2 Kinds of Models

ADAM MORTON AND MAURICIO SUÁREZ
University of Bristol, UK

2.1 THE M WORD

'Model' is a term of the working scientist's self-explanatory and self-justifying vocabulary. 'Here is my model of the phenomenon', 'it follows from our model that . . .', 'our model does not capture the following aspects of the data, but this is no problem since it is just a model . . .', 'to model the data we have made the following assumptions . . .'. ('We are modelling the river as a drunken snake', 'the 3000% discrepancy between predicted sedimentation and observation is very satisfactory for a purely theoretical model.') In some such assertions the word 'model' could be replaced with 'theory' or 'hypothesis' with no big loss of meaning. But in many it could not. There are examples of both in the chapters in this volume, as we argue in Section 2.3. When scientists describe their creations as models they often intend to take advantage of some of the following features of models, as opposed to theories:

- Two models can be inconsistent with each other, and both can be good models.
- A model can contradict some aspects of the observed phenomena and not be refuted.
- A model can contain assumptions which there are theoretical reasons to believe to be false.
- A model can contain assumptions which observation shows to be false.

That makes it sound as if making models is just doing sloppy science. But models are also supposed to have these further features:

- Models are evaluated in accordance with the available data, and rejected if they are inadequate.
- A good model informs us about important properties of its subject matter.

How can we have the features on the second list and also those on the first? How do we get our cake and eat it? It may not be as difficult as it seems, depending on how we interpret the ideas of evaluation (and its variants confirmation and validation or vindication – see Section 2.2) and explanation (and its cousins prediction and derivation). One aspect of modelling builds on the familiar idea of a harmless idealisation, with objects as point masses in Newtonian mechanics, gases as homogeneous fluids or as collections of randomly distributed point masses. It is clear that we can use such idealisations in formulating hypotheses which have explanatory value and can be tested. But some of the false consequences of such hypotheses are not to count as refuting them: those that are direct results of the idealisation rather than of those aspects of the hypothesis that is intended as a description of the subject matter.

Model Validation: Perspectives in Hydrological Science. Edited by M.G. Anderson and P.D. Bates.
© 2001 John Wiley & Sons, Ltd.

The danger now is that all theories will be models, that the theory–model distinction will collapse. Instead of an interesting category of intellectual constructs called 'models' we will have an account of the conditions under which a false consequence does not refute a hypothesis. (This would be worth having, but it could be had without any use of the M word.)

We strongly agree with Bates and Anderson's arguments in this volume (chapter 13) that models are typically propositional and truth-valued – and thus that they can and must be empirically tested and verified. It is very important to remind ourselves of that. An old philosophical tradition used to take it that models were merely interpretations of theories, and therefore could neither be verified nor falsified empirically. Models could not be tested, only theories could. This is a tradition that we oppose, and that we have independently criticised elsewhere. Typically, in putting forward a model, we claim, a scientist (among other things) puts forward a hypothesis – i.e. a set of claims that are subject to empirical scrutiny. But we do not believe that 'model' *just* means 'theory' (or 'hypothesis', or 'assumption'). Modelling is a distinct activity from mere theorising: in addition to putting forward hypotheses, in talking of models, scientists usually are signalling that they are making some very specific uses of idealising assumptions. The problem is that there are many different such signals they can be sending. Here are some worthwhile non-redundant claims that can be intended when a scientists describes her creation as a model:

(a) *Models as tamed theories.* Often there are theoretical reasons for thinking that the laws governing a certain domain should take a certain form, but the result is a theory that is very hard to work with. The difficulty can take the form of problems in deriving practical consequences or problems deriving the kinds of observable predictions which could be evidence for the theory. Then simplifications or idealisations are needed. Fluid mechanics is the obvious example. Newtonian physics conceives of physical systems as complexes of discrete particles, so that when we apply this way of thinking to continua it is not surprising that we get such intractable monsters as the Navier–Stokes equations. (So modelling in hydrology inevitably runs into the hard questions.) A model in this case will be obtained by taking the background theory and changing some of the assumptions that make it badly behaved or hard to understand, hoping that the particular changes made will not produce inaccurate results in the particular application that one has in mind.

In fact there are two distinct directions in which one can tame a theory. One direction is towards intelligibility or intuitive understanding. For example, the equations of general relativity permit mind-bogglingly many solutions. So expositions of the theory usually subtly restrict the range of solutions so that only those that are cosmologically plausible or geometrically manageable are considered. The other direction is towards the deduction of observable consequences. This can be motivated by the need to test the background theory or by practical applications. Most of the models in this book that are tamed models are of this second kind, which we shall refer to as *theory-based models.*

(b) *Models as analogies with other (real) systems.* We model the atom as a planetary system (warily); we model gases as fluids; we model an economy as a collection of independent self-interested perfectly rational agents. In each of these cases we know perfectly well that our assumptions are false. Electrons are much more unlike planets than they are like them, and electrostatic force and gravitation are fundamentally different. Gases are composed of molecules with empty space between them. Economies consist of people of limited rationality who group into coalitions and care very much about the welfare of particular others. We make these false assumptions in part to reduce the possibilities enough that we can begin to describe and explain, in part to make it possible to apply known theories or theory-making techniques, and in part to stimulate our imaginations. The result is then sometimes a fairly simple account, too simple to account for all the data. So as with the models discussed under (a) there is a crucial and problematic distinction between disconfirming evidence which merely shows the obvious fact that the modelling assumptions are not literally true, and more worrying evidence which shows that even taken on its own terms there is something wrong with the model.

Sometimes, though, the attempt to model one kind of phenomena in terms of ideas taken from a different domain produces a result that is far from simple. Thinking in Newtonian terms about fluids results in the Navier–Stokes equations. So then a model of the (a) kind may be needed to tame the results of (b)-type modelling.

(c) *Models as images of data.* Just as a theory can be intractable taken on its own, so can the data that a theory is supposed to explain. A macroeconomic theory, for example, may have as its empirical base the entire statistical corpus of a nation's financial data. There is too much there to begin to explain. So one has to pre-digest the data into a manageable form. Two forms of data models which are so familiar that they might escape notice are generalisations and statistics. A theory might be asked to explain why all businesses over a certain size in a certain sector of the economy were subject to take-over bids, or why the variance of the life-expectancy of new firms in a sector had declined. At first these look like raw data, but in fact they are several steps removed from it. In fact, such assertions are rarely literally true. Thus we say that elementary physics explains why all unsupported bodies near the surface of the earth accelerate downwards at 980 m/sec^2, quite ignoring birds and balloons. We are not lying; we do not really mean that absolutely literally all bodies behave in this way. We mean that for the purposes of understanding the world through physical theory it is a not too misleading summary of the data to represent it in this way. Similarly, *mutatis mutandis*, for sophisticated statistical summaries of data.

Some data models work in a much less obvious way. Quantities that cannot be directly observed, organised in ways that beg questions about their relations, are essential to much of science (see Suppes 1969).

(d) *Models as instances of theories.* Only the very simplest theories are ever completely stated. Put the equations or the axioms, even assuming they exist in a standard form, in front of an intelligent person who knows nothing about the area, and they will usually fall far short of grasping the theory, without all the necessary background facts and unstated assumptions. And there is no uniform way of stating these. They may not all even be stateable in any familiar language. So what is a theory? One fashionable suggestion – see Giere (1988) – is to take a theory to be a collection of abstract objects, such that the theory would be true if the real world resembles any of them. So, for example, we can take a theory in dynamics to consist of a set of trajectories in an abstract space, which is true of real objects in real space if they actually follow trajectories corresponding to those in one model of (or, constituting) the theory. (The model in question may not be one that anyone has described or can imagine, perhaps resulting from initial conditions and forces that are beyond our ken.)

These models are like (a)-models in that they represent reality (and like all representations are selective in what they represent), and like (c)-models in that they are abstract objects, typically mathematical entities, rather than material objects. But in fact (d)-models are very unlike the other kinds of models. They are not shaped by theories – they are (constitutive of) theories. They do not abstract from data – they represent the causal facts not the available observations.

Who cares about a word? These four kinds of models are very different, and used for very different scientific purposes. So there is room for a verbal debunking strategy. That line would run: there are at least four different ways to use the word 'model', with very little in common; it is a linguistic accident that we use the one word in these different ways, and it would be less confusing if we used four different words.

This is a line to take seriously. Science consists of a lot more than theories bearing explanatory relations to bare data. And each of the purposes that one of the model types (a) to (d) serves is a real and necessary one. But that does not mean that the *same* objects do all these things, or even that it is possible for the same objects to do them. The constraints are too different.

It is possible that this is the end of the story. We should stop studying modelling techniques and validation of models in general, perhaps, and start differentiating between the techniques and validations appropriate to different kinds of models. We think there is a lot of truth in this moral. But it is also an exaggeration. To bring out the ways in which it is exaggerated consider the connections between the kinds of models.

One connection has already been mentioned. (b)-type modelling (model-as-analogy) may result in a complexity which needs to be simplified by an (a)-type modelling (model-as-tamed-theory). Moreover, a need for (c)-type modelling (data models) can be created by (a)- or (b)-type modelling. The structures that result from such modelling will by their very nature simplify, abstract away from, or simply ignore some of the causal processes at work in the given domain. As a result there will almost always be evidence with which they are not consistent. So it would be inappropriate to apply to them a 'one strike and you're out' evidential methodology (such as one might get from the writings of Karl Popper or other classic philosophers of science). Instead, the evidence will have to be massaged before it can fairly be compared with the model. It has to be structured in such a way that some of the phenomena, that the model aims to capture, are highlighted, and others hidden. So starting with (a) or (b) we are pushed towards (c).

There are links between other kinds of modelling and the rather dissimilar (d)-type modelling (models-as-instances), too. The range of models that constitute a theory is typically enormous and hard to grasp. So as an approximation to understanding the theory (rather than using it as in (a) or testing it as in (a) or (c)) one takes a small number of relatively easily described models and uses them as paradigms. So with mechanics one uses planetary systems and harmonic oscillators. But even these may be hard to handle or grasp, or more importantly hard to apply to a specific subject matter, so one can make explicit simplifications or idealisations, with a particular application or form of evidence in mind. Therefore, thinking of theories in terms of (d)-type models invites us to make (a)-type models.

Modelling is like sin. Once you begin with one form of it you are pushed to others. In fact, as with sin, once you begin with one form you *ought* to consider other forms. (You have lied to your best friend, to avoid hurting his feelings. Now you have to break a promise to tell him what a third person said about him, in order to keep up the benevolent deception.) But unlike sin – or at any rate unlike sin as a moral purist conceives of it – modelling is the best reaction to the situation in which we find ourselves. Given the meagreness of our intelligence in comparison with the complexity and subtlety of nature, if we want to say things which are true, as well as things that are useful and things which are testable, then we had better relate our bids for truth, application and testability in some fairly sophisticated ways. This is what modelling does.

2.2 MODEL VALIDATION

We now turn to the concept of validation. Oreskes and Belitz suggest in this volume (chapter 3) that the notion is unhelpful, and that the term 'validation' should be abandoned in favour of more neutral terms such as 'evaluation' and 'assessment'. All scientific knowledge is provisional, and at most we can aspire to link a particular model to current modelling practice elsewhere, or to justify the model by appeal to our best theories of natural processes, or to ground the model on a rich and highly constraining set of abundant data. (See Oreskes and Beltz, this volume, chapter 3).

Our claims in this paper are in a sense complementary to theirs: we agree that scientific knowledge is provisional, and that this makes the notion of validation initially suspect. But instead of rejecting the notion, we prefer to fine-grain, by analysing the notion of validation down into five different categories, and showing all of them to be in one way or another legitimate.

Broadly speaking, validation has two different meanings. Validating a model is sometimes taken to involve calibrating it to fit a particular set of data, or physical system, and sometimes it is taken to involve providing a vindication of the model by appeal to high-level theory. Then model calibration can take one of two forms: fixing the parameters, and refining the description. And there are three types of vindication of a model by theory: logical deduction from theory, physical derivation from theory and consistency with theory.

2.2.1 Calibration

A model may be calibrated in order to apply it to either (i) a concrete physical system, or (ii) a specific phenomenon. Examples of calibration in this volume are provided by, among others, Dietrich, Germann and Refsgaard (chapters 6, 10 and 18). Calibration may take one or both of two mathematical forms. The first consists of fixing the parameters in the model in order to fit the data that result from experiments or observations on the system, or to fit the main characteristics of interest of the phenomenon in question. This is typically just a matter of specifying the boundary conditions for the equations in the model, and it is relatively straightforward. It is arguable that the data can never suffice to fix all parameters, especially so when the data is statistical, but it nonetheless constrains the values that the parameters can take. So this type of calibration, let us call it *calibration-1*, consists of explicitly introducing formal constraints on the values of the parameters in the model.

The second form that calibration may take is the introduction of correction factors in the model to account for causal variables known to be part of the physical system, or to be efficacious in the phenomenon. This involves actually tinkering with the equations in the model, or at least making additional assumptions beyond those strictly speaking contained in the model about the real system to be modelled. The goal of this process is a *refinement* of the model, in order to bring it closer to the physical situation at hand.

This second form of calibration, or *calibration-2*, is philosophically and methodologically the most interesting. It involves a process of concretisation: an abstract model is made more concrete by introducing causal factors that permit its application to a particular system or specific phenomenon. The important fact to note, however, is that for a correct calibration these correction factors need not be suggested by any well-established theory; they may instead be suggested by tacit knowledge of past model-building practice, or specific knowledge of the causal properties of the system or phenomenon at hand. (For example, Germann – this volume, Section 10.3 – in discussing the 'validation of momentum dissipation approach' introduces a number of causal hypothesis, such as 'recognising that some of the input water may get stuck between the surface and the depth of observation').

It is important to distinguish here between idealisation and abstraction (see Cartwright 1989, chapter 5). A model is idealised if it says something literally false about some of the features of the system that it is intended for, and it is abstract if it remains silent about a number of its features, and says nothing literally false. Everything turns on the presuppositions: from the formal point of view these two models may look identical. Consider some classical mechanical models of a pendulum. The simple harmonic oscillator, for instance, has no term to represent the slowing down of the pendulum due to the friction in the pendulum's motion through air. If, in entertaining this model, we presuppose that a pendulum suffers no air-friction at all, we idealise. If, on the other hand, we simply ignore the question of friction, and refuse to make any assumptions regarding it, we abstract.

A useful way to characterise this difference is due to Ronald Giere (see Giere 1988, chapter 3, and Giere 1999, chapter 6). We say that an idealised model *speaks about the real world*, and that it makes some false claims about it. The classical mechanical model of a simple harmonic oscillator, the SHO, if an idealisation, makes claims about real-life pendulums, but not all the claims it makes are true – for instance it makes the false claim that real-life pendulums are subject to no air-friction. An abstract model, on the other hand, does not lie. But this is *because it does not speak about any object in the real world*, only about an abstract object, the one implicitly defined by the equations in the model. Considered as an abstraction, thus, the SHO model would speak truly about an abstract object, namely the simple harmonic oscillator, defined implicitly by its equation of motion.

We can employ Giere's framework to make further headway on the relation between models and theories, and to understand the kind of procedure involved in calibration. Idealisation requires that there be a (true) theory, available to us, describing the corrections needed to apply the model to every concrete system and specific phenomenon. For only then can we be sure that the model makes *false* claims about the system, and can we try to correct these claims. Classical mechanics, for instance, recommends modelling air-friction as a linear function of velocity. So it tells us exactly the way in which the SHO idealises. It is because of this that the SHO is, properly speaking, an idealisation, not an abstraction.

But in most cases of calibration – such as those discussed by Dietrich, Germann and Refsgaard in this volume – no such theory is available. The corrections come from a varied mixture of theoretical and non-theoretical modelling techniques. We have no established body of knowledge to compare the model with, and thus we have no means to tell if the model is speaking falsely about the world. It is the availability of such background knowledge that turns an equation into an idealisation; as we have seen there is nothing in the formal description a model M that will tell us that M is an idealisation of some phenomenon P. So, properly speaking, most of the models that require validation in the calibration-2 sense are not idealised, but abstract. And calibration-2 is a process that makes an abstract model concrete.

2.2.2 Theory Vindication

A model can also be said to be validated if it can be shown to be a consequence (logical or a physical consequence) of some established theory, or if it can be shown to be consistent with some established theory.

Let us carefully define these terms. A logical consequence of a theory is a statement that is consistent with the theory and whose contrary is inconsistent with the theory. In other words, as a theory logically dictates all of its logical consequences, a contradiction can only be a logical consequence of an inconsistent theory. Logical deduction provides the strongest type of theoretical vindication; let us call this *vindication-1*. But genuine examples of vindication-1 in science are surprisingly hard to find. Astronomy provides some: Kepler's laws, which describe the ratios of the planets' periods and distances from the sun, for instance, follow deductively from Newton's theory, and they are thus validated by Newtonian mechanics.

On the other hand, the contrary of a theory's physical consequence, as we wish to define the notion here, is not necessarily inconsistent with the theory. The theory cannot logically dictate, but merely suggest, its physical consequences, and a contradiction may be a 'physical consequence' of a perfectly logically consistent theory.

This may seem at first counterintuitive, but it is in fact only counterintuitive in light of the assumptions of determinism and completeness that many of us have been schooled into. The assumption of determinism states that at its most fundamental level nature is always deterministic, statistics only enters the picture in order to account for our lack of knowledge. But consider genuinely irreducible statistical theories. Quantum theory, for instance, even if complete, is only able to tell us that the electron just created in a cascade experiment in the laboratory will be found at end E of the lab with probability p, or at end E* with probability p*; where, say, p = 1 − p*. Then the statements 'the electron will be at E' and 'the electron will be at E*' are both physical consequences of the theory, even if, of course, when the experiment is performed, the electron appears either at E or at E*, but never simultaneously at E and E*. Thus a model of the electron in E may be vindicated by showing it to be a physical consequence of the fundamental quantum theory, even if another model of the inconsistent claim that the electron is at E* can similarly be vindicated.

We claim that this type of vindication of a model as 'physical consequence' of a theory (let us call it *vindication-2*) is not unique to statistical theories. It is by now common lore that the relation of theory and evidence is rather more subtle and complex than simple logical deduction, even for deterministic theories. There is a gap between theory and evidence that, we claim, the notion of physical consequence can fill. This is because the theory does not describe all the facts that are conditions for the occurrence of its physical consequences. The theory, in other words, is not complete. As a result of this, the theory cannot logically dictate its physical consequences but, at most, suggest them. The assumption of completeness would rule out theories that are not able to describe all the relevant facts in their domain. Combined with the assumption of determinism this gives us the result that all consequences of a theory must be consistent with each other. But the result does not follow if any of these two assumptions is abandoned. And there are good reasons to think that our best theories reject not just determinism but also completeness. (For arguments against determinism, see, for instance, Dupre (1993) and Suppes (1984). For arguments against completeness see Cartwright (1999).)

Why is this relevant? Vindication of a model by a theory of which it is shown to be a logical consequence provides confirmation for the theory. In those instances, the model, if true, confirms the theory, and, if false, falsifies it. This is a strong relation, and difficult to ever find in practice. When it is found, the model that the theory vindicates provides a resounding verification (or falsification, if the model turns out to be incorrect) of the theory, and vice versa. The observational correctness of Kepler's laws verifies Newtonian mechanics, and Newtonian mechanics lends confirmation to Kepler's laws. However, very often a theory does not logically dictate a model of a phenomenon; at most, and with some luck, it suggests it. The model may then increase the probability of the theory being true. If it does (i.e. if Prob (T/M) > Prob (T/not M)), we can still say that the model confirms the theory, although only by constituting *evidence* for T. Bayesianism explains well how this works. The theory is more likely to be true if the model is true (and vice versa). But if the model turns out to be false the theory could still be true: the theory and the contrary of the model are not logically inconsistent. (This is so unless (i) Prob (T/not M) = 0 and (ii) Prob (T) > 0.) The quantum mechanical example of the position of an electron is arguably such a case. Both the observation that 'the electron is at E' and the observation that 'it is at E*' can constitute positive evidence for the theory, even if they are contrary.

The logical consequences of a theory may increase its confirmation. But so can the physical consequences of a theory. Both can increase the degree of confirmation of the theory. These possibilities are well known to philosophers, and they are well understood. But sometimes, perhaps often, the truth of the model does not increase the probability of the theory being true even if the theory vindicates, and in this sense validates, the model. Even if a theory can be applied to a model of a phenomenon, it may not be able to suggest the corrections required for the application. (See Suárez (1999), where an example from the physics of superconductivity is discussed.) If the theory does not even suggest a model, then the success of the model will not provide any confirmation for the theory.

To many philosophers this is deeply puzzling – how is it possible? The contrast that we drew earlier between idealisation and concretisation may help here. A model M that can only be applied to a phenomenon P by introducing corrections into M that are not suggested by any established theory T is not an idealisation of P (or not an idealisation relative to P). M is, at best, an abstraction relative to P. And the process of correction of the model is therefore not a de-idealisation, but a concretisation of the model. It follows that the model does not speak about P, and therefore T cannot be confirmed by M, either by strict verification or by the Bayesian methods of probabilistic confirmation. What this entails is that, in this case, M is neither a *logical* consequence nor a *physical* consequence of T. The most we can do is to show that M and T are consistent. This proof of consistency is, we claim, a legitimate form of theoretical vindication – and in fact a most interesting and common one. Let us refer to it as *vindication-3*. Examples of vindication-3 in the present volume are the non-cohesive sediment transport models by Garcia (chapter 15), the more general among the snow models described by Davis et al. (chapter 11), and some of the simulations in Booltink (chapter 9). The conceptual and physics-based models of Young (chapter 7) may be employing vindication-3 or vindication-2.

This completes our analysis of the notion of model validation. We have seen that validation may take one of five forms: (i) calibration by fixing boundary conditions; (ii) calibration by refinement, (iii) vindication by logical deduction (from a theory), (iv) vindication by physical derivation (from a theory), and (v) vindication by proof of consistency (with a theory).

2.2.3 When Vindication is Calibration

Even if these five forms of validation are clearly distinct, presenting different practical difficulties and requiring different solutions, they have often been conflated. We think that we can diagnose why. It turns out that the most common and complex form of calibration, calibration-2, is methodologically very similar to the most common and complex form of vindication, vindication-3. Vindication-3 and calibration-2 are alike in that they do not involve idealisation, but concretisation. Even if there is a body of theoretical knowledge available to us that we can compare our model to, we are nonetheless unable to use it in order to correct our model to make it applicable to the phenomenon. The corrections required to apply the model come from other sources. This is true of calibration-2 by definition. It turns out to be true also of vindication-3. Remember that in vindicating-3 a model by appeal to a theory, we show the model and theory to be consistent because we cannot show the theory to suggest the model, or the corrections required to apply the model to a particular physical system. It follows that the corrections required to apply the model to the physical system are suggested by other sources.

In both cases the same broad methodology is used: consider the model, consider the object to be modelled, and with patience and ingenuity bring to bear on the modelling situation whatever knowledge you may have – knowledge that is typically not already contained in the high-level theory assumed to be true in the domain. So, even if fluid dynamics is the underlying true theory of hydrological phenomena, validation in hydrological models must follow a course that essentially ignores fluid dynamics, and takes it at most as one more tool in model-building. And what we have been arguing is that this is neither exceptional nor uncommon in the practice of model-building in science at large.

2.3 SIMULATION VERSUS ESTIMATION

This chapter has been an exercise in distinction-making. There are distinct things one can mean by 'model' and distinct enterprises of model validation. In fact there are five of each. So any particular model validation should find a place in one of the cells of a five by five grid. And – if our account of modelling is along the right lines – the characteristic difficulties of modelling should respect the distinctions between the cells. Difficulties characteristic of one category should – if there is anything to what we are saying – be largely absent in another. (Just about any data can be squeezed and shoved into just about any purely descriptive framework; a stricter test is whether the framework captures significant causal generalisations.) Let us consider some data: the chapters on hydrological modelling in this book. But before we confront the data, a simplification of the account is in order.

To begin, as remarked in Section 2.1, models in hydrological science are very rarely tamed theories where the taming is motivated by the need for understanding or exposition. For hydrology does not aim to challenge fluid dynamics. It assumes fluid dynamics is fundamentally right, though not much use in its pure form for predicting and explaining actual flows of water in actual geological conditions. Moreover, hydrological scientists rarely worry about models as instances of theories. That is not to say that philosophers or psychologists of science may not interpret their patterns of thought in terms of them. But the scientists themselves rarely take such things as conscious objects of concern. (Nor should they.) And while models as analogies may play an important role in the exposition and motivation of hydrological theories, hydrological scientists do not spend a lot of time trying to fix their exact form or decide between rivals. If metaphorical models do their job of suggesting or communicating, they will be used, and questions of their accuracy or uniqueness will not arouse much concern. And this, too, is as it should be, as uniqueness can be expected to reside only in what actual water flows do and in what the underlying laws of fluid motion are.

That leaves us with just two large categories of models: what in Section 2.1 we called theory-based models, models as theories tamed for evidence and prediction, and data models. If what we are saying is right, there should be different kinds of problems which arise in the four main possibilities: theory-based/calibration, theory-based/vindication, data model/calibration and data model/vindication. And indeed there do seem to be two characteristic methodological problems, each of which arises in just one cell of the grid.

The first problem is that of simulation. As argued in the chapters by Beven and by Young (chapters 4 and 7), one frequent hydrological situation is that one has a predictive model that does not generate good simulations. That is, one has a model in which the values of crucial parameters have been fixed in such a way that it predicts the output of a particular physical system quite accurately. But when one tries to generate data for different related systems problems arise. Either the predictions that emerge are clearly wrong, or there is no clear way of determining the appropriate values of the parameters. It is as if the predictively adequate model works because of a specific combination of parameter values, but once one parameter is varied suitable combinations of the others are not forthcoming.

The simulation problem is a crucial hazard of data models, as is clear from Beven and Young. With theory-based models, though simulation may be problematic, there will not typically be a prediction/simulation asymmetry. Neither need be easy, but the problems should be the same. As Germann's chapter (10) argues, the problem that arises with theory-based models is not comprehensive prediction for one system that does not transfer to others, but adequate prediction and simulation for one kind of data that does not transfer to other kinds of data.

The simulation problem is moreover a problem primarily of vindication rather than of calibration. As applied to a given physical system the given parameter values are fine, but if they do not generate reasonable *counterfactual* consequences — specifications of what the system would have done under other circumstances — then they cannot be taken as a way of describing the real causes of the phenomena. (In this connection see Section 1.1 of Dietrich's chapter (6).) Without some simulatory power a power is 'only' a description of the data. The other characteristic problem is best illustrated by Garcia's chapter (15), on sediment transport (though particular models described in several other chapters would also have made the point). Here the model consists of a highly theoretical treatment of a widespread phenomenon and the problem is to make a connection between the some undetermined theoretical quantity and some quantity whose value can be estimated from the behaviour of a particular system, using either raw data or a suitable data model. The problem is essentially a statistical one, of inferring the true value of some quantity which, if the theoretical assumptions are right, will correspond to an unknown in an equation derived by artful simplification from the underlying physics. Call this the problem of truthful estimation.

Truthful estimation problems will occur primarily in connection with theory-based models. For given that the underlying theory characterises correctly the underlying reality, the value of the unknowns to be filled in is what they actually are (or what an infinite set of them are, as when one is specifying a boundary condition). And such problems will occur primarily when the aim is calibration rather than vindication. For vindication is conceptually simple with theory-based models: it consists of showing that the model is deducible from, physically derivable from, or consistent with the underlying physics. These may not be simple tasks, but their nature is clear, and they do not involve any pure estimation.

The four cells are thus characterised as follows:

- theory-based/calibration = (+ truthful estimation, − simulation)
- theory-based/vindication = (− truthful estimation, − simulation)
- data model/calibration = (− truthful estimation, − simulation)
- data model/vindication = (− truthful estimation, + simulation)

Of the four combinations of kinds of model and kinds of validation, two receive unique characterisations in this way, and two are not distinguished. Appeal to these two kinds of problem does not distinguish 'theory-based/vindication' from 'data model/calibration'. But those two are distinguished in other ways. For, as remarked above, the characteristic problems of 'theory-based/vindication' are those of theoretical ingenuity and mathematical sophistication, while those of 'data model/calibration' are those of physical intuition, and experimental skill and ingenuity. In other words, in the first case one is looking for assurance that a model is an accurate portrayal of reality as characterised by a background theory, so the basic question is just 'does it follow?' (or, since these are models 'how closely does it follow?'), and in the second one the basic question is just 'how can we find out?'. (Note that the situation described is one where one already has a data model and is trying to calibrate its parameters. In the construction of data models, on the other hand, no end of theoretical ingenuity and mathematical sophistication can enter.) So, representing this dimension, admittedly rather crudely, as the relevance or not of empirical data, we can complete the grid as follows.

- theory-based/calibration = (+ truthful estimation, − simulation, + empirical)
- theory-based/vindication = (− truthful estimation, − simulation, − empirical)
- data model/calibration = (− truthful estimation, − simulation, + empirical)
- data model/vindication = (− truthful estimation, + simulation, + empirical)

As model of modelling, this can only be a first attempt. The most general conclusion to draw is that modelling in any scientific domain is a very varied as well as a very challenging activity. The variability is as important as the challenge, since the kinds of difficulties that modellers face are not randomly distributed. Different kinds of models lead to different kinds of challenge.

REFERENCES

Cartwright, N. 1989. *Nature's Capacities and their Measurement*. Oxford University Press, Oxford.

Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge University Press, Cambridge.

Dupre, J. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Harvard University Press, Cambridge, MA.

Giere, R. 1988. *Explaining Science: A Cognitive Approach*. University of Chicago Press, Chicago.

Giere, R. 1999. *Science Without Laws*. University of Chicago Press, Chicago.

Suárez, M. 1999. The role of models in the application of scientific theories: epistemological implications. In: Morrison, M. and Morgan, M. (eds), *Models as Mediators*. Cambridge University Press, Cambridge, 168–196.

Suppes, P. 1969. Models of Data. In: P. Suppes, *Studies in the Methodology and Foundations of Science*. Reidel Publishing Company, Dordrecht, 24–35.

Suppes, P. 1984. *Probabilistic Metaphysics*. Blackwell, Oxford.