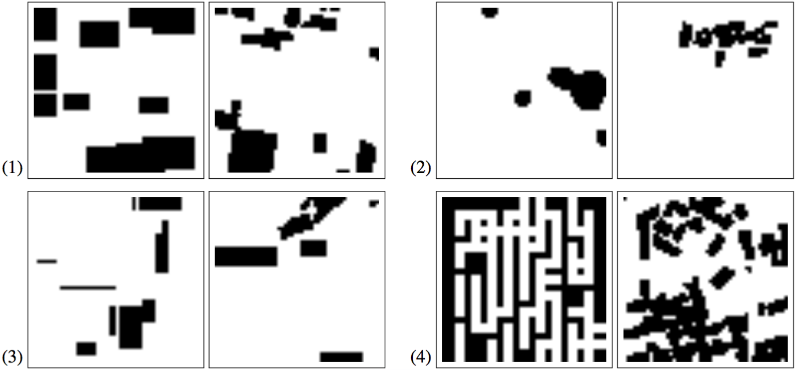
Example of validation methods included in OpenMOLE:
Calibration: Evolutionary (GA) and Bayesian (ABC) methods

Diversity Search: unveil the variety of obtainable patterns in output space: can the model produce unexpected patterns, and if so what does it means for its mechanisms ?

Origin Search: inverse problem, tackling the problem of equifinality

Spatial sensitivity analysis techniques

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



Multi-modeling

$\rightarrow$ Validation of submodels to foster diverse questions and approaches

$\rightarrow$ Validation of coupled models remains an open question (e.g. error propagation techniques)

$\rightarrow$ Comparison of the model with alternative formalisms: for example agent-based modeling against differential equations

$\rightarrow$ Importance of systematic model benchmarks/classifications

$\rightarrow$ Occam's razor and parsimony plays a certain role for model validity in this context

Varenne's model function families [Varenne, 2017]:

- **Perception and observation**: perception medium, visualization, experimental medium
- **Understanding**: description, prediction, explication, comprehension
- **Theory construction**: interpretation of a theory, test of internal coherence, applicability, co-computability
- **Communication**: scientific communication, stakeholders involvement
- **Decision-making**: planning, decision-making, self-fulfilling system prescription

Validation and model functions

- **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.)?

Validation and model type

- **Statistical**: model fit/statistical power
- **Machine learning**: predictive power
- **Analytical**: level of resolution, genericity
- **Simulation/generative**: model behavior, sensitivity analysis, pattern reconstruction, causal processes
- **Operational**: planning/decision-making relevance
- ...

Rq: classification of "model types" can neither be exhaustive nor consistent

Social aspects of validation

- Acceptance and impact within the discipline/specific subject of study
- Impact in other disciplines
- Impact outside of science
- Interdisciplinary/bridging/integrative role [Raimbault and Pumain, 2019]
- Different dimensions: complex and multidimensional nature of scientometrics [Raimbault and Pumain, 2019] [Cronin and Sugimoto, 2014]
- ...

Validation within a knowledge framework

OpenMOLE

Epistemological foundations of a knowledge framework for integrated approaches to complex systems [Raimbault, 2017], coined by [Raimbault and Pumain, 2019] as **Applied Perspectivism**:

Giere's cognitive approach to science [Giere, 1990] : cognitive agents have *perspectives* on aspects of the real world.

Scientific perspectivism [Giere, 2010] : *cognitive agents* use *media*, the models, to represent something with a certain purpose.

[Varenne, 2017]'s classification of main model functions : perception and observation, understanding, theory building, communication, decision making.

Knowledge domains

OpenMOLE

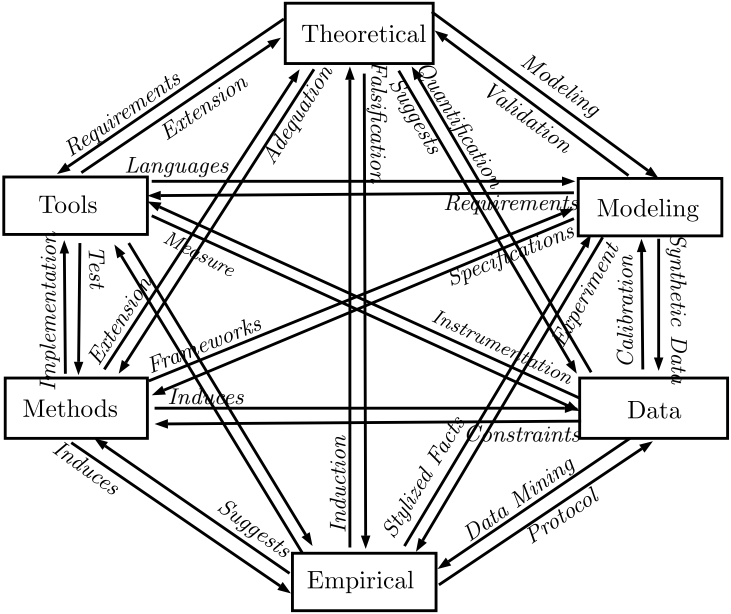
Definition of Knowledge Domains :

- **Empirical.** Empirical knowledge of real world objects.
- **Theoretical.** Conceptual knowledge, implying cognitive constructions.
- **Modeling.** The model as the formalized *medium* of the perspective.
- **Data.** Raw information that has been collected.
- **Methods.** Generic structures of knowledge production.
- **Tools.** Implementation of methods and supports of others domains.

Description of the Knowledge Framework :

1. Any scientific knowledge construction on a complex system can be understood as a perspective, decomposed into knowledge domains.
2. Contents within domains *coevolve* [Holland, 2012] between themselves and with other elements of the perspective (including cognitive agents and the purpose).
3. It implies weak emergence [Bedau, 2002] what is consistent with the existence of bodies of knowledge.

OpenMOLE


Theoretical
Tools
Modeling
Methods
Data
Empirical
Requirements
Extension
Modeling
Validation
Adequation
Falsification
Suggests
Quantification
Languages
Requirements
Measure
Specifications
Implementation
Test
Extension
Frameworks
Instrumentation
Experiment
Calibration
Synthetic Data
Induces
Constraints
Induces
Suggests
Induction
Stylized Facts
Data Mining
Protocol

Validation within the knowledge framework

OpenMOLE

$\rightarrow$ Role and type/method of validation are proper to each perspective

$\rightarrow$ Links and interaction between domains are part of the model/theory construction process and thus of validation of the perspective

$\rightarrow$ Intrinsically iterative nature of validation

$\rightarrow$ Cannot be dissociated (at least for the study of complex systems) to new methods and tools

Conclusion

→ Meaning of "model validation" is indeed strongly dependant on its properties, including type, function, context of application, discipline

→ *Obvious?* Not for all seeing some debates/questions here and there. **Interdisciplinarity requires an opening to other standards/definitions/viewpoints**

→ Validation within the Applied Perspectivism knowledge framework: validation proper to each perspective and to the coupling of perspectives, intrinsically iterative

→ Construction of integrative theories and models implies this multiple view of model validation and the variety of methods and tools, in particular in the case of simulation models

Arthur, W. B. (2013).
Complexity economics.
Oxford University Press, Oxford.

Barlas, Y. and Carpenter, S. (1990).
Philosophical roots of model validation: two paradigms.
System Dynamics Review, 6(2):148–166.

Bedau, M. (2002).
Downward causation and the autonomy of weak emergence.
Principia: an international journal of epistemology, 6(1):5–50.

Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., and Walsh, A. (2018).
Machine learning for molecular and materials science.
Nature, 559(7715):547.

References II

Bzdok, D., Altman, N., and Krzywinski, M. (2018).
Points of significance: statistics versus machine learning.

Cariboni, J., Gatelli, D., Liska, R., and Saltelli, A. (2007).
The role of sensitivity analysis in ecological modelling.
Ecological modelling, 203(1-2):167–182.

Cronin, B. and Sugimoto, C. R. (2014).
Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact.
MIT Press.

Epstein, J. M. and Axtell, R. (1996).
Growing artificial societies: social science from the bottom up.
Brookings Institution Press.

Giere, R. N. (1990).
Explaining science: A cognitive approach.
University of Chicago Press, Chicago, ISBN: 9780226292069.

Giere, R. N. (2010).
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.

Holland, J. H. (2012).
Signals and boundaries: Building blocks for complex adaptive systems.
MIT Press.

Landry, M., Malouin, J.-L., and Oral, M. (1983).
Model validation in operations research.
European journal of operational research, 14(3):207–220.

Legates, D. R. and McCabe Jr, G. J. (1999).
Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation.
Water resources research, 35(1):233–241.

References V

Park, B. and Schneeberger, J. (2003).
Microscopic simulation model calibration and validation: case study of vissim simulation model for a coordinated actuated signal system.
Transportation Research Record, 1856(1):185–192.

Raimbault, J. (2017).
An applied knowledge framework to study complex systems.
In *Complex Systems Design & Management*, pages 31–45.

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

Raimbault, J. and Pumain, D. (2019).
Exploration methods for simulation models.
arXiv preprint arXiv:1905.07160.

References VI

Saltelli, A. (2019).
A short comment on statistical versus mathematical modelling.

Nature communications, 10(1):1–3.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010).
Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index.
Computer Physics Communications, 181(2):259–270.

Sargent, R. G. (2010).
Verification and validation of simulation models.
In *Proceedings of the 2010 Winter Simulation Conference*, pages 166–183. IEEE.

van Vliet, J., Bregt, A. K., Brown, D. G., van Delden, H., Heckbert, S., and Verburg, P. H. (2016).
A review of current calibration and validation practices in land-change modeling.
Environmental Modelling & Software, 82:174–182.

Varenne, F. (2017).
Théories et modèles en sciences humaines. Le cas de la géographie.
Editions Matériologiques.