A Bibliography of ABM Research Explicitly Comparing Real and Simulated Data for Validation: Version 1

Edmund Chattoe-Brown
School of Media, Communication and Sociology
University of Leicester
ecb18@le.ac.uk

2 June 2020

CPM Report CPM-20-216

Note: Please cite this document as "version 1" with the appropriate publication date as the intention is that there will be updated versions with later publication dates.

Introduction

The aim of this discussion paper is to provide the first draft of a bibliography for published ABM that attempt validation of their models against real data.[1] The broader scientific advantages for constructing such a bibliography will be discussed shortly but the initial aim is just that this bibliography should serve as a resource to the wider ABM community (and thus be developed at least in part collectively through successive versions).

Methodology

Because (as I shall shortly show) explicitly validated ABM are still rather rare (and thus resource intensive to find as well as to analyse) I began with three survey articles that I knew categorised (in at least some sense) research by its validation status. The most recent of these (Angus and Hassani-Mahmooei 2015, see Appendix 2) is not specifically designed to identify validated ABM but does so indirectly by reporting the proportion of articles (in the *Journal of Artificial Societies and Social Simulation* – hereafter *JASS*) which display both simulated and empirical "data objects" (graphs, tables and so on) between 2001 and 2012. These authors kindly provided their raw analysis spreadsheets which allowed me to identify seven specific examples of articles which all of which did, in fact, prove to attempt validation. However, because their article only considers *JASS* over about a decade, the obvious next step was to complete the sample from that journal by adding the years 1998-2000 and 2013-2019 (excluding the currently incomplete year 2020.) This involved reading each article (nothing was excluded unless it was clearly a book review so the sample

---

[1] It thus excludes "validation" against other models (see the following discussion on the challenges of useful definitions).

includes "think pieces" without a model, purely technical articles and so on) and looking primarily for real and simulated data either adjacently or displayed through the same graph/table. In addition, where these sorts of data were not immediately visible, the visual assessment was cross checked by searching the text term "validat*" (which will pick "validated" and "validation" for example) and looking at the context in which it was used.[2] This reduced the risk that non-visual attempts at validation (like reporting summary statistics) were missed. This strategy, combined with the seven instances identified by Angus and Hassani-Mahmooei, along with a small number of examples I was already aware of for other reasons, gave rise to an initial bibliography of 68 items.

Apart from creating a document of suitable length for an initial version of the bibliography, the other two review sources created practical problems that (at least at this stage) discouraged me from attempts to analyse them in full. The second most recent source (Heath *et al.* 2009, see Appendix 2) lists their sample in full and provides percentages by validation status but I was unable to contact any of the authors to discover which specific articles were validated. Thus although the sample is probably somewhat "enriched" with validation examples relative to a random search (35% of their sample are reported as both conceptually and operationally validated which turns out to be quite a lot higher than *JASSS*) there would still be a lot of effort to acquire many sources that would turn out not to be relevant. By contrast Dutton and Starbuck (1971, see Appendix 2), which deserves much better recognition for its early and encyclopaedic treatment of the field, lists every article surveyed according to several different calibration and validation categories. The problems here are that, owing to its age, many of its sources are now very hard to access (technical reports, dissertations, journals that have not been transferred to electronic format that far back) and that the survey reports on simulation as a whole (including methods of limited relevance to ABM) and examples that are not at all social (like simulations of individual decision making in an "Artificial Intelligence" vein). Thus, while this source is again considerably more productive than a random search, its nature still builds in the possibility of significant frustrating wasted effort.[3]

Empirical Context

As can be seen from the figure below, there has been an unsteady but relatively clear trend towards increased attempts at ABM validation during the life of *JASSS* (fractions are used because the number of articles in an average yearly volume have increased considerably since 1998). Taking the whole completed run of the journal (1998-2019) about 7.5% of articles attempt validation but the great majority of those are post 2010 (which was also the last year in which an annual volume apparently did not contain a single article that attempted validation). That said, however, the fraction of articles including attempted validation has barely exceeded 20% in *any* year to date.

---

[2] For example, a major use of the search term "validation" in context was discovered to be something along the lines of "we hope to achieve validation of the model in future research". The number of articles making no mention of validation at all (particularly in models of opinion dynamics or game theory) was also notable.
[3] Study of the JASSS sample also suggested that search terms for actual validation would be ineffective. Although no articles that attempt validation fail to mention the term, most articles that do *not* attempt it mention it too!

Fraction of Non Book Reviews Attempting Empirical Validation in JASSS by Year
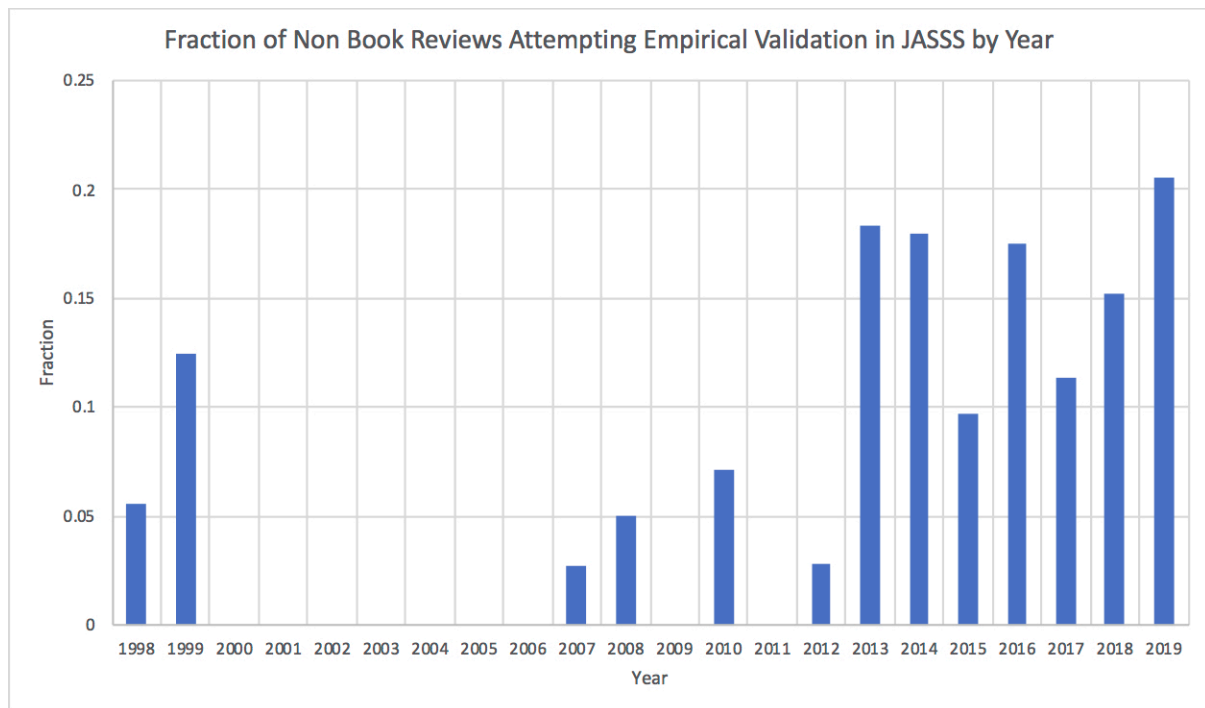
Scientific Motivation

The situation of data in ABM is uneasy. There seems to be a general approval of (and recognition of the need for) empirical grounding of models as well as widespread endorsement of the empirically based methodology laid out by Gilbert and Troitzsch (2005). Nonetheless, as the graph above shows, data in modelling (at least for validation purposes) still appears to be much more common in the breach than in the observance.

Not all reasons for this situation are necessarily problematic. It is often argued (though seldom in published articles and never, as far as I have been able to discover, in combination with an alternative methodology to replace the empirical one proposed by Gilbert and Troitzsch) that some legitimate uses of ABM (like "thought experiments") simply do not require data. It seems to me that this position could be questioned (Chattoe-Brown 2019, Edmonds *et al.* 2019) but, coming out at it does in conference debates and referee reports, it is hard to pin down for analysis. Nonetheless, it would be reasonable to say that the ideal amount of validation would perhaps never be 100% (particularly since the JASSS sample is known to contain purely technical articles, "think pieces" and so on.) Nonetheless, it would be equally fair to suggest that taking two decades to reach merely 20% attempted validation could imply a significant and extended gap between aspiration and performance. That gap could therefore legitimately be narrowed further. (Many articles at least imply that their models should be suitable for validation or that they intend to attempt it at some point.)

Given this situation, there are a number of different reasons for compiling and improving a bibliography of validation attempts in ABM:

- Without attempting to resolve the contested issue of when models "should be" validated (or calibrated), such a bibliography can help to build community of empirical modellers at least consisting in the first instance of others who have thought it worthwhile to attempt validation. The first stage in the construction of any community is the awareness that there are actually a good number of "others" out there (and knowing who they all are).

- Once one has a sample to work with (given that this is relatively difficult to build for empirical ABM), one can examine cases to identify and clarify best practice. Which studies have made the "the best" attempts at validation and what does that mean? What technical tools have been used by some researchers that perhaps should be known to all or become standard reporting techniques for certain types of analysis?

- Conversely, a sample of adequate size may allow us to identify instances of validation that work less well and understand why. Good practice often needs the contrast of poorer practice (and its harmful consequences) to support and reinforce it.

- A sample constructed by researchers largely in isolation (i. e. without the benefit of a self-conscious empirical modelling community) also helps us to clarify the definition of validation by showing us when we are comfortable (or uncomfortable) with the way the term is used and why (for example, suggesting that further distinctions and terminology may need to be developed). Issues arising from this aspect of the bibliography will be discussed in more detail in the next section.

- An adequate sample of instances allows us to systematically analyse how validation is actually done (and thus compare it with claims from various articles about how it "ought" to be done – see Appendix 3 for examples). This ensures that general advice on validation is actually feasible and does indeed deliver convincing results when followed.

- From a sample of different kinds of validation, we can Identify where suitable techniques exist (and are usable) and where they don't (or aren't) for different kinds of data. For example, Thorngate and Edmonds (2013) present a technique for comparing time series that has not been cited in *JASSS* by any of the attempted validations presented here (nor by the extensive survey of methods for "output analysis" provided in Lee *et al.* 2015).[4] Those attempting validation should know the appropriate and workable techniques for justifying results based on their particular types of data (and if they don't then their reviewers should be able to put them straight authoritatively).

- A sample of instances also allows us to identify research approaches or topics that are particularly weak or strong in validation and to reflect on these differences and perhaps affect "knowledge transfer" between more and less developed areas. For example validation is noticeably weak in the opinion dynamics field and for game theoretic approaches (unless these involve laboratory experiments) but much stronger for demography and transportation research. This is another area in which unconscious differences between research cultures may benefit from a good dose of data.

---

[4] Though obviously at least a few validation attempts predate the publication of these articles and so can hardly be criticised for not citing them!

- More generally, examples of validation enable the possibility for progressive model building: You can only do "better" in modelling if there is some relatively robust comparison measure to begin with. To my knowledge there has never yet been an attempt by one team of researchers to outperform the validation success – or lack of it – of another[5] (and certainly if this approach has been used it cannot be found in *JASSS*). Further, one can only really discuss what "better" means effectively once one has concrete examples to work from. How should we rank a model with better match but more tuneable parameters for example?

- An adequate sample also gives us a basis to judge (apart from any technical tools designed to support this judgement) whether (and when) some validations are more "believable" than others (relative to models that have not had their parameters tuned, relative to statistical predictions and so on.) Do intuitive notions of match agree with technical ones and if not, does this tell us something about problems with intuition or with the technical tools? For example, do we rate a model with a very large discrepancy in one year better or worse than a model with smaller discrepancies every year (even if the sum of the discrepancies is similar) and on what basis?

- Proper science always has to start somewhere in terms of a clear statement of position. I am actively hoping that people will query, criticise and improve on the work presented here. However bad it is, my aim is that it should be systematic and clearly justified enough to act as a solid foundation for whatever comes next. But currently the alternative to this bibliography (whatever its failings) is scratching round with difficulty to find relatively few (and possibly unrepresentative) examples of attempted validation (and this simply doesn't seem to be something that most modellers attempt unless they want to write "advice" pieces on validation and sometimes not even then.)

- A set of examples may help to promote validation where appropriate. Do we find that validated models are more cited than unvalidated ones? (If not, we might be rather worried about the scientific status of ABM?) Might people discover that at least attempting validation is not as difficult as they feared (or can be done better by building on what is already known rather than having to "start from scratch" each time?) Might reviewers be encouraged to think that validation is actually something it is reasonable to ask authors to do in certain circumstances (and thus incrementally raise modelling standards?)

Given these diverse justifications, I now turn to a limited analysis of wider issues arising from the construction of the bibliography so far.

Some Tentative Analysis

<u>Don't Say, Do!</u> There is much discussion of calibration and validation in general and of how validation might be done in principle (see Appendix 3). But the process of actually carrying out an analysis has made these relationships much clearer in practical terms. It is fairly obvious when real and simulated data are compared (particularly visually). Having identified

---

[5] The closest I have been able to find, which is still weaker than what I am proposing, is my own attempt (Chattoe-Brown 2014).

such examples, it is only then that we have to decide what to make of them. (For example, are we more impressed by match generated by independent calibration or by parameter tuning? Why?) This suggests that a corresponding study of data use for calibration (and the very small number of researchers that attempt both to any significant extent) can be carried out somewhat independently. This is a not a result that might have been expected from the theoretical discussions of calibration and validation which tend to consider them together (but see Chattoe-Brown for a discussion of the implications of different combinations of calibration or validation and their absence.) Similarly, a real sample makes it much clearer how many techniques we need to use or develop to evaluate validation attempts. The great majority of validations involve distributions or time series so these should be where we focus our attention on establishing best practice. Very few involve spatial distributions and network structures so it may be here that techniques need to be developed or improved. There is nothing like the analysis of a sample of real examples to focus effort effectively.

The Devil is in the Detail: Although the great majority of independent cases appear to agree implicitly on what validation is (comparison of "equivalent" real and simulated data, ideally on the same graph or table – thus following Gilbert and Troitzsch closely), there are some cases (listed in Appendix 4) which challenge us to consider what we are comfortable calling validation. Some of the issues involved are mundane. For example, are we prepared to accept a verbal report that a model has been validated elsewhere as evidence (particularly when, as it happens, the link to that source is broken?) Are we happy calling the comparison of two different *model* outputs validation (or should that be called replication so it is clear there need be no empirical content?) What do we make of outputs that are not, in fact, presented in a way that is directly comparable (for example in terms of time scales) or are only comparable on the basis of proxies of unknown quality (for example, does the "pedestrian footfall" of a commercial property actually tell us what its rental price is likely to be or is there only a weak correlation?) Does it make sense to talk about validating *predictions* or does that oblige us to compare models rather than actual data? (And do we then need to distinguish carefully between genuine prediction of future events and various forms of retrodiction and back casting where the data is actually known already.) Does comparison with hypotheses constitute validation or is there a risk that these could be arbitrarily selected because, whether wittingly or unwittingly, they happened to endorse a model? (By comparison, nobody would sensibly attempt to claim that a model fitted just some data points well!) It would not be appropriate for me to "adjudicate" what counts as "proper" validation (since this should be a collective decision arrived at after reasoned argument) but we do need to develop our terminology and arguments so we can say clearly what the ABM community considers to be appropriate and inappropriate use of these (and related) terms and why relevant distinctions need to be made.[6] To start this process off, however, we need some substantive examples to feed on.

---

[6] Unfortunately, the normative value attached to validation makes some authors attempt to suggest it has been done when it appears not to have been and to use terms which carry the scientific glamour of the concept without the associated effort. For example, if validation is comparison of real and simulated data, it is really not clear what "internal validation" can be (unless is means verification). One simply cannot substitute confronting a model with data by any amount of examining its abstract properties no matter how grandly one names it or with what technical skill one executes it. Other examples of this problem occur. For example, a "simulation experiment" implies a relationship with constraining reality (as in a laboratory experiment) which does not actually exist. Nobody would expect to be taken seriously if they said they were performing a "tap experiment" or a "sofa experiment".

<u>The Domain of Application</u>: Perhaps because this sample is mainly constructed from *JASSS*, there is strong agreement about the subject matter of the resulting ABM. Nearly all the models clearly deal with aspects of human sociality. The only real exceptions deal with *non-human* sociality and what one might call a "null hypothesis" approach to understanding (where the role of the simulation is not to describe a social process but to explore the effects of a particular assumption – that real marriages are contracted "randomly" – against which real data can then be compared.) As with the issue of definitions, it is not appropriate to say that either use is "wrong", only to draw attention to the value these potentially borderline cases may have in sharpening our intuition about the aims and evaluation of validation research. (Combining real data and simulation strongly supports the claim that marriage is not random but then did anybody ever suppose it was! I shall return to this point in the next subsection.)

<u>Technical Issues</u>: Real examples also concentrate the mind on the technicalities of validation. At first we simply need to have effective measures of the differences between two time series for example (see Thorngate and Edmonds 2013 and Lee *et al*. 2015) but then we also need some systematic way of deciding when this difference is small enough to constitute some sort of "result". For example, some attempted validations in the bibliography report the Kolmogorov-Smirnov (KS) test for comparing distributions (Massey 1951). However, this test is designed to show that two distributions come from the same underlying sample and it isn't totally clear that this is appropriate (or even what it means) when one distribution is based on data and the other is based on the output of a processual (not statistical) ABM. This raises another major challenge for the evaluation of validation which is being able to devise sensible null hypotheses against which results can be compared. Statistical inference, for example, can sensibly ask if there is a real difference between sample means (or an apparent difference that occurred "by chance"). But what is a useful equivalent of "by chance" for a social network? Technically, it is easy to compare an actual network with a random one but since no real network would ever be random, it isn't clear that this proves much. So then the question is, how "much match" do we need before we think a correspondence is a genuine result rather than just something that could have arisen "accidentally" (and what, therefore, can "accidentally" sensibly mean) and do existing techniques make assumptions (for example about samples) that disqualify them from evaluating comparisons of real data and model outputs? Such questions will be taken much more seriously once it is clear that ABM validation has a practical problem to solve to establish the credibility of its outcomes.

Limitations

Given the labour intensive nature of this bibliography (and the hope that it will improved collectively), it was more important that it be done at all than that it be almost perfect on the first attempt. Two obvious limitations of the methodology used suggest themselves.

Firstly, it is noticeable that all counts of validation attempts are higher in years I analysed myself than years reported by Angus and Hassani-Mahmooei. The most obvious suggestion

is therefore that our analysis methods are not comparable.[7] However, having direct access to their raw analysis classifications, it is not clear what additional categories I should have considered to find further validation attempts. Articles only displaying simulation data (or no data at all) would clearly be unlikely to qualify. Nonetheless, it is at least possible that their emphasis on objects like tables or graphs might have excluded some validations presented in the text as summary statistics. However, this raises a practical point. If one does not trust the analysis of Angus and Hassani-Mahmooei to some extent (and there seems no good reason not to), one simply has to reanalyse the entire set of articles without any particular expectation of a large number of additional cases.[8] (The annual trend in validation showed in the graph above suggests that in the early years the number of cases would still be small even if it turned out it wasn't actually zero in a number of years as reported.)

Secondly, the workload involved in analysing all these articles probably biases the sample towards the most visually salient validations. I would be more likely to miss a summary statistic embedded in text, two comparable graphs that were not close together and so on. But recall that this is a robust first attempt at a sample and no claim of perfection is implied. If I have made mistakes, I will be delighted to have them pointed out for correction in the interests of steadily improving quality. But without something definite "out there", there is nothing to criticise and make better.

Some limitations of the research almost go without saying. Perhaps *JASSS* is unrepresentatively good (or poor) in attempting validation. Perhaps other academic sources (like books) try harder. Maybe certain fields with their own specialist journals are closer to the cutting edge of validation. But one has to start somewhere.

Future Work

Despite their discouraging features, I still plan to analyse the sources from Heath *et al.* (2009) and Dutton and Starbuck (1971) for successive versions of this bibliography. (Ideally I would make contact with the authors of Heath *et al.* directly if anyone can help with that.) In addition, it makes sense to identify and analyse any more survey articles that categorise published research by its validation status (or empirical status more generally) and I would appreciate references to these.

As "curator" of the bibliography (balancing the quality control of a fixed and citable document version against the strengths – and weaknesses – of "pure" collective authorship) I will also welcome all forms of additions, deletions and amendments. If authors feel I have been unfair in characterising their articles then I should amend these. If I have left out what I should have included or vice versa then I will change anything provided supporting

---

[7] For example, it is not clear why Janssen (2009, for example Figures 2 and 4) is not counted as combining empirical and simulation data. This could be a simple oversight (the search term "Janssen" does not seem to appear in their raw analysis file) or it could reflect a deliberate sampling strategy (since this is a restudy of an existing model rather than a novel contribution).

[8] In fact, Heath *et al.* (2009) also have a significant number of *JASSS* articles in their sample and their timescale for sampling is "within" that of Angus and Hassani-Mahmooei. Looking at these articles will thus provide an interesting cross check on their categorisation.

arguments are offered (though the decision of the curator remains, as they say, final). I freely admit that in a rather small number of cases I found myself unable to be sure exactly what authors were claiming about validation. I will add references that I (or others) find with full credit. Contributors who take the trouble to add substantially and systematically to the size of the bibliography will be eligible for co-authorship of later versions.

I am also keen to investigate the extent to which advice on validation (and formal techniques) overlap with actual examples of attempted validation and whether this casts light on the value and viability of the advice/techniques or on the examples. How feasible is it to do what is recommended and, if it is done, how much does it add to the credibility of the validation? The gap between examples and "principles" in terms of citation and discussion is notable (as the example of Thorngate and Edmonds illustrates).

I am also starting to collect notes on an equivalent bibliography on calibration and would welcome collaborators for that project.

References

Chattoe-Brown, Edmund (2014) 'Using Agent Based Modelling to Integrate Data on Attitude Change', *Sociological Research Online*, **19**(1), February, article 16. doi:0.5153/sro.3315

Chattoe-Brown, Edmund (2019) 'Agent Based Models', in Atkinson, Paul, Delamont, Sara, Cernat, Alexandru, Sakshaug, Joseph W. and Williams, Richard A. (eds.) *SAGE Research Methods Foundations*. doi:10.4135/9781526421036836969

Edmonds, Bruce, le Page, Christophe, Bithell, Mike, Chattoe-Brown, Edmund, Grimm, Volker, Meyer, Ruth, Montañola-Sales, Cristina, Ormerod, Paul, Root, Hilton and Squazzoni, Flaminio (2019) 'Different Modelling Purposes', *Journal of Artificial Societies and Social Simulation*, **22**(3), June, article 6. doi:10.18564/jasss.3993

Gilbert, Nigel and Troitzsch, Klaus G. (2005) *Simulation for the Social Scientist*, second edition (Milton Keynes: Open University Press).

Janssen, Marco A. (2009) 'Understanding Artificial Anasazi', *Journal of Artificial Societies and Social Simulation*, **12**(4), October, article 13. <http://jasss.soc.surrey.ac.uk/12/4/13.html>

Lee, Ju-Sung, Filatova, Tatiana, Ligmann-Zielinska, Arika, Hassani-Mahmooei, Behrooz, Stonedahl, Forrest, Lorscheid, Iris, Voinov, Alexey, Polhill, Gary, Sun, Zhanli and Parker, Dawn C. (2015) 'The Complexities of Agent-Based Modeling Output Analysis', *Journal of Artificial Societies and Social Simulation*, **18**(4), October, article 4. <http://jasss.soc.surrey.ac.uk/18/4/4.html> doi:10.18564/jasss.2897

Massey, Frank J. Junior (1951) 'The Kolmogorov-Smirnov Test for Goodness of Fit', *Journal of the American Statistical Association*, **46**(253), xx, pp. 68-78. doi:10.1080/01621459.1951.10500769

Thorngate, Warren and Edmonds, Bruce (2013) 'Measuring Simulation-Observation Fit: An Introduction to Ordinal Pattern Analysis', *Journal of Artificial Societies and Social Simulation*, **16**(2), March, article 4. <http://jasss.soc.surrey.ac.uk/16/2/4.html> doi:10.18564/jasss.2139

Appendix 1: Bibliography of ABM Attempting Validation

In addition to a full citation, this bibliography provides a brief summary of what real and simulated data are compared and where in the article this comparison can be found. Please correct descriptions that seem to be factually inaccurate – particularly if you are one of the authors of the article – or do not stand alone sufficiently to explain the attempted validation without having to read the article itself.

Unpaginated sources use table/figure numbers to locate comparisons. Paginated sources give page numbers and table/figure numbers where both are available.

Abdou, Mohamed and Gilbert, Nigel (2009) 'Modelling the Emergence and Dynamics of Social and Workplace Segregation', *Mind and Society*, **8**(2), December, pp. 173-191. Special Issue: Social Simulation. doi:10.1007/s11299-009-0056-3 [Compares observed and simulated values for social segregation, workplace segregation and unemployment rates for Muslims and Copts. Table 3.]

Adam, Carole and Gaudou, Benoit (2017) 'Modelling Human Behaviours in Disasters from Interviews: Application to Melbourne Bushfires', *Journal of Artificial Societies and Social Simulation*, **20**(3), June, article 12. <http://jasss.soc.surrey.ac.uk/20/3/12.html> doi:10.18564/jasss.3395 [Compares real and simulated rates of death by cause. Figure 4.]

Ahrweiler, Petra, Schilperoord, Michel, Pyka, Andreas and Gilbert, Nigel (2015) 'Modelling Research Policy: Ex-Ante Evaluation of Complex Policy Instruments', *Journal of Artificial Societies and Social Simulation*, **18**(4), October, article 5. <http://jasss.soc.surrey.ac.uk/18/4/5.html> doi:10.18564/jasss.2927 [Compares a variety of real and simulated network properties in the research/funding community, a wide range of real and simulated system attributes, real and simulated project sizes and project contributions. Figures 5-11. Tables 2 and 3.]

Amin, Engi, Abouelela, Mohamed and Soliman, Amal (2018) 'The Role of Heterogeneity and the Dynamics of Voluntary Contributions to Public Goods: An Experimental and Agent-Based Simulation Analysis', *Journal of Artificial Societies and Social Simulation*, **21**(1), January, article 3. <http://jasss.soc.surrey.ac.uk/21/1/3.html> doi:10.18564/jasss.3585 [Compares experimental game theory data with simulation over time in a contribution game. Figure 1.]

Barthelemy, Johan and Toint, Philippe (2015) 'A Stochastic and Flexible Activity Based Model for Large Population. Application to Belgium', *Journal of Artificial Societies and Social Simulation*, **18**(3), June, article 15. <http://jasss.soc.surrey.ac.uk/18/3/15.html> doi:10.18564/jasss.2819 [Compares real and simulated house departure times by purpose and real and simulated travel distance by purpose. Figures 3-6.]

Bergman, Noam, Haxeltine, Alex, Whitmarsh, Lorraine, Köhler, Jonathan, Schilperoord, Michel and Rotmans, Jan (2008) 'Modelling Socio-Technical Transition Patterns and Pathways', *Journal of Artificial Societies and Social Simulation*, **11**(3), June, article 7. <http://jasss.soc.surrey.ac.uk/11/3/7.html> [Graph compares real and simulated UK tonnage of sail and steam vessels 1850-1914. Figure 11.]

Boero, Riccardo, Bravo, Giangiacomo, Castellani, Marco and Squazzoni, Flaminio (2010) 'Why Bother with What Others Tell You? An Experimental Data-Driven Agent-Based Model', *Journal of Artificial Societies and Social Simulation*, **13**(3), June, article 6. <http://jasss.soc.surrey.ac.uk/13/3/6.html> doi:10.18564/jasss.1620 [Graphs real and simulated data on endowment, best discovered and final profit over experiment rounds and also compares summary statistics for these. Table 3 and Figures 1-3.]

Bravo, Giangiacomo, Squazzoni, Flaminio and Boero, Riccardo (2012) 'Trust and Partner Selection in Social Networks: An Experimentally Grounded Model', *Social Networks*, **34**(4), October, pp. 481-492. doi:10.1016/j.socnet.2012.03.001 [Compares fractional distributions of investments and returns for a laboratory experiment and an ABM. Also reports average investment and return values for experiment and variant simulations with different network structures. Figure 2, Figure 4 and Table 2.]

Buttò, Michele, Pereira, Carlos and Taylor, Matthew (2014) 'Sunshine or Shield? Secret Voting Procedures and Legislative Accountability', *Journal of Artificial Societies and Social Simulation*, **17**(4), October, article 10. <http://jasss.soc.surrey.ac.uk/17/4/10.html> doi:10.18564/jasss.2620 [Compares certainty value against probability of vote for simulated data and a statistical model for a range of voting outcomes. Figure 2.]

Cardona, Andrés (2014) 'Parental Choices and Children's Skills: An Agent-Based Model of Parental Investment Behavior and Skill Inequality Within and Across Generations', *Journal of Artificial Societies and Social Simulation*, **17**(4), October, article 8. <http://jasss.soc.surrey.ac.uk/17/4/8.html> doi:10.18564/jasss.2595 [Compares simulated model trends – for example negative sloped relationships between variables – with corresponding plots of real data by country. Figures 4-6.]

Carley, Kathleen M., Prietula, Michael J. and Lin, Zhiang (1998) 'Design Versus Cognition: The Interaction of Agent Cognition and Organizational Design on Organizational Performance' *Journal of Artificial Societies and Social Simulation*, **1**(3), June, article 4. <http://jasss.soc.surrey.ac.uk/1/3/4.html> [Compares team performance as percentages and "scores" between human experiments and various simulations. Tables 2 and 3.]

Cedeno-Mieles, Vanessa, Hu, Zhihao, Ren, Yihui, Deng, Xinwei, Adiga, Abhijin, Barrett, Christopher, Contractor, Noshir, Ekanayake, Saliya, Epstein, Joshua M., Goode, Brian J., Korkmaz, Gizem, Kuhlman, Chris J., Machi, Dustin, Macy, Michael W., Marathe, Madhav V., Ramakrishnan, Naren, Ravi, S. S., Saraf, Parang and Self, Nathan (2020) 'Networked Experiments and Modeling for Producing Collective Identity in a Group of Human Subjects Using an Iterative Abduction Framework', Social Network Analysis and Mining, 10, 7 January, article 11. doi:10.1007/s13278-019-0620-8 [Compares model and experimental density data on distribution of replies received, replies sent, requests received, requests

sent and words submitted in a cooperative problem solving task. Does this for a baseline model and two variants. Figure 9, 11 and 14.]

Chattoe-Brown, Edmund (2014) 'Using Agent Based Modelling to Integrate Data on Attitude Change', *Sociological Research Online*, **19**(1), February, article 16. doi:0.5153/sro.3315 [Compares graphs of the distributions of positive, neutral and negative attitudes over time for homosexual sexual relations, the death penalty and the reunification of Ireland with the corresponding distributions in a simulation of attitude change. Figures 1-3 and 5.]

Christensen, Keith and Sasaki, Yuya (2008) 'Agent-Based Emergency Evacuation Simulation with Individuals with Disabilities in the Population', *Journal of Artificial Societies and Social Simulation*, **11**(3), June, article 9. <http://jasss.soc.surrey.ac.uk/11/3/9.html> [Compares simulated number of evacuees (some of whom have disabilities) from different exits in a building with corresponding real "evacuation drill" data and also compares the total time taken in simulated and "drill" evacuations.]

Civico, Marco (2019) 'The Dynamics of Language Minorities: Evidence from an Agent-Based Model of Language Contact', *Journal of Artificial Societies and Social Simulation*, **22**(4), October, article 3. <http://jasss.soc.surrey.ac.uk/22/4/3.html> doi:10.18564/jasss.4097 [Compares real and simulated percentage of Romansh and German speakers over time. Figure 4.]

Cottineau, Clémentine, Chapron, Paul and Reuillon, Romain (2015) 'Growing Models from the Bottom Up. An Evaluation-Based Incremental Modelling Method (EBIMM) Applied to the Simulation of Systems of Cities', *Journal of Artificial Societies and Social Simulation*, **18**(4), October, article 9. <http://jasss.soc.surrey.ac.uk/18/4/9.html> doi:10.18564/jasss.2828. [Compares real and simulated rank size distributions of cities for two model variants. Figures 3 and 6.]

Cuenda, Sara, Fernández, Maximiliano, Galeano, Javier and Capitán, José A. (2018) 'A Minimal Agent-Based Model Reproduces the Overall Topology of Interbank Networks', *Journal of Artificial Societies and Social Simulation*, **21**(1), January, article 2. <http://jasss.soc.surrey.ac.uk/21/1/2.html> doi:10.18564/jasss.3562 [Compares plots of cumulative distribution against indegree, outdegree, in transactions and out transactions for simulations and real data from several countries. Figure 5.]

Czupryna, Marcin, Oleksy, Paweł, Przybek, Piotr and Kamiński, Bogumił (2018) 'Agent-Based Modelling of Viticulture Development in Emerging Markets: The Case of the Małopolska Region', *Journal of Artificial Societies and Social Simulation*, **21**(3), June, article 6. <http://jasss.soc.surrey.ac.uk/21/3/6.html> doi:10.18564/jasss.3726 [Compares simulation, regression and real data for number of vineyards over time. Figure 9.]

Demarest, Jeffrey, Pagsuyoin, Sheree, Learmonth, Gerard, Mellor, Jonathan and Dillingham, Rebecca (2013) 'Development of a Spatial and Temporal Agent-Based Model for Studying Water and Health Relationships: The Case Study of Two Villages in Limpopo, South Africa', *Journal of Artificial Societies and Social Simulation*, **16**(4), October, article 3.

<http://jasss.soc.surrey.ac.uk/16/4/3.html> doi:10.18564/jasss.2169 [Compares real and simulated child height increase with age. Figure 6.]

Dobbie, Samantha, Schreckenberg, Kate, Dyke, James G., Schaafsma, Marije and Balbi, Stefano (2018) 'Agent-Based Modelling to Assess Community Food Security and Sustainable Livelihoods', *Journal of Artificial Societies and Social Simulation*, **21**(1), January, article 9. <http://jasss.soc.surrey.ac.uk/21/1/9.html> doi:10.18564/jasss.3639 [Compares "diagonal plots" of observed versus predicted data for number of months of household food insecurity over the last 12 months and proportion of land households allocate to grain crops. Figure 6.]

Dolado, Ruth, Beltran, Francesc S. and Quera, Vicenç (2014) 'Modeling the Emergence of Social Structure from a Phylogenetic Point of View', *Journal of Artificial Societies and Social Simulation*, **17**(1), January, article 8. <http://jasss.soc.surrey.ac.uk/17/1/8.html> doi:10.18564/jasss.2356 [Borderline case for human sociality. Compares a wide variety of real and simulated data for four simulation conditions against observed – non-human – primate behaviour. Table 8.]

Duggins, Peter (2017) 'A Psychologically-Motivated Model of Opinion Change with Applications to American Politics', *Journal of Artificial Societies and Social Simulation*, **20**(1), January, article 13. <http://jasss.soc.surrey.ac.uk/20/1/13.html> doi:10.18564/jasss.3316 [Compares real and simulated opinions across issues and across time for the US. Figures 9 and 10.]

Fitzpatrick, Ben, Martinez, Jason, Polidan, Elizabeth and Angelis, Ekaterini (2015) 'The Big Impact of Small Groups on College Drinking', *Journal of Artificial Societies and Social Simulation*, **18**(3), June, article 4. <http://jasss.soc.surrey.ac.uk/18/3/4.html> doi:10.18564/jasss.2760 [Compares simulated group size data with secondary observational data in related domains.]

Gargiulo, Floriana, Lenormand, Maxime, Huet, Sylvie and Espinosa, Omar Baqueiro (2012) 'Commuting Network Models: Getting the Essentials', *Journal of Artificial Societies and Social Simulation*, **15**(2), March, article 6. <http://jasss.soc.surrey.ac.uk/15/2/6.html> doi:10.18564/jasss.1964 [Compares log-log plot of network indegree and probability of that indegree (and the same kind of plot separately for outdegree and traffic) for each of three values of a simulation parameter along with corresponding real data from a case study in the Auvergne. The same comparisons between the simulation and other case study sites (Bretagne, Altmark and Notts and Derbyshire) are also made. Figures 5 and 6.]

Geard, Nicholas, McCaw, James M., Dorin, Alan, Korb, Kevin B. and McVernon, Jodie (2013) 'Synthetic Population Dynamics: A Model of Household Demography', *Journal of Artificial Societies and Social Simulation*, **16**(1), January, article 8. <http://jasss.soc.surrey.ac.uk/16/1/8.html> doi:10.18564/jasss.2098 [Compares real and simulated age distributions, final household sizes, household type by age group and household type transitions. Also compares real and simulated age distributions and final household sizes for countries with different demographic situations. Figures 2, 4, 9, 12. Table 4.]

Giulioni, Gianfranco, Di Giuseppe, Edmondo, Toscano, Piero, Miglietta, Francesco and Pasqui, Massimiliano (2019) 'A Novel Computational Model of the Wheat Global Market with an Application to the 2010 Russian Federation Case', *Journal of Artificial Societies and Social Simulation*, 22(3), June, article 4. <http://jasss.soc.surrey.ac.uk/22/3/4.html> doi:10.18564/jasss.4063 [Compares real and simulated trade network structure for imports and exports, time series of wheat bought by a specific country, aggregate wheat prices over time. Figures 11, 14-16.]

Gooding, Tim (2014) 'Modelling Society's Evolutionary Forces', *Journal of Artificial Societies and Social Simulation*, **17**(3), June, article 3. <http://jasss.soc.surrey.ac.uk/17/3/3.html> doi:10.18564/jasss.2497 [Compares real and simulated Gini coefficients over time though the exact comparability of the time dimension is hard to establish from the relevant figures. Figures 5 and 6.]

* Hägerstrand, Torsten (1965) 'A Monte Carlo Approach to Diffusion', *European Journal of Sociology/Archives Européennes de Sociologie*, **6**(1), May, pp. 43-67. Special Issue: Simulation in Sociology. doi:10.1017/S0003975600001132 [Compares real number of adopters and three simulation runs, p. 61, real and simulated numbers of cells reaching a certain adoption percentage, p. 62, and spatial matches in adoption levels, pp. 64 and 65.]

Hayes, Roy and Hayes, Reginald (2014) 'Agent-Based Simulation of Mass Shootings: Determining How to Limit the Scale of a Tragedy', *Journal of Artificial Societies and Social Simulation*, **17**(2), March, article 5. <http://jasss.soc.surrey.ac.uk/17/2/5.html> doi:10.18564/jasss.2421 [Compares death toll from a real incident and a spatial simulation of the setting where that incident occurred.]

Heinke, Dietmar, Carslaw, Gregory and Christian, Julie (2013) 'An Agent-Based Simulation of Destigmatization (DSIM): Introducing a Contact Theory and Self-Fulfilling Prophecy Approach', *Journal of Artificial Societies and Social Simulation*, **16**(4), October, article 10. <http://jasss.soc.surrey.ac.uk/16/4/10.html> doi:10.18564/jasss.2244 [Shows that model output is compatible with survey data displaying a negative relationship between size of group and level of stigmatisation but does not do an explicit comparison. Figure 3. Table 1.]

Hoffmann, A. O. I., Jager, W. and Von Eije, J. H. (2007) 'Social Simulation of Stock Markets: Taking It to the Next Level', *Journal of Artificial Societies and Social Simulation*, **10**(2), March, article 7. <http://jasss.soc.surrey.ac.uk/10/2/7.html> [Compares graphs of prices, returns and autocorrelations of returns and bar charts of returns distribution for two simulated network architectures and equivalent Dutch stock market data 1987-2005. Also presents comparison of summary statistics from analysis. Figures 4-15 and Table 3.]

Houssou, Noudéhouénou Lionel Jaderne, Cordero, Juan Durango, Bouadjio-Boulic, Audren, Morin, Lucie, Maestripieri, Nicolas, Ferrant, Sylvain, Belem, Mahamadou, Sanchez, Jose Ignacio Pelaez, Saenz, Melio, Lerigoleur, Emilie, Elger, Arnaud, Gaudou, Benoit, Maurice, Laurence and Saqalli, Mehdi (2019) 'Synchronizing Histories of Exposure and Demography: The Construction of an Agent-Based Model of the Ecuadorian Amazon Colonization and Exposure to Oil Pollution Hazards', *Journal of Artificial Societies and Social Simulation*, **22**(2), March, article 1. <http://jasss.soc.surrey.ac.uk/22/2/1.html> doi:10.18564/jasss.3957

[Compares deforestation in simulation over time against satellite imaging of land cover. Table 4.]

Huynh, Nam, Barthelemy, Johan and Perez, Pascal (2016) 'A Heuristic Combinatorial Optimisation Approach to Synthesising a Population for Agent Based Modelling Purposes', *Journal of Artificial Societies and Social Simulation*, **19**(4), October, article 11. <http://jasss.soc.surrey.ac.uk/19/4/11.html> doi:10.18564/jasss.3198 [Compares counts of demographic categories i. e. number of children with corresponding census data, fraction of census districts matched by simulation for different demographic categories, real and simulated household sizes, real and simulated counts of males, females and family households by family household type, distribution of males and females by household relationship, distribution of households by household type, distribution of females and males by household type, spatial population density. Tables 4-6. Figures 2, 5 and 6.]

Ibrahim, Azhar Mohd, Venkat, Ibrahim and De Wilde, Philippe (2019) 'The Impact of Potential Crowd Behaviours on Emergency Evacuation: An Evolutionary Game-Theoretic Approach', *Journal of Artificial Societies and Social Simulation*, **22**(1), January, article 3. <http://jasss.soc.surrey.ac.uk/22/1/3.html> doi:10.18564/jasss.3837 [Compares escape times from an evacuation experiment with different exit widths against a variety of simulations. Table 3.]

Kaye-Blake, Bill, Schilling, Chris and Post, Elizabeth (2014) 'Validation of an Agricultural MAS for Southland, New Zealand', *Journal of Artificial Societies and Social Simulation*, **17**(4), October, article 5. <http://jasss.soc.surrey.ac.uk/17/4/5.html> doi:10.18564/jasss.2564 [Compares real and simulated conversion to dairy farming over time. Figure 7.]

Kim, Hyeyoung, Xiao, Ningchuan, Moritz, Mark, Garabed, Rebecca and Pomeroy, Laura W. (2016) 'Simulating the Transmission of Foot-And-Mouth Disease Among Mobile Herds in the Far North Region, Cameroon', *Journal of Artificial Societies and Social Simulation*, **19**(2), March, article 6. <http://jasss.soc.surrey.ac.uk/19/2/6.html> doi:10.18564/jasss.3064 [Compares the percentage of simulated disease outbreaks with more than one peak under variant assumptions with survey data. Table 2. Figure 10.]

Krebs, Friedrich (2017) 'An Empirically Grounded Model of Green Electricity Adoption in Germany: Calibration, Validation and Insights into Patterns of Diffusion', *Journal of Artificial Societies and Social Simulation*, **20**(2), March, article 10. <http://jasss.soc.surrey.ac.uk/20/2/10.html> doi:10.18564/jasss.3429 [Compares real and simulated numbers of customers in the green electricity market over time in Germany. Figure 4.]

Lamy, Francois, Quinn, Brendan, Dwyer, Robyn, Thomson, Nicola, Moore, David and Dietze, Paul (2016) 'TreatMethHarm: An Agent-Based Simulation of How People Who Use Methamphetamine Access Treatment', *Journal of Artificial Societies and Social Simulation*, **19**(2), March, article 3. <http://jasss.soc.surrey.ac.uk/19/2/3.html> doi:10.18564/jasss.3069 [Compares real and simulated population sizes at different stages of use/treatment. Table 4.]

Lee, Keeheon, Kim, Shintae, Kim, Chang Ouk and Park, Taeho (2013) 'An Agent-Based Competitive Product Diffusion Model for the Estimation and Sensitivity Analysis of Social Network Structure and Purchase Time Distribution', *Journal of Artificial Societies and Social Simulation*, **16**(1), January, article 3. <http://jasss.soc.surrey.ac.uk/16/1/3.html> doi:10.18564/jasss.2080 [Compares real and simulated product market shares over time. Figure 5.]

Lenormand, Maxime and Deffuant, Guillaume (2013) 'Generating a Synthetic Population of Individuals in Households: Sample-Free Vs Sample-Based Methods', *Journal of Artificial Societies and Social Simulation*, **16**(4), October, article 12. <http://jasss.soc.surrey.ac.uk/16/4/12.html> doi:10.18564/jasss.2319 [Compares real and simulated demographic distributions but it is not so easy to tell exactly what of from the figures themselves. Figures 1 and 2.]

Ma, Yan, Shen, Zhenjiang and Nguyen, Dinh Thanh (2016) 'Agent-Based Simulation to Inform Planning Strategies for Welfare Facilities for the Elderly: Day Care Center Development in a Japanese City', *Journal of Artificial Societies and Social Simulation*, **19**(4), October, article 5. <http://jasss.soc.surrey.ac.uk/19/4/5.html> doi:10.18564/jasss.3090 [Compares real and simulated numbers of people attending centres. Table 3.]

Ma, Yan, Shen, Zhenjiang and Kawakami, Mitsuhiko (2013) 'Agent-Based Simulation of Residential Promoting Policy Effects on Downtown Revitalization', *Journal of Artificial Societies and Social Simulation*, **16**(2), March, article 2. <http://jasss.soc.surrey.ac.uk/16/2/2.html> doi:10.18564/jasss.2125 [Compares ratios of household types by year. Table 3.]

Marchione, Elio, Johnson, Shane D. and Wilson, Alan (2014) 'Modelling Maritime Piracy: A Spatial Approach', *Journal of Artificial Societies and Social Simulation*, **17**(2), March, article 9. <http://jasss.soc.surrey.ac.uk/17/2/9.html> doi:10.18564/jasss.2477 [Compares real and simulated attack numbers over time and spatial "heat maps" of the distribution of these attacks. Figure 8.]

Mercuur, Rijk, Dignum, Virginia and Jonker, Catholijn M. (2019) 'The Value of Values and Norms in Social Simulation', *Journal of Artificial Societies and Social Simulation*, **22**(1), January, article 9. <http://jasss.soc.surrey.ac.uk/22/1/9.html> doi:10.18564/jasss.3929 [Compares secondary experimental game theory data with a variety of simulation types. Figures 1 and 4-7. Tables 4-6.]

Muis, Jasper (2010) 'Simulating Political Stability and Change in the Netherlands (1998-2002): An Agent-Based Model of Party Competition with Media Effects Empirically Tested', *Journal of Artificial Societies and Social Simulation*, **13**(2), March, article 4. <http://jasss.soc.surrey.ac.uk/13/2/4.html> doi:10.18564/jasss.1482 [Compares votes fraction for various Dutch parties with results from a set of variant simulations at several political "junctures". Tables 3, 6 and 9.]

Owadally, Iqbal, Zhou, Feng and Wright, Douglas (2018) 'The Insurance Industry as a Complex Social System: Competition, Cycles and Crises', *Journal of Artificial Societies and*

Social Simulation, **21**(4), October, article 2. <http://jasss.soc.surrey.ac.uk/21/4/2.html> doi:10.18564/jasss.3819 [Compares actual and simulated loss ratios from the UK property insurance industry. Figure 6.]

Perello-Moragues, Antoni, Noriega, Pablo and Poch, Manel (2019) 'Modelling Contingent Technology Adoption in Farming Irrigation Communities', *Journal of Artificial Societies and Social Simulation*, **22**(4), October, article 1. <http://jasss.soc.surrey.ac.uk/22/4/1.html> doi:10.18564/jasss.4100 [Compares simulated and real adoption curves for farm modernisation over time. Figure 10.]

Pizzitutti, Francesco, Mena, Carlos F. and Walsh, Stephen J. (2014) 'Modelling Tourism in the Galapagos Islands: An Agent-Based Model Approach', *Journal of Artificial Societies and Social Simulation*, **17**(1), January, article 14. <http://jasss.soc.surrey.ac.uk/17/1/14.html> doi:10.18564/jasss.2389 [Compares real and simulated instances of tourist activities i. e. visiting a particular town or taking a boat trip. Table 1.]

Rakowski, Franciszek, Gruziel, Magdalena, Krych, Michal and Radomski, Jan P. (2010) 'Large Scale Daily Contacts and Mobility Model: An Individual-Based Countrywide Simulation Study for Poland', *Journal of Artificial Societies and Social Simulation*, **13**(1), January, article 13. <http://jasss.soc.surrey.ac.uk/13/1/13.html> doi:10.18564/jasss.1529 [Compares real and simulated population density for an area of Poland and bar charts of real and simulated number of children in family and age gap between spouses. Figures 4, 5a and 5b. Figure 7 also appears to include real and simulated data on travel to college but it doesn't appear to be comparable and it isn't clear how it should be interpreted.]

Rand, William, Herrmann, Jeffrey, Schein, Brandon and Vodopivec, Neža (2015) 'An Agent-Based Model of Urgent Diffusion in Social Media', *Journal of Artificial Societies and Social Simulation*, **18**(2), March, article 1. <http://jasss.soc.surrey.ac.uk/18/2/1.html> doi:10.18564/jasss.2616 [Compares models to Twitter data about message diffusion. Figures 5-12.]

Rangoni, Ruggero and Jager, Wander (2017) 'Social Dynamics of Littering and Adaptive Cleaning Strategies Explored Using Agent-Based Modelling', *Journal of Artificial Societies and Social Simulation*, **20**(2), March, article 1. <http://jasss.soc.surrey.ac.uk/20/2/1.html> doi:10.18564/jasss.3269 [Compares field experiment data with simulation results. Figure 3.]

Sajjad, Mazhar, Singh, Karandeep, Paik, Euihyun and Ahn, Chang-Won (2016) 'A Data-Driven Approach for Agent-Based Modeling: Simulating the Dynamics of Family Formation', *Journal of Artificial Societies and Social Simulation*, **19**(1), January, article 9. <http://jasss.soc.surrey.ac.uk/19/1/9.html> doi:10.18564/jasss.2988 [Compares real and simulated data for men and women: age at marriage. Marriages for different 5 year periods as probabilities. The same comparisons again with the model corrected for education levels, though this time with five yearly probabilities separated by gender. Figures 2-6.]

Salgado, Mauricio, Marchione, Elio and Gilbert, Nigel (2014) 'Analysing Differential School Effectiveness Through Multilevel and Agent-Based Modelling', *Journal of Artificial Societies and Social Simulation*, **17**(4), October, article 3. <http://jasss.soc.surrey.ac.uk/17/4/3.html>

doi:10.18564/jasss.2534 [Compares parameter values for a Multi-Level Model and ABM on educational attainment. Table 3.]

Scalco, Andrea, Macdiarmid, Jennie I., Craig, Tony, Whybrow, Stephen and Horgan, Graham W. (2019) 'An Agent-Based Model to Simulate Meat Consumption Behaviour of Consumers in Britain', *Journal of Artificial Societies and Social Simulation*, **22**(4), October, article 8. <http://jasss.soc.surrey.ac.uk/22/4/8.html> 10.18564/jasss.4134 [Compares real and simulated UK meat prices over time. Figure 2.]

Scott, Nick, Livingston, Michael, Hart, Aaron, Wilson, James, Moore, David and Dietze, Paul (2016) 'SimDrink: An Agent-Based NetLogo Model of Young, Heavy Drinkers for Conducting Alcohol Policy Experiments', *Journal of Artificial Societies and Social Simulation*, **19**(1), January, article 10. <http://jasss.soc.surrey.ac.uk/19/1/10.html> doi:10.18564/jasss.2943 [Compares real data on distribution of number of drinks, amount spent and time on a drinking "event" with simulated equivalents. Comparison of real and simulated verbal violence incidence. Figures 1 and 2.]

Silverman, Eric, Bijak, Jakub, Hilton, Jason, Cao, Viet Dung and Noble, Jason (2013) 'When Demography Met Social Simulation: A Tale of Two Modelling Approaches', *Journal of Artificial Societies and Social Simulation*, **16**(4), October, article 9. <http://jasss.soc.surrey.ac.uk/16/4/9.html> doi:10.18564/jasss.2327 [Compares real and simulated population pyramids for 2011 in the UK and real and simulated hazard rates for partnership formation. Figures 2 and 3.]

Sobkowicz, Pawel (2017) 'Utility, Impact, Fashion and Lobbying: An Agent-Based Model of the Funding and Epistemic Landscape of Research', *Journal of Artificial Societies and Social Simulation*, **20**(2), March, article 5. <http://jasss.soc.surrey.ac.uk/20/2/5.html> doi:10.18564/jasss.3399 [Compares real and simulated plots of log binned distribution of research topics against number of projects devoted to a research topic. Figures 12 and 13.]

Spaiser, Viktoria and Sumpter, David J. T. (2016) 'Revising the Human Development Sequence Theory Using an Agent-Based Approach and Data', *Journal of Artificial Societies and Social Simulation*, **19**(3), June, article 1. <http://jasss.soc.surrey.ac.uk/19/3/1.html> doi:10.18564/jasss.3084 [Compares real and simulated "democracy levels", real and simulated data on changes in resources by changes in democracy levels, real and simulated changes in democracy levels by Gini coefficient. Figures 4, 5 and 12.]

Sullivan, Abigail, An, Li and York, Abigail (2018) 'Which Perspective of Institutional Change Best Fits Empirical Data? An Agent-Based Model Comparison of Rational Choice and Cultural Diffusion in Invasive Plant Management', *Journal of Artificial Societies and Social Simulation*, **21**(1), January, article 5. <http://jasss.soc.surrey.ac.uk/21/1/5.html> doi:10.18564/jasss.3611 [Compares survey data on plant management practices with simulations of dominant strategies under different theories of institutional change. Table 2. Figure 3.]

Tucker, Peter and Smith, Duncan (1999) 'Simulating Household Waste Management Behaviour', *Journal of Artificial Societies and Social Simulation*, **2**(3), October, article 3.

<http://jasss.soc.surrey.ac.uk/2/3/3.html> [Compares real and simulated quantities of recyclables put out for special collection weekly under two "promotional interventions". Also compares this for different neighbourhoods to evaluate social influence effects. Compares monthly participation rates. Compares effects of collection frequency change. Compares putting out rates for real and simulated streets (though these are apparently not directly matched). Also evaluates variability in simulation runs against actual data for putting out rates. Figure 7-10. Tables 3 and 4.]

Urquhart, Neil, Powers, Simon, Wall, Zoe, Fonzone, Achille, Ge, Jiaqi and Polhill, J. Gary (2019) 'Simulating the Actions of Commuters Using a Multi-Agent System', *Journal of Artificial Societies and Social Simulation*, **22**(2), March, article 10. <http://jasss.soc.surrey.ac.uk/22/2/10.html> doi:10.18564/jasss.4007 [Compares real transport mode distributions in journeys at three locations and corresponding simulated distributions. Table 3.]

Walker, Lyndon and Davis, Peter (2013) 'Modelling "Marriage Markets": A Population-Scale Implementation and Parameter Test', *Journal of Artificial Societies and Social Simulation*, **16**(1), January, article 6. <http://jasss.soc.surrey.ac.uk/16/1/6.html> doi:10.18564/jasss.2106 [Compares actual and simulated numbers of homogamous and mixed marriages in various New Zealand cities in different years. Tables 6-8.]

Wei, Xiaochao, Hu, Bin and Carley, Kathleen M. (2013) 'Combination of Empirical Study with Qualitative Simulation for Optimization Problem in Mobile Banking Adoption', *Journal of Artificial Societies and Social Simulation*, **16**(3), June, article 10. <http://jasss.soc.surrey.ac.uk/16/3/10.html> doi:10.18564/jasss.2222 [Compares real and simulated investments by different adoption drivers. Figure 10.]

White, Andrew (2017) 'A Model-Based Analysis of the Minimum Size of Demographically-Viable Hunter-Gatherer Populations', *Journal of Artificial Societies and Social Simulation*, **20**(4), October, article 9. <http://jasss.soc.surrey.ac.uk/20/4/9.html> doi:10.18564/jasss.3393 [Compares a variety of simulated demographic variables – such as total fertility rates, infant mortality rate and age at first marriage – with ranges from ethnographic data. Figure 3.]

White, Douglas R. (1999) 'Controlled Simulation of Marriage Systems', *Journal of Artificial Societies and Social Simulation*, **2**(3), October, article 5. <http://jasss.soc.surrey.ac.uk/2/3/5.html> [A borderline case. Compares real ethnographic marriage data with a simulation of a random matching assumption i. e. effectively a kind of null hypothesis and one suspects not a claim about any real social system *per se*. Compares structural endogamy, marriage among consanguineal kin, relinking frequencies and the existence of "sidedness rules" for several ethnographic groups. Tables 5-14.]

Wright, Mason and Sengupta, Pratim (2015) 'Modeling Oligarchs' Campaign Donations and Ideological Preferences with Simulated Agent-Based Spatial Elections', *Journal of Artificial Societies and Social Simulation*, **18**(2), March, article 3. <http://jasss.soc.surrey.ac.uk/18/2/3.html> doi:10.18564/jasss.2736 [Compares opinion poll data and model outputs. Table 4.]

Xiong, Hang, Payne, Diane and Kinsella, Stephen (2018) 'Identifying Mechanisms Underlying Peer Effects on Multiplex Networks', *Journal of Artificial Societies and Social Simulation*, **21**(4), October, article 6. <http://jasss.soc.surrey.ac.uk/21/4/6.html> doi:10.18564/jasss.3797 [Compares real and simulated data on farming innovation adoption over time for two models of adaptation and 10 villages. Table 3.]

Zimmermann, Albert, Möhring, Anke, Mack, Gabriele, Ferjani, Ali and Mann, Stefan (2015) 'Pathways to Truth: Comparing Different Upscaling Options for an Agent-Based Sector Model', *Journal of Artificial Societies and Social Simulation*, **18**(4), October, article 11. <http://jasss.soc.surrey.ac.uk/18/4/11.html> doi:10.18564/jasss.2862 [Compares retrospective predictions from different models with now known data to investigate the effects of different upscaling strategies for models. Tables 4 and 5.]

Appendix 2: Survey/Review Sources Classifying Samples of ABM Research by Validation Status

Angus, Simon D. and Hassani-Mahmooei, Behrooz (2015) '"Anarchy" Reigns: A Quantitative Analysis of Agent-Based Modelling Publication Practices in JASSS, 2001-2012', *Journal of Artificial Societies and Social Simulation*, **18**(4), October, article 16. <http://jasss.soc.surrey.ac.uk/18/4/16.html> doi:10.18564/jasss.2952

Dutton, John M. and Starbuck, William H. (1971) 'Computer Simulation Models of Human Behavior: A History of an Intellectual Technology', *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-1**(2), April, pp. 128-171. doi:10.1109/TSMC.1971.4308269

Heath, Brian, Hill, Raymond and Ciarallo, Frank (2009) 'A Survey of Agent-Based Modeling Practices (January 1998 to July 2008)', *Journal of Artificial Societies and Social Simulation*, **12**(4), October, article 9. <http://jasss.soc.surrey.ac.uk/12/4/9.html>

Appendix 3: Highly Provisional Bibliography of Sources Advocating Good Practice or Techniques Relevant to Validation

This bibliography is only provided because it could be constructed almost "for free" from the existing analysis of articles for the bibliography. Additions (or completions of existing references) are also welcomed.

Amblard, Frédéric, Bommel, Pierre and Rouchier, Juliette (2007) 'Assessment and Validation of Multi-Agent Models', in Phan, Denis and Amblard, Frédéric (eds.) *Multi-Agent Modelling and Simulation in the Social and Human Sciences*, GEMAS Studies in Social Analysis (Oxford: Bardwell Press), pp. 93-114.

Axtell, Robert L. and Epstein, Joshua M. (1994) 'Agent-Based Modeling: Understanding Our Creations', *Bulletin of the Santa Fe Institute*, **9**(2), Winter, pp. 28–32. [Distinguishes four levels of validation.]

Brown, Daniel G., Page, Scott, Riolo, Rick, Zellner, Moira and Rand, William (2005) 'Path Dependence and the Validation of Agent-Based Spatial Models of Land Use', *International Journal of Geographical Information Science*, **19**(2), xx, pp. 153-174. doi:10.1080/13658810410001713399

David, Nuno (2013) 'Validating Simulations', in Edmonds, Bruce and Meyer, Ruth (eds.) *Simulating Social Complexity*, Understanding Complex Systems Series (Berlin: Springer), pp. 135-171. doi:10.1007/978-3-540-93813-2_8 [There was a new edition of this book in 2017 which may contain an updated version of this chapter.]

Dridi, Mohamed H. (2014) 'Pedestrian Flow Simulation Validation and Verification Techniques', arXiv Preprint, arXiv:1410.0603v1. <https://arxiv.org/abs/1410.0603>

Fortino, G. and North, M. J. (2013) 'Simulation-based Development and Validation of Multi-agent Systems: AOSE and ABMS Approaches', *Journal of Simulation*, **7**(xx), xx, pp. 137-143. doi:10.1057/jos.2013.12

Garcia, Rosanna, Rummel, Paul and Hauser, John (2007) 'Validating Agent-Based Marketing Models Through Conjoint Analysis', *Journal of Business Research*, **60**(8), August, pp. 848-857. doi:10.1016/j.jbusres.2007.02.007

Klüpfel, H. (2007) 'The Simulation of Crowd Dynamics at Very Large Events: Calibration, Empirical Data, and Validation', in Waldau, Nathalie, Gattermann, Peter, Knoflacher, Hermann and Schreckenberg, Michael (eds.) *Pedestrian and Evacuation Dynamics 2005* (Berlin: Springer), pp. 285-296. doi:10.1007/978-3-540-47064-9_25

Moss, Scott (2008) 'Alternative Approaches to the Empirical Validation of Agent-Based Models', *Journal of Artificial Societies and Social Simulation*, **11**(1), January, article 5. <http://jasss.soc.surrey.ac.uk/11/1/5.html>

Moss, Scott and Edmonds, Bruce (2005) 'Sociology and Simulation: Statistical and Qualitative Cross-Validation', *American Journal of Sociology*, **110**(4), January, pp. 1095-1131. doi:10.1086/427320

Ngo, The An and See, Linda (2012) 'Calibration and Validation of Agent-Based Models of Land Cover Change', in Heppenstall Alison J., Crooks, Andrew T., See, Linda M. and Batty Michael (eds.) *Agent-Based Models of Geographical Systems* (Berlin: Springer), pp. 181-197. doi:10.1007/978-90-481-8927-4_10

Pontius, Robert Gilmore Junior, Boersma, Wideke, Castella, Jean-Christophe, Clarke, Keith, de Nijs, Ton, Dietzel, Charles, Duan, Zengqiang, Fotsing, Eric, Goldstein, Noah, Kok, Kasper, Koomen, Eric, Lippitt, Christopher D., McConnell, William, Mohd Sood, Alias, Pijanowski, Bryan, Pithadia, Snehal, Sweeney, Sean, Trung, Tran Ngoc, Veldkamp, A. Tom and Verburg, Peter H. (2008) 'Comparing the Input, Output, and Validation Maps for Several Models of Land Change', *The Annals of Regional Science*, **42**(1), xx, pp. 11-37. doi:10.1007/s00168-007-0138-2

Rouchier, Juliette (2007) 'Data Gathering to Build and Validate Small-Scale Social Models for Simulation', in Rennard, Jean-Philippe (ed.) *Handbook of Research on Nature-Inspired Computing for Economics and Management* (Hershey, PA: Idea Group Reference), pp. 198-210.

Sargent, R. G. (2005) 'Verification and Validation of Simulation Models', *Proceedings of the 37th Conference on Winter Simulation*, pp. 130-143. doi:10.1109/wsc.2005.1574246

Sargent, R. G. (2013) 'Verification and Validation of Simulation Models', *Journal of Simulation*, **7**(xx), xx, pp. 12-24. doi:10.1057/jos.2012.20 [Updated version of the previous item.]

Troitzsch, Klaus G. (2004) 'Validating Simulation Models', in *Proceedings of 18th European Simulation Multiconference, Magdeburg*, pp. 98-106. <http://scs-europe.net/services/esm2004/pdf/esm-43.pdf>

Windrum, Paul, Fagiolo, Giorgio and Moneta, Alessio (2007) 'Empirical Validation of Agent-Based Models: Alternatives and Prospects', *Journal of Artificial Societies and Social Simulation*, **10**(2), March, article 8. <http://jasss.soc.surrey.ac.uk/10/2/8.html>

Xiang, Xiaorong, Kennedy, Ryan, Madey, Gregory and Cabaniss, Steve (2005) 'Verification and Validation of Agent-Based Scientific Simulation Models', in *Agent-Directed Simulation Conference*, pp. 47-55.

Appendix 4: Cases That May Challenge a Unitary Definition of Validation

Adelt, Fabian, Weyer, Johannes, Hoffmann, Sebastian and Ihrig, Andreas (2018) 'Simulation of the Governance of Complex Systems (SimCo): Basic Concepts and Experiments on Urban Transportation', *Journal of Artificial Societies and Social Simulation*, **21**(2), March, article 2. <http://jasss.soc.surrey.ac.uk/21/2/2.html> doi:10.18564/jasss.3654 [Validation is asserted but not actually visible for checking in the source itself.]

Bloomquist, Kim M. and Koehler, Matt (2015) 'A Large-Scale Agent-Based Model of Taxpayer Reporting Compliance', *Journal of Artificial Societies and Social Simulation*, **18**(2), March, article 20. <http://jasss.soc.surrey.ac.uk/18/2/20.html> doi:10.18564/jasss.2621 [Refers the reader to validation done elsewhere – which happens to be a broken link!]

D' Orazio, Paola and Giulioni, Gianfranco (2017) 'From Micro Behaviors to Macro Dynamics: An Agent-Based Economic Model with Consumer Credit', *Journal of Artificial Societies and Social Simulation*, **20**(1), January, article 9. <http://jasss.soc.surrey.ac.uk/20/1/9.html> doi:10.18564/jasss.3260 [Does show real and simulated income distributions by country but they do not appear to be directly comparable or to be explicitly compared.]

Drchal, Jan, Čertický, Michal and Jakob, Michal (2016) 'VALFRAM: Validation Framework for Activity-Based Models', *Journal of Artificial Societies and Social Simulation*, **19**(3), June, article 5. <http://jasss.soc.surrey.ac.uk/19/3/5.html> doi:10.18564/jasss.3127 [Although the

article talks about use of "raw" data for validation, all the analysis appears to involve comparison between models in practice.]

Fonseca, Fernando, Ramos, Rui António Rodrigues, and da Silva, Antônio Nélson Rodrigues (2015) 'An Agent-Based Model to Assess the Attractiveness of Industrial Estates', *Journal of Artificial Societies and Social Simulation*, **18**(4), October, article 13. <http://jasss.soc.surrey.ac.uk/18/4/13.html> doi:10.18564/jasss.2893 [Validated against potentially arbitrary hypotheses.]

Ge, Jiaqi and Polhill, Gary (2016) 'Exploring the Combined Effect of Factors Influencing Commuting Patterns and CO2 Emissions in Aberdeen Using an Agent-Based Model', *Journal of Artificial Societies and Social Simulation*, **19**(3), June, article 11. <http://jasss.soc.surrey.ac.uk/19/3/11.html> doi:10.18564/jasss.3078 [Turns out to validate against another model.]

Gore, Ross J., Lemos, Carlos, Shults, F. LeRon and Wildman, Wesley (2018) 'Forecasting Changes in Religiosity and Existential Security with an Agent-based Model', *Journal of Artificial Societies and Social Simulation*, **21**(1), January, article 4. <http://jasss.soc.surrey.ac.uk/21/1/4.html> doi:10.18564/jasss.3596 ["Validates" model predictions against "no change" and regression models. Not clear how this should be interpreted or in what sense it ought to be considered validation.]

Lee, SeHoon, Hong, Jeong Hee, Bae, Jang Won and Moon, Il-Chul (2015) 'Impact of Population Relocation to City Commerce: Micro-Level Estimation with Validated Agent-Based Model', *Journal of Artificial Societies and Social Simulation*, **18**(2), March, article 5. <http://jasss.soc.surrey.ac.uk/18/2/5.html> doi:10.18564/jasss.2719 [Does not make a direct comparison but compares real data with a potential model proxy of unknown quality – customer "footfall" against market rent.]

Putra, Handi Chandra, Zhang, Haiyan and Andrews, Clinton (2015) 'Modeling Real Estate Market Responses to Climate Change in the Coastal Zone', *Journal of Artificial Societies and Social Simulation*, **18**(2), March, article 18. <http://jasss.soc.surrey.ac.uk/18/2/18.html> doi:10.18564/jasss.2577 [Analysis of model results is talked of as "validation" but there is no clear sign of any real data.]

Vanhée, Loïs and Dignum, Frank (2018) 'Explaining the Emerging Influence of Culture, from Individual Influences to Collective Phenomena', *Journal of Artificial Societies and Social Simulation*, **21**(4), October, article 11. <http://jasss.soc.surrey.ac.uk/21/4/11.html> doi:10.18564/jasss.3881 [Validates against what could be an arbitrary set of hypotheses.]