Different Modelling Purposes

Bruce Edmonds, Christophe le Page, Cristina Montanola, Paul Ormerod, Flaminio Squazzoni, and Hilbert Root[1]

Abstract:

How one builds, checks, validates and interprets a model depends on its 'purpose'. This is true even if the same model code is used for different purposes. This means that a model built for one purpose but then used for another needs to be re-justified for the new purpose and this will probably mean it also has to be re-checked, re-validated and maybe even re-built in a different way. Here we review some of the different purposes for a simulation model of complex social phenomena, focussing on seven in particular: prediction, explanation, description, theoretical exploration, illustration, analogy, and Social Interaction. The paper looks at some of the implications in terms of the ways in which the intended purpose might fail. This analysis motivates some of the ways in which these 'dangers' might be avoided or mitigated. It also looks at the ways that a confusion of modelling purposes can fatally weaken modelling projects, whilst giving a false sense of their quality. These distinctions clarify some previous debates as to the best modelling strategy (e.g. KISS and KIDS). The paper ends with a plea for modellers to be clear concerning which purpose they are justifying their model against.

Introduction

A common view of modelling is that one builds a 'life-like' reflection of some system, which then can be relied upon to act like that system. This is a correspondence view of modelling where the details in the model correspond in a one-one manner[2] with those in the modelling target – as if the model were some kind of 'picture' of what it models. However, this view can be misleading since models always differ from what they model, so that they will capture some aspects of the target system but not others. With complex phenomena, especially social phenomena, it is inevitable that any model is, at best, a very partial picture of what it represents – in fact I suggest that this picture analogy is not appropriate when it comes to the justification or judging of models[3].

Rather, here I suggest a more pragmatic approach, where models are judged as *tools* designed for specific purposes – the goodness of the model against how good it is for its declared purpose[4]. Although a model designed for one purpose may turn out to be OK for another, it is more productive to use a tool designed for the job in hand. One may be able to use a kitchen knife for shaping wood, but it is much better to use a chisel. In particular, I argue that even when a model (or model component) turns out to be useful for more than one purpose it needs to be re-justified with respect to *each* of the claimed purposes *separately*. To extend the previous analogy: a tool with the blade of a chisel but the handle of a kitchen knife may satisfy some of the criteria for a tool to carve wood and some of the criteria for a tool to carve cooked meat, but not be much good for anything. If one *did* come up with a new tool that is good at both, this would be because is could be justified for each purpose separately.

In his paper 'Why Model?', Epstein (2008) lists 17 different reasons[5] for making a model: from the abstract, 'discover new questions', to the practical 'educate the general public'. This illustrates both the usefulness of modelling but also the potential for confusion. As Epstein points out, the power of modelling comes from making an informal set of ideas formal. That is, they are made precise using unambiguous code or mathematical symbols. This lack of ambiguity has huge benefits for the process of science, since it allows researchers to share, critique, and improve models without transmission errors (Edmonds 2010). However, in many social simulation papers the purpose that its model was developed for, or more critically, the purpose under which it is being presented is often left implicit or confused. Maybe this is due to the prevalence of the 'correspondence picture' of modelling discussed above, maybe the

authors optimistically conceive of their creations being useful in many different ways, or maybe they simply developed the model without a specific purpose in mind. However, regardless of the reason, the consequence is that readers do not know how to judge the model when presented. This has the result that models might avoid proper judgement – demonstrating partial success in different ways with respect to a number of purposes, but not adequacy against any. The recent rise of standard protocols like ODD (Grimm et al. 2006) for describing published models may help improving this inconvenient situation: stating the purpose of the model constitutes the first part of the framework.

Our use of language helps cement this confusion: we talk about a 'predictive model' as if it is something in the code that makes it predictive (forgetting the role of the mapping between the model and what it predicts). Rather I am suggesting a shift from the code as a thing in itself, to code as a tool for a particular purpose. This marks a shift from programming, where the focus is on the nature and quality of the *code*, to modelling, where the focus is on the *relationship* of the behaviour of some code to what is being modelled. Using terms such as 'explanatory model' is OK, as long as we understand that this is shorthand for 'a model which establishes an explanation' etc.

Producing, checking and documenting code is labour intensive. As a result, we often wish to reuse some code produced for one purpose for another purpose. However, this often causes as much new work as it saves due to the effort required to justify code for a new purpose, and – if this extra work is not done – the risk that time and energy of many researchers are wasted due to the confusions and false sense of reliability that can result. In practice, I have seen very little code that does not need to be re-written when one has a new purpose in mind. Ideas can be transferred and well-honed libraries for very well defined purposes, but not the core code that makes up a model of complex social phenomena[6]. This is partly because different uses imply different assumptions (Bell & al. 2015) and be open to different risks (as we will argue below).

In this paper, I will look at seven common modelling purposes:
- *prediction,*
- *explanation,*
- *description,*
- theoretical exploration,
- illustration,
- analogy[7],
- and social interaction[8].

The first three of these are empirical – they have a well-defined relationship with observed evidence – the others are not directly related to what is observed, but concern ways of thinking about things, theoretical properties or communication. There are, of course, many other reasons for modelling (Epstein 2008) but the above seven seem to be the main ones that concern the papers in the field of social simulation (and JASSS). Given the pluralist and pragmatist intent of this paper, we have avoided trying to theorise too much as to how these purposes relate or cluster, but rather simply aim to distinguish these and discuss the mostly practical aspects that relate to each. There will be other purposes for modelling that we have not covered, or even thought of – this is fine, the main aim of the paper is to get authors to make their purpose crystal clear so that readers know how on what basis it should be judged.

Each of these purposes is discussed, in turn, below. For each purpose a brief 'risk analysis' is presented – some of the ways one might fail to achieve the stated purpose – along with some ways of mitigating these risks. In the penultimate section, some common confusions of purpose are illustrated and discussed, before the paper concludes with a discussion on modelling strategies, a brief summary and plea to make one's purpose clear.

Prediction

Motivation

If one can reliably predict anything that is not already known, this is undeniably useful regardless of the nature of the model (e.g. whether its processes are a reflection of what happens in the observed system or not[9]). For instance, the gas laws (stating e.g. that at a fixed pressure, the increase in volume of gas is proportional to the increase of temperature) were discovered long before the reason why they worked.

However, there is another reason that prediction is valued: it is considered the gold standard of science – the ability of a model or theory to predict is taken as the most reliable indicator of a model's truth. This is done in two principle ways: (a) model A fits the evidence better than model B – a comparative approach[10] or (b) model A is falsified (or not) by the evidence – a falsification approach. In either, the idea is that, given a sufficient supply of different models, better models will be gradually selected over time, either because the bad ones are discarded or outcompeted by better models.

Definition

> *By 'prediction', we mean* the ability to reliably anticipate data that is not currently known to a useful degree of accuracy *via computations using the model.*

Unpacking this definition:

- It has to do it *reliably* – that is under some known (but not necessarily precise) conditions the model will work; otherwise one would not know when one could use it.
- The data it anticipates has to be *unknown* to the modeller. 'Predicting' out-of-sample data is not enough, since pressures to re-do a model and get a better fit are huge and negative results are difficult to publish.
- The anticipation has to be to a *useful* degree of accuracy. This will depend upon the purpose to which it is being put, e.g. as in weather forecasting.

Unfortunately, there are at least two different uses of the word 'predict'. Almost all scientific models 'predict' in the weak sense of being used to calculate some result given some settings or data, but this is different from correctly anticipating *unknown* data. For this reason, some use the term 'forecast' for anticipating unknown data and use the word 'prediction' for almost any calculation of one aspect from another using a model[11]. However, this causes confusions in other ways so this does not necessarily make things clearer. *Firstly*, 'forecasting' implies that the unknown data is in the future (which is not always the case) and, *secondly*, large parts of science use the word 'prediction' for the process of anticipating unknown data. For example, if a modeller says their model 'predicts' something when they simply mean that it calculates it, then most of the audience may misunderstand and assume the author is claiming more utility than is intended.

As Watts (2014) points out, useful prediction does not have to be a 'point' prediction of a future event. For example, one might predict that some particular thing will not happen, the existence of something in the past (e.g. the existence of Pluto), something about the shape or direction of trends or distributions, or even qualitative facts. The important fact is that what is being predicted is not known beforehand by the modeller and that it can be unambiguously checked when it is known.

An Example

Nate Silver aims to predict social phenomena, such as the results of elections and the outcome of sports competitions. This is a data-hungry activity, which involves the long-term development of simulations that carefully see what can be inferred from the available data. As well as making predictions, his unit tries to establish the level of uncertainty in those predictions – being honest about the probability of those predictions coming about given the

likely levels of error and bias in the data. As described in his book (Silver 2012) this involves a number of properties and activities, including:

- Repeated testing of the models against unknown data;
- Keeping the models fairly simple and transparent so one can understand clearly what they are doing (and what they do not cover);
- Encoding into the model aspects of the target phenomena that one is relatively certain about (such as the structure of the US presidential electoral college);
- Being heavily data-biased, requiring a lot of data to help eliminate sources of error and bias;
- Producing probabilistic predictions, giving a good idea about the level of uncertainty in any prediction;
- Being clear about what kind of factors are not covered in the model, so the predictions are relative to a clear set of declared assumptions and one knows the kind of circumstances in which one might be able to rely upon the predictions.

Post hoc analysis of predictions – explaining why it worked or not – is kept distinct from the predictive models themselves – this analysis may inform changes to the predictive model but is not then incorporated into the model. The analysis is thus kept independent of the predictive model so it can be an effective check. Making a good predictive model requires a lot of time getting it wrong with real, unknown data, and trying again before one approaches qualified successful predictions.

Risks

Prediction (as we define it) is *very* hard for any complex social system. For this reason, it is rarely attempted[12]. Many re-evaluations of econometric models against data that has emerged since publication have revealed a high rate of failure (e.g. Meese & Rogoff 1983) – 37 out of 40 model failed completely. Clearly, although presented as being predictive models, they did not actually predict unknown data. Many of these used the strategy of first dividing the data into in-sample and out-of-sample data, and then parameterising the model on the former and exhibiting the fit against the latter. Presumably, the apparent fit of the 37 models was not simply a matter of bad luck, but that all of these models had been (explicitly or implicitly) fitted to the out-of-sample data, because the out-of-sample data was known to the modeller before publication. That is, if the model failed to fit the out-of-sample data the first time the model was tested, it was then adjusted until it did work, or alternatively, only those models that fitted the out-of-sample data were published (a publishing bias). Thus, in these cases the models were not tested against predicting the out-of-sample data even though they were presented as such. Fitting known data is simply not a sufficient test for predictive ability.

There are many reasons why prediction of complex social systems fails, but two of the most prominent are: (1) it is unknown what processes are needed to be included in the model and (2) a lack of enough quality data of the right kinds. We will discuss each of these in turn.

1. In the physical sciences, there are often well-validated micro-level models (e.g. fluid dynamics in the case of weather forecasting) that tell us what processes are potentially relevant at a coarser level and which are not. In the social sciences this is not the case – we do not know what the essential processes are. Here, it is often the case that there are other processes that the authors have not considered that, if included, would completely change the results. This is due to two different causes: (a) we simply do not know much about how and why people behave in different circumstances and (b) different limitations of intended context will mean that different processes are relevant.
2. Unlike in the physical sciences, there has been a paucity of the kind of data we would need to check the predictive power of models. This paucity can be due to (a) there is not enough data (or data from enough independent instances) to enable the iterative checking and adapting of the models on new sets of unknown data each time we need to, or (b) the data is not of the right kind to do this. What can often happen is that one has partial sets of data

that require some strong assumptions in order to compare against the predictions in question (e.g. the data might only be a proxy of what is being predicted, or you need assumptions in order to link sets of data). In the former case, (a), one simply has not enough to check the predictive power in multiple cases so one has to suspend judgement as to whether the model predicts in general, until the data is available. In the latter case, (b), the success at prediction is relative to the assumptions made to check the prediction.

A more subtle risk is that the conditions under which one can rely upon a model to predict well might not be clear. If this is the case then it is hard to rely upon the model for prediction in a new situation, since one does not know its conditions of application.

Mitigating Measures

To ensure that a model does indeed predict well, one can seek to ensure the following:
- That the model has been tested on several cases where it has successfully predicted data unknown to the modeller (at the time of prediction);
- That information about the following are included: exactly what aspects it predicts, guidelines on when the model can be used to predict and when not, some guidelines as to the degree or kind of accuracy it predicts with, any other caveats a user of the model should be aware of;
- That the model code is distributed so others can explore when and how well it predicts.

Explanation

Motivation

Often, especially with complex social phenomena, one is particularly interested in understanding *why* something occurs – in other words, explaining it. Even if one cannot predict something before it is known, you still might be able to explain it afterwards. This distinction mirrors that in the physical sciences where there are both *phenomenological* as well as *explanatory* laws (Cartwright 1983) – the former matches the data, the latter explains why that came about. Often we have good predictive or good explanatory models but not both. For example, the gas laws that link measurements of temperature, pressure and volume were known before the explanation in terms of molecules of gas bouncing randomly around, and similarly Darwin's explanatory Theory of Natural Selection was known long before the Neo-Darwinist synthesis with genetics that allowed for predictive models.

Understanding is important for managing complex systems as well as understanding when predictive models might work. Whilst generally with complex social phenomena explanation is easier than prediction, sometimes prediction comes first (however if one can predict then this invites research to explain *why* the prediction works).

If one makes a simulation in which certain mechanisms or processes are built in, and the outcomes of the simulation match some (known) data, then this simulation can support an explanation of the data using the built-in mechanisms. The explanation itself is usually of a more general nature and the traces of the simulation runs are examples of that account. Simulations that involve complicated processes can thus support complex explanations – that are beyond natural language reasoning to follow. The simulations make the explanation explicit, even if we cannot fully comprehend its detail. The formal nature of the simulation makes it possible to test the conditions and cases under which the explanation works, and to better its assumptions.

Definition

> By 'explanation' we mean establishing a possible causal chain from a set-up to its consequences *in terms of the mechanisms in a simulation.*

Unpacking some parts of this:

- The *possible causal chain* is a set of inferences or computations made as part of running the simulation – in simulations with random elements, each run will be slightly different. In this case, it is either a possibilistic explanation (A could cause B), in which case one just has to show one run exhibiting the complete chain, or a probabilistic explanation (A probably causes B, or A causes a distribution of outcomes around B) in which case one has to look at an assembly of runs, maybe summarising them using statistics or visual representations.

- For explanatory purposes, the structure of the model is important, because that limits what the explanation consists of. If, for example, the model consisted of mechanisms that are known *not* to occur, any explanation one established would be in terms of these non-existent mechanisms – which is not very helpful. If one has parameterised the simulation on some in-sample data (found the values of the free parameters that made the simulation fit the in-sample data) then the explanation of the outcomes is also in terms of the in-sample data, mediated by these 'magic' free parameters[13].

- The *consequences* of the simulations are generally measurements of the outcomes of the simulation. These are compared with the data to see if it 'fits'. It is usual that only some of the aspects of the target data and the data the simulation produces are considered significant – other aspects might not be (e.g. might be artefacts of the randomness in the simulation or other factors extraneous to the explanation). The kind of fit between data and simulation outcomes needs to be assessed in a way that is appropriate to what aspects of the data are significant and which are not. For example, if it is the level of the outcome that is key then a distance or error measure between this and the target data might be appropriate, but if it is the shape or trend of the outcomes over time that is significant then other techniques will be more appropriate (e.g. Thorngate & Edmonds 2013)

Example

Stephen Lansing spent time in Bali as an anthropologist, researching how the Balinese coordinated their water usage (among other things). He and his collaborator, James Kramer, build a simulation to show how the Balinese system of temples acted to regulate water usage, through an elaborate system of agreements between farmers, enforced through the cultural and religious practices at those temples (Lansing & Kramer, 1993). Although their observations could cover many instances of localities using the same system of negotiation over water, they were necessarily limited to all their observations being within the same culture. Their simulation helped establish the nature and robustness of their explanation by exploring a close universe of 'what if' questions, which vividly showed the comparative advantages of the observed system that had developed over a considerable period. The model does not predict that such systems will develop in the same circumstances, but it substantially adds to the understanding of the observed case.

Risks

Clearly, there are several risks in the project of establishing a complex explanation using a simulation – what counts as a good explanation is not as clear-cut as what is a good prediction.

Firstly, the fit to the target data to be explained might be a very special case. For example, if many other parameters need to have very special values for the fit to occur, then the explanation is, at best, brittle and, at worst, an accident.

Secondly, the process that is unfolded in the simulation might be poorly understood so that the outcomes might depend upon some hidden assumption encapsulated in the code. In this case, the explanation is dependent upon this assumption holding, which is problematic if this assumption is very strong or unlikely.

Thirdly, there may be more than one explanation that fits the target data. So although the simulation establishes one explanation it does not guarantee that it is the only candidate for this.

Mitigating Measures

To improve the quality and reliability of the explanation being established:

- Ensure that the mechanisms built into the simulation are plausible, or at least relate to what is known about the target phenomena in a clear manner;
- Be clear about which aspects of the outcomes are considered significant in terms of comparison to the target data – i.e. exactly which aspects of that target data are being explained;
- Probe the simulation to find out the conditions for the explanation holding using sensitivity analysis, addition of noise, multiple runs, changing processes not essential to the explanation[14] to see if the results still hold, and documenting assumptions;
- Do experiments in the classic way, to check that the explanation does, in fact, hold for your simulation code – i.e. check your code and try to refute the explanation using carefully designed experiments with the model.

Description

Motivation

An important, but currently under-appreciated, activity in science is that of description. Charles Darwin spent a long time sketching and describing the animals he observed on his travels aboard the HMS Beagle. These descriptions and sketches were not measurements or recordings in any direct sense, since he was already selecting from what he perceived and only recording an abstraction of what he thought of as relevant. Later on, these were used to illustrate and establish his theoretical abstraction – his theory of evolution of species by natural selection.

One can describe things using natural language, or pictures, but these are inadequate for dynamic and complex phenomena, where the essence of what is being described is how several mechanisms might relate over time. An agent-based simulation framework allows for a direct representation (one agent for one actor) without theoretical restrictions. It allows for dynamic situations as well as complex sets of entities and interactions to be represented (as needed). This can make it an ideal complement to scenario development because it ensures consistency between all the elements and the outcomes. It is also a good base for future generalisations when the author can access a set of such descriptive simulations.

Definition

> *A description (using a simulation) is an attempt to* partially represent *what is important of* a specific observed case *(or small set of closely related cases).*

Unpacking some of this:

- This is not an attempt to produce a 1-1 representation of what is being observed but only of the features thought to be relevant for the intended kind of study. It will leave out some features, however a description tends to err on the side of including aspects rather than excluding them (for simplicity, communication etc.).
- It is not in any sense general, but seeks to capture a restricted set of cases – it is specific to these and no kind of generality beyond these can be assumed.
- The simulation has to relate in an explicit and well-documented way to a set of evidence, experiences and data. This is the opposite of theoretical exposition and should have a direct and immediate connection with observation, data or experience.
- This kind of purpose may involve the integration, within a single picture, of many different kinds of evidence, including: qualitative, expert opinion, social network data, survey data, geographic data, and longitudinal data (Edmonds .

Examples

In 1993, François Bousquet and colleagues (Bousquet & al. 1993) developed a simulation[15] whose purpose was described as follows:

> *"We have developed a simulator to represent both social, economic and ecological knowledge to contribute to a synthesis of the multidisciplinary knowledge … As a result of the simulations, focus can be put on the relation between space sharing rules and the evolution of the ecological equilibrium. The simulator is considered as a discussion tool to lead to interdisciplinary meetings."*

Thus this is not designed to explain or predict but to represent and formalise, aiming to formalise some of the relationships and dynamics of the situation.

In (Moss 1998), Scott Moss describes a model that captures some of the interactions in a water pumping station during crises. This came about through extensive discussions with stakeholders within a UK water company about what happens in particular situations during such crises. The model sought to directly reflect this evidence within the dynamic form of a simulation, including cognitive agents who interact to resolve the crisis. This simulation captured aspects of the physical situation, but also tackled some of the cognitive and communicative aspects. To do this, he had represent the problem solving and learning of key actors, so he inevitably had to use some existing theories and structures – namely, Alan Newell and Herbert Simon's general problem solving architecture (Newell & Simon 1972) and Cohen's endorsement mechanism (Cohen 1984). However, this is all made admirably explicit in the paper. The paper is suitably cautious in terms of any conclusions, saying that the simulation *indicate[s] a clear need for an investigation of appropriate organizational structures and procedures to deal with full-blown crises.*

Risks

Any system for representation will have its own affordances – it will be able to capture some kinds of aspect much more easily than others will. This inevitably biases the representations produced, as those elements that are easy to represent are more likely to be captured than those which are more difficult. Thus, the medium will influence what is captured and what is not.

Since agent-based simulation is not theoretically constrained[16], there are a large number of ways in which any observed phenomena could be expressed in terms of simulation code. Thus, it is almost inevitable that any modeller will use some structures or mechanisms that they are familiar with in order to write the code. Such a simulation is, in effect, an *abduction* with respect to these underlying structures and mechanisms – the phenomena are seen through these and expressed using them.

Finally, a reader of the simulation may not understand the limitations of the simulation and make false assumptions as to its generality. In particular, the inference within the simulations may not include all the processes that are in what is observed – thus it cannot be relied upon to either predict outcomes or justify any specific explanation of those outcomes.

Mitigating Measures

As long as the limitations of the description (in terms of its selectivity, inference and biases) are made clear, there are relatively few risks here, since not much is being claimed. If it is going to be useful in the future as part of a (slightly abstracted) evidence base, then its limitations and biases do need to be explicit. The data, evidence or experience it is based upon also needs to be made clear. Thus, good documentation is the key here – one does not know how any particular description will be used in the future so the thoroughness of this is key to its future utility. Here it does not matter if the evidence is used to specify the simulation or to check it afterwards in terms of the outcomes, all that matters is that the way it relates to evidence is well documented. Standards for documentations (such as the ODD and its various extensions (Grimm & al. 2006, 2010) help ensure that all aspects are covered.

Theoretical exposition

Motivation

If one has a mathematical model, one can do analysis upon its mathematics to understand its general properties. This kind of analysis is both easier and harder with a simulation model – to find out the properties of simulation code one just has to run the code – but this just gives one possible outcome from one set of initial parameters. Thus, there is the problem that the runs one sees might not be representative of the behaviour in general. With complex systems, it is not easy to understand how the outcomes arise, even when one knows the full and correct specification of their processes, so simply knowing the code is not enough. Thus with highly complicated processes, where the human mind cannot keep track of the parts unaided, one has the problem of understanding how these processes unfold in general.

Where mathematical analysis is not possible, one has to explore the theoretical properties using simulation – this is the goal of this kind of model. Of course, with many kinds of simulation one wants to understand how its mechanisms work, but here this is the only goal. Thus, this purpose could be seen as more limited than the others, since some level of understanding the mechanisms is necessary for the other purposes (except maybe black-box predictive models). However, with this focus on just the mechanisms, there is an expectation that a more thorough exploration will be performed – *how* these mechanisms interact and *when* they produce different kinds of outcome.

Thus the purpose here is to give some more general idea of how a set of mechanisms work, so that modellers can understand them better when used in models for other purposes. If the mechanisms and exploration is limited this would greatly reduce the usefulness of doing this. General insights are what is wanted here.

In practice, this means a mixture of inspection of data coming from the simulation, experiments and maybe some inference upon or checking of the mechanisms. In scientific terms, one makes a hypothesis about the working of the simulation – why some kinds of outcome occur in a given range of conditions – and then tests that hypothesis using well-directed simulation experiments.

The complete set of simulation outcomes over all possible initialisations (including random seeds) does encode the complete behaviour of simulation, but that is too vast and detailed to be comprehensible. Thus some general truths covering the important aspects of the outcomes under a given range of conditions is necessary – the complete and certain generality established by mathematical analysis might be infeasible with many complex systems but we would like something that approximates this using simulation experiments.

Definition

> *'Theoretical exposition' means discovering then* establishing (or refuting) hypotheses *about the* general behaviour *of a set of mechanisms (using a simulation).*

Unpacking some key aspects here.
- One may well spend some time illustrating the discovered hypothesis (especially if it is novel or surprising), followed by a sensitivity analysis, but the crucial part is showing these hypotheses are refuted or not by a sequence of simulation experiments.
- The hypotheses need to be (at least somewhat) general to be useful.
- A use of theoretical exposition can be to refute a hypothesis, by exhibiting a concrete counter-example, or to establish a hypothesis.
- Although any simulation has to have some meaning for it to be a model (otherwise it would just be some arbitrary code) this does not involve any other relationship with the observed world in terms of data or evidence.

Example

Deffuant et al. (20002) study the behaviour of a class of continuous opinion dynamic models. This paper does multiple runs, mapping the outcomes over the parameter space using some key indicators (e.g. relative mean influence of 'extremist agents'). From these it summarises the overall behaviour of this class in qualitative terms, dividing the outcomes into an number of kinds of category. Later Deffuant and Weisbuch (2008) use probability distribution models to try and get a study the behaviour of a class of continuous opinion dynamic models, approximating the agent-based models using equations. These analytic models where then tested against the agent-based models. Although the work is vaguely motivated with reference to observed phenomena, the described research is to understand the overall behaviour of a class of models.

Risks

In theoretical exposition one is not relating simulations to the observed world, so it is fundamentally an easier and 'safer' activity[17]. Since a near complete understanding of the simulation behaviour is desired this activity is usually concerned with relatively simple models. However, there are still risks – it is still easy to fool oneself with one's own model. Thus the main risk is that there is a bug in the code, so that what one thinks one is establishing about a set of mechanisms is really about a different set of mechanisms (i.e. those including the bug).

A second area of risk lies in a potential lack of generality, or 'brittleness' of what is established. If the hypothesis is true, but only holds under very special circumstances, then this reduces the usefulness of the hypothesis in terms of understanding the simulation behaviour.

Lastly, there is the risk of over-interpreting the results in terms of saying anything about the observed world. The model might suggest a hypothesis about the observed world, but it does not provide any level of empirical support for this.

Mitigating Measures

The measures that should be taken for this purpose are quite general and maybe best understood by the community of simulators.

- One needs to check ones code thoroughly – see (Izquierdo & al 2017) for a review of techniques).
- One needs to be precise about the code and its documentation – the code should be made publically available
- Be clear as to the nature and scope of the hypotheses established.
- A very thorough sensitivity check, trying various versions with extra noise added, testing for extreme conditions etc.
- It is good practice to illustrate the simulation so that the readers understand its key behaviours but then follow this with a series of attempted refutations of the hypotheses about its behaviour to show its robustness.
- Be very careful about not claiming that this says anything about the observed world.

Illustration

Motivation

Sometimes one wants to make an idea clear, and an illustration is a good way of doing this. It makes a more abstract theory or explanation clear by exhibiting a concrete example that might be more readily comprehended. Complex systems, especially complex social phenomena, can be difficult to describe, including multiple independent and interacting mechanisms and entities. A well-crafted simulation can help people see these complex interactions at work and hence appreciate these complexities better. As with description, this purpose does not claim a great deal. If the theory is already instantiated as a simulation (e.g. for theoretical exposition or explanation) then the illustrative simulation might well be a simplified version of this.

Illustration goes to the heart of the power of formal modelling, since it is the instantiation of a set of ideas into a inspectable example. This example can then be used to make precise the meaning of the terms used to express that idea. It also shows the possibility of the process being illustrated (but nothing about its likelihood). One powerful use of an illustration is as a counter-example to a set of assumptions.

Definition

> *An illustration (using a simulation) is to communicate or make clear an idea, theory or explanation.*

Unpacking this.
- Here the simulation does not have to fully express what it is illustrating, it is sufficient that it gives a simplified example. So it may not do more than partially capture the idea, theory or explanation that it illustrates, and it cannot be relied upon for the inference of outcomes from any initial conditions or set-up.
- The clarity of the illustration is of over-riding importance here, not its veracity or completeness.
- An illustration should not make any claims, even of being a description. If it is going to be claimed that it is useful as a theoretical exposition, explanation or other purpose then it should be justified using those criteria – that it seems clear to the modeller is not enough.

Example

Schelling developed his famous model for a theoretical purpose. He was advising the Chicago district on what might be done about the high levels of segregation there. The assumption was that the sharp segregation observed must be a result of strong racial discrimination by its inhabitants. Schelling's model (Schelling 1969, 1971) showed that segregation could result from just weak preferences of inhabitants for their own kind – that even, a wish for 30% of people of the same trait living in the neighbourhood could result in segregation. This was not obvious without building a model. The model was a clear counter example to the assumption.

What the model did not do is say anything about what actually caused the segregation in Chicago – it might well be the result of strong racial prejudice. The model did not predict anything about the level of segregation, nor did it explain it. All it did was provide a counter-example to the current theories as to the cause of the segregation, showing that this was not *necessarily* the case.

Risks

The main risk here is that you might deceive people using the illustration into reading more into the simulation than is intended. There is also a risk of confusion if it is not clear which aspects are important to the illustration and which are not. A simulation for illustration will show the intended behaviour, but (unlike when its theory is being explored) it has been tested only for a restricted range of possibilities, indeed the claimed results might be quite brittle to insignificant changes in assumption.

Mitigating Measures

Be very clear in the documentation that the purpose of the simulation is for illustration only, maybe giving pointers to fuller simulations that might be useful for other purposes. Also be clear in precisely what idea is being communicated, and so which aspects of the simulation are relevant for this purpose.

Analogy

Motivation

Analogy is a powerful way of thinking about things. Roughly, it applies ideas or a structure from one domain and projects it onto another (Hofstader 1995). This can be especially useful for

thinking about new or unfamiliar phenomena or as a guide to the direction of more rigorous thought. Sometimes it can result in new insights that are later established in a more rigorous way. There is some evidence that analogical thinking is very deeply entrenched in the whole way we think and communicate (Lakoff 2008). Almost anything can be used as an analogy, including simulation models.

Playing about with simulations in a creative but informal manner can be very useful in terms of informing the intuitions of a researcher. In a sense, the simulation has illustrated an idea to its creator. One might then exhibit a version of this simulation to help communicate this idea to others. However, this does not mean that the simulation achieves any of the other purposes described above, and it is thus doubtful whether that idea has been established to be of public value (justifying its communication in a publication) until this happens. This is not to suggest that analogical thinking is not an important process in science. Providing new ways of thinking about complex mechanisms or giving us new examples to consider is a very valuable activity. However, this does not imply its adequacy for any other purpose.

Definition

> *An analogy (using a simulation) is when processes illustrated by a simulation are used as a way of thinking about something in an informal manner.*

Figure 1, Figure 2 and Figure 3 illustrate the difference between common-sense, scientific (i.e. empirical) and analogical understanding. The difference between empirical and analogical relationships is that in the latter there is no direct relationship between the analogy (in this case the simulation) and the evidence, but only indirectly through one's intuitive understanding. A empirical relationship may be staged (model⇔data⇔observations) but these are well-defined and testable.
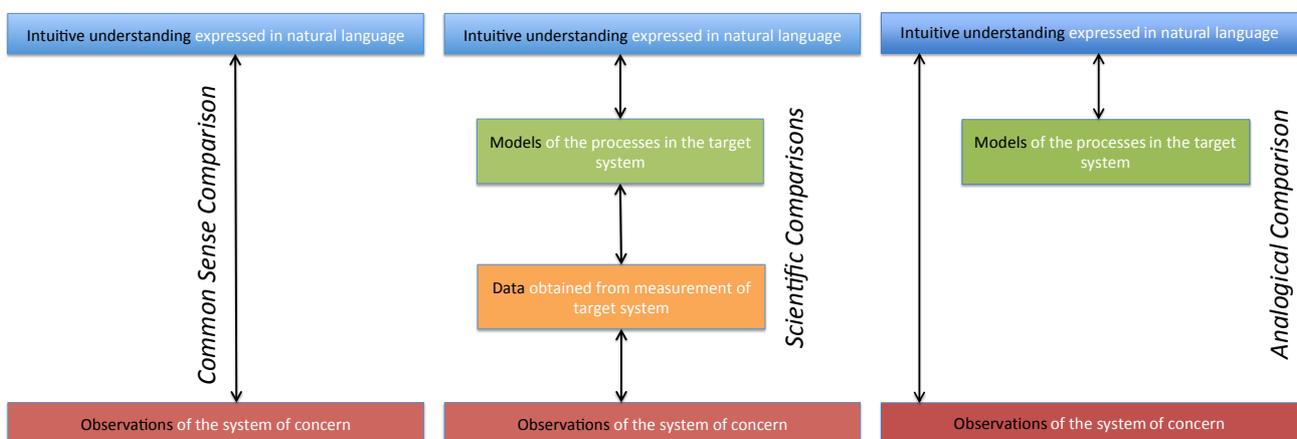

Figure 1. Common-Sense   Figure 2. Scientific   Figure 3. Analogical

It is not fully understood what the mind does when using an analogy to think, but does involve applying the structure of some understanding from one domain and applies it to another. Humans have an innate ability to think using analogies and often do it unconsciously. Here the relationship between simulation and the observed is re-imagined for each domain in which it is applied[18] – thus this relationship is not well defined but flexible and creative.

The modelling purposes of analogy and illustration are very similar and often overlap. A simulation that illustrates an idea can then be used as an analogy (applying that idea to a new domain). However, the purposes are subtly different – illustration is to make an idea clear and give a specific example of it; an analogy is when an idea is applied to other domains in an informal manner. Thus, there are substantive differences: for example, an illustrated idea might be then tested in an explanatory model in a well-defined manner against some data, or

the workings of an analogy go well beyond that required of an illustration in terms of mapping many possibilities but without any well-defined relationship with the real world.

Specifically:

- An analogy might not be clear or well defined but an illustration should be,
- An illustration can be a single instance or example but an analogy might cover a whole class of processes or phenomena (albeit informally),
- A good analogy suggests a useful way of thinking about things, maybe producing new insights, a good illustration does not have to do this,
- An illustration can be used as a counter example to some widely-held assumptions, an analogy can not be so used.

Example

In his book, Axelrod (1984) describes a formalised computational 'game' where different strategies are pitted against each other, playing the iterated prisoner's dilemma. Some different scenarios are described, where it is shown how the tit for tat strategy can survive against many other mixes of strategies (static or evolving). The conclusions are supported by some simple mathematical considerations, but the model and its consequences were not explored in any widespread manner[19]. In the book, the purpose of the model is to open up a new way of thinking about the evolution of cooperation. The book claims the idea 'explains' many observed phenomena[1], but in an analogical manner – no precise relationship with any observed measurements is described. There is no validation of the model here or in the more academic paper that described these results (Axelrod & Hamilton 1981). In the academic paper there are some mathematical arguments which show the plausibility of the model but the paper, like the book, progresses by showing the idea is coherent with some reported phenomena – but it is the ideas rather than the model that are so related. Thus, in this case, the simulation model is an analogy to support the idea, which is related to evidence in a qualitative manner – the relationship of the model to evidence is indirect (Edmonds, 2001). Thus the role of the simulation model is that of a way of thinking about the processes concerned and does not qualify for either explaining specific data, predicting anything unknown or exploring a theory.

Risks

These simulations are just a way of thinking about other phenomena. However, just because you can think about some phenomena in a particular way does not make it true. The human mind is good at creating, 'on the fly', connections between an analogy and what it is considering – so good that it does it almost without us being aware of this process. The danger here is of confusing being able to think of some phenomena using an idea, and that idea having any force in terms of a possible explanation or method of prediction. The apparent generality of an analogy tends to dissipate when one tries to precisely specify the relationship of a model to observations, since an analogy has a different set of relationships for each situation it is applied to – it is a supremely flexible way of thinking. This flexibility means that it does not work well to support an explanation or predict well, since both of these necessitate an explicit and fixed relationship with observed data.

Mitigating Measures

The obvious mitigation is to simply make clear when one is using a model as an analogy and when a rigorous relationship with evidence is intended. Analogies can be a very productive way of thinking as long as one does not infer anything more from them – analogies can be very misleading.

---

[1] For example, the spontaneous outbreaks of non hostility across the trenches in World War I.

Social Interaction

Motivation

Sometimes our simulations are not about the observed world at all, but (at least substantially) designed to reflect an actor's view of that world. Such simulations can be developed collectively by a group of people. In this case, the model can act as a mediator between members of the group and can result in a shared understanding of the world (or at least a clear idea where the members agree and differ). This shared understanding is explicitly encapsulated in the model for all to discuss, avoiding the misunderstandings that can come from more abstract discussion.

When the participatory dimension of the modelling process is to the fore, the overriding purpose can be referred to promoting social learning, i.e. increasing the knowledge of individuals through social interactions and processes within a social network and favouring the acquisition of collective skills (Reed et al., 2010). To build and use models in a collaborative way is here a means for: generating a shared level of information among participants, to create common knowledge, explore common goals, and understand the views, interests, and rationale of opposing parties about the target system. Approaches, such as mediated modelling (Van den Belt, 2004) and companion modelling (Barreteau et al., 2003; Etienne, 2014), have been applied to pursuing this goal in the domain of environmental sciences, and more specifically to tackle problems related to the collective management of renewable resources. As pointed out by Elinor Ostrom (2009), when users of a given socio-ecological system (SES) share common knowledge of relevant SES attributes as well as how their actions affect each other, they can more easily self-organise. There is here a radical shift from the purpose of using models for prediction, related to the command-and-control posture seeking an optimal solution, to the purpose of learning (Brugnach, 2010), related to the concept of adaptive co-management which leitmotiv is *learning to manage and managing to learn* (Olsson et al., 2004).

Definition

> *A simulation is a tool for social learning when it encapsulates a shared understanding (or set of understandings) of a group of people.*

This is distinct from when one uses people as a source of evidence in the form of their expert or personal opinion for a descriptive of explanatory model. In this case, although the source is human and might be subjectively biased, the purpose is to produce an objectively justifiable model. Although a model produced as part of a participatory Social Interaction process might have other uses, these are not what such a process is designed for.

Developing models to support social learning in the domain of environmental management is a kind of participatory approach that is very time-consuming (Barreteau et al., 2013). This is related to the wickedness of problems such as those that are involved in environmental management, which are characterized by a high degree of scientific uncertainty and a profound lack of agreement on values (Balint et al., 2011). To confront such complex policy dilemmas requires developing learning networks (Stubbs and Lemon 2001) of stakeholders to create a cooperative decision-making environment in which trust, understanding, and mutual reliance develop over time.

Scientists are part of the group of people. To make the most of collaboration between them and other stakeholders, their role needs to be redefined toward scientists as stakeholders and participants, as has been suggested by Ozawa (1991), rather than scientists as an objective third party.

Using models to promote social learning implies a need recontextualize uncertainty in a broader way — that is, relative to its role, meaning, and relationship with participants in decision making (Brugnach et al., 2008). Instead of trying to eliminate or reduce uncertainties by collecting more data, the post-modern view of science (Funtowicz and Ravetz, 1993)

suggests that we should try to illuminate them (one of the purpose mentioned by Epstein in his 2008 paper), for their better recognition and acceptance by stakeholders then enabled to make informed collective decisions.

Ambiguity is a distinct type of uncertainty that results from the simultaneous presence of multiple valid, and sometimes conflicting, ways of framing a problem. As such, it reflects discrepancies in meanings and interpretations. Under the presence of ambiguity it is not clear what problem is to be solved, who should be involved in the decision processes or what is an appropriate course of action (Brugnach and Ingram 2012). The co-design of a model may be a way to explore ambiguity among stakeholders through a constructivist approach (Salliou et al, 2017).

Example

To mitigate the looming conflict in land use between government agencies seeking to rehabilitate the forest cover in upper watersheds of northern Thailand and smallholders belonging to ethnic minorities exploiting grasslands/fallows with an extensive cattle rearing systems a participatory modelling approach was applied. This involved an iterative series of collaborative agent-based simulation activities was implemented in Nan Province (Dumrongrojwatthana et al., 2011). The exchange and integration of local empirical knowledge on the vegetation dynamics and extensive cattle rearing system was fostered by the co-design of the interactive simulation tools used to explore innovative land use options consistent with the dual interests of the parties in conflict. Herders requested the researchers to modify the model to explore new cattle management techniques, especially rotations of 'ruzi grass' (*Brachiaria ruziziensis*) pastures. These collaborative modelling activities, implemented with a range local stakeholders (some of whom initially refused to even meet the herders), established a communication channel between herders, foresters and rangers. The dialogue led to an improved mutual understanding of their respective perceptions of land-use dynamics, objectives and practices. The improvement of trust between the villagers and the forest conservation agencies was also noticeable and translated into the design of a joint field experiment on artificial pastures.

Risks

Finding the right way to engage stakeholders in co-designing a model of a complex system is tricky. Using easily accessible tools like role-playing games (Bouquet et al., 2002) can be better suited to grassroots people than for policy-makers (who tend to discard them as not serious enough). On the other hand, using conceptual designs and computer implemented simulations with participants who did not receive any formal education may prove to be difficult. Beyond the risk of not selecting appropriate tools to support the co-design. For example, there is also a risk that the models remain too generic and too abstract to get the interest of participants concerned by real-life problems. On the other hand if one sticks to some particularities of a specific situation without stepping back sufficiently may not be sufficient to facilitate constructive discussions.

Without paying attention to power asymmetries, there is a risk that the most powerful stakeholders will have greater influence on the outcomes of the participatory modelling process than marginalized stakeholders.

In the same line, when there is an initial demand expressed by one of the stakeholders, there is a risk that the modellers bias the implementation of the process to fit their own perspectives (manipulation).

Mitigating Measures

Barnaud and Van Paassen (2013) advocates a critical companion posture, which strategically deals with power asymmetries to avoid increasing initial power imbalances. The goal of this is to shift more power from the modellers to the stakeholders. This posture suggests that designers of participatory modelling processes, intending to mitigate conflicts ,should make

explicit their assumptions and objectives regarding the social context so that local stakeholders can choose to accept them as legitimate or to reject them.

Some Confusions of Purpose

It should be abundantly clear by now, that establishing a simulation for one purpose does not justify it for another, and that any assumptions to the contrary risk confusion and unreliable science. However, the field has many examples of such confusions and conflations, so I will discuss this a bit further. It is true that a simulation model justified for one purpose might be used as *part* of the development of a simulation model for another purpose – this can be how science progresses. However, just because a model for one purpose *suggests* a model for another, does not mean it *is* a good model for that new purpose. If it is being suggested that a model can be used for a new purpose, it should be separately justified for this new purpose. To drive home this point further, we look at some common confusions of purpose. Each time some code is mistakenly relied upon for a purpose other than has been established for it.

1. ***Theoretical exposition → Explanation***. Once one has immersed oneself in a model, there is a danger that the world looks like this model to its author. This is a strong kind of Kuhn's 'Theoretical Spectacles'[20], and results from the intimate relationship that simulation developers have with their model. Here the temptation is to jump from a theoretical exposition, which has no empirical basis, to an explanation of something in the world. A simulation can provide a way of looking at some phenomena, but just because one can view some phenomena in a particular way does not make it a good explanation. Of course, one can form a hypothesis from anywhere, including from a theoretical exposition, but it remains only a hypothesis until it is established as a good explanation as discussed above (which would almost certainly involve changing the model).

2. ***Description → Explanation***. In constructing a simulation for the purpose of describing a small set of observed cases, one has deliberately made many connections between aspects of the simulation and evidence of various kinds. Thus, one can be fairly certain that, at least, some of its aspects are realistic. Some of this fitting to evidence might be in the form of comparing the outcomes of the simulation to data, in which case it is tempting to suggest that the simulation supports an explanation of those outcomes. The trouble with this is twofold: (a) the work to test *which* aspects of that simulation are relevant to the aspects being explained has not been done; (b) the simulation has not been established against a range of cases – it is not general enough to make a good explanation. An explanation that only explains aspects of a small number of cases using a complex simulation is a bad explanation since there will be many other potentialities in the simulation that are not used for these few cases.

3. ***Explanation → Prediction***. A simulation that establishes an explanation traces a (complex) set of causal steps from the simulation set-up to outcomes that compare well with observed data. It is thus tempting to suggest that one can use this simulation to predict this observed data. However, the process of using a simulation to establish and understand an explanation inevitably involves iteration between the data being explained and the model specification – that is the model is fitted to that particular set of data. Model fitting is not a good way to construct a model useful for prediction, since it does not distinguish between what is essential for the prediction and the 'noise' (what cannot be predicted). Establishing that a simulation is good for prediction requires its testing against unknown data several times – this goes way beyond what is needed to establish a candidate explanation for some phenomena. This is especially true for social systems, where we often cannot predict events, but we can explain them after they have occurred.

4. ***Illustration → Theoretical exposition***. A neat illustration of an idea suggests a mechanism. Thus, the temptation is to use a model designed as an illustration or playful exploration as

being sufficient for the purpose of a theoretical exposition. A theoretical exposition involves the extensive testing of code to check the behaviour and the assumptions therein, an illustration, however suggestive, is not that rigorous. For example, it may be that an illustrated process is a very special case and only appears under very particular circumstances, or it may be that the outcomes were due to aspects of the simulation that were thought to be unimportant (such as the nature of a random number generator). The work to rule out these kinds of possibility is what differentiates using a simulation as an illustration from a theoretical exposition.

5. ***Illustration → Prediction***. 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, illustrating how this 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 prediction and the model criticised on these grounds[21]. Dennis Meadows did not stop asserting that the purpose was only predictive.. As it says in the book (p.92-92) "*... in the most limited sense of the word. These graphs are not exact predictions of the values of the variables at any particular year in the future. They are indications of the system's behavioural tendencies only*" (Meadows et al. 1972 pp. 92-93). However the book also often conflates talk about the variables in their model and what those variables represent, e.g. "*Our world model was built specifically to investigate five major trends of global concern-accelerating industrialization, rapid population growth, widespread malnutrition, of non-renewable resources, and a deteriorating environment.*" (p. 21). It does claim that the model was not purely theoretical, since it says "*We feel that the model described here is already sufficiently developed to be of some use to decision-makers*", however the way in which it would be use to decision-makers is not clear. This was an illustration of a possible future and not a prediction[22].

6. ***Social learning → Prediction.*** Having being involved into a companion modelling process implemented in the National Park of Cevennes in South of France (Etienne et al., 2003), a group of farmers, who had been sensitized to the process of pine encroachment of their pastures through participating to role-playing game sessions, decided to engage into a incentivized collective action plan. To specify the management plan, they requested to use as a decision support tool the agent-based model that had been designed to support the role-playing game sessions. The model was clearly not meant for such a use. A new version of the model was then designed, with a refined spatial resolution (1 Ha instead of 4 Ha).

There is a natural progression in terms of purpose attempted as understanding develops: from illustration to description or theoretical exposition, from description to explanations and from explanations to prediction. However, each stage requires its *own* justification and probably a complete re-working of the simulation code for this new purpose. It is the lazy assumption that one purpose naturally or inevitably follows from another that is the danger.

Modelling Strategies

As well as a declared purpose for a model, most modellers will also employ a number of *modelling strategies* to guide the model development to achieve this purpose. A modelling strategy is a 'rule of thumb' that can be used as a guide as to *how* to develop the model. It is an informal meta-model for the modelling processes, a story to help the modellers cut down the overwhelming number of choices presented to them down to something more manageable. A modelling strategy is about the process of model *development*, the modelling purpose is about its *justification*. This is essentially the same as the classic distinction between in the philosophy of science between the 'context of discovery' and the 'context of justification' (????). Since, usually, we are not privy to the process of model development but only the justification of models we often do not know which strategies people are using[23].

Some of these strategies are as follows.
- **Keep it Simple Stupid (KISS)**. This is the strategy of starting as simple as one can get away with and then only adding complication when this is shown to be needed. Thus one starts with relatively simple models and develops from there. This is a well established principle in engineering since it maintains the maximum control and understanding of the system as it is develops, and hence reduces the possibility of making mistakes.
- **Mimimise the Number of Parameters (MNP)**. This is where one has as few parameters for the model as possible. This is slightly different from KISS both in terms of purpose and practice. Under MNP you can have complex processes as long as the number of parameters is small because the point is to reduce the space of output behaviours. A random number generator may be implemented using a complex process, but has only one parameter (its seed). A linear equation with many variables is simple but has many parameters.
- **Mimimise the Number of Free Parameters (MNFP)**. A free parameter is one that is not determinable using measurement from the target of modelling, even in principle. These might be scaling factors (e.g. relating different time scales) or represent unknown aspects (e.g. the parameters of some learning algorithm that an agent uses). Having fewer free parameters makes observed data difficult to 'fit' just by adjusting parameters. Free parameters are sometimes estimated by fitting a model to in-sample data, but then one is not clear which aspects of that data are being used and that in-sample data becomes (in effect) a huge set of parameters.
- **Keep it Descriptive Stupid (KIDS)**. This is the strategy of starting from the evidence[24] – if there is evidence for some aspect of the phenomena of interest then you include this in your model by default. One then starts the model exploration from this point, possibly trying the model without some of the elements and seeing if it makes any significance difference to the model behaviour. However, as with the KISS strategy, one might equally find one needs even more detail. This is described in (Edmonds & Moss 2005).
- **Make the Model More Like the Phenomena (MTMMLTP[25])**. This is the strategy of making the model more like how we think things are. In other words to use the correspondence view of modelling (discussed at the beginning) as a guide to model development. We suspect that most of us use this strategy to some extent, whether we are aware of it or not. This, we claim, is the source of much misunderstanding – just because we *think* of a model as being a kind of simularculum of reality does not *make it* good for all purposes (or indeed good for any particular purpose).
- **Keep It a Learning Tool (KILT)**. This is the strategy of maximising the relevance and accessibility of a model so that participants can engage with the model and be willing to help specify or improve it (Le Page and Perrotton, 2017). This is somewhat different to the other strategies here, since it is so tied in with the purpose of Social Interaction. Social Interaction is a case where the process of model building is its point, and not its subsequent justification.

These strategies are not something that can ever be proved or disproved, but rather only shown as helpful or otherwise during the processes of modelling – especially the extent to which a model's purpose is achieved and demonstrated using them. To some extent, these are a matter of taste and personal style, but I argue that they are also differently relevant to different modelling purposes. I consider each purpose in turn.

Prediction

With prediction it actually does not matter *why* your model predicts, just that it does. However, one does need some understanding of the code, because in practice the only way to find a predictive model is a process of trial and error, adapting the code until it predicts – if one does not understand the code it is hard to know how to adapt this to improve it for the next trial. Thus KISS plays a role here. What is more important to this purpose is (a) being able to set the parameters correctly for a case where one wants to predict and (b) that the sensitivity of the

key outcomes to those parameters can be established so that one knows to the likely accuracy of the prediction. For (a), MNFP is important so the model can be set up correctly for a prediction. For (b), MNP is important so that the sensitivity (including joint sensitivity) of the predictions to the parameters can be established using a Monte Carlo approach (or similar). If there is very reliable evidence for key parts of the process one is simulating (e.g. the electoral college method of electing a president of the US) then it might be sensible to fix that in your code. Thus KIDS does have a small role here. MTMMLTP is not relevant here and could result from adding more parameters which would be counter-productive.

Explanation

When explaining some measured phenomena using a model, one does so in terms of the processes and structures that are built into the model. If those process and structures are not strongly related to those observed one gets an explanation in terms that are not so relevant to what is being explained. Rather for explanation, it is vital that those aspects of the phenomena that we want the explanation to be in terms of, are in the model. Explanation differs from that of analogy because it has a well-defined relationship between the model and evidence of what is observed, so that the KIDS strategy is very important here. It is desirable to have few free parameters, as otherwise there is a danger that the 'fit' to observed outcomes was due to these having special values, so MNFP plays a role. Social phenomena is often quite complex so explanatory models are often complex (hence KISS is not relevant) and do often have many parameters, but these should be based on evidence so there should not be many of these, so that MNP is not so relevant here. MTMMLTP might lead one to include aspects for which there is no evidence, which would increase the number of free parameters, so MTMMLTP could also be counter-productive here.

Description

In a description you are just trying to record the observed aspects of some phenomena for latter use and inspection. Ideally you should not be projecting what you 'think' is there but focussing on what is observed – i.e. what there is evidence for. Thus it is KIDS and not MTMMLTP that is more relevant here. The number of parameters (MNP) is not important, but having lots of free parameters is an indication you are including unobserved aspects into your description, so MNFP is. A description should be as simple or as complex as what is being described so KISS is not relevant here either.

Theoretical exposition

When trying to systematically characterise the behaviour of model, there are essentially two techniques:
1. to systematically map out the space of behaviours by running the model many times
2. to approximate the behaviour with a simpler model (e.g. an analytic model, a social network analysis or a simpler simulation)

The *first* of these becomes more difficult with an increasing number of parameters that need to be mapped. Thus MNP is core here. Whether the parameters are free or not is not really relevant to a theoretical exposition (so MNFP is not very important here). The second of these is harder the more complicated and complex the original model is, thus KISS is also key. Although theoreticians may have an eye to the eventual applicability of their work, the relationship of their model to what is observed is not important. Whilst, evidence from observed systems might be a helpful guide to whether one is on the right track, it is well known that trying to add too many features just because that is how one thinks of a system can be fatal to a good exposition (thus MTMMLTP is unhelpful here).

Illustration

A good illustration is as simple as possible to capture what is needed. Thus KISS has to be the over-riding principle here and it probably does not have many parameters. It can help if the

illustration is plausible (MTMMLTP). An illustration is often of an idea rather than an observed process, so KIDS does not usually come into this.

Analogy

A good analogy is relatively simple (KISS) but is flexible as to its relationship with what is observed. Indeed some of the point of an analogy is that it can be applied to completely unforeseen situations (so not KIDS). It helps if it is plausible (MTMMMLTP). Since analogies are often used in Social Interaction KILT is relevant here (although we mediate so naturally using analogies that there is less need for a 'science' of this).

Social Interaction

Regarding the suitability of the tools and the degree of model complication (Sun et al., 2016), Le Page and Perrotton (2017) advocates using an intermediate level of stylized but empirically based models and proposed the motto Keep It a Learning Tool (KILT) to stress the importance of sticking to the announced purpose of promoting social learning among participants. The KISS approach consists in initiating the process with an over-simplified stylized yet empirically grounded model that enables tackling the complexity of the target socio-ecosystem with a tool that has the status of a sketch. It provides the main features of the final version; however, it is clearly unfinished: there remains an important work of progressive shaping and improvement so that it acquires its final form and becomes usable with people who were not involved in its design. The crudeness of the initial version of the model is actually a strategy to stimulate the involvement of the participants targeted as potential co-designers.

This strategy includes elements of MTMMLTP since aspects of what is being represented in the model (those assumed by all those involved) need to be recognisable by the stakeholders, especially if they structure the more contentious elements (e.g. geographical layout, or seasons of the year) – there does not *have* to be evidence for these since it is the perception of participants that counts, but there might be. The processes involved do need to be comprehensible enough to be understood and criticisable by the participants, so KISS does play a role here but not a central one. On the other hand, having many parameters (free or otherwise) can be an advantage since these might be adjusted to enhance the plausibility or 'realism' of a simulation, so MNP is not helpful here.

Summary of Strategies

Table 1. Relevance of strategies for the development of models with various purposes

Modelling Purpose	Important	Can be Helpful	Possibly Unhelpful
Prediction	MNFP, MNP	KISS, KIDS	MTMMLTP
Explanation	KIDS	MNFP	MTMMLTP
Description	KIDS	MTMMLTP, MNFP	KISS
Theoretical	MNP, KISS	KIDS	MTMMLTP
Illustration	KISS	MTMMLTP, MNP	
Analogy	KISS	MTMMLTP, KILT	KIDS
Social Interaction	KILT	MTMMLTP, KISS	MNP

Table 1 summarises the above discussion in terms of the rough importance of the various strategies for achieving the various model purposes. This is **not** definitive, since strategy is partly a matter of taste – a rule of thumb that maybe be helpful.

Using Models with Different Purposes Together

Clearly one sometimes needs to be able to achieve more than one of the above purposes with respect to some phenomena. In such circumstances one might be tempted to imagine a model that mirrors reality so well that it can be used for many purposes (e.g. to both predict and explain). In the above I have argued that, even if that were achievable, you would need to justify the model for each purpose separately.

However, I *further* suggest that this is usually better done with separate, but related models. The reason for this should be have become clear during the discussion as to what modelling strategies might be helpful to achieve each purpose – each purpose had different concerns and hence different strategies were helpful. Even if one had a model of something as simple as an ideal gas, with a different entity for each gas particle, it would not be a very practical model for prediction. Rather one would probably use the gas laws that connect volume, temperature and pressure that approximate this individual-based model.

Cartwright (1983) argues that, in physics, there is a distinction between explanatory and phenomenological laws. That is, in the above language, the models that explain and the models that predict are different. However, in a mature science, the relationship between these different models is well understood. Cartwright describes some examples where an explanatory theory is used as the starting point for generating a number of alternative predictive models – the predictions only trusted when the alternate routes to that prediction roughly agree.

In the social sciences, the phenomena is more complex than those usually tackled in physics. For this reason we should expect there to be many different models that achieve different purposes in respect to some phenomena, different versions of these models for different assumptions one might make, and maybe different versions to address different aspects of that phenomena. Thus we might expect that a whole cluster of different models might be used to get a better picture of any particular phenomena, each with different purposes. The idea that we might have one model that is adequate for one target is probably infeasible.

An example of using different models for different purposes is when models of different levels of abstraction are used to 'stage' abstraction. Following a KIDS kind-of strategy, one can start with a relatively complex or descriptive model and then seek to approximate this with simpler models that one can get a better theoretical understanding of. This was tried in (Fieldhouse et al. 2016, Lafuerza et al. 20161, 2016b).

By keeping models for different purposes separately but defining/describing the relationship between them, we can get many of the advantages of combining models but with less chance of confusion of purpose.

Conclusion

In Table 2 we summarise the most important points of the above distinctions. This does not include all the risks of each kind of model, but simply picks the most pertinent ones.

Table 2. A brief summary of the discussed modelling purposes

Modelling Purpose	Essential features	Particular risks (apart from that of lacking the essential features)
Prediction	Anticipates *unknown* data	Conditions of application unclear
Explanation	Uses *plausible* mechanisms to match outcome data in a *well-defined* manner	Model is brittle, so minor changes in the set-up result in bad fit to explained data, bugs in the code
Description	Relates directly to evidence for a set of cases	Unclear provenance; over generalisation from cases described
Theoretical exposition	Systematically maps out or establishes the consequences of some mechanisms	Bugs in the code; inadequate coverage of possibilities
Illustration	Shows an idea clearly as a particular example	Over interpretation to make theoretical or empirical claims, vagueness
Analogy	Provides a way of thinking about something, gives insights	Taking it seriously for any other purpose
Social Interaction	Facilitates communication or agreement	Lack of engagement; confusion with objective modelling

As should be clear from the above discussion, being clear about one's purpose in modelling is central to how one goes about developing, checking and presenting the results. Different modelling purposes imply different risks, and hence different activities to avoid these. If one is intending the simulation to have a public function (in terms of application or publication), then one should have taken action to minimise the particular risks to achieving its purpose, so that others can rely on the model for this purpose (at least to some extent).

This does not include private modeling, whose purpose is maybe playful or exploratory and done only between consenting adults (or alone). However, in this case one should not present the results or model as if they have achieved anything more than an illustration, otherwise one risks wasting other people's time and attention. If one finds something of value in the exploration it should then be re-worked so it can be justified for a particular purpose to be sure it is worth public attention.

This is not to say that models that have been published without a clear purpose should not have been published, but that now the field has matured we should aim for higher and more scientifically productive standards, and one of these standards is that an audience should know under what purpose a model is being justified, so they know how to judge it (and hence in what ways they might rely on it in their work).

A confused, conflated or unclear modelling purpose leads to a model that is hard to check, can create misleading results, and is hard for readers to judge – in short, it is a recipe for **bad science***.*

Acknowledgements

discussions helped formulate these ideas. Others who come to mind include: Guillaume Deffuant, Mario Paolucci, and Rosaria Conte. If you are not on this list you should have argued harder.

References

Axelrod, R. (1984), The Evolution of Cooperation, Basic Books,

Axelrod, R. & Hamilton, WD (1981) The Evolution of Cooperation. Science, 211:1390-1396.

Balint, P. J., R. E. Stewart, A. Desai & L. C. Walters (2011). Wicked Environmental Problems: Managing Uncertainty and Conflict. Washington, USA, Island Press.

Barnaud, C. & A. Van Paassen (2013). Equity, Power Games, and Legitimacy: Dilemmas of Participatory Natural Resource Management. Ecology and Society 18(2):21.

Barreteau, O., M. Antona, P. D'Aquino, S. Aubert, S. Boissau, F. Bousquet, W. s. Daré, M. Etienne, C. Le Page, R. Mathevet, G. Trébuil & J. Weber (2003). Our Companion Modelling Approach. Journal of Artificial Societies and Social Simulation 6(2): 1.

Barreteau, O., P. Bots, K. Daniell, M. Etienne, P. Perez, C. Barnaud, D. Bazile, N. Becu, J.-C. Castella, W. s. Daré & G. Trébuil (2013). Participatory approaches. Simulating Social Complexity: A Handbook. B. Edmonds and R. Meyer, Springer: 197-234.

Bell, A. R., D. T. Robinson, A. Malik & S. Dewal (2015). Modular ABM development for improved dissemination and training. Environmental Modelling & Software 73: 189-200.

Boschetti, F., N. J. Grigg & I. Enting (2011). Modelling = conditional prediction. Ecological Complexity 8(1): 86-91.

Bousquet, F., C. Cambier, C. Mullon, P. Morand, J. Quensière & A. Pavé (1993). Simulating the interaction between a society and a renewable resource. Journal of Biological Systems 1(2): 199-214.

Bousquet, F., O. Barreteau, P. D'Aquino, M. Etienne, S. Boissau, S. Aubert, C. Le Page, D. Babin & J.-C. Castella (2002). Multi-agent systems and role games: collective learning processes for ecosystem management. Complexity and ecosystem management. The theory and practice of multi-agent systems. M. A. Janssen. Cheltenham, Edward Elgar Publishing: 248-285.

Brugnach, M. (2010). From prediction to learning: the implications of changing the purpose of the modelling activity. Proceedings of the iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software (iEMSs 2010). International Environmental Modelling and Software Society, Ottawa, Canada.

Brugnach, M. & H. Ingram (2012). Ambiguity: the challenge of knowing and deciding together. Environmental Science & Policy 15(1): 60-71.

Brugnach, M., A. Dewulf, C. Pahl-Wostl & T. Taillieu (2008). Toward a relational concept of uncertainty: about knowing too little, knowing too differently, and accepting not to know. Ecology and Society 13(2): 30.

Cartwright, N. (1983). How the laws of physics lie. Oxford University Press.

Cohen, P. R. (1984). Heuristic reasoning about uncertainty: an artificial intelligence approach.

Cohen, P. R. (1984). Heuristic reasoning about uncertainty: an artificial intelligence approach. Pitman Publishing, Inc. Marshfield, MA.

Deffaunt, G. & Weisbuch , G. (2008) Probability Distribution Dynamics Explaining Agent Model Convergence to Extremism. In Edmonds, B., Hernandez, C. and Troitzsch, K. G.

(eds.) (2007) Social Simulation: Technologies, Advances and New Discoveries. IGI Publishing, .

Dumrongrojwatthana P., Le Page C., Gajaseni N. & Trébuil G. 2011. Co-constructing an agent-based model to mediate land use conflict between herders and foresters in northern Thailand. In: Journal of Land Use Science, 6 (2): 101-120.

Edmonds, B. (2001) The Use of Models - making MABS actually work. In. Moss, S. and Davidsson, P. (eds.), Multi Agent Based Simulation, Lecture Notes in Artificial Intelligence, 1979:15-32.

Edmonds, B. (2010) Bootstrapping Knowledge About Social Phenomena Using Simulation Models. Journal of Artificial Societies and Social Simulation 13(1)8. (http://jasss.soc.surrey.ac.uk/13/1/8.html)

Edmonds, B., Lucas, P., Rouchier, J. & Taylor, R. (2017) Understanding Human Societies. In Edmonds, B & Meyer, R (2017) Simulating Social Complexity – a handbook. Springer, 763-800. http://doi.org/10.1007/978-3-319-66948-9_28

Epstein, J. M. (2008). Why model?. Journal of Artificial Societies and Social Simulation, 11(4), 12. http://jasss.soc.surrey.ac.uk/11/4/12.html

Etienne, M. (Ed.). (2014). Companion Modelling: A Participatory Approach to Support Sustainable Development. Dordrecht: Springer.

Etienne, M., C. Le Page & M. Cohen (2003). A Step-by-step approach to building land management scenarios based on multiple viewpoints on multi-agent system simulations. Journal of Artificial Societies and Social Simulation 6(2):2.

Fieldhouse, E; Lessard-Phillips, L; & Edmonds, B. (2016) Cascade or echo chamber? A complex agent-based simulation of voter turnout. Party Politics. 22(2):241-256. http://doi.org/10.1177/1354068815605671

Funtowicz, S. O. & J. R. Ravetz (1993). Science for the post-normal age. Futures, 25(7): 739-755.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: a review and first update. Ecological modelling, 221(23), 2760-2768.

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

Guillaume Deffuant, G., Amblard, F., Weisbuch, G. & Faure, T (2002). How can extremism prevail? A study based on the relative agreement interaction model. Journal of Artificial Societies and Social Simulation, 5(4):1. http://jasss.soc.surrey.ac.uk/5/4/1.html

Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., Santos, J.I., del Olmo, R. & López-Paredes, A. (2017) Checking Simulations: Detecting and Avoiding Errors and Artefacts. In Edmonds, B & Meyer, R (2017) Simulating Social Complexity – a handbook. Springer, 119-140. http://doi.org/10.1007/978-3-319-66948-9_7

D.R. Hofstadter, FARG (The fluid analogies research group) (1995), Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought, Basic Books, New York.

Kuhn, TS (1962) The structure of scientific revolutions, University of Chicago Press.

Lakoff, G. (2008). Women, fire, and dangerous things. University of Chicago press.

Lafuerza LF, Dyson L, Edmonds B, McKane AJ (2016a) Staged Models for Interdisciplinary Research. PLoS ONE, 11(6): e0157261. http://doi.org/10.1371/journal.pone.0157261

Lafuerza, LF, Dyson, L, Edmonds, B & McKane, AJ (2016b) Simplification and analysis of a model of social interaction in voting, European Physical Journal B, 89:159. http://doi.org/10.1140/epjb/e2016-70062-2

Lansing, JS & Kramer, JN (1993) Emergent Properties of Balinese Water Temple Networks: Coadaptation on a Rugged Fitness Landscape, American Anthropologist, (1):97-114.

Le Page, C. and A. Perrotton (2017). KILT: A Modelling Approach Based on Participatory Agent-Based Simulation of Stylized Socio-Ecosystems to Stimulate Social Learning with Local Stakeholders. Autonomous Agents and Multiagent Systems: AAMAS 2017 Workshops, Visionary Papers, São Paulo, Brazil, May 8-12, 2017. G. Sukthankar and J. A. Rodriguez-Aguilar (Eds.). Cham, Springer International Publishing: 31-44.

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

Meese, RA & Rogoff, K (1983) Empirical Exchange Rate models of the Seventies - do they fit out of sample? Journal of International Economics, 14:3-24.

Minsky, M. (1965). Matter, mind and models. http://dspace.mit.edu/bitstream/handle/1721.1/6119/AIM-077.pdf

Morgan, M. S. and M. Morrison (1999). Models as mediators : perspectives on natural and social sciences. Cambridge University Press.

Moss, S. (1998) Critical Incident Management: An Empirically Derived Computational Model, Journal of Artificial Societies and Social Simulation, 1(4):1, http://jasss.soc.surrey.ac.uk/1/4/1.html

Newell, A., & Simon, H. A. (1972) Human problem solving Englewood Cliffs, NJ: Prentice-Hall.

Norling, E., Meyer, R & Edmonds, B. (2017) Informal Approaches to Developing Simulations. This volume.

Olsson, P., C. Folke and F. Berkes (2004). Adaptive Comanagement for Building Resilience in Social–Ecological Systems. Environmental Management 34(1): 75-90.

Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. Science 325(5939): 419.

Ozawa, C. P. (1991). Recasting science. Consensual Procedures in Public Policy Making. Boulder, USA, Westview Press.

Reed, M. S., A. Evely, G. Cundill, I. Fazey, J. Glass, A. Laing, J. Newig, B. Parrish, C. Prell, C. Raymond & L. C. Stringer (2010). What is social learning? Ecology and Society 15(4).

Salliou, N., C. Barnaud, A. Vialatte & C. Monteil (2017). A participatory Bayesian Belief Network approach to explore ambiguity among stakeholders about socio-ecological systems. Environmental Modelling & Software 96: 199-209.

Schelling, T. C. (1969). Models of segregation. The American Economic Review, 59(2), 488-493.

Schelling, T. C. (1971). Dynamic models of segregation. Journal of mathematical sociology, 1(2), 143-186.

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

Stubbs, M. & M. Lemon (2001). Learning to Network and Networking to Learn: Facilitating the Process of Adaptive Management in a Local Response to the UK's National Air Quality Strategy. Environmental Management 27(8): 321-334.

Sun, Z., I. Lorscheid, J. D. Millington, S. Lauf, N. R. Magliocca, J. Groeneveld, S. Balbi, H. Nolzen, B. Müller, J. Schulze & C. M. Buchmann (2016). Simple or complicated agent-based models? A complicated issue. Environmental Modelling & Software 86: 56-67.

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. http://jasss.soc.surrey.ac.uk/16/2/4.html

Van den Belt, M. (2004). Mediated modeling: a system dynamics approach to environmental consensus building. Washington, Island press.

Watts, D. J. (2014). Common Sense and Sociological Explanations. American Journal of Sociology, 120(2), 313-351.

Zeigler, B. P. (1976). Theory of Modeling and Simulation. Wiley Interscience, New York.

---

[1] The paper was written by Bruce Edmonds and Christophe le Page, with suggestions and additions by the other authors.

[2] Zeigler (1976) describes this as not one-one but as a homomorphism (a one-many) relationship so that parts of the model correspond to sets of elements in what is observed, but this is effectively the same.

[3] With the exception of the purpose of description where a model is *intended* to reflect what is observed.

[4] We prefer the idea of a tool to that of a mediator (Morgan & Morrison 1999) since a model as a mediator is too vague and frustrates any assessment of goodness for purpose (except for the purpose of social interaction where people are good at assessing the quality of communication between themselves, and hence how useful a model was in facilitating this).

[5] He discusses 'prediction' and then lists 16 other reasons to model.

[6] We are not ruling out the possibility of re-usable model components in the future using some clever protocol, we have now read many proposals improving the re-usability of model code, but not seen any cases where this works for anyone except the team making these proposals.

[7] For thinking about things analogically would be more precise but that is too long.

[8] Short alternatives for this might be social learning, collaborative learning or the French word concertation. We went with the broadest term.

[9] It would not really matter even if the code had a bug in it, if the code reliably predicts (though it might impact upon the knowledge of *when* we can rely upon it or not).

[10] Where model B may be a random or null model but also might be a rival model.

[11] There is a sense in which any model output can be re-cast as a conditional prediction (Boschetti & al. 2011) but in many cases (e.g. where a model is not empirically based) the conditions are so unlikely as to be meaningless. Thus we are not convinced that this interpretation of models is very useful – many models are simply not intended to predict.

[12] To be precise, some people have claimed to predict various social phenomena, but there are very few cases where the predictions are made public *before* the data is known and where the number of failed predictions can be checked. Correctly predicting events after they are known is much easier! The European Social Simulation Association is planning a prediction competition at their conference in Stockholm in 2018, it will be interesting to see how this goes.

[13] We are being a little disparaging here, it may be that these have a definite meaning in terms of relating different scales or some such, but too often they do not have any clear meaning but just help the model fit stuff.

[14] Whilst ideally, processes that are not thought essential to the explanation should not be included in models designed for the purpose of supporting explanations, in practice it is often impossible not to include some, if only to make the simulation run.

[15] The core of this was essentially François' doctoral research.

[16] To be precise, it does assume there are discrete entities or objects and that there are processes within these that can be represented in terms of computations, but these are not very restrictive assumptions.

[17] In the sense of not being vulnerable to being shown to be wrong later.

[18] To be exact, the relationship is between some part of the understanding of the working of a simulation to the domain of application.

[19] Indeed the work spawned a whole industry of papers doing just such an exploration.

[20] Kuhn (1962) pointed out the tendency of scientists to only see the evidence that is coherent with an existing theory – it is as if they have 'theoretical spectacles' that filter out other kinds of evidence.

[21] For example http://foreignpolicy.com/2009/11/09/the-dustbin-of-history-limits-to-growth/ which mocks the quality of predictions.

[22] We are aware that there are some that argue that, in fact, it was pretty good as a predictive model (e.g. https://www.theguardian.com/commentisfree/2014/sep/02/limits-to-growth-was-right-new-research-shows-were-nearing-collapse), but the model only 'predicts' in some and not other aspects over a short time frame with no major turning points.

[23] (Grimm et al. 2014) argues that we should record and make public the process of model development in 'TRACE' documents.

[24] We must admit that the acronym is a bit of a misnomer since here it is evidence rather than representation that is key to it, however KIEB (Keep It Evidence Based) was not such a good acronym and would have ruined the title of (Edmonds & Moss 2005).

[25] OK, we admit it – we made this acronym long and unpronounceable so that it would not catch on.