



Making Modelling Assumptions More Explicit

Bruce Edmonds^{1,2,3} and many others...

¹Centre for Policy Modelling, UK.

²UiT Norway University of the Arctic, Tromsø.

³Umeå University, Sweden.

The other authors...

These ideas have been developed by the **Open Modeling Foundation**'s standards sub-group on **Assumptions**, including (in no particular order):

Tati Micheletti, Christian Vincenot, Maria E. Pierce, Carsten Lemmen, Volker Grimm, Geerten Hengeveld, Hsiao-Hsuan Wang, Tomasz E. Koralewski, William E. Grant, Caroline Rosello, Alexey Voinov, Janneke Remmers, Lieke Melsen, Serena Hamilton, and Stephen Osakpolor

The Problem

“We still have a culture where it is considered more original and valuable to develop new models than to start from existing ones”

(Grimm, 2023)

- Resulting in an overabundance of models – the “YAAWN” syndrome (O’Sullivan et al. 2016)
- Partly due to the necessary complexity of models
- but also due to inadequate documentation...
- We focus on *listing all the assumptions in a model*

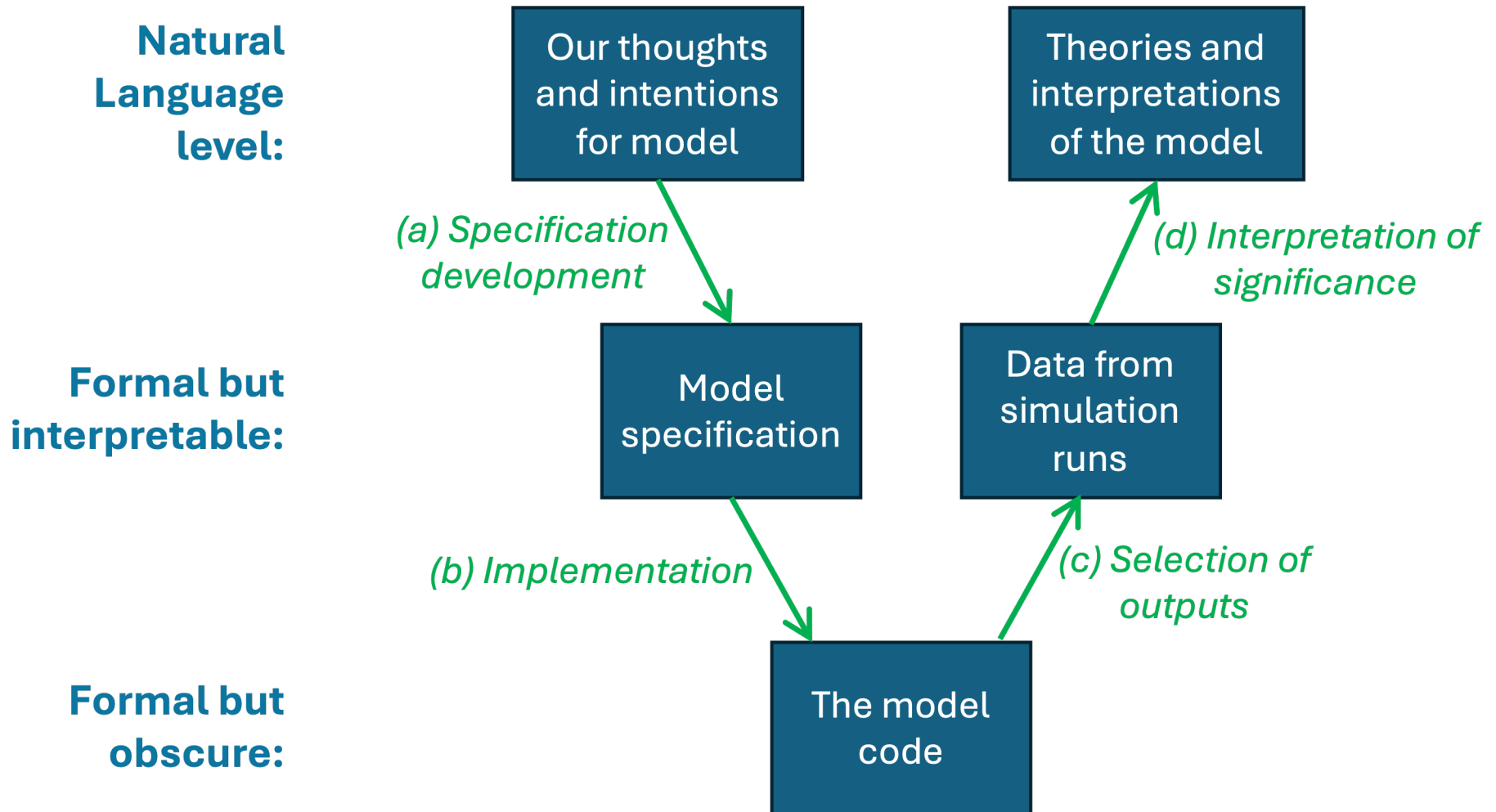
Unearthing Assumptions

- Unearthing all the assumptions in a formal logical or mathematical system is relatively easy
- Doing the same with informal reasoning (i.e. thought or natural language) is very hard (with assumptions behind assumptions behind assumptions etc.)
- The case for complex computational modelling is somewhere in between:
 - it is possible to discover all their assumptions but this is much harder if one only has: a narrative account plus the code to work with (even if one has an ODD)

In complex computational modelling...

- (a) the assumptions actually in our code might be different from what we have in our mind or our specification,
- (b) we can look at the code of a simulation and directly analyse it for its effective assumptions, but this is hard to do – and almost impossible in the absence of good, multilevel documentation of that code (Polhill & Edmonds 2015)
- (c) we can inspect the code and analyse runs of the simulation and so find out their effect in an experimental manner

Key stages in the modelling process where assumptions may be introduced



The practical outcomes are...

- 1) assumptions are distributed across sections and documents
- 2) they are often implicit, so one has to read between the lines of a paper or disambiguate parts,
- 3) many are omitted for a variety of reasons – some may just seem obvious to the authors, or they forgot that they made these choices
- 4) they have been incorporated into a model without the authors even being aware of them.

Our Proposal

*Authors of publications should
list all assumptions in a single,
accessible location...*

...to facilitate comparison,
replication, and extension of models

At one of three levels

- *Bronze level*: All the assumptions are listed in one place in the documentation (e.g. Appendix or SI).
- *Silver level*: As above, but as a table with columns for:
 - a) a short description of the assumption
 - b) a label indicating where the assumption came
 - c) the importance of the assumption for the described conclusions (high/medium/low)
 - d) the reliability of the assumption – can we rely upon it? (high/medium/low)
- *Gold level*: As for Silver, but also describing more about the provenance of the assumption (for example: at which stage of modelling process did the assumption emerge or was discussed, what was the context of that decision, and what were its presumed implications)

Advantages

- a) it would make it easier for readers to assess the scope of a presented model (the set of circumstances in which the model holds)
- b) easier to reproduce or extend models
- c) it would facilitate model comparison
- d) it would encourage modellers to document and think carefully about the assumptions they use
- e) help formulate communication to stakeholders (e.g. using “PAVE” , Bell & al. 2026)

Feasibility

It is very hard to list *all* assumptions BUT there are several accessible sources to look for these:

- a) decisions as to what to include in the conceptual modelling/specification development phase
- b) implementation changes and decisions (e.g. in a “TRACE” doc, Grimm et al. 2015)
- c) any assumptions that are widely discussed and/or contentious in the field
- d) which output measures/visualisations were chosen in the description of the results
- e) assumptions behind any explicit theories being used or tested
- f) the reasons behind the choice of experiments

References

- Bell, A. R. & al. (2026) The PAVE pathway: how to talk about models with people who built them, and people who didn't, BioScience, biag010, DOI: 10.1093/biosci/biag010
- Edmonds, B. & Polhill, G. (2015) Open Modelling for Simulators. In Terán, O. & Aguilar, J. (Eds.) Societal Benefits of Freely Accessible Technologies and Knowledge Resources. IGI Global, 237-254. DOI: 10.4018/978-1-4666-8336-5
- Grimm, V. & al. (2015) Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. Ecological Modelling, 280, 129–139. DOI: 10.1016/j.ecolmodel.2014.01.018
- Grimm, V. (2023). Ecology needs to overcome siloed modelling. Trends in Ecology & Evolution, 38(12), 1122-1124. DOI: 10.1016/j.tree.2023.09.011
- Moreau, L. & al. (2011). The open provenance model core specification (v1. 1). Future generation computer systems, 27(6), 743-756. DOI: 10.1016/j.future.2010.07.005
- O'Sullivan, & al. (2016). Strategic directions for agent-based modeling: avoiding the YAAWN syndrome. Journal of land use science, 11(2), 177-187. DOI: 10.1080/1747423X.2015.1030463

The End!



These slides at: <http://cfpm.org/slides>

Bruce Edmonds' Publications:
<http://bruce.edmonds.name/pubs.html>