Asking the Oracle: Introducing Forecasting Principles into Agent-Based Modelling

Samer Hassan[1], Javier Arroyo[1], José M. Galán[2], Luis Antunes[3] and Juan Pavón[1]

[1] GRASIA, Universidad Complutense de Madrid, Spain {samer, javier.arroyo, jpavon}@fdi.ucm.es
[2] INSISOC, Universidad de Burgos, Burgos, Spain jmgalan@ubu.es
[3] GUESS, LabMAg, Universidade de Lisboa, Portugal, xarax@di.fc.ul.pt

Abstract. A general question that often appears when working with agent-based modelling and simulation for social systems is whether it is possible to make predictions with some degree of confidence. Although many consider that agent-based models are not meant for prediction, some claim that they are mature enough to be able to predict outcomes of social processes, as it happens in other fields. This paper first reviews the current state of this debate. Afterwards, it considers how core principles from the field of forecasting can be applied in agent-based modelling. This approach intends to be useful to those modellers who look for the predictive power demanded by stake-holders and policy makers.

Key words: Agent-based modelling; Forecasting; Prediction

1 Introduction

Agent-Based Modelling (ABM) has become a widely used technique for research in Social Sciences [?], especially for understanding social phenomena or to validate social theories. Given its ability to show the evolution of complex systems, can ABM support forecasting? This question arises quite often (cf. the recent debate in SIMSOC[1], triggered by Scott Moss).

Many researchers, such as Epstein [?], place prediction as a secondary objective, arguing that there are many other possible reasons to build models different than prediction. In fact, he lists 16 of them, including explanation, guiding data collection, raise new questions or suggest analogies. He stresses his point stating that ‘Explanation does not imply Prediction’, the same way as Tectonics explains earthquakes but cannot predict them. An interesting reply to these arguments, by Thompson and Derr [?], considers that ‘good explanations predict’, as explanatory models must appropriately predict real behaviours if they seek to be considered valid. Troitzsch [?] joins the debate with an important clarification on the meanings of prediction, arguing that Epstein and Thompson discuss over different concepts. He defines three levels of prediction:

[1] SIMSOC is a mailing list for the Social Simulation field: https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=SIMSOC

1. Prediction of the kind of behaviour of a system, under arbitrary parameter combinations and initial conditions: *"Earthquakes occur because X and Y."*
2. Prediction of the kind of behaviour of a system in the near future: *"Region R is likely to suffer earthquakes in the following years because X and Y."*
3. Prediction of the state a system will reach in the near future: *"Region R will suffer an earthquake of power P in expected day D with confidence C."*

Troitzsch argues that explanation does not have to imply the 3rd level prediction (Epstein's statement refined), but that good explanations usually imply 1st and even 2nd level predictions (Thompson and Derr's point refined). In fact, Heath et al. [**?**] reach to a similar classification, which can be redefined as follows: *Generators* are models whose aim is 1st level prediction (theoretical understanding); *Mediators* are those whose aim is 2nd level prediction (insight of behaviour); and *Predictors* are models seeking 3rd level prediction (estimation). From this approach, Moss's debate in SIMSOC can be seen as the difficulty to find a *Predictor* model that has been already applied for 3rd level prediction with success. In fact, in some complex systems that present chaotic regimes, arbitrarily small variations in initial conditions can lead to very different trajectories. This implies that in those cases is demonstrably impossible to create models of the 3rd level of prediction proposed by Troitzsch. However, for the rest of complex systems, how can the 3rd level of prediction be reached?

Forecasting is a field which focuses on the study of prediction, specially the aforementioned 3rd level. It has been applied in many contexts for more than 30 years, and this experience has driven to the establishment of a set of principles that could be reviewed for ABM, if this is to be considered as a tool for making predictions. This paper explores how these principles can be applied in ABM.

2 Setting up a forecasting experiment

Forecasting is the process of making statements about future events. When there is uncertainty about a future outcome, formal forecasting procedures can help to reduce the uncertainty in order to make better decisions, especially if poor forecasts can lead to disastrous decisions. In the forecasting literature, simulation models are not usually regarded as forecasting tools (see, for instance, the taxonomy of forecasting methods in [**?**]). However, and following the discussion in section **??**, they may be considered as forecasting tools. Therefore, we believe that the standard procedure applied by the forecasting community [**?**] should guide the construction of *Predictor* models and the experiments carried out with them. This procedure should mainly consists on the following points:

- Split the data into two sets: the estimation (or training or in-sample) set, which is used to adjust the model, and the forecasting (or test or out-of-sample) set, which is used to truly check the forecasting ability of the model.
- Use an objective error measure: the forecast error in time t (e_t) is the difference between the forecasted value ($\hat{x}_t$) and the actual value (x_t) in time t, i.e. ($e_t = \hat{x}_t - x_t$). The time series of forecasted values can be obtained,

except for the case of some path-dependant models, as the mean of a sufficient number of model runs. In order to aggregate the error in t along time, many measures can be used [?], such as the Root Mean Square Error, the Mean Absolute Error and the Mean Absolute Percentage Error.
- Compare your model: if there are simulation models or forecasting methods that are known to provide reliable forecasts, they should be included in the comparison. If not, at least the naive method which assumes that the future value of a time series will be equal to the current value, should be included.
- Establish a fair comparison:
  - Use an adequate, representative and large enough sample of forecasts
  - Compare on the basis of ex-ante (each forecast only uses information available by the time of such forecast) and out-of-sample performance

These guidelines are useful for building any simulation model, but they are specially indicated when forecasting is one of the goals of the simulation. In this case, it may happen that our simulation model is not the best approach to forecast. It may be discouraging, but does not invalidate the use of the resulting model as a proof of concept to explain the dynamics of the studied phenomenon.

3 Guidelines for Forecasting with ABM

The reference book *Principles of Forecasting* [?] summarises the forecasting practice along the years and translates the findings into principles. These principles should guide a forecasting process to make it more effective. We have selected a subset of them and have adapted them to ABM, adding when possible some pertinent references from the ABM literature to illustrate the point. The selected principles mainly deal with six important issues: modelling process, use of data, space of solutions, stake-holders, validation and replication. These should serve as guidelines for forecasting with ABM[2].

Modelling Process

- **Decompose the problem into parts** (g.2.3). Do a bottom-up approach and then combine results. This synthetic approach is inherent to ABM. This decomposition is again risky, as it might not be unique, and the synthesis might be a harder problem than the target problem itself. However, the idea of approaching a problem with a computational stance necessarily implies some decomposition and synthesis.
- **Structure problems that involve causal chains** (g.2.6). Sometimes it is possible to use the results of some models as inputs to other models, and this allows for better accuracy than simulating everything simultaneously. Surely, causality can bring on deeper problems than it solves, but some kind of naïveté can be of use when delving through the web of intricate relations than build up social reality. Some ABM follow this guideline structuring the

[2] The numeration of the volume is included for direct matching, as (g.X.Y).

models in different coupled layers, each level implementing a submodel which is used as an input for the others [**?**].
- **Consider the use of adaptive forecasting models** (g.16.1). ABM is essentially an answer to this principle. Everything about ABM is an exercise of adaptability, and one that responsibly considers every component of a system that addresses a complex issue, including the human factor and the field methodologies and practices to be used.

Data-driven Modelling

- **Use theory to guide the search for information on explanatory variables** (g.3.1). This in turn allows complexity to be cut down by limiting the design space in advance. Even in utmost complex problems, there are some 'truths' and 'facts' that can be established and used to progress towards a deeper knowledge of a problem and contribute to its solutions. No truths are definitive, but that does not mean that we never know anything (e.g. Benenson's work empirical work over a previous stress resistance theoretical model [**?**]). However, this does not mean to restrict the modelling foundations just to theoretical literature [**?**].
- **Use diverse sources of data** (g.3.4). This enhances data reliability. Different sources will bring on different errors, but also different models for data collection, different methodologies for elicitation, different structures and representations, different error measures, different stances. If a compromising error is nested inside the data, it is more likely that it will be detected by explicit incoherence, than by sheer chance in the absence of different views on the data.
- **Select simple methods unless empirical evidence calls for a more complex approach** (g.6.6). This can be seen as application of the data-driven *Deepening KISS* [**?**] strategy of tackling complexity. In sum, build up a 'broad but shallow' approach to the ABM, keeping components as simple as you can, but not so simple that they become redundant. Then, introduce complexity (deeper views) by demand, both as a consequence of real complexity and as a tool to explore the space of designs. In the same line, Cioffi-Revilla proposes a systematic developmental sequence of models, increasing successively details and complexity in the models [**?**].
- **Keep forecasting methods simple** (g.7.1). Complex methods may include errors that propagate through the system or mistakes that are difficult to detect. Again, KISS, or at least, as Simple as Possible. Deepening is always possible later, either driven by demand or as an exploration strategy.

Space of Solutions

- **Identify possible outcomes prior to making forecasts** (g.2.1). This aids to know the boundaries of the space of possibilities and to structure the approach in situations where outcomes are not obvious. The goal is to avoid introducing a bias in the model, caused by overlooking a possible outcome.

- **Adjust for events expected in the future** (g.7.5). So, reduce complexity by adjusting what-if questions and perform sensitivity analysis driven by expectability. Again, when such complexity is at stake, the use of common-sense to explore the design space around what is believed to be the most probable cases is a parsimonious strategy.
- **Design test situations to match the forecasting problem** (g.13.3). Put forward scenarios to rehearse policies. Of course, for the sake of parsimony, these situations can be designed to follow the aforementioned expectability, that is, the what-if scenarios of higher likelihood. A significant example showing simulated scenarios is [**?**].

Stake-holders and Policy Makers

- **Obtain decision makers' agreement on methods** (g.1.5). Stake-holders should agree on the premisses and methods to be deployed. Their involvement in the ABM deployment and exploration is a keystone of the methodology, namely in what involves both the notion of truth/usefulness and the trust placed in the outcomes obtained. This process is usually carried out through participatory processes in modelling and validation, e.g. [**?**].
- **Ask unbiased experts to rate potential methods** (g.6.2). This emphasises the role of stake-holders and their special relation to experimenters. The involvement of stake-holder is important not only for the development and deployment of the ABM, but also for the usefulness of its outcomes, conclusions, and its permanence as a tool future decisions. Furthermore, some agent based models [**?**] are conceived as iterative projects where stake-holders and domain experts act as validation loops for each modelling iteration.
- **Test the client's understanding of the methods** (g.13.11). The role of stake-holders, especially important in participatory simulation [**?**]. But in any ABM, there should be a clear statement about what the model can and cannot yield. Never should models be sold as the ultimate solution to any problem, but rather as a tool that can and should be used to provide a deeper understanding of the problem and its foreseeable solutions. Stake-holders involvement is key to the success and usefulness of ABMs.
- **Establish a formal review process to ensure that forecasts are used properly** (g.16.4). Again, policy deployment should be controlled previously to check for appropriateness. When policies are offered to politicians, the danger is that the full consequences of the models (and their contingent nature) cannot be fully grasped. A formal procedure to be followed for use of the ABMs and their outcomes will be decisive to ensure its proper use and an adequate interpretation of its yieldings.

Validation

- **List all the important selection criteria before evaluating methods** (g.6.1). The relevant criteria should be specified at the start of the evaluation process. Although this should be obvious, the exploratory nature of ABM makes this a difficult directive. Nevertheless, to list criteria beforehand is important, as it allows for subsequent revision. The real danger comes form having no criteria whatsoever and giving in to the temptation of defining criteria later to fit the outcomes [?].

- **Use objective tests of assumptions** (g.13.2). Use quantitative approaches to test assumptions when possible. Known quantitative methods, especially statistical and stochastic methods, will strengthen the confidence in the outcomes. Once again, we try to produce as powerful a model as possible, and known objective tests will not only better support the model, but also the trust that the experimenters can have in it and its outcomes [?].
- **Use extensions of evaluations to better generalise about what methods are best for what situations** (g.13.14). Generalisation enhances the scope of applicability. These can be based in the what-if scenarios previously put forward, but should be carefully performed in order not to introduce extrapolation (or indeed interpolation) errors. Full and responsible involvement of all participants, including stake-holders is decisive, as generalisations are the first step for policy prescriptions [?].
- **Use error measures that adjust for scale in the data** (g.13.20). Error measuring is as important as accuracy of data. To know a model is to know its limitations. Errors in data are just one possible cause of trouble. In fact, this recommendation can be increased in level of abstraction and be applied to programming errors, design errors, biases, wrong objectives, etc.
- **Establish a formal review process for forecasting methods** (g.16.3). This is especially important to perform the sensitivity analysis of policies derived from ABM results. Formalism will ensure verifiability and replicability, and so will not only increase trust in the field and its methods, but can also open the gate for further scientific developments, including automatisms.

Replication

- **Compare track records of various forecasting methods** (g.6.8). This pinpoints the role of replication for verification in ABM. In such complex systems, error can come from several sources, and it is important that design and programming errors are discarded as soon as possible. Replication by different teams is a simple way of ensuring some degree of validity, and its starting to become a standard in the field. At least, the level of description of a system should be detailed enough so a replication can be built. The need for replication is well illustrated in the re-implementation that located serious weakness in the influential Axelrod's norms model [?].
- **Assess acceptability and understandability of methods to users** (g.6.9). Sharing of models and code between practitioners is facilitated by this [?]. Not all the field researchers have computer science education, so the development of workbenches and languages in which ABMs can be easily developed, debugged, tested and used is key to the development and spread of the field [?]. As a matter of fact, in such a multi-disciplinary area, accept-

ability and understandability endowing sharing might well have been the sparkle behind the huge growth of ABM.

- **Describe potential biases of forecasters** (g.13.6). Again the role of experimenter. A good encompassing model will allow for the experimenter to be represented in it, and the weaknesses of the models must be represented. Biased forecasting (as stubborn stake-holders, as...) are a liability of any model, so there should be a description of what those biases could be and how sensitive is the model to biases coming from those people involved.

4 Concluding Remarks

Many of the ABMs in the literature have explanatory vocation; they search explicit causality from the generative mechanisms that govern phenomena. In fact, the mere choice of this modelling approach implicitly indicates that finding good correlations is not enough for the modeller. If someone interested in the evolution of a given macro-variable chooses the use of ABM as modelling tool, he will be obliged to break up the system in entities and interactions, to create a detailed simulation model as inference tool and to finally aggregate again in order to analyse the dynamics of the variable. If the model is not theoretical, the researcher will also need to collect data to parametrise the sub-processes and interactions that take place in the model, which is a very demanding task and not always possible to perform. Such effort is justified only if the modeller is interested in "how" the phenomenon occurs and not just in "what" is going to happen. This implies that if the prediction power of the results relies on data, it is intrinsically more difficult to parametrise correctly an agent based model than a model simply aimed at forecasting through correlation.

Apart from that, there are additional problems to make predictions in social systems where many of the processes are modelled stochastically. A prediction in a non-deterministic model is a probability function associated with the solution space. To know if a model predicts properly, we should compare that function with the frequency of occurrence of events under the same assumptions. The problem lies in the fact that often in many social systems, as in many forecasting problems, we just have a unique realisation of the event and hence it is difficult to distinguish if the forecast matches accordingly to the confidence intervals or if the model is simply not valid.

Those two problems of forecasting with ABM, together with the behaviour of chaotic systems, are just a sample of the complexity of the target systems that we usually face with and the difficulty of making prediction in those contexts. Notwithstanding, in our opinion this is not an excuse for not introducing to our field the best-practises of other disciplines, such as forecasting. Our community should carry on applying rigour to understand "how" the phenomena occur, where complex systems sciences in general and agent based modelling in particular are leading many advances. However, we should also incorporate as far as possible those guidelines provided by the methodologies specialised in answer to "what".

References

1. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist. Open University Press (1999)
2. Epstein, J.M.: Why model? JASSS **11**(4) (2008)
3. Thompson, N.S., Derr, P.: Contra epstein, good explanations predict. JASSS **12**(1) (2009)
4. Troitzsch, K.G.: Not all explanations predict satisfactorily, and not all good predictions explain. JASSS **12**(1) (2009)
5. Heath, B., Hill, R., Ciarallo, F.: A survey of agent-based modeling practices (january 1998 to july 2008). JASSS **12**(4) (2009)
6. Armstrong, J.S., ed.: Principles of Forecasting: A Handbook for Researchers and Practitioners. Kluwer Academic Publishers, Boston (2001)
7. Chatfield, C.: Time-series Forecasting. 1 edn. Chapman and Hall/CRC Press, Boca Raton (2001)
8. Hyndman, R.J., Koehler, A.B.: Another look at measures of forecast accuracy. International Journal of Forecasting **22**(4) (2006) 679–688
9. Athanasiadis, I.N., Mentes, A.K., Mitkas, P.A., Mylopoulos, Y.A.: A hybrid agent-based model for estimating residential water demand. Simulation **81**(3) (2005) 175–187
10. Benenson, I., Omer, I., Hatna, E.: Entity-based modeling of urban residential dynamics: the case of Yaffo, Tel Aviv. Environment & Planning B: Planning & Design **29**(4) (2002) 491–512
11. Moss, S., Edmonds, B.: Towards good social science. JASSS **8**(4) (2005)
12. Hassan, S., Antunes, L., Arroyo, M.: Deepening the demographic mechanisms in a data-driven social simulation of moral values evolution. Multi-Agent-Based Simulation IX (2009) 167–182
13. Cioffi-Revilla, C.: A Methodology for Complex Social Simulations. JASSS **13**(1) (2010)
14. Becu, N., Perez, P., Walker, A., Barreteau, O., Page, C.: Agent based simulation of a small catchment water management in northern Thailand: Description of the CATCHSCAPE model. Ecological Modelling **170**(2-3) (2003) 319–331
15. Ramanath, A., Gilbert, N.: The Design of Participatory Agent-Based Social Simulations. JASSS **7**(4) (2004)
16. Pahl-Wostl, C.: Towards sustainability in the water sector-the importance of human actors and processes of social learning. Aquatic Sciences **64**(4) (2002) 394–411
17. Guyot, P., Honiden, S.: Agent-Based participatory simulations: Merging Multi-Agent systems and Role-Playing games. JASSS **9**(4) (2006)
18. Troitzsch, K.G.: Validating simulation models. Networked Simulations and Simulated Networks (2004) 265–270
19. Windrum, P., Fagiolo, G., Moneta, A.: Empirical validation of agent-based models: Alternatives and prospects. JASSS **10**(2) (2007)
20. Moss, S.: Alternative approaches to the empirical validation of agent-based models. JASSS **11**(1) (2008)
21. Galan, J., Izquierdo, L.: Appearances can be deceiving: Lessons learned re-implementing Axelrod's' evolutionary approach to norms'. JASSS **8**(3) (2005)
22. Polhill, J., Edmonds, B.: Open Access for Social Simulation. JASSS **10**(3) (2007)
23. Hassan, S., Fuentes-Fernández, R., Galán, J.M., López-Paredes, A., Pavón, J.: Reducing the modeling gap: On the use of metamodels in Agent-Based simulation. In: 6th Conference of European Social Simulation Association, Guildford, UK (2009)