Brief Note on the Logic of Replicating Implementations Before and After Publishing a Model

Nuno David

Department of Information Science and Technology
Instituto Superior de Ciências do Trabalho e da Empresa (ISCTE)
Lisboa, Portugal
Nuno.David@iscte.pt

Abstract. This short paper introduces a hypothesis complementary to the current logic of replicating computerized models in social simulations. It is submitted that the logic of cumulative knowledge provided by the process of peer reviewing may become more resourceful once the effort of replication becomes focused on producing multiple computerized versions of the same conceptual model *before* actually submitting it for peer reviewing, in contrast to the usual practice of having other teams aligning and replicating models after they have already been reviewed, accepted and published. Note that while this is a similar approach to Edmonds and Hales' proposal, it differs from a methodological point of view, insofar as they double replicated an already published model. While this approach would require more manpower in simulation projects, it is argued here that the resulting benefits may override the costs, particularly in participative-based simulations, where potential under-verified simulations may lead, nevertheless, to the actual implementation of policies with unpredictable effects in the economic and social life of stakeholders.

1 The Logic of Replication and Challenges in Simulation

The goal of this short article is to contribute a few reflections for the problem of replicating social science simulations. A hypotheses is introduced which may be both complementary and alternative to the current logic of replicating computerized models in social simulations, namely: that the logic of cumulative knowledge provided by the process of peer reviewing may become more resourceful once the effort of replication becomes focused on producing multiple computerized versions of the conceptual model before submitting the latter for peer reviewing. This contrasts the usual practice of having other teams replicating models after they have already been reviewed, accepted and published.

This proposal draws on methodological and epistemological arguments, as well as one technological argument:

a) Rather than the computerized model, it is the conceptual model that is susceptible to being confirmed or infirmed by the scientific community. It is

the conceptual model that is internalized, transformed, generalized and connected to other disciplines and methods, and eventually incorporated into a community consciousness,[1] in contrast to the inability of computer programs to become scrutinized by the scientific community.

b) Only very simple and highly abstract models – in the sense of the Axelrod's KISS motto – are susceptible to being re-implemented; however, if one adopts a KIDS philosophy for constructing simulations,[2] which is a typical approach in participative-based simulations, computerized models are hardly susceptible to becoming re-implemented by other teams, given the context dependency of individual and institutional stakeholders' opinions; however, these models can – and, it is argued, should – be subject to replication by the team proposing the model, *before* publishing or applying its results in the relevant field. This seems particularly important in policy making, where un-replicated simulations in complex domains run the risk of being under verified, while leading to the implementation of specific policies which influence the economic and social context of individual and institutional stakeholders.

c) Often, the correspondence among conceptual and computerized models is not only established formally and empirically, but also intentionally by researchers and stakeholders. Insofar as the semantic richness of social processes surpasses the empirical expressiveness of computer programs, the intended meanings ascribed to computer programs are evaluated experimentally in an intentional way. This means that the conceptual model is actually the one being internalized by the researchers and stakeholders along the executions of the computerized model, whereas the syntactic structure and operational semantics of the computerized model plays a secondary role in interpreting experiments.[3]

d) Articles on replication seem to have a low academic return (Galan et al. 2002). Most replications seem to be presented in forums especially dedicated to replication[4], rather than resulting from spontaneous confirmation or infirmation by other scientists according to the logic of peer reviewing and cumulative science, except when the goal is to extend the original conceptual model.

e) The increasing portability of programming languages for the Web, such as Java and XML, provides technologies for executing and visualizing models through Applets or Servlets on the Internet. These technologies may be able to facilitate the comparative analysis of different implementations of the

[1] C.f. De Millo et al. (1979).
[2] See Edmonds and Moss (2005).
[3] See David et al. (2005;2007).
[4] For instance, the well known Model-to-Model workshops.

same model, and are susceptible to being submitted to peer-review together with the description of the conceptual model.

2. Replication through N-Version in the Development Process

Given all these reasons, it is suggested that the following topics be introduced into the methodological debate:

1) A distinction between replicating a model before and after it has been submitted for peer reviewing and published for the first time;

2) Two or more replicas of the computerized model should be submitted to the peer-review process along with the conceptual model, in order to confirm the verifiability of the results obtained, under the penalty of having the article rejected for reviewing and publication.

Like Edmonds and Hales (2003), the author believes that "an unreplicated simulation is an untrustworthy simulation – one should not rely on their results, they are almost certainly wrong," in the sense that the computerized model differs from what was intended or assumed in the conceptual model. This is the well-known problem of verification.[5]

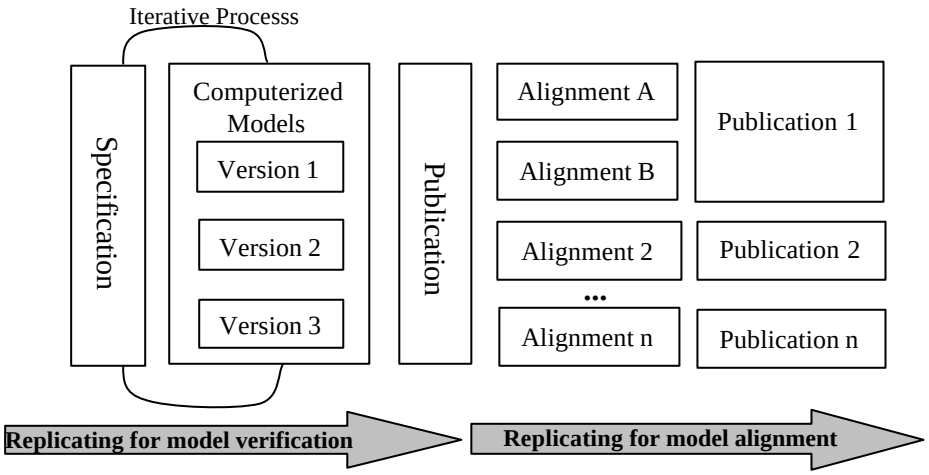

Fig. 1. Replication through multiple versions *vs*. replication for model alignment.

The report of Edmonds and Hales is indeed particularly informative in this context. In order to align with a published conceptual model, they had to rely on a double

[5] For a definition of the term "verification" in the context of simulation see David (2007).

replication process, insofar as the first replication did not seem to result in alignment. It was only after implementing a second replication of the original model that the authors were sure that the results reported in the original model did not align with their own results, and hence concluded that the original implementation was leading to possibly misleading results.

The essence of the present argument is that the burden of verifying computerized models through multiple implementations, on different platforms and by different people in the same team, should be carried out before the model is published. Consider Figure 1 above. The left part illustrates a methodology where a common specification should produce a number of different implementations carried out by different persons on the same team. The right part of the figure illustrates the usual approach to replication. Both kinds of replications have a useful role, and may be complementary.

As for the left part of Figure 1, the implementation of multiple versions before publication should be attained by an iterative process, whereby inconsistencies in the different versions would result in new corrected versions or a re-specification of the model. If all versions result in the same outputs, then there are reasonable grounds for confidence in the results obtained, and the model is ready to be published. A similar approach can be traced back to the so called N-Version Programming technique in software engineering.[6]

All versions should be made accessible to the reviewers, along with the description of the conceptual model, and reports on the equivalence of versions should be provided. The fact that the model is only published after going through a process of program verification via replication should increase overall confidence in the results obtained.

3. Prospects on N-version Programming in Simulation

It is submitted that the proposed methodology would have the following benefits:

a) It would foster the replication of simulations, regardless of the types of models and targets. Whereas the replication of participative-based simulations is rarely reported in the literature, this methodology would foster the provision of multiple implementations of the model before peer-reviewing even in the case of participative-based simulation, thereby increasing overall confidence in the model proposed, and the results obtained, whereas an un-replicated simulation in complex domains would run the risk of being under-verified but lead, nevertheless, to the actual implementation of policies

b) The burden of showing that the simulation is verified would be on the team proposing the simulation, rather than on the recipient researchers. Conversely, it would be the role of the recipient researchers to confirm or refute the validation and usefulness of the conceptual model, rather than to verify the computerized model.

[6] See Avizienis (1995).

It is also believed that the effort in replicating computerized models before submitting the corresponding conceptual model to publication comes with a significant overhead. In particular, it requires that the team enrols in multiple implementations, requiring more manpower in projects.

It is submitted, however, that the benefits obtained may override the costs, particularly in the context of participative-based simulations. If the dissemination of knowledge obtained via simulation is to acquire a reliable status of cumulative science, it is necessary to make sure that the verification stage of the simulation development process is rigorously carried out before publishing the results, and that this process can be adequately scrutinized by the scientific community after the model is published.

References

1. Avizienis, A. (1985) The N-version approach to fault-tolerant software. *IEEE Transactions on Software Engineering*, 11(12), 1491-1501.
2. David, N., Sichman, J. and Coelho, H. (2005). "The Logic of the Method of Agent-Based Simulation in the Social Sciences: Empirical and Intentional Adequacy of Computer Programs". *Journal of Artificial Societies and Social Simulation* 8(4).
3. David, N. (2007). "Verification and Validation in Social Simulation: Patterns and Clarification of Terminology Use". Article from the workshop Epistemological Perspectives on Simulation, Italy, submitted to Synthese Book Series. Draft version available at: http://iscte.pt/~nmcd/papers/drafts/PatternsNDavid.pdf.
4. David, N., Sichman J. and Coelho H. (2007. "Simulation as Formal and Generative Social Science: The Very Idea" In C. Gershenson, Aerts, D. and B. Edmonds (editors) *Worldviews, Science and Us: Philosophy and Complexity*, World Scientific, 266-284.
5. DeMillo R., Lipton R. and Perlis A. (1979). Social Processes and proofs of theorems and programs, *Communications of the ACM* 22, 5 (May), 271-280.
6. Edmonds, B. and Hales, D. (2003). Replication, replication and replication: Some hard lessons from model alignment. *Journal of Artificial Societies and Social Simulation* 6(4).
7. Edmonds, B. and Moss, S. (2005). From KISS to KIDS – an 'anti-simplistic' modelling approach. *Multi Agent Based Simulation*. Springer, Lecture Notes in Artificial Intelligence, **3415**:130–144.

6. J.S. Galán, J.M. Tom Downing, A. López-Paredes (2002) "A standardization framework for replicating agent-based models. A case study SDML vs Repast.". First conference of the European Social Simulation Association, September 18-21, Groningen, The Netherlands.