Modelling = Conditional Prediction

Fabio Boschetti[1,4], Nicky Grigg[2], Ian Enting[3]

[1] Corresponding author. CSIRO Marine and Atmospheric Research, Private Bag 5, Wembley WA 6913,Australia. Fabio.Boschetti@csiro.au, phone +61 8 9333 6563

[2]CSIRO Land and Water, GPO Box 1666, Canberra ACT 2601, Australia. Nicky.Grigg@csiro.au, +61 2 6246 5569

[3]MASCOS, 139 Barry St, The University of Melbourne, Victoria 3010, Australia. ienting@unimelb.edu.au

[4]School of Earth and Environment, The University of Western Australia , 35 Stirling Highway, Crawley WA 6009, Australia

Abstract

We emphasise the benefits of viewing the output of any numerical model as a conditional prediction. More specifically, this is a view that the characterisation and communication of model context warrants as much attention as the technical description of the model problem and solution. Drawing on modelling examples from physical, ecological and social systems, we explore the potential benefits of adopting this practice more widely and more explicitly in modelling domains where the conditional nature of the models is not often emphasised. Given the growing reliance on numerical models, particularlyin informing policies for natural resource management, we suggest the 'Modelling = Conditional Prediction' perspective offers a useful lens through which to view the results of these models. It is a view which provides clarity about the role of modelling within the larger research scope and can facilitate communication between model users from different disciplines.

Keywords: modelling; prediction; model conditioning; model interpretation

1 Introduction

It is increasingly common for teams of scientists and practitioners to address problems which encompass physical, ecological and social aspects. Often numerical models are used as tools in the approach, either as individual components or as an avenue to link the knowledge residing in the multi-disciplinary team.

These large, often complex problems face several challenges, including finding a common framework for discussion among specialists in different disciplines who may hold different assumptions and expectations. We have witnessed several such discussions and realised that a shared common understanding is not readily available even in dealing with numerical models: different groups, and often different individuals within the same group, may have a different idea of what a numerical model provides and how its output should be interpreted. Questions such as whether a numerical model represents a simulation of reality, a virtual laboratory, a platform to test a researcher's idea or simply an avenue for communication and education need to be addressed before multiple teams can work with the same model or before multiple models may be used effectively within the same project. Similarly, the question whether a model outcome represents a prediction of possible future events, the testing of a scenario or simply a tool to generate stakeholder feedback needs to be addressed before such outputs can be shared and used, for example, as input for other models or for decision making. This issue is of particular relevance as there is no consensus on theoretical approaches to the role of models in science; for a discussion see (Morgan and Morrison, 1999; Barreteau *et al.*, 2003; Aumann, 2007).

The purpose of this work is to propose a framework which may contribute to a shared understanding; it interprets the output of a computer model as a conditional prediction, that is a prediction which depends crucially on the conditioning (assumptions) imposed or implicit in the model as well as on the purpose of the model in the context of the problem at hand.

This framework is based on three concepts: first, a computer model implements a set of rules; second, a comparison between the output of a model and real phenomena needs to be possible (at least conceptually); third, the input, output and logical steps in the computer code need to find an interpretation in the real world for the exercise to have a meaning, this last step being very different for different disciplines and imposing an increasing number of approximations going from physical to ecological and social problems.

Rather than adopting a technical definition of prediction, in this work we prefer to conform to a broad everyday meaning: a prediction is an act of anticipating consequences of a process, an event or an interaction by formulating an expectation of what may happen. In practice, for most problems this means being able to anticipate limits on the expected system behaviour rather than exact system trajectories, and the reliability of a prediction is commonly understood to be scale-dependent. 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 Tucson, Arizona, in August could be 40°C or -40°C with equal probability; as a result no one would travel to Tucson in August with a ski jumper. The same reasoning applies to most systems for which predictability depends on time scales and resolution (Israeli and Goldenfeld, 2004).

Our main argument is that any prediction is carried out within a context and that the context is given by the conditioning of the prediction. In the above example, the conditioning is given by our understanding of the current status of the global atmospheric circulation; should this status change, the prediction would no longer hold and would require updating.

In the rest of this document we will discuss in more detail why it is a useful discipline to interpret a model output as a prediction, why this prediction is conditional and why the explicit acknowledgement of this conditioning is essential. We will give examples of conditioning in modelling exercises of differing complexity and subjectivity and conclude by explaining why we believe this framework is a useful avenue for communication between modellers of different backgrounds.

Before proceeding, some further clarifications are useful in order to define the scope of this work. First, we focus on the conditioning on the *prediction* of a model, not on the possible conditioning on the *model*; in other words, we assume that a model is available, it has been already chosen according to meaningful criteria and is appropriate for the problem at hand. Second, we limit our discussion to numerical models, either as algorithms written in a computer program or as closed form mathematical equations which can be solved analytically; while some of our arguments could be extended to non computational models, for sake of clarity we do not address them in this paper. Third, we are mostly concerned with the use of fairly complex, ‘large’ numerical models, such as the types of models commonly employed in real world engineering, ecological, economical and management problems. Once again, while some of our arguments could be extended to any kind of numerical model, fundamental issues related to the general applicability of models to natural phenomena and their ability to discriminate and represent effective causation are widely addressed in the philosophy of science literature, to which we refer the reader. Our concern is to project these issues into the realm of the modelling practitioner and researcher, clarify them by examples and highlight the relevance to real world applications and the ethical issues they arise.

2 Why prediction

Pragmatically, numerical modelling appears to be a relatively simple endeavour: we represent our understanding of a process in a computer program, we choose a suitable input, we run the code on some test cases and if the results are consistent with observations we extend the use of the code to explore other situations.

Theoretically, modelling of any kind poses questions at the core of the philosophy of science, since building working (mental or logic) models of reality is part of the foundation of science. When the problem we try to address via modelling is complex, these arcane philosophical questions acquire practical relevance: the authors have witnessed several times modellers of different disciplines and backgrounds come together and find it difficult even to agree on what modelling is for, let alone what it is. What may be obvious to a physical scientist is not to an ecologist nor to a social scientist and vice versa. This makes building interdisciplinary models which address problems spanning physical, biological and social fields very difficult.

We do not address these philosophical problems here; rather we focus on the notion that the output of numerical models of any process need to be able to be interpreted as a prediction in order to be relevant to the scientific method.

We restrict this discussion to models which are computer programs or closed form mathematical equations[1]. A computer program, as well as an equation, is a closed system to which the rules of formal logic apply: it takes an input and manipulates it via the set of transformation rules coded in the algorithm. The output of the model is a logical consequence of, thus equivalent to, the input given the rules of the algorithm (Boschetti et al., 2008). Consequently, technically it does not contain any information not already coded in the input and the rule set. Within this framework, the output of a model is not a prediction but rather a logic inference. Note that in this framework there is no link between the model and the real world.

In practice, a model output is informative because the modeller often is not able to infer all the possible consequences of the rule set. Within this framework, the output of a model can help the modeller to learn and to develop intuition by playing out different rule sets and initial conditions and exploring their consequences. Here, the only link between the model and the real world lies within the modeller; that is in the effort to infer the working of the model.

[1] A claim is occasionally encountered in the complex system literature according to which purely computational models are potentially able to simulate the action of 'primitive' elements of the process under analysis, while closed form equation (often in the form of differential equations) are better suited to model 'emergent' larger scale behaviours, which are insensitive to the details of the lower level processes. For example, the law of perfect gas described in Section 4.1 can be seen as emerging from the interaction of many randomly-moving point particles, whose individual interaction could be modelled as primitives in a cellular automaton. For our discussion, both computational models and closed form equations are example of algorithms, that is, set of instructions which can be followed mechanically once suitable initial conditions are defined. As a result our discussion applies unaltered to both.

Despite the weak link, prediction already appears in the picture as the process of discovering or learning involves a form of comparison between prior and newly informed understandings.

If the purpose of the model is to mimic a dynamical process or to represent correlations between data, the link between the model and the real world becomes much stronger: the input needs to have some resemblance with the initial conditions of the process and the rules need to have some relation to the process or to observed correlations. Since a perfect knowledge of Nature's working can not be achieved, the resemblance between the rules and the process in Nature can be accessed only experimentally by matching the model output to the observations or, to be stricter, by trying to falsify the relation. This imposed relation between the model and the real world turns the model outcome into a prediction. Notice also that the outcome of the model here has the same logic role as the outcome of a theory (indeed they are logically equivalent (Boschetti and Gray, 2007)).

Obviously not all models need to predict: a video-game is a model whose purpose is to create settings which may or may not resemble reality; but the aim of the video-game is to entertain not to do science. Similarly, symbolic equation solvers do not carry out predictions, rather they perform logical transformations so that equations can be expressed in alternative ways. Since the determination to *winnow out belief* and confront hypotheses with data are key features which sets science apart from other forms of human knowledge (Hilborn and Mangel, 1997; Bradbury, 1999), we suggest that only models that generate predictions are amenable to scientific enquiry.

To conclude, we do not claim that the purpose of a model needs necessarily to be to perform a prediction, rather that for a model to be used within a scientific framework – even if its purpose is one of 'intuition pumping' or learning – it is a useful discipline to interpret each model realisation as a conditional prediction. Doing so sets up the model realisations as hypotheses amenable to confrontation with data and it provides clarity for assessing whether the model is appropriate for the purpose assigned to it. This step is essential also when the purpose of a model is to suggest some explanatory hypotheses as in the case of the ability of to explain past events or capture likely past causal relations, since the reliability of the explanation of the past cannot be assessed independently of the ability of predict the future; for a discussion on the relation between prediction and retrodiction see (Ellison *et al.*, 2009).

3 Why conditional

As discussed above, the output of a model is a logical consequence of the model input and the rule set. The output depends on the model input and rule set and thus is conditional on them. This conditioning is so strong it represents a logical tautology (Boschetti and Gray, 2007).

To present some practical examples and discuss the implication for the use of modelling in real world problems, we find it convenient to distinguish several types of conditions:

- from the model perspective, we define *hard* conditioning as conditions explicitly *coded* in the input or in the rule set. We define *soft* conditioning, as the human interpretation of the hard conditioning within the context of the modelling exercise;
- from the modeller's perspective, the conditioning can be *explicit* or *implicit*, depending on whether the modeller is aware of this conditioning while using or implementing model.

This classification is not unique and different schemes may be considered; it is used here merely out of convenience in order to describe a number of possible scenarios which we addressed below.

3.1 Hard conditioning

Hard conditioning is imposed directly in the input or in the rule set (following the discussion above, this can be interpreted as a *logic* conditioning). As an example, setting the parameter for the growth rate of phytoplankton in an ecological model represents a hard conditioning in the working of the model, because it determines the exact numerical result of the model. Similarly, approximating a missing data point via a spline interpolation assumes continuity and differentiability in the data distribution. If the user has provided the growth parameter or coded the spline interpolation, we assume (s)he is also aware of its use and these are examples of *explicit* conditioning.

Let's now assume that the algorithm performs a Fourier transform in its inner working. This imposes further conditions in the assumption of periodic and infinite data replication past the calculation domain as well as analytical continuity between samples. The user may not be aware of the presence of the Fourier transform in the code. Alternatively, (s)he may not be aware of the assumptions behind the use of the Fourier transform. In these cases we say that the conditioning is *implicit*: the user will run the model without being aware of these hidden hard assumptions.

3.2 Soft conditioning

Let's refer to the above example of setting a parameter for the growth rate of phytoplankton in an ecological model. Beyond the precise numerical estimate, a value may hold a certain meaning in a specific discipline: for example, a specific phytoplankton growth rate may imply a range of temperatures characteristic of a tropical environment. This can be interpreted as a soft constraint: using a precise numerical value for phytoplankton growth rate imposes a soft conditioning on the kind of scenario modelled (e.g. it may not be appropriate for modelling phytoplankton growth in polar conditions). Modellers often need to perform this interpretation step: in order to run a model, they may need to input the precise numerical value of parameters they are unable to measure, they have no indirect observation or may be impossible to observe even in principle; in this case they may resort to tables of commonly used parameters representative of the broad scenario the model is applied to (for example, expected phytoplankton growth rate for polar, temperate or tropical environments). Clearly the same conditioning can thus be interpreted as hard or soft depending on the context and expected accuracy of the simulation.

The user may or may not be aware of the implication of the input or rule set in terms of soft conditioning and thus this conditioning can also be either explicit or implicit.

3.3 Implicit soft (hidden) conditioning

Conditionings which are both soft and implicit may be troublesome since they are hidden from the user and often from the model developer as well; when made explicit, they may alter considerably the interpretation of a model result.

One type of hidden conditioning may need to be addressed separately: it includes all the *events we expect will not happen* during the process under study. Implementing a computer program involves writing rules for processes we anticipate being relevant to the problem; rules for processes considered irrelevant are usually not coded. To give an example, global models of human greenhouse gas emissions and climate change include hidden assumptions such as the following: a massive meteorite will not hit and destroy most industrialised countries; a new pandemic will not decimate world population; or global political or terrorist unrest will not seriously impact global business cycles. The list can be endless and as fanciful and entertaining as we wish, but this aspect of modelling conditioning is a useful reminder of the constraints implicit in our assumptions (see also (Isham, 1995), pp 55-56).

This type of conditioning broadly relates to what is defined as *ceteris paribus* clause in the philosophy of science, that is to all the factors which either we are not aware of, or we assume do not interfere with the process we study, or we assume are constant during the time of the analysis. This clause is essential in order to isolate the variables we want to study from the external environment and are the cornerstone of traditional scientific data collection and inference. This issue is particularly relevant to the use of large numeral models because one of the tenets of complex system science lies exactly on discussing the applicability of the *ceteris paribus* clause (Ashby, 1956; Kauffman, 2000; Laughlin *et al.*, 2000).

Table 1 summarises the type of conditioning described above.

Figure 1 summarises the framework we propose. On the right hand side we have a computer model; this is a closed structure implementing a logic system. It is made up of three components: the input, the computer code and the output. Notice that the distinction between input and computer code is in fact fairly arbitrary since any input could in principle be incorporated into the code as a parameter. In the picture we refer to them jointly as hard conditioning. On the left hand side we have the natural process we want to study. This has a very different nature from the model: first it is an open system and as a result can not be properly defined. Second, input and output are simply two different states of a dynamics and acquire a special significance only in relation to the problem we address. Third, because the system is open, input and output themselves cannot be finitely defined, but only approximated. The figure also includes the three steps we discussed above: in step 1 we merely have a formal system in which the input, rules and output are linked by logical inevitable consequences. Step 2 provides the interpretation of the model in terms of a natural process, thus defining the meaning of the model and of its conditioning. The transition from hard to soft conditioning often happens at this stage, via the use of problem-specific knowledge. Step 3 allows the interpretation of model output as a conditional prediction amenable to comparison with observations from the natural system. Note that although Figure 1portrays input and output as system states at different times, our definition of prediction need not limit us only to models representing dynamics in time. For example, a model may predict what ocean pH will be when in equilibrium with a particular atmospheric concentration of carbon dioxide, and so not require any reference to dynamics in time.

4 Some examples

In this section we give some examples of conditioning by analysing different classes of problems and models used to address them: a closed physical problem, an open physical problem and a social problem. The different classes represent increasingly complex problems, where complexity is seen as departure from pure mechanistic behaviour, as discussed below. Since a computer program implements a purely mechanical process (Boschetti and Gray, 2007), the different classes also represent an increasing gap between the dynamics occurring in the process and its representation in the model and thus imply an increasing amount of conditioning on the prediction carried out by the model. Table 2 summarises this discussion by showing examples of conditioning which apply to the following different modelling cases.

4.1 A closed physical system

In high-school physics we learnt the relation $P = \frac{nR}{V}T$. This relation can be coded easily into a computer program which outputs values of *P* given values assigned to the other variables. So far, our code simply carries out a computation, whose result is an inevitable consequence.

Once we decide to interpret *P* as the pressure of a gas, *V* its volume, *T* its temperature, *n* its amount in moles and we assign the proper value to the constant *R*, then our code becomes a model of a natural process, and it allows us to predict the behaviour of a perfect gas. Now, and only now, the model has a relation to outside reality and we can attempt to verify whether this relation holds by testing its predictions against real data.

Doing so shows us that the relation does not hold exactly. Apart from measurement errors, the behaviour of a real gas does not match the equation accurately: it is not a perfect gas. The prediction of the model is *conditional* on the gas behaving as a perfect gas and consequently the predictions carried out by the model are accurate only to the extent that the real gas approximates the assumed conditions; departure from this explicit condition will result in prediction errors.

4.2 An open physical system

Most of us are indirect users of numerical models, since most of us check the weather forecast and take some decision based upon it. Weather forecasts are carried out via extremely complex models which are conditioned not only on certain assumptions about atmospheric and oceanic circulation but also on a large amount of measurements: given similar physical settings, the more measurements are available the more accurate the forecast is likely to be.

The general public has a fairly good understanding of the mechanics of these conditional forecasts: people understand that weather can be easier to predict under stable conditions and that tomorrow's forecasts are more reliable than next week's ones. Furthermore, no one would take seriously forecasts not conditioned on real data. The general public not only understands the uncertainty which is inherent in any prediction but, importantly, is able to make a decision by accounting for it (this is what we do when we decide whether or not to take an umbrella and whether or not to aim for a week-end destination given a weather forecast).

From this perspective weather forecasts offer a very useful example of the framework we propose in this paper. The general public is accustomed to a) judging the weather forecast provided by a certain agency as more reliable that the one from another agency, supposedly because provided by (i.e. conditioned upon) better models, better data or better scientists, b) accepting the level of uncertainty on the forecast without dismissing its usefulness and c) making a decision based on the forecast as well as its uncertainty. We suggest this represents a useful analogue for the use of any numerical model: uncertainty and conditioning do not undermine the usefulness of a model, provided they are both fully understood.

4.3 A social system

It is becoming increasing common to model social behaviour via numerical modelling, agent based modelling and game theory being two popular examples. The purpose of these models usually is to study the consequences of decisions taken by individuals and how these lead to different patterns of group behaviour. Given the complexity of understanding human decision making from a psychological perspective and the effect of human relationships from a social perspective, it is natural to question the interpretation of the output of social models as firm predictions. To this we need to add the difficulty of collecting reliable input data (reliable conditioning), as well as to falsify the model outcome by comparison with characteristics seen in real social systems.

Here is where the framework 'modelling = conditional forecast' is probably most useful, since it focuses our attention on the component which is weakest: the reliable conditioning. Model outcomes are of little validity if the input data has not been verified or justified. However, very often attention is focused mostly on the rule setting, that is on the algorithm, under the common criticism that no rule set can encompass all facets of human behaviour. The 'modelling = conditional forecast' does not circumvent the problem, but suggests how to interpret it. In an agent based model, for example, an agent is an automaton following prescribed rules; as a result, the output of the model can reproduce only the features of human behaviour which can be encapsulated by rules. Human behaviours which are not characterisable by rules (e.g. some emotional responses, contradictory choices or moral decisions) cannot be accounted for in a model. An agent based model can then be seen as (at best) an approximation of what human processes would be if human behaviour was purely rule based. The model represents a projection of human processes into a sub-space of rule-based behaviours, which we believe are only a subset of all behaviours available to humans. As discussed in the previous section, this does not undermine the usefulness of agent based models, provided their assumptions are understood, are reasonable for the problem at hand and the model output is interpreted within the context determined by the conditioning.

5 Why is this useful

Modellers with a background in information theory and Bayesian statistics may find the framework we propose trivial, so here we highlight why we believe this approach is useful.

The main benefit, in our opinion, comes from imposing a certain discipline both in the modelling exercise and in the delivery of the results. First, accepting that any interpretation of model output is a prediction enforces an avenue for falsification and accountability. Both are inevitable steps in a scientific endeavour and too often modellers fall into the temptation to circumvent this responsibility.

Second accepting that the prediction is conditional imposes the responsibility of making this conditioning explicit: delivering a modelling result without information on the underlying assumptions is equivalent to assuming the prediction is not conditional but inevitable, which modelling results rarely, if ever, are.

Furthermore, seeing modelling as conditional prediction is beneficial because it highlights both similarities and differences among different modelling approaches. Analytical equation solving, numerical modelling of physical processes, numerical modelling of social and ecological processes all give results whose usefulness depends on the soundness of assumptions, data and rationale of the model.

While we believe that physical processes and ecological/social process are fundamentally different, it is useful to remember that the *numerical modelling* of physical or ecological/social processes is logically equivalent. A numerical model is an algorithm, whose rules are given a priori. While it is not clear to what extent a physical open system can be modelled via a closed formal system, it is largely assumed that a social/ecological model can not be (Rosen, 2001). As discussed in Section 4.3, we suggested that modelling ecological/social processes implies treating them as *physical processes,* in which ecological/social units act as rule-following automata. The usefulness of such numerical modelling results then depends crucially on the meaning ascribed to them. This step, deciding what the results *mean,* is where difficulties arise. It requires a tangible path between model assumptions and model interpretation. In some systems (particularly physical systems) the system characteristics being modelled are often non-controversial and it is easy to agree on common metrics for assessing model worth. In ecological and social systems this step is more difficult and inherently more debateable, but necessary nevertheless. In both physical and living systems alike we suggest that the 'modelling=conditional forecasting' perspective provides a useful avenue to clarify this step.

6 Who is a model for? User as further conditioning

It is a fairly common practice for researchers to build models for others to run; indeed in many groups the word *modeller* refers to an expert in *running* specific models, rather than in developing them. Given that the *purpose* of a model can not be coded, and that the results depend crucially on the assumptions in the code (some implicit and others hidden in remote, often undocumented and possibly even uncommented lines of code) there are obvious dangers in interpreting results (*predictions*) when the *conditions* they depend on are not fully transparent.

It is in considering the relation between model developer and model user that we encounter the age-old debate of model simplification and how the ‘right’ level of simplification can be achieved. Occasionally a view is given that the ‘right’ level of simplification depends on the user, not on the problem. Whether this view is appropriate or not depends crucially on the purpose of the model, which, interestingly, is both user *and* problem dependent.

Here we wish to place emphasis of two observations: first, an oversimplified model is a badly conditioned one; second, oversimplified results equate to *badly conditioned predictions;* third, the interpretation of badly conditioned predictions may lead to unfortunate decisions, which inevitably echoes Feynman’s famous “*for a successful technology, reality must take precedence over public relations, for Nature can not be fooled*”. It’s a plain acknowledgment that in many situations model accuracy is vitally important and this needs to take precedence over user-friendliness, convenience, lack of time or lack of funding. Among the many constraints on the path to well informed decision making we can list a) complexity of the problem, b) lack of data, c) lack of understanding, d) lack of time, will or money to investigate sufficiently. Of all of them, *the complexity of the problem is the one we have the least control on*. If dismissing this complexity is done in the name of user-friendliness, this effectively equates to *yet another conditioning* in the model; prediction will depend upon it and it needs to be properly acknowledged at the delivery of the results.

7 Conclusions

Modellers are familiar with the constraints imposed on the quality of model results by the quality of input data and of the algorithm. In this work we propose that a set of further constraints should be accounted for on an equal basis. These constraints include the interpretation of the model meaning with respect to the system under analysis, plus the acknowledgment of other requirements during model development and use, such as user friendliness, simplicity and communicability of results. More importantly, we propose that all these constraints should be seen as conditioning on the output of a model, and the output itself viewed as a prediction (i.e. an expectation of what would happen if the conditions implied in the constraints did occur). These notions clarify the minimum requirements for modelling to be consistent with scientific method. Furthermore, making such conditioning explicit can facilitate discussion of the validity and meaning of a model when integrating modelling results with other knowledge to inform decision-making.

References

Ashby, W. R. (1956). *An Introduction to Cybernetics*, ISBN

Aumann, C. A. (2007). "A methodology for developing simulation models of complex systems," *Ecological Modelling*, ISSN 202(3-4): 385-396.

Barreteau, O., M. Antona, et al. (2003). "Our companion modelling approach," *Journal of artificial societies and social simulation*, ISSN 6(1).

Boschetti, F. and R. Gray (2007). "Emergence and Computability," *Emergence: Complexity and Organization*, ISSN 9(1-2): 120-130.

Boschetti, F., D. McDonald, et al. (2008). "Complexity of a modelling exercise: a discussion of the role of computer simulation in Complex System Science,," *Complexity*, ISSN 13(6): 21-28.
Bradbury, R. H. (1999). "Just what is science anyway?," *Nature & Resources*, ISSN 35(4): 9-11.
Ellison, C. J., J. R. Mahoney, et al. (2009). "Prediction, Retrodiction, and The Amount of Information Stored in the Present," *Journal of Statistical Physic*, ISSN 136(6): 1005-1034.
Hilborn, R. and M. Mangel (1997). *The Ecological Detective: Confronting Models with Data* ISBN
Isham, C. J. (1995). *Lectures on Quantum Theory: Mathematical and Structural Foundations*, ISBN
Israeli, N. and N. Goldenfeld (2004). "Computational Irreducibility and the Predictability of Complex Physical Systems," *Physical Review Letters*, ISSN 00319007, 92(7): 074105-1-074105-4.
Kauffman, S. (2000). *Investigations*, ISBN
Laughlin, R., D. Pines, et al. (2000). "The Middle Way," *Proc. Natl. Acad. Sci.*, ISSN 57: 32.
Morgan, M. S. and M. Morrison (1999). *Models as mediators : perspectives on natural and social sciences/ edited by Mary S. Morgan and Margaret Morrison*, ISBN 0521650976 (hb)
0521655714 (pb)
Rosen, R. (2001). *Life Itself, A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*, ISBN

Table 1. Schematic summary of the combinations of hard vs soft and explicit vs implicit conditioning.

Conditioning	**Hard**	**Soft**
Explicit	Expressed as a functional relation or in the input data; the user is aware of its existence and implications	Implied in the interpretation of the hard conditioning; the user is aware of its meaning
Implicit	Expressed as a functional relation or in the input data; the user is *not* aware of its existence or implications	Implied in the interpretation of the hard conditioning or in the absence of possible processes; the user is not aware of the implications

Table 2. Examples of hard, soft and hidden conditioning in the modelling examples discussed in this paper.

Model	**Hard conditioning**	**Soft conditioning**	**Hidden conditioning**
Closed physical model (Perfect Gas Law)	Input data and gas law equation	Specific input data refer to STP	Real gas assumed to be ‘perfect’
Open physical model (weather forecast)	Input data, specific physical equations	Tropical vs Polar conditions, Winter vs Summer, etc	Absence of large perturbation (no volcanic eruptions), no major suddenly human effect on

			climate, etc
Social model (agent based modelling)	Input data, rule-based behaviour	Rules carry an interpretation, eg. economically rational, individual profit-maximimising vs cooperative, etc.	Human behaviour can be encapsulated by rules

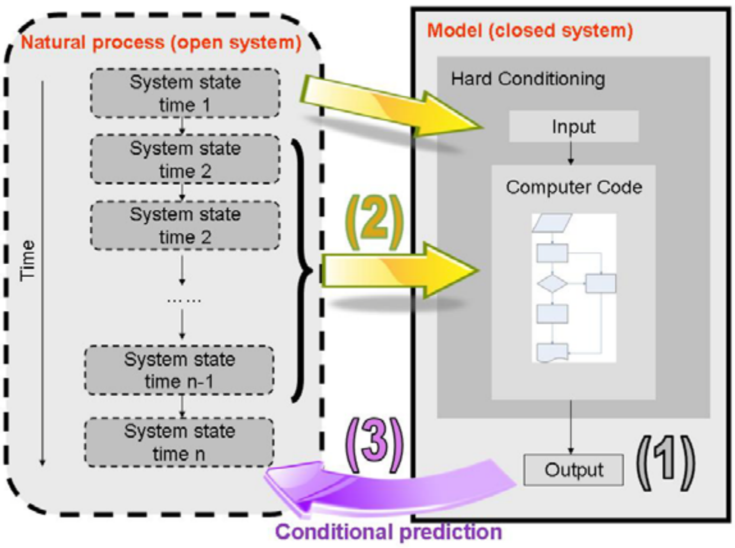

Figure 1. Schematic representation of the concept of conditional prediction and the relation between a model and the natural process under analysis.