
International Journal of Social Research Methodology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsrm20

# The practice and rhetoric of prediction – the case in agent-based modelling

Bruce Edmonds

**To cite this article:** Bruce Edmonds (2023) The practice and rhetoric of prediction – the case in agent-based modelling, International Journal of Social Research Methodology, 26:2, 157-170, DOI: 10.1080/13645579.2022.2137921

**To link to this article:** https://doi.org/10.1080/13645579.2022.2137921

INTERNATIONAL JOURNAL OF SOCIAL RESEARCH METHODOLOGY
2023, VOL. 26, NO. 2, 157–170
https://doi.org/10.1080/13645579.2022.2137921



# The practice and rhetoric of prediction – the case in agent-based modelling

Bruce Edmonds

Faculty of Business and Law, Centre for Policy Modelling, Manchester Metropolitan University, Manchester, UK

**ABSTRACT**
This paper looks at the tension between the desire to claim predictive ability for Agent-Based Models (ABMs) and its extreme difficulty for social and ecological systems, suggesting that this is the main cause for the continuance of a rhetoric of prediction that is at odds with what is achievable. Following others, it recommends that it is better to avoid giving the impression of predictive ability until this has been iteratively and independently verified, due to the danger of suggesting more than is empirically warranted, especially in non-modellers. It notes that there is a restricted and technical context where prediction is useful, that of meta-modelling – when we are trying to explain and understand our own simulation models. If one is going to claim prediction, then a lot more care needs to be taken, implying minimal standards in practice and transparent honesty about the empirical track record – the over-enthusiastic claiming of prediction in casual ways needs to cease.

**KEYWORDS**
prediction; rhetoric; meta-modelling; agent-based model; practice

## Introduction

Anyone who has sought to buy a residential property will have come across the following phenomenon: that the boundaries of desirable areas tend to expand whilst those not in demand, contract. This is presumably because people want to be in the high status areas and so real estate agents adjust the definitions of areas to give people what they want – a property they can think of as having status. This desire for a high status attribution also can be seen in the attributions that researchers seek for their achievements. This paper looks at the claim that a model is able to predict aspects of complex ecological or social phenomena using agent-based models. I strongly suspect that the points made here also apply to other kinds of modelling. However, I am not an expert in those and so restrict my argument to the agent-based modelling of complex phenomena only. The conclusion will be that, like desirable residential areas, the usage of 'prediction' has weakened – allowing a wider variety of achievements to be included under this label – thus undermined some useful distinctions in science.

The paper starts by looking at some of the different uses of the word 'predict' in the agent-based modelling (ABM) literature to illustrate the range of claims being made under this label. It then contrasts such exercises with some of the other legitimate and productive uses of ABMs. It then deals with a couple of issues that can complicate the argument, namely: the various moderating factors that impact upon prediction and the difficulties involved in proving prediction attempts. The desirability of being able to predict is then contrasted with the difficulty in proving such an ability and the sheer difficulty in achieving this. It is argued that the tension between its desirability

**CONTACT** Bruce Edmonds bruce@edmonds.name Centre for Policy Modelling, Manchester Metropolitan University, Manchester, England

and its difficulty has combined with other factors (such as tradition) to result in a rhetoric of prediction that is in contrast to what the wider public (and policy actors) might expect of such claims. I then look at a limited and technical area where prediction is desirable – in the understanding of our own models. The practical aspects of prediction are briefly discussed, arguing for higher standards if prediction is claimed. The paper then concludes arguing for a shift away from a misleading rhetoric of prediction.

## Some uses of 'predict' in the ABM literature

This section looks at some of the ways in which the word 'prediction' is used in reports of ABM-based research. It does not intend to be a comprehensive review, because the point is to illustrate some of the variety of uses in order to motive the discussion that follows. If there are other kinds I have missed, then this merely reinforces our main point – that there is confusion around use of the word.[1]

Some researchers distinguish between 'forecasting' and 'prediction' – the former indicating an attempt to correctly anticipate characteristics of future events, and the latter anything that a model might indicate in terms of its results, regardless of whether the model can be relied upon to do this in a way that corresponds to what is modelled. For example, Bonnasse-Gahot et al. () use the word 'predict' in the latter sense of a model output. In this sense, the model predicts in a similar sense to the prognostications of a fortune teller – it can indeed be interpreted as a prediction but not one upon which one can rely.

The former sense of forecasting is closer to a common-sense understanding of 'prediction'. If a modeller claims to be able to use their model to predict where crimes tend to occur, then a wider public audience will tend to think that this will be useful to the organisation of the police (for example, in planning the routes of patrols). They would probably be disappointed if it turns out that these were only the areas an unvalidated model indicated, and could not be relied upon to usefully inform planning (Edmonds, 2017).

However, the word 'forecasting' implies a temporal ordering that may not be present. The existence of Pluto was inferred from an examination of the orbit of Neptune, before any observations of Pluto had occurred. Calculations based on data about Neptune were then used to predict where Pluto should be and observations made to check this. Clearly, there is a sense in which the existence of Pluto was predicted on the basis of observations of Neptune and gravitational theory, and this is an example of strong prediction, because what is predicted was not known at the time of the prediction. Having what is being predicted to be in the future is an easy way of preventing 'cheating' in prediction, because it is impossible to adjust a model to fit future data and thus the field of forecasting focuses on this kind of prediction, but useful prediction is not limited to being forward in time.

Another sense of 'predict' seems to be merely as an indication that a model can be related to empirical evidence. In other words, that it is not just a purely theoretical entity. Thus, a model is sometimes said to 'predict' some data that is already known to the modeller, e.g. (Alsaedi et al., 2017) where they correctly recognise past known events using Twitter data. Clearly, this is a weaker claim than predicting something unknown since, in this case, the modeller has ample chance to adjust the model to fit the target data. This is essentially explanatory modelling but, for reasons of tradition, the model fitting the data well is called 'prediction'.

A similar case is where a model suggests some pattern that is perceivable in the results, and then the modellers have then searched for similar observed phenomena where this pattern has occurred. An example where this seems to be the case is Lambert & Rochard (2007). In some sense, the model can be said to have predicted this pattern, a pattern that was unknown at the time of model building and running. This is a useful thing to do; however, but this is more in the line of discovering that the model has a wider explanatory scope than expected. There is no possibility of the model being decisively disconfirmed if these were not found, since one simply might not have searched in the right phenomena and one could simply continue searching until such a pattern was discovered somewhere. Thus, this is, again, a weaker sense of 'predict' than a useful anticipation of a targeted unknown pattern.

In other papers, prediction is mentioned, but as an aspiration – they believe that their model will enable prediction, but have not yet proved this. For example, Zhang et al., (2016) report the design of a model (and its validation against data) with a view to (future) epidemic prediction. They do not claim they have proven such an ability at this stage, as will be clear to a modeller reading the paper, but a wider public might well be confused by the language. It does seem that some fields, such as epidemiology do have a tradition of saying a model 'predicts the data' when a less confusing description might be that it merely matches or fits the data. Perhaps, fields that come from a more positivist tradition where the use of deterministic mathematical modelling techniques are common, tend to use the word 'predict' like this.

Language is flexible, often imparting its meaning in analogical ways (Lackoff ????). However, the wider public might expect scientists to be more precise, and to go to some lengths to avoid giving false impressions of what their models can achieve. In particular, whilst they may be tolerant of such language between expert modellers (who know how to interpret them), they may rightly be annoyed when this impacts negatively upon them – either as a result of wrongly choosing a policy that affects their lives on the basis it was backed up by such 'predictions' or giving public money to an application that claims to be able to deliver reliable predictions when this is highly unlikely.

## Other reasons for modelling than prediction

As Epstein (2008) points out, there are many other reasons to develop and use ABMs than prediction – he lists 15 other purposes. ABM is such a flexible tool that it can be used for very different ends, so we should not expect any uniformity or general rules that cover all of them. (Edmonds et al. 2019) discusses seven scientific purposes, including prediction and explanation – all legitimate uses for a model but all with their own dangers, limitations and standards for adequacy. In some cases, another modelling goal or achievement might be more appropriate and more honest than claiming predictive ability. For this reason, we now discuss a few other modelling purposes to stress that these are valid and valuable scientific purposes and encourage modellers to aim their modelling research in these directions instead. The reason for doing this is to encourage modellers to frame their claims in terms of these other uses – empirical science progresses by other many other activities than prediction. However, a lot of the philosophical discussion has been centred around prediction and so this needs dealing with here.

There has been discussion on the epistemological nature of prediction vs. explanation in ABM (Thompson & Derr, 2009; Troitzsch, 2009), distinguishing different kinds of prediction. Elsenbroich (2012) summarises the arguments concerning prediction well, arguing for the mechanism-based approach for ABM. Here, I am interested in the practical differences between these caused by the fact that the exercises of predicting and establishing an explanation have different goals. When establishing an explanation, one sees the target output data before the model is run and in prediction one checks this afterwards. The management of uncertainty is also different – in prediction on needs to minimise the uncertainty in the model, whilst when establishing a possible complex explanation, one merely has to find credible parameter values and inputs so the outcomes match the target data (Breiman, 2001). This also means that the need for good data is less for establishing a possible explanation than making a good prediction.

However, more fundamentally, the kind of model that produces a useful explanation is different from those that predict well, because to support an explanation in terms of certain processes or structures, these have to be explicitly represented in the model. Models of ideal gases neatly illustrate this point; if one uses the gas laws, an explanation of an increase in gas pressure is simply that the temperature has increased or the volume decreased (or both) – this is an explanation of an empty and not very useful kind, because it does not tell one why this happens. The particle model of ideal gases gives a physical mechanism for what this happens in terms of the momentum of 'billiard ball' particles that bounce against each other and the gas container, but do not interact otherwise. Adding in the necessary detail to support a better explanation might mean the introduction of more noise into the model, making it less predictively accurate.

In a mature science, explanatory and predictive models (what Cartwright (1983) calls 'explanatory' and 'phenomenological' laws) are related to each other, e.g. by formally relating them to each other using reasonable assumptions and approximations. However, in developing sciences it can happen that predictive models come long before good explanations (as in the Gas laws) or the other way around (as in the theory of Natural selection). Furthermore, even when we can formally 'bridge' between explanatory and predictive models, this does not mean that we merge them all into a super model – rather we often maintain a family of related models, each useful for a different purpose (Giere, 1988).

ABM is a versatile tool that can be used for a variety of purposes. All of them are necessary parts of the process of science as it develops in all its messy and practical ways. Thus, unlike others, I do not see the utility of putting all models on a single dimension or hierarchy of quality, with predictive models (or models that integrate explanatory and predictive aspects as in (Hofman et al., 2021)) at the top. From considering the different purposes of models, it follows that they will each have different criteria for success, hence different threats to that success, so different activities to mitigate those threats, resulting in different kinds of model (Edmonds, le Page et al., 2019). Maybe it is simply that we tend to think of all models as pictures of an observed reality, rather than different scientific tools, that leads us in this direction (and hence judge them using the same criteria).

The conclusion of this section is that prediction is just one of the activities that makes up a science, I suspect that many of the claims for prediction made of ABMs could be more informatively recast as one of these other activities. Such more accurate labelling would ultimately make ABM-based research more useful and productive, even if maybe sounding less impressive in the short term. The philosophy seems to over-emphasise the importance of prediction, not due to it being essential to progress in science but because it is essential for supporting certain philosophical conceptions of science (i.e. realism).

## Kinds of prediction

There are many factors that affect the kind of prediction that is made using a model, in other words: the extent, the usefulness and the provenance of it. These complicate the argument but dealing with these allows us to rule out certain kinds of trivial, non-useful kinds of prediction as well as bringing to light some practicalities that are needed if prediction is to be actually useful.

Even if prediction is claimed in the strong sense of anticipating some aspect of what is modelled – an aspect that is unknown to the modeller at the time using the model – there is a range of precisions to which this can be claimed. As Watts (2014) points out, predictions can be of a variety of forms, and is not restricted to the prediction of a numerical value. For example, one might predict that the results will fall within a range of values, that the outcomes will conform to some named distribution or specified trend (Thorngate & Edmonds, 2013), that the outcomes will tend to increase over a specified period of time, that a particular outcome is possible (Edmonds & Adoha, 2019) or even that some particular described occurrence will not occur. These are all valid predictions, varying in terms of the accuracy and kind of what is predicted.

On the other hand, prediction that 'anything might happen', does not narrow down the possibilities, and so is an empty prediction – it is bound not to be falsified and it does not tell us anything useful. It is implicit in most work that if a prediction is mentioned then it is a useful one otherwise it would not be worth mentioning. If I predict that next week's lottery number will be a finite sequence of integers, I am not doing prediction in any useful manner. Generally, the more a prediction constrains the possibilities the more useful it is, but in cases where some outcomes are catastrophic just knowing a particular outcome is possible might still be useful. If one does not have any idea of an expected margin of error or what is ruled out by a prediction, then it is not useable in practice.

Even if a prediction is accurate, in the sense that it excludes most of the possibilities, the prediction may not be at all surprising. As Scott Moss and others pointed out (Moss et al., 1994) the main national economic forecasting models did OK when nothing much was changing but the models missed all the turning points in the economy. It is easy to predict when current trends are roughly continuing, but if your forecast is no better than an obvious null model (such as to predict the same value as it was last time), then this is not a very useful prediction. Indeed, if it predicts when nothing fundamental changes but misses these when they occur, then this could give false confidence in the results. Clearly, any useful prediction has to do better than the obvious null models, such as 'key values will be much the same as last time' or 'current linear trends will continue'. Prediction can only be judged against a specific null model, e.g. better than random (for discrete choices) or better than a linear extrapolation over the past N data points (if continuous).

The fact is that most models do not literally predict in all its detail, but that only some aspects of its outputs are deemed predictive, means that there is an intermediate step between model results and a meaningful prediction. For example, most agent-based models have some stochastic component in them, some random elements, so that each time they are run they may trace out a slightly different path. Whilst deterministic mathematical models are still common, I know of hardly any deterministic agent-based models. Any model that uses a random seed, or does not produce exactly the same output each time it is run with the same parameter values and inputs has some stochastic element. In this case, some modellers will run the model many times to get a distribution of outcomes; here the prediction is either the expected value of this distribution (the 'Monte Carlo' method) or is the distribution itself. Clearly, explicitly or otherwise, there is some idea from the modeller about what is significant in the output and what is not. At the minimum, not all the detailed 'noise' in the output is considered as significant. To be useful, it has to be clear which elements of a model can be expected to reliably predict and in what way.

More broadly, a model might be only one source for a prediction, which might be formulated based on a mixture of kinds of evidence and not just from the results of a model. For example, the results of a model might simply confirm an existing prediction – given those making the prediction extra confidence or understanding of the underlying processes. Furthermore, in such a case, it might be that although a prediction was formulated primarily on other bases that a model is presented as the basis for the prediction because this may give the prediction more credibility.

For the above reasons, the following kinds of 'prediction' are so unuseful that other words should be used to describe them.

Empty 'anything may happen' kinds of prediction that do not narrow down the possibilities at all

Predictions that do not do any better than the obvious null models that could be postulated with zero knowledge of the phenomena concerned

Cases where it is claimed that a model predicts, but where how it predicts is left so vague that one could not tell what would count as a successful prediction and what not

Where the contribution of a model to a prediction is not clear due to the prediction being the result of a mix of sources other than the model – what the model adds should be clear

## Proving an ability to predict

Proving whether a model predicts is not straightforward, for a number of reasons, which I now discuss. This difficulty allows many claims of predictive ability to survive – since, in these cases, it is almost impossible to refute them. This is important since claims of predictive ability can have negative impacts upon policy if this turns out to be unfounded or simple brittle and unreliable.

As I discussed above, it is natural for those developing a model to project upon it an optimistic view of its powers, so one would not expect the original modellers to be the best people to evaluate their own model's predictive ability. Thus, it is best if a claim of predictive ability is verified as independently, at a minimum by checking against new data – data that is not known at the time of publication. This is for two reasons. Firstly, if one's model does not fit the data, then it is almost impossible for the modeller to resist changing the model to improve the data fit. Nobody throws away a model because it does not fit first time. Secondly, there is a publication bias against negative results, so a report of research that fits data is far more likely to get published than one that does not. If one is checking an already published model against new data, then this bias is much reduced.

Ideally, a model's predictive ability should be checked by researchers that are independent to the team that developed the model. This reduces the temptation to adjust a model to fit the target data, but it can also be helpful in revealing what conditioning is necessary to get a model to predict – separating out the contribution of the model from the skill of the modeller using it. An independent check on predictive ability can thus help to reveal any unconsciously applied adjustments that are necessary – the existence of these do not necessarily mean there is no predictive power, but does help reveal the totality of the process that is necessary for this. For example, it may be that a background parameter (one that does not represent an explicit input) be adjusted depending on some factor or other (e.g. a smoothing parameter based on volatility). For this to be able to happen the modelling would need to be 'open', as described in (Polhill & Edmonds, 2007).

If one does have a model with some predictive abilities, then it is also important to know the conditions under which it can be relied upon. The gas laws predict the temperature, pressure or volume of a gas when it approximates an ideal gas, namely when the chemical interaction between molecules is very weak. This condition about weak interaction between molecules gives us a good indication of when we might rely on the laws. A publication bias for only positive results reduces the information in the public domain about the conditions of application, because negative cases are not reported. If one has a predictive model but do not know when it is applicable, then this is much less useful since one does not know when one can safely use it and when not. There is a whole other problem of using the results of a model out of the original context, but that is the topic for a whole essay in itself and I will not deal with that here.

It is very rare that predictions about complex systems are 100% reliable. Thus, a single correct prediction (or indeed a single incorrect one) may well not give a complete picture about its general predictive power. Rather, a series of (preferably independent) attempts at predicting using a model is preferable before coming to an assessment as to its power in this regard. Such iterated testing may also provide useful information about the model's conditions of application for reliable prediction.

For all these reasons, claims of predictive ability that are not independently verified need to be taken very cautiously – for even if they turn out to be true, it is likely that information concerning this ability will be incomplete. One is right to distrust claims for predictive ability that have not been so verified. Similarly, if you are suggesting that your model can predict, or even might be able to predict in the future, this needs to be backed up with evidence that this is, in fact, the case and not merely wishful thinking.

## The desirability of prediction

Prediction has a high status in science, and there are some good reasons for this. I now briefly discuss these in order to understand some of the motivation for claiming this use of models.

If a model can reliably predict something before it is known, then this is undeniably useful. For example, predicting the short-term impact upon tax revenue due to a change in the tax rules enables governments to compare possible changes in the rules and budget for the impact of any one that is enacted. Thus, the outside world (governments, the public, funding authorities, etc.) will often want researchers to develop models for the purposes of prediction. This may well be in the form of wanting to know which of several competing policies they should take, since this involves predicting the effects of each and comparing these. Such 'what if' pronouncements from a model are effectively a conditional prediction of some kind. Reliable prediction of the impact of policies would make their decision-making easier and more defensible. Academics are often accused of living in an 'ivory tower', and this is a way of demonstrating a useful connection with the real world.

However, prediction is also valued by researchers for epistemic reasons. If one can reliably use a model to predict what is currently unknown, then this conclusively shows that at least some aspects of your model are correct (in some sense). However, it may be that what is correct is not very profound. The gas laws accurately predict the future characteristics of a gas under a wide range of conditions, but do not say anything about why they work so well and their predictions are not at all surprising (since their results were qualitatively known before – e.g. that pressure squeezes gas, or that heating it expands it). Popper (1963) discusses the importance of making risky predictions.

A model that correctly predicts surprising or unusual things gives us more information about the world because it helps us update our understanding from that where these did not occur. The model seems abIe to tell us more about we observe than our common sense. In other words, it changes our view of the world. It is also more impressive; in that it is much easier to predict something expected rather than something out of the ordinary. It is hard for modellers not to design and condition their models to present results they consider plausible (deliberately or unconsciously), so a strange prediction is less likely to have arisen from such adaptation, but seems to come from the model itself.

Of course, it may be that other good things come out of the process of trying to predict, as Polhill (2018) describes. For example, prediction competitions may motivate researchers as well as be very revealing about the assumptions researchers do and do not share, even if none of the predictions are much good (Janssen & Rollins, 2012). However, this is distinct from the main purpose for prediction – athletic competitions may spur the development of better running shoes, but this is not the point of them.

## The difficulty of prediction in complex social or ecological systems

Making successful predictions about social or ecological systems is very, very hard. I now briefly discuss some of the reasons why this might be. The reason for doing this is to emphasise the a priori reasons why one should not claim this and why one might not believe such claims, shifting the burden of proof to those that make these claims. It is, of course, impossible to say in general why prediction is hard, since systems vary so much, so the following has to be merely a sketch of some of the possible reasons. The point is not that modellers are doing it wrong (though they may be) but that prediction for these systems might be inherently hard, however one might try to do it.

### *Incompleteness*

Complicated systems (those with many processes or kinds of part) are almost never completely modelled, with some elements deemed not crucial to the outcomes not explicitly included. These are often modelled as random processes, so that each time one runs a model a slightly different outcome results, with what is significant about the outcomes is the ensemble of outcomes, rather than those from a particular run (often this is summarised as an average or a distribution of outcomes). This makes it more difficult to know if a model definitely predicts or not, and so makes any process of iterative model improvement harder.

### *Chaos*

Many complex systems are chaotic, in the sense that what is significant about outcomes is indefinitely sensitive to the value of an input or parameter (Gleick, 2011). If inputs or parameters cannot be measured indefinitely accurately, this means that what is significant cannot itself be predicted in critical regions. This means that either prediction is constrained to only those regions that are not critical, or some other feature of the outcomes is predicted (e.g. characteristics of the distribution of outcomes).

### Complication

Models of complicated systems might well have many variables and parameters. This is not, in itself, a bad thing and may be necessary, but each of these is a potential source of noise causing increased uncertainty in what is being predicted. Thus, for the purpose of prediction, adding in extra detail, even if one knows this to reflect the system being modelled, may not help the accuracy of predictions. However, omitting such features is an implicit assumption that these features will not have a significant effect on outcomes or that the effects average out in some way.

### Non-modularity

When building any software system, it is sensible to make the system as modular as possible. This allows for independent testing and evaluation of modules before they interact, increasing the chance that the developer understands the system as a whole. However, real-world systems, especially the biological, social and the ecological, are often not structured in such a neat way (Wimsatt, 1972). The result is that the effects of these different processes are not possible to analyse separately but have to be explored together. This means that we often do not fully understand our own models (e.g. as shown in Edmonds & Hales, 2003), and the search space for looking at all the possible interactions too big to do, even automatically, in most models (since it goes with the power of the number of parameters). In such a system, even if a model does seem to predict reliably, the lack of understanding of why it does means that there are no theoretical indications of when it will be reliable and when not.

### Context specificity

Many processes and characteristics of ecological and social systems are that these are highly context-specific. That is, they occur or hold for some kinds of situation and not others (Edmonds et al., 2020a). The problem here is that the characterisation of relevant kind of situation can be implicit, or even only unconsciously recognised. If a model is developed and tested within one context, or one kind of context, then it may be that that factors within the context are part of what makes prediction work (e.g. by ruling out certain kinds of process) but that these are not explicitly recognised. In that case, the prediction may unexpectantly not work for another context.

### Emergence

All of the above reasons can contribute to the phenomena of emergence – that is the appearance of patterns one might not expect from the make-up of the system to start with (because they are higher-order, more complex or just odd). Emergence is bread-and-butter for ABMs because they are ideally suited to representing this kind of situation (Bankes, 2002). An exact definition is contested, but one that is clear is that of 'weak emergence' (Bedau, 1997), which is where there is not a 'short cut' to determining the (significant aspects of) the outcomes from the set-up without running the simulation. In this sense, unpredictability is not a result of emergence, but part of its cause. However, there is also the possibility of fundamentally new elements appearing in what one is trying to simulate – elements that go beyond what is in any simulation representing it as (Polhill et al., 2021) point out.

The above factors do not mean that it is impossible to make any predictions in these domains, just that it is likely to be difficult and so one has to be careful when attempting or claiming this. Nate Silver and his team regularly publish predictions about US elections and sporting events before they occur (Silver, 2012), but they do not use agent-based simulations and are very careful about the claims they make (although they do use disaggregated simulations with stochastic perturbations of the input data; Edmonds, le Page et al., 2019).

## The rhetoric of prediction

The tension between the desirability of claiming prediction and the difficulty of achieving it is at the root of the fundamental danger here. Given its desirability researchers want to be able to suggest that their models have predictive ability, but what is the track record of such claims in this domain? We cannot do a systematic review, checking up on all such claims (that would be a very long term, community-wide project) but we can present some negative evidence and examples illustrating the dangers.

In April 2009, Scott Moss asked the SimSoc email distribution list whether there had ever been an ABM making a successful prediction of a policy implementation. The answer was 'no' – many of the examples turned out to not be prediction in any strong sense. In 2019, I repeated this challenge in the slightly different format of asking the question: 'Are there any documented examples of models that predict useful aspects of complex social systems?' (Edmonds, Polhill et al., 2019). Again, although there were examples of empirical modelling sent in, none of these were such that the prediction was documented before the event happened. Rather, some were more about establishing the empirical adequacy of models as evidence to support an explanation and others merely expressing the hope that their model would predict.

The undesirability of claiming predictive ability when this has not been convincingly verified or merely allowing an audience to think a model has this ability is illustrated in the following two examples.

The ecological modelling of the cod fisheries off Newfoundland and Labrador, which were used to judge the sustainability of its fishing quotas (essentially predictions of fish stocks), was part of the fisheries management system that led to its collapse. Indeed, a number of Commissions investigating the status of the cod, at the behest of the fishermen's concerns, failed to make adequate inferences about the ailing stock health. These Commissions were buoyed by model-based assessments that predicted an increasing resource base. These predictions supported the expansion of the fishery throughout much of the 1980s. During this time, these figures were consistently challenged by inshore fishermen, but it was not until 1989 that the erroneous forecasting of fish stocks was corrected (Finlayson 1994). The subsequent Harris Commission after its collapse (Harris, 1990) said: ' *... scientists ... overconfident of the validity of their predictions, failed to recognize the statistical inadequacies in their bulk biomass model and failed to ... recognize the high risk involved with state-of-stock advice based on relatively short and unreliable data series.*' (p. 2). Although one cannot know for sure the extent to which the policy-makers believed the predictions, they undoubtedly were a factor in its collapse. This is discussed more by Edmonds & Adoha (2019).

In 1972, a group of academics under the auspices of 'The Club of Rome' published a systems dynamics model that showed how laggy feedback between key variables could result in a catastrophe (in terms of a dramatic decrease in the variable for world population; Meadows et al., 1972). Unfortunately, this stark and much needed illustration was widely taken as a prediction and the model criticised on these grounds. The book did not make it explicit that the purpose of this model was not predictive, with such phrases as '*We feel that the model described here is already sufficiently developed to be of some use to decision-makers*'. This left the book open to attack on the grounds of model fragility – the model was not suitable for prediction due to the sensitivity of its outputs to some of its unknown inputs. Although it is highly likely that vested interests were the main reason for these attacks, not making the nature of the model clear made it more vulnerable, blunting the impact of the overall message.

These examples illustrate the potential danger of implying prediction – policy makers seem to be particularly asked for prediction, so any suggestion that a model might predict will be grabbed upon by them. However, (as Scott Moss used to point out) it is immoral to allow a stakeholder to get the wrong impression as to what can actually be achieved – however much they want this. The above example from the fishing industry vividly illustrates the harm that such misperceptions can cause.

Thus we, with others (De Matos Fernandes & Keijzer, 2020; Steinmann et al., 2020), ask researchers to avoid claiming prediction for their models, at least until this ability has been independently verified several times. Whilst an enthusiasm for one's own model might lead one to hope it can do all kinds of useful things (including prediction), it is part of scientific rigour to not speculate where this will mislead others. Others might be tempted to do so out of tradition or terminological laziness where it benefits them to do so, but if one aspires to being a 'scientist' one would do better to avoid this (Edmonds, 2017).

## A special case: Prediction in meta-modelling

There is one area where epistemological prediction is not only more feasible, but positively helpful – in the understanding of our own simulation models. Due to the kind of reasons discussed above in the section on the difficulty of prediction, many of our models are inevitably, themselves, complex. Here, the model is an intermediate stage between us and what we are trying to understand, giving us 'leverage' over the phenomena that we could not have without the model. In such cases, it is desirable to try and understand the model as much as possible, especially if the model is deemed interesting enough to present to others. Whilst a sensitivity analysis is start in such understanding, it is not enough – what is needed is a good explanation of what in the model causes the significant results. In other words, we need a good meta-model.

It is good practice to check whether an explanation concerning what a model is doing is a good one by trying to refute it using simulation experiments. In other words, to make predictions, based upon the explanation and then see if the predictions hold up in the face of various interventions to the model (e.g. turning off a particular component process). Here one is not making predictions about observed phenomena using the model, but rather making predictions about the model. Making predictions about a model is far more feasible, since some of the difficulties that underly the prediction of observed complex phenomena do not hold: we know the micro-foundations of the system (since we can inspect the code), we can indefinitely inspect and experiment with it (so we have good data about it), we can map out any chaotic elements and test this to a very fine level of precision, we know exactly where the stochasticity is (and so can temporarily turn it off to make the system deterministic), we can experimentally determine the conditions of application of the explanation, and although different views about the model are still possible (due to our incomplete understanding of it), all such views can be grounded in definite referents (the elements in the model). Thus, epistemological prediction can play a vital part in checking our understanding of our own creations.

Of course, if one has a good predictive meta-model, then the temptation is to project this onto what the model is representing – the observed phenomena. Maybe due to the effect of the 'Kuhnian Spectacles' – whereby one filters one's perceptions of the phenomena using our understanding of the model (Kuhn, 1962) – modellers are sometimes lazy about distinguishing between their model and what it represents – conflating or confusing them (Edmonds, 2020b). However, the huge difficulties in predicting observed phenomena remain and one should still be careful that progress in the technical matter of understanding one's own model does not lead one to premature anticipation of an ability to predict the phenomena that the model represents.

## Taking prediction seriously

As I have hinted above, there are two distinct purposes for prediction: as a forecast of the unknown that might be useful to someone else (e.g. in decision-making or the development of technology) or as an epistemological test of a model (Sarewitz & Pielke, 1999).

In terms of useful prediction of the unknown, the practicalities of doing it well have been covered by others. Armstrong (2001) provides a comprehensive guide to forecasting, whose lessons are usefully extracted for ABM in (Hassan et al., 2013) and (Steinmann et al., 2020) review approaches suitable for situations of deep uncertainty. However, I would like to highlight three areas here – the slippage areas where the rhetoric and practice might be confused.

Firstly, there needs to be honest transparency concerning any attempts at prediction, documenting: what is being predicted and to what level of accuracy or probability, the context and purpose for the prediction (who it is intended to be useful to and in what way), at least a sketch of the likely conditions of application (under what conditions it is expected to work), and the data and model conditioning processes that are involved in getting the model to predict well.

The power of any predictive ability needs to be compared to the obvious null models. The null model consisting of 'the value at the next step is the same as the present value' predicts surprisingly well in many cases, but does not tell us much except that the temporal granularity is sufficiently fine that outcomes do not tend to change much. A useful predictive ability needs to do more than this.

Due to the inevitable biases that creep in (e.g. the publication bias against unsuccessful predictions) establishing a process where predictions are tried many times by independent teams. Not only does this provide a more reliable proof of the predictive ability, but is also likely to result in a refined understanding of when, how and why the prediction works. It is only such a transparent process by a community of scientists that could provide any level of public confidence in the ability (Bithell & al. 2022). If one is serious about learning to predict, then the failures need to be recorded as much as the successes.

Finally, great care should be taken in the context and manner in which predictive claims are made. What might be acceptable within a room of modellers who can distinguish between sheer enthusiasm and actual model ability, is not acceptable to use language that could deceive when presenting to policy actors, funders in grant applications or the wider public. Guidance as to the assumptions behind the model and its proper interpretation is essential but, more fundamentally, all claims need to err on the side of caution – the public expects us only to make pronouncements when we are pretty sure of them.

In terms of epistemological prediction, that process is similar to that of checking an explanation encapsulated in a model, but one that avoids the particular danger of conditioning a model to fit the target data. To aid this, predictions and the accompanying documentation should be publicly archived before the predictions can be checked in a similar manner to that of experiments in psychology following their replication crisis (e.g. Spellman et al., 2017). However, simple good practice in explanatory modelling (where the target data is known by the modeller) can largely ameliorate these dangers – particularly in terms of transparent honesty about the process (maybe via the 'Trace' documentation proposed in Grimm et al., 2014).

## Conclusion

A lot has been made of the epistemological implications, and hence status, of prediction. Prediction is still seen as the 'gold standard' of science. However, a mature field of science is able to relate the different uses for modelling – bridging between explanation, prediction, description, theoretical explorations and mental models for phenomena. These are not within a hierarchy of desirability, with prediction at the top, but simply different uses of models.

Regardless of whether, in an abstract or epistemological sense, prediction is linked in terms of structure to other uses for modelling, in practice it is very different – different in terms of: goals, kinds of model, levels of uncertainty, process, temporal sequence, possible mistakes (and how to mitigate against these) and the steps necessary to check it. The complexity and sheer messiness of many observed social and ecological systems means that, although to be able to claim prediction is attractive, it is very, very hard to achieve. For these reasons, there are a priori grounds to suspect that many such claims will turn out to be other than a reliable and useful ability to anticipate unknown data when checked against independent data and by independent researchers. Predicting something useful is a messy, pragmatic affair characterised by its own goals, activities, checks and dangers.

The conclusions of this paper are fourfold: (a) that we should be very careful to ensure that the wider public, but particularly policy actors and funding reviewers, understand the nature of any claims made, preferably avoiding the word 'prediction' altogether in that context (b) that other purposes for modelling, such as supporting complex explanations of empirical data, are distinguished and valued, (c) that epistemological prediction does have a valuable role in the understanding of our own models and (d) if one is going to attempt prediction, there should be minimum standards that apply, so as to make any such attempts as useful as possible, both to science and a wider set of stakeholders.

Claiming a model 'predicts' something when really all it does is fit some known data, or implying a model will predict without good evidence that this may be so is misleading to others and, hence, should not be acceptable behaviour. As a modelling community, we need to 'up our game', being more transparent and modest in the short term, in order to ensure our collective reputation and facilitate progress in the longer term – an empty rhetoric of prediction damages both. Prediction of complex systems is very hard and, if attempted, should be subject to high standards of rigour and proof.

## Notes

1. For an arbitrary but extensive list of papers that mention 'prediction' in this context, I suggest you do a Google Scholar search for: *predict 'agent-based'* after you have read the paper to verify its claims.

## Acknowledgments

I would like to thank all those with whom I have discussed these issues. These are far too numerous to remember – I have bent the ear of too many people for too long in this regard. However, most significantly these include Scott Moss, who was pointing out the immorality of claiming prediction when one cannot deliver it, way before I even entered academia, as well as long and sometimes useful conversations with Juliette Rouchier, David Hales and Gary Polhill. The author acknowledges funding from the ESRC, project number ES/S015159/1.

## Disclosure statement

No potential conflict of interest was reported by the author.

## Notes on contributor

*Bruce Edmonds* is the Director of the Centre for Policy Modelling and Senior Research Fellow at the Manchester Metropolitan University. His first degree was in Mathematics, and his PhD on "Syntactic Measures of Complexity" in the philosophy of science. He now works in the field of social simulation, but particularly using agent-based modelling techniques to understand complex social phenomena using the maximum amount of empirical evidence possible. He publishes widely across the fields of social science, cognitive science, computer science and philosophy.

## References

Alsaedi, N., Burnap, P., & Rana, O. (2017). Can we predict a riot? Disruptive event detection using twitter. *ACM Trans. Internet Technol*, *17*(2), Article 18. https://doi.org/10.1145/2996183

Armstrong, J. S. (2001). *Principles of forecasting: A handbook for researchers and practitioners*. Kluwer Academic Publishers. https://doi.org/10.1007/978-0-306-47630-3

Bankes, S. C. (2002). Agent-based modeling: A revolution? *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7199–7200. https://doi.org/10.1073/pnas.072081299

Bedau, M. A. (1997). Weak emergence. *Philosophical Perspectives*, *11*, 375–399. https://www.jstor.org/stable/2216138

Bithell, M., Chattoe-Brown, E., & Edmonds, B. (2022). The Large-Scale, Systematic and Iterated Comparison of Agent-Based Policy Models. In: M. Czupryna & B. Kamiński (Eds.) Advances in Social Simulation. Springer Proceedings in Complexity. Cham: Springer https://doi.org/10.1007/978-3-030-92843-8_28https://doi.org/10.1007/978-3-030-92843-8_28, .

Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Stat. Sci*, *16*, 199–231. https://doi.org/10.1214/ss/1009213726

Cartwright, N. (1983). *How the laws of physics lie*. Clarendon Press.

de Matos Fernandes, C. A., & Keijzer, M. A. (2020) *No one can predict the future: More than a semantic dispute*. Review of Artificial Societies and Social Simulation. 15th April 2020. https://rofasss.org/2020/04/15/no-one-can-predict-the-future/

Edmonds, B. (2017) The post-truth drift in social simulation. *Social Simulation Conference (SSC2017)*, Dublin, Ireland. http://cfpm.org/discussionpapers/195

Edmonds, B. (2020a). How social simulation could help social science deal with context. In H. Verhagen (Eds.) *Advances in social simulation springer - looking in the mirror. Springer proceedings in complexity*, Springer, (pp. 1–14). https://doi.org/10.1007/978-3-030-34127-5_1

Edmonds, B. (2020b) *Basic modelling hygiene - keep descriptions about models and what they model clearly distinct*. Review of Artificial Societies and Social Simulation. 22nd May 2020. https://rofasss.org/2020/05/22/modelling-hygiene/

Edmonds, B., & Adoha, L. (2019). Using agent-based simulation to inform policy – What could possibly go wrong? In P. Davidson & H. Verhargen (Eds.), *Multi-agent-based simulation XIX, 19th International Workshop, MABS 2018* (pp. 1–16). Springer. https://doi.org/10.1007/978-3-030-22270-3_1

Edmonds, B., & Hales, D. (2003). Replication, replication and replication - Some hard lessons from model alignment. *Journal of Artificial Societies and Social Simulation*, *6*(4), 11. http://jasss.soc.surrey.ac.uk/6/4/11.html

Edmonds, B., le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different modelling purposes. *Journal of Artificial Societies and Social Simulation*, *22*(3), 6. https://doi.org/10.18564/jasss.3993

Edmonds, B., Polhill, G., & Hales, D. (2019) *Predicting social systems – A challenge*. Review of Artificial Societies and Social Simulation. 4th June 2019. https://rofasss.org/2018/11/04/predicting-social-systems-a-challenge

Elsenbroich, C. (2012). Explanation in agent-based modelling: Functions, causality or mechanisms? *Journal of Artificial Societies and Social Simulation*, *15*(3), 1. https://doi.org/10.18564/jasss.1958

Epstein, J. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, *11*(4), 12. https://www.jasss.org/11/4/12.html

Finlayson, A. C. (1994). *Fishing for truth : A sociological analysis of northern cod stock assessments from 1977 to 1990*. Institute of Social and Economic Research, Memorial University of Newfoundland.

Giere, R. N. (1988). *Explaining science - A cognitive approach*. University of Chicago Press.

Gleick, J. (2011). *Chaos: Making a new science*. Open Road Media.

Grimm, V., Augusiak, J., Focks, A., Frank, B. M., Gabsi, F., Johnston, A. S., Liu, C., Martin, B. T., Meli, M., Radchuk, V., Thorbek, P., & Railsback, S. F. (2014). Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. *Ecological Modelling*, *280*, 129–139. https://doi.org/10.1016/j.ecolmodel.2014.01.018

Harris, L. (1990) *Independent review of the state of the northern cod stock executive summary and recommendations*. https://waves-vagues.dfo-mpo.gc.ca/Library/114276.pdf

Hassan, S., Arroyo, J., Galán Ordax, J. M., Antunes, L., & Pavón Mestras, J. (2013). Asking the oracle: Introducing forecasting principles into agent-based modelling. *Journal of Artificial Societies and Social Simulation*, *16*(3), 13. https://doi.org/10.18564/jasss.2241

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature*, *595*(7866), 181–188. https://doi.org/10.1038/s41586-021-03659-0

Janssen, M. A., & Rollins, N. (2012). Predicting behavior in new behavioral experiments: Outcomes of a modeling competition. *Journal of Artificial Societies and Social Simulation*, *15*(3), 5. https://doi.org/10.18564/jasss.2015

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Lackoff, A. (2008) Women, fire and dangerous things. University of Chicago press.

Lambert, P., & Rochard, E. (2007). Identification of the inland population dynamics of the European eel using pattern-oriented modelling. *Ecological Modelling*, *206*(1–2), 166–178. https://doi.org/10.1016/j.ecolmodel.2007.03.033

Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972). *The limits to growth*. Universe Books.

Moss, S., Artis, M., & Ormerod, P. (1994). A smart automated macroeconometric forecasting system. *Journal of Forecasting*, *13*(3), 299–312. https://doi.org/10.1002/for.3980130305

Polhill, G. (2018) *Why the social simulation community should tackle prediction*. Review of Artificial Societies and Social Simulation, 6th August 2018. https://rofasss.org/2018/08/06/gp/

Polhill, J. G., & Edmonds, B. (2007). Open access for social simulation. *Journal of Artificial Societies and Social Simulation*, *10*(3). https://www.jasss.org/10/3/10.html

Polhill, G., Hare, H., Anzola, D., Bauermann, T., French, T., Post, H., & Salt, D. (2021). Using agent-based models for prediction in complex and wicked systems. *Journal of Artificial Societies and Social Simulation*, *24*(3), 2. https://doi.org/10.18564/jasss.4597

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Routledge.

Sarewitz, D., & Pielke, R. (1999). Prediction in science and policy. *Technology in Society*, *21*(2), 121–133. https://doi.org/10.1016/S0160-791X(99)00002-0

Silver, N. (2012). *The signal and the noise: The art and science of prediction*. Penguin UK.

Spellman, B., Gilbert, E. A., & Corker, K. S. (2017). Open science: What, why, and how. https://doi.org/10.31234/osf.io/ak6jr.

Steinmann, P., Wang, J. R., van Voorn, G. A. K., & Kwakkel, J. H. (2020) *Don't try to predict COVID-19. If you must, use deep uncertainty methods*. Review of Artificial Societies and Social. Simulation, 17th April 2020. https://rofasss.org/2020/04/17/deep-uncertainty/

Thompson, N. S., & Derr, P. (2009). Contra Epstein, good explanations predict. *Journal of Artificial Societies and Social Simulation*, *12*(1), 9. https://jasss.soc.surrey.ac.uk/12/1/9.html

Thorngate, W., & Edmonds, B. (2013). Measuring simulation-observation fit: An introduction to ordinal pattern analysis. *Journal of Artificial Societies and Social Simulation*, *16*(2), 14. https://jasss.soc.surrey.ac.uk/16/2/4.html

Troitzsch, K. G. (2009). Not all explanations predict satisfactorily, and not all good predictions explain. *Journal of Artificial Societies and Social Simulation*, *12*(1), 10. https://jasss.soc.surrey.ac.uk/12/1/10.html

Watts, D. J. (2014). Common sense and sociological explanations. *American Journal of Sociology*, *120*(2), 313–351. https://www.jstor.org/stable/10.1086/678271

Wimsatt, W. (1972). Complexity and organisation. In K. Schaffner & R. Cohen (Eds.), *Studies in the philosophy of sciences* (pp. 67–86). Reidel.

Zhang, M., Verbraeck, A., Meng, R., Chen, B., & Qiu, X. (2016). Modeling spatial contacts for epidemic prediction in a large-scale artificial city. *Journal of Artificial Societies and Social Simulation*, *19*(4), 3. https://doi.org/10.18564/jasss.3148