Chapter 8

What is a model, why people don't trust them, and why they should

Fabio Boschetti
CSIRO Marine Research

Elizabeth A. Fulton
CSIRO Marine Research

Roger Bradbury
Crawford School of Public Policy, Australian National University,

John Symons
Department of Philosophy, University of Kansas

It is easier to make one's way in the world if one has some sort of expectation of the world's future behaviour. Even when facing a very complex problem, we are rarely in a state of full ignorance: some expectations of system behaviour and the level of risk arising from uncertainty are usually available and it is on the basis of these expectations that most decisions are taken. Humans use models, which are mental or formal representations of reality, to generate these expectations, employing an ability that is shared more or less by all forms of life. Whether it is a tree responding to shortening day length by dropping its leaves and preparing its metabolism for the winter ahead or a naked Pleistocene ape storing food in advance of winter for the same reasons, both are using models. This view leads to two outcomes. The first is that predictions, seen as an expectation of ranges of future behaviours, are not just desirable, but necessary for decision-making. The often-asked question 'do models provide reliable predictions?' then shifts to 'given a certain problem, what type of models provide the most useful and reliable prediction?' The second outcome is that modelling is no longer a scientist's activity but is instead a social process. Different types of models can be employed to ensure that all available information is included in model building and that model results are understood, trusted and acted upon.

1 Introduction

For a discussion of models we begin from what may seem an unusual point: a definition of life. Rosen [1] introduced the concept of anticipatory systems, suggesting that a defining distinction between living and non-living systems lies in the need for the living to anticipate the future behaviour of their environment and the likely outcome of their interaction with it. Loosely speaking, a ball rolling towards a wall is bound to hit the wall, while a living being provided with perception can detect the presence of the wall, anticipate the impact and, if convenient, plan for its avoidance.

This idea is at the core of much work in artificial intelligence as well as in complex system science, as formalised in the computational mechanics literature [2–5]: agents store information from the past and from it extract regularities. These regularities are a 'model' of the environment, which is used to anticipate (predict) its future behaviour. The number and sophistication of the regularities the agent is able to store and process are measures of a model's complexity. Forms of life at different levels of the evolutionary tree are able to use models of higher and higher complexity. At one extreme, bacteria hardwire simple models in their biological structure, while at the other extreme humans employ conscious mental processes and store formal mathematical tools in books and computers. Nonetheless, they both 'model' and 'predict'.

From this perspective, a computational Earth System model running on a supercomputer is 'just' a sophisticated solution humans have evolved to address their need to interact with their environment. More important for our discussion, according to this approach modelling (including computational modelling) is not only a natural but also an essential activity.

This view of the role and purpose of models may not match our intuition, according to which highly complex processes are extremely hard to understand. Also our experience tells us that complex dynamics often appear to be controlled by surprises rather than regularities. This has led many authors to claim that the use of computer modelling to study and predict complex processes is unwarranted. This criticism takes many forms, which for the sake of conciseness we summarise in three points: a) computational models have a very poor prediction track record [6, 7]; b) most model predictions are not testable because of their conditional nature [8–11]; and c) behind an appearance of objectivity, model outcomes reflect the subjective beliefs and assumptions of model users [12, 13]. According to this criticism, the benefit of modelling is limited to one or more of these activities: explanation of past events, understanding of natural processes, learning [14] or simply providing an avenue for communications [14, 15].

This criticism is very important. However, it is based on a crucial assumption and a misunderstanding. The assumption is that a prediction is *desirable* but not *necessary*; that is, a prediction is an ideal or discretionary input to, not a requirement for, decision-making. Our discussion above suggests the opposite: if we accept that a prediction is *essential* to *any* decision-making, then the question would shift from 'can model predictions be trusted?' to 'how do models compare to other approaches to prediction?' We thus need to address this crucial question: is prediction desirable or necessary for planning and decision-making?

The misunderstanding has to do with the habit scientists have of using whatever is available—in this case, models—to do science. While prediction as we are describing it requires models, models can be used for more than prediction—they can be used for exploration and understanding and even control [16]. Prediction need not necessarily be scientific—although we argue that, at its best, it is.

2 About prediction

Three concepts are fundamental to our discussion. First, for prediction we do not intend the anticipation of an exact behaviour or event, rather an estimation of its likely limits. In other words, a prediction should not be understood as a prophecy [11]. For example, while it is widely known that weather forecasts are not reliable past 5–6 days, no one would believe that the temperature in Darwin in summer could be 40°C or –40°C with equal probability; as a result no one would travel to Darwin in January with a ski jumper. The limited predictability past 5–6 days still has allowed a certain level of effective planning by avoiding carrying unnecessary clothing. Second, predictions are conditional: any prediction is carried out within a context and is valid only within that context. In the above example the conditioning is given by our understanding of tropical climate; should this change, the prediction would no longer hold and would require updating [16]. Finally, the effectiveness of a prediction is scale-dependent [17]. For example, while the geophysical community is today sceptical about its ability to provide accurate prediction on where and when large earthquakes can occur, it is nevertheless able to predict the broad geographical areas in which large earthquakes can be expected. While this kind of predictability seems to offer little to planning [18], it still has considerable practical impact in deciding, for example, in which geographical areas expensive antiseismic constructions methods are necessary and where they are not. Once understood in these terms prediction becomes an integral part of any decision-making process: formulating a plan implies choosing among potential alternatives and envisaging (= predicting) which one is more likely to deliver desired outcomes [19].

If we accept that prediction is necessary for planning and decision-making, then it is important to next ask what tools provide the most reliable prediction given the problem at hand. Notice that this question is problem-dependent, not only because different problems may require different approaches, but also because the most accurate prediction is not necessarily the most reliable. Together with using numerical models or other computational tools predictions can be provided by experts, local knowledge or participatory settings. It is thus important to compare the predictive performance of computational models against alternative approaches on the core items of criticisms discussed above: a) prediction track record; b) lack of testability due to their conditional nature; and c) inherent subjectivity.

We are not aware of any large-scale comparisons of the predictive accuracy of models and alternative methods. However, the available literature on the logical and attitudinal fallacies that even experts display for simple dynamical problems should warn us that it is probably unwarranted to expect humans to mentally predict the behaviour of highly complex systems in a consistent and reliable manner [20-28]. Predictions provided by experts, local knowledge and participatory settings will also be dependent on both explicit and implicit conditioning, including tacit information and hidden assumptions. A discussion of different forms of conditioning and its impact on modelling can be found in [29]. As nicely argued in [30], modelling offers an avenue for making such conditioning and assumptions explicit by coding and documenting the model, which may be sidestepped or not considered necessary in alternative approaches. Naturally, the same reasoning applies to the subjective nature of predictions.

So we arrive at a more expansive view of prediction for humans than just ‘expectation of the future’. We need to include the human knowledge of the past (culture), the human need to explore and understand (science) and the human bias to act (policy and intervention). Together these colour our approach to prediction and the sorts of models we tend to use.

3 Types of models

Accepting that models are a necessary component of a decision-making process does not imply believing that a) modelling reduces to running a single sophisticated computational model; b) that modelling is something only scientists do; and c) that the outcome of a model should be trusted uncritically. In fact, much of our work has been based on preaching and practising the opposite: a) models need to be built by teams including scientists and model recipients because much of the information needed to implement a model is implicit or tacitly held; b) model results need to be carefully explained and understood in order to be trusted and acted upon by decision makers; and c) information about uncertainty in the model outcome is crucial to formulate an effective plan.

As a result, our approach to modelling focuses on treating 'building a model' as the catalyst rather than the final aim of the process. In other words, extensive interactions between scientist, decision-makers and model recipients to introduce, showcase, discuss and tune the model used for final decision-making become both a requirement and an opportunity to ensure model relevance and acceptance and increase the broader understanding of the system's function. To fulfil these roles we develop five broad classes of models: conceptual, toy, single-system, shuttle, and full-system [also see Volume 1, Chapter 5].

In *conceptual models* the main drivers of a system are highlighted for subsequent representation as components of the model. This first kind of model is usually expressed as a conceptual diagram summarising our understanding of system function. In *toy models* a problem is simplified in such a way that only a handful of components are included. Their purpose is mostly educational: we want to understand how each component affects the problem and in order to achieve this we temporarily renounce a satisfactory understanding of the overall problem. In *single-system models* we include a fairly detailed representation of a single component of the system. These can be used to introduce stakeholders to modelling, provide results from the study of a single activity (addressing sector-specific issues) and feed into the development of a final full-system (multisector) model. In *shuttle models* [31] (or minimum realistic models), we include the minimum number of processes we believe are crucial for a basic understanding of the overall problem. We know these models are simpler than they need to be for full-system description, but they provide a sufficient understanding to enable us to contemplate many questions with existing (often incomplete) datasets. These models can also be a useful stepping stone to building, using and correctly interpreting the more complex models needed to check for unexpected outcomes resulting for feedbacks buried in a full-problem description. The term 'shuttle' refers to taking us from a minimum to a full description of the problem, a journey that is necessary both to developers in model definition and parameterisation, and to stakeholders in the interpretation of the final full-system model results. Finally, the *full-system model* includes all information collected for a region and addresses all scenarios of stakeholders concern, whose definition has been greatly eased by using the 'simpler' models.

As an example, a conceptual model may identify population growth and industrialisation as two of the main drivers for climate change; a toy model may describe how emissions affect peak temperature; a single-system model may include the effect of regulations on national emissions; a shuttle-model may include a simplified representation of the interaction between economic growth, population dynamics and warming. This will gradually 'take' us to comprehend the 'full' model, which may include trade, financial market dynamics, sequestration, geo-engineering etc. Figure 1 summarises the stages at which different model types are employed, the role they play and how they can support the development and use of a full-system model.

All of the model types have their own value and the full set need not always be employed—in some cases enough is learnt from conceptual models to improve predictions, in other cases the feedbacks captured in shuttle models are an effective means of refining predictions, while in others the system of interacting pressures and actors is sufficiently complex that only a full-system exposition can guide decision-makers through the complex network of feedbacks and unexpected outcomes of interactions.

For complex issues, 'full models' may represent the crucial component of the 'anticipatory system' according to Rosen's definition: it is the tool that the decision-making team, as a unified agent, employs in order to explore and evaluate options for actions. The other model types allow for engagement of different parties involved in the decision-making team, including researchers, formal decision-makers and stakeholders at large. In other words, they help the decision-making team work as a team. In particular, they can: a) allow a less-biased interpretation of available information; b) allow for learning of specific skills and attitudes needed when facing complex problems; and c) provide an avenue for communication and collaboration.

Allowing a less biased interpretation of available information is important because people with different world views may interpret and draw very different conclusions from the same information [32–40]. Research on attitudes to climate change, nanotechnology and vaccination, among other issues, shows how world views affect policy support more than available information does because these views filter how such information is processed. Accounting for such biases in a model (by parameterising the model according to different worldviews, [41]) may be a way to highlight the issue and the potential inconsistencies that may arise from it and move the discussion from perspectives to mechanisms in the hope that this may reduce the influence of biased information interpretation.

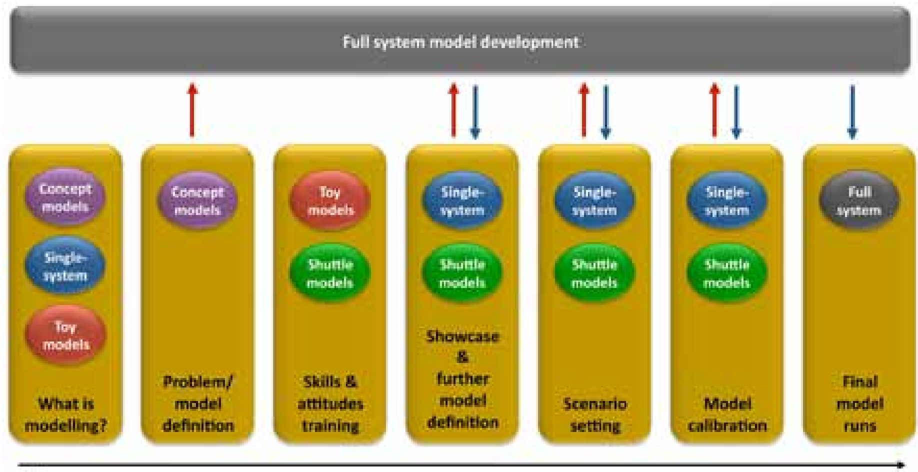

Figure 1: Relation between a) a modelling project (top dark rectangle); b) different types of models (coloured ovals); and **c)** stakeholders engagement phases (yellow rectangles). The arrow at the bottom suggests an approximate chronological order.

Conclusion

We think that prediction is what living things do and that they do it through models. But we also think that the relationships between models and predictions are varied, and—importantly in the case of humans—dependent on the context. We also think that rather than diminishing or muddying the matter of prediction, this enriches it and places it in a very human context. Models could be better used and their results more trusted if this was better understood.

Thus we argue that training on specific skills and attitudes is needed to help us face complex problems. This is important because scientific insights risk being lost unless they are understood by those making and supporting decisions: recent studies have shown that human cognition and psychology affect decision-making at least as much as the complexity of the problem at hand [21, 23, 25, 26, 42]. Computer models resembling flight simulators can be designed to train individuals to better understand the basic processes of real-world significance for decision-making, including management of limited resources and unexpected feedbacks. The belief underneath this approach is that managing and predicting complex behaviours can be learned and that models can represent systems in a manner appropriate for learning and training.

Not only cognitive skills but also cognitive attitudes are crucial to effective decision-making about complex problems [24, 43, 44]—the behavioural attributes and habits we bring into a problem, the way we formulate goals, interpret outcomes against expectations and balance emotional responses like humility, curiosity, frustration and blame-shifting have a significant influence on how effectively we deal with complex situations [43]. Tangible, constructive means to improving problem solving in complex settings can be identified and trained via computer models [43, 45]. Interestingly, some of the most effective cognitive approaches (including tolerating high levels of uncertainty, acknowledging mistakes, searching for counterevidence, self-reflection etc.) can be in direct opposition to behaviours rewarded in political and management roles. More widespread awareness of what makes an effective decision-maker, possibly leading to improvements in training programs, may have an immense impact on a wide variety of real-world issues.

References:

1. Rosen R (1985) *Anticipatory systems. Philosophical, mathematical and methodological foundations* (Pergamon Press, Oxford).

2. Crutchfield JP (1994) The calculi of emergence: computation, dynamics, and induction. *Physica D* 75:11–54.

3. Crutchfield JP, Young K (1989) Inferring statistical complexity. *Physical Review Letters* 63:105–108.

4. Shalizi C (2001) Causal architecture, complexity and self-organization in time series and cellular automata, PhD (Santa Fe Institute, Santa Fe).

5. Shalizi C, Shalizi K, Haslinger R (2004) Quantifying self-organization with optimal predictors. *Physical Review Letters* 93(11):118701

6. Ascher W (1993) The ambiguous nature of forecasts in project evaluation: diagnosing the over-optimism of rate-of-return analysis. *International Journal of Forecasting* 9(1):109–115.

7. Brunner R (1999) Predictions and policy decisions. *Technological Forecasting and Social Change* 62:73–86.

8. Ascher W (1989) Beyond accuracy. *International Journal of Forecasting* 5(4):469–484.

9. Oreskes N (2000) Why believe a computer? Models, measures, and meaning in the natural world, in ed. Schneiderman J *The Earth around us: maintaining a livable planet*, (WH Freeman and Co, San Francisco), pp 70–82.

10. Beven K (2006) A manifesto for the equifinality thesis. *Journal of Hydrology* 320 (1-2):18–36.

11. Beven KJ (2002) Towards a coherent philosophy for environmental modelling. *Proceedings of the Royal Society A* 458:2465–2484.

12. Aligica PD (2003) Prediction, explanation and the epistemology of future studies. *Futures* 35(10):1027–1040.

13. Oreskes N (2001) Philosophical issues in model assessment, in eds Anderson MG, Bates PD *Model validation: perspectives in hydrological science*, (John Wiley and Sons Ltd, London), pp 23–41.

14. Brugnach M (2010) From prediction to learning: the implications of changing the purpose of the modelling activity, in *International Congress on Environmental Modelling and Software* (Ottowa, Canada).

15. D'Aquino P, Le Page C, Bousquet F, Bah A (2003) Using self-designed role-playing games and a multi-agent system to empower a local decision-making process for land use management: the SelfCormas experiment in Senegal. *Journal of Artificial Societies and Social Simulation* 6(5):3.

16. Bradbury RH (2002) Futures, predictions and other foolishness, in ed. Elgar E *Complexity and ecosystem management: the theory and practice of multi-agent systems*, (Cheltenham), pp 48–62.

17. Israeli N, Goldenfeld N (2004) Computational irreducibility and the predictability of complex physical systems. *Physical Review Letters* 92(7):074105-074101-074105-074104.

18. Pielke RA (2003) The role of models in prediction for decision, in eds Canham CD, Cole JJ, Lauenroth WK *Models in ecosystem science*, (Princeton University Press, Princeton and Oxford), pp 111–135.

19. Bradbury RH, R.E. R, Green DG (1985) Policy = f(theory, data, models, tools): Rational solutions to coral reef conservation. *Fifth International Coral Reef Congress*, pp 247–251.

20. Boschetti F, Hardy P, Grigg N, Horowitz P (2011) Can we learn how complex systems work? Emergence: complexity and organization. In print.

21. Moxnes E (1998) Overexploitation of renewable resources: the role of misperceptions. *Journal of Economic Behavior and Organization* 37(1):107–127.

22. Moxnes E (2000) Not only the tragedy of the commons: misperceptions of feedback and policies for sustainable development. *System Dynamics Review* 16(4):325–348.

23. Moxnes E, Saysel AK (2009) Misperceptions of global climate change: information policies. *Climatic Change* 93:15–37.

24. Stanovich K (1999) *Who is rational? Studies of individual differences in reasoning* (Lawrence Erlbaum Associates).

25. Sterman JD, Sweeney LB (2002) Cloudy skies: assessing public understanding of global warming. (Translated from English) *System Dynamics Review* 18(2):207–240 (in English).

26. Sterman JD, Sweeney LB (2007) Understanding public complacency about climate change: adults' mental models of climate change violate conservation of matter. *Climatic Change* 80(3-4):213–238.

27. Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. *Science* 185:1124–1131.

28. Tversky A, Kahneman D (1983) Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review* 90:293–315.

29. Boschetti F, Grigg NJ, Enting I (2011) Modelling = conditional prediction. *Ecological Complexity* 8(1):86–91.

30. Meadows DL, Behrens III WW, Meadows DH, Naill RF, Randers J et al. (1974) *Dynamics of growth in a finite world* (Wright-Allen Press, Cambridge, Massachusetts).

31. Perez P (2011) Science to inform and models to engage, in *Australia 2050: achieving an environmentally sustainable and socially equitable way of living*, (Australian Academy of Science, Canberra).

32. Duckitt J, Sibley CG (2009) *A dual-process motivational model of ideology, politics, and prejudice.* (Psychology Press), pp 98–109.

33. Duckitt J, Wagner C, du Plessis I, Birum I (2002) The psychological bases of ideology and prejudice: testing a dual process model. *Journal of Personality and Social Psychology* 83(1):75–93.

34. Heath Y, Gifford R (2006) Free-market ideology and environmental degradation. *Environment and Behavior*, 38(1):48–71. doi: 10.1177/0013916505277998..

35. Kahan DM, Braman D, Slovic P, Gastil J, Cohen GL (2007) The second national risk and culture study: making sense of—and making progress, in *The American culture war of fact*. (SSRN).

36. Leviston Z, Walker IA (2010) Baseline survey of Australian attitudes to climate change, in *Climate adaptation flagship*, (CSIRO, Perth).

37. Lewandowsky S, Stritzke WGK, Oberauer K, Morales M (2005) Memory for fact, fiction, and misinformation. *Psychological Science*, 16:190–195.

38. Mirisola A, Sibley CG, Boca S, Duckitt J (2007) On the ideological consistency between right-wing authoritarianism and social dominance orientation. *Personality and Individual Differences* 43(7):1851–1862.

39. O'Riordan T, Jordan A (1999) Institutions, climate change and cultural theory: towards a common analytical framework. *Global Environmental Change* 9(2):81–93.

40. Unger RK (2002) Them and us: hidden ideologies—differences in degree or kind? *Analyses of Social Issues and Public Policy* 2:43–52.

41. Boschetti F (2011) A simple model of how people reason about climate change. Submitted.

42. Sterman JD (2008) Risk communication on climate: mental models and mass balance. *Science* 322(5901):532–533.

43. Dorner D (1996) *The logic of failure: recognizing and avoiding error in complex situations* (Metropolitan Books, New York).

44. Sorrentino R, Roney C (2000) *Uncertain mind: individual differences in facing the unknown* (Psychology Press, Philadelphia).

45. Boschetti F, Hardy P, Grigg N, Horowitz P (2010) Can we learn how complex systems work? in *Emergence: complexity and organization.* In print.

Authors

Fabio Boschetti (Fabio.Boschetti@csiro.au) is a research scientist, CSIRO, Marine Research.

Elizabeth A. Fulton (Beth.Fulton@csiro.au) is a science fellow at CSIRO Marine Research.

Roger Bradbury (roger.bradbury@anu.edu.au) is adjunct professor, Resource Management in Asia-Pacific Program at the Crawford School of Public Policy, the Australian National University.

John Symons (johnsymons@ku.edu) is the chair of the philosophy department, University of Kansas.