Reproducibility in Practice : Loosing your Balance on the Shoulder of Giants

Working Paper

Juste Raimbault

Monday 9th March

Abstract

The strength of science comes from the cumulative and collective nature of research, as progresses are made as Newton said "standing on the shoulder of giants", meaning that the scientific enterprise at a given time relies on all the work done before and that advances would not be possible without constructing on it. It includes development of new theories, but also extension, testing or falsifiability of previous ones. In that context

1 Introduction

As scientific reproducibility is an essential requirement for any study, its practice seems to be increasing [Stodden, 2010] and technical means to achieve it are always more developed (as e.g. ways to make data openly available, or to be transparent on the research process such as `git` [Ram, 2013], or to integrate document creation and data analysis such as `knitr` [Xie, 2013]), at least in the field of numerical modeling and simulation. However, the devil is indeed in the details and obstacles judged at first sight as minor become rapidly a burden for reproducing and using results obtained in some previous researches. We describe two cases studies where models of simulation are apparently highly reproducible but unveil as puzzles on which research-time balance is significantly under zero, in the sense that trying to exploit their results may cost more time than developing from scratch similar models.

2 On the Need to Explicit the Model

A current myth is that providing entire source code and data will be a sufficient condition for reproducibility. It will work if the objective is to produce exactly same plots or statistical analysis, assuming that code provided is the one which was indeed used to produce the given results. It is however not the nature of reproducibility. First, results must be as much implementation-independent as possible for clear robustness purposes. Then, in relation with the precedent point, one of the purposes of reproducibility is the reuse of methods or results as basis or modules for further research (what includes implementation in another language or adaptation of the method), in the sense that reproducibility is not replicability as it must be adaptable [Drummond, 2009].

Our first case study fits exactly that scheme, as it was undoubtedly aimed to be shared with and used by the community since it is a model of simulation provided with the Agent-Based simulation platform NetLogo [Wilensky, 1999]. The model is also available online [De Leon et al., 2007] and is presented as a tool to simulate socio-economic dynamics of low-income residents in a city based on a synthetic urban environment, generated to be close in stylized facts from the real town of Tijuana, Mexico. Beside providing the source code, the model appears to be poorly documented in the literature or in comments

and description of the implementation. Comments made thereafter are based on the study of the urban morphogenesis part of the model (setup for the "residential dynamics" component) as it is our global context of study [Raimbault, 2014]. In the frame of that study, source code was modified and commented, which last version is available on the repository of the project[1].

Rigorous Formalization   An obvious part of model construction is its rigorous formalization in a formal framework distinct from source code. There is of course no universal language to formulate it [Banos, 2013], and many possibilities are offered by various fields (e.g. UML, DEVS, pure mathematical formulation). No paper nor documentation is provided with the model, apart from the embedded NetLogo documentation since it only thematically describes in natural language the ideas behind each step without developing more and provides information about role of different elements of the interface.

This formulation is a key for it to be understood, reproduced and adapted ; but it also avoids implementation biases such as

- Architecturally dangerous elements : in the model, world context is a torus and agents may "jump" in the euclidian representation, what is not acceptable for a 2D projection of real world. To avoid that, many tricky tests and functions were used, including unadvised practices (e.g. dead of agents based on position to avoid them jumping).

- Lack of internal consistence : the example of the patch variable `land-value` used to represent different geographical quantities at different steps of the model (morphogenesis and residential dynamics), what becomes an internal inconsistence when both steps are coupled when option `city-growth?` is activated.

- Coding errors : in an untyped language such as NetLogo, mixing types may conduct to unexpected runtime errors, what is the case of the patch variable `transport` in the model (although no error occurs in most of run configurations from the interface, what is more dangerous as the developer thinks implementation is secure). Such problems should be avoided if implementation is done from an exact formal description of the model.

Transparent Implementation   A totally transparent implementation is expected, including ergonomics in architecture and coding, but

Expected Model Behavior   Whatever the definition, a model can not be reduced to its formulation and/or implementation, as expected model behavior or model usage can be viewed as being part of the model itself. In the frame of GIERE's perspectivism [Giere, 2010], the definition of model includes the purpose of use but also the agent who aims to use it. Therefore a minimal explication of model behavior and exploration of parameter roles is highly advised to decrease chances of misuses or misinterpretations of it. It includes simple runtime charts that are immediate on the NetLogo platform, but also indicators computations to evaluate outputs of the model. It can also be improved visualizations during runtime and model exploration, such as showed in Fig. 1.

---

[1]at `https://github.com/JusteRaimbault/CityNetwork/tree/master/Models/Reproduction/UrbanSuite`
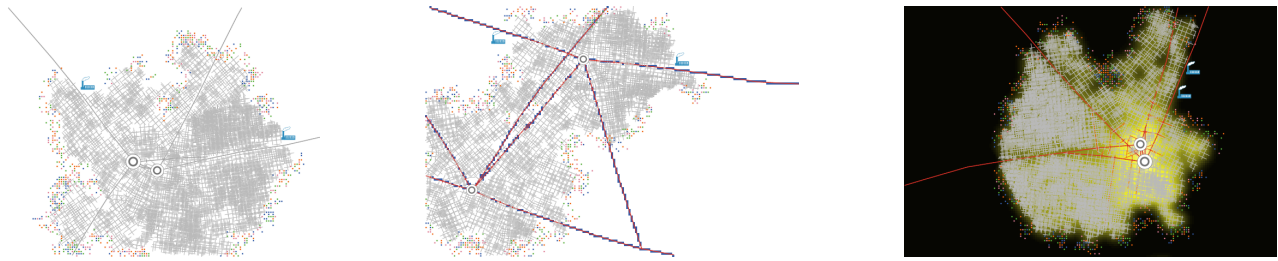
Figure 1: Example of simple improvement in visualization that can help understanding mechanisms implied in the model. *Left :* example of original output ; *Middle :* visualization of main roads (in red) and underlying patches attribution, suggesting possible implementation bias in the use of discretized trace of roads to track their positions ; *Right :* Visualization of land values using a more readable color gradient. This step confirms the hypothesis, through the form of value distribution, that the morphogenesis step is an unnecessary detour to generate a random field for which simple diffusion method should provide similar results, as detailed in the paragraph on implementation.

3   On the Need of Exactitude in Model Implementation

4   Conclusion

References

[Banos, 2013] Banos, A. (2013). Pour des pratiques de modélisation et de simulation libérées en géographies et shs. *HDR. Université Paris*, 1.

[De Leon et al., 2007] De Leon, F., Felsen, M., and Wilensky, U. (2007). Netlogo urban suite-tijuana bordertowns model. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.*

[Drummond, 2009] Drummond, C. (2009). Replicability is not reproducibility: nor is it good science.

[Giere, 2010] Giere, R. N. (2010). *Scientific perspectivism*. University of Chicago Press.

[Raimbault, 2014] Raimbault, J. (October 2014). Vers des modèles couplant développement urbain et croissance des réseaux de transports, phd project description. Technical report, Géographie-Cités UMR CNRS 8504/LVMT UMR-T IFSTTAR 9403.

[Ram, 2013] Ram, K. (2013). Git can facilitate greater reproducibility and increased transparency in science. *Source code for biology and medicine*, 8(1):7.

[Stodden, 2010] Stodden, V. (2010). The scientific method in practice: Reproducibility in the computational sciences.

[Wilensky, 1999] Wilensky, U. (1999). Netlogo.

[Xie, 2013] Xie, Y. (2013). knitr: A general-purpose package for dynamic report generation in r. *R package version*, 1(7).