# 1

# Modelling and Model Building

MARK MULLIGAN AND JOHN WAINWRIGHT

*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. . . . 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 which 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.*

(Morton and Suárez, 2001: 14)

## 1.1 THE ROLE OF MODELLING IN ENVIRONMENTAL RESEARCH

### 1.1.1 The nature of research

Research is a means of improvement through understanding. This improvement may be personal, but it may also be tied to broader human development. We may hope to improve human health and well-being through research into diseases such as cancer and heart disease. We may wish to improve the design of bridges or aircraft through research in materials science, which provides lighter, stronger, longer-lasting or cheaper bridge structures (in terms of building and of maintenance). We may wish to produce more or better crops with fewer adverse impacts on the environment through research in biotechnology. In all of these cases, research provides in the first instance better understanding of how things are and how they work, which can then contribute to the improvement or optimization of these systems through the development of new techniques, processes, materials and protocols.

Research is traditionally carried out through the accumulation of observations of systems and system behaviour under 'natural' circumstances and during experimental manipulation. These observations provide the evidence upon which hypotheses can be generated about the structure and operation (function) of the systems. These hypotheses can be tested against new observations and, where they prove to be reliable descriptors of the system or system behaviour, then they may eventually gain recognition as tested theory or general law.

The conditions, which are required to facilitate research, include:

1. a means of observation and comparative observation (measurement);
2. a means of controlling or forcing aspects of the system (experimentation);
3. an understanding of previous research and the state of knowledge (context);
4. a means of cross-referencing and connecting threads of 1, 2 and 3 (imagination).

### 1.1.2 A model for environmental research

What do we mean by the term *model*? A model is an abstraction of reality. This abstraction represents a complex reality in the simplest way that is adequate for the purpose of the modelling. The best model is always that which achieves the greatest realism (measured objectively as agreement between model outputs

*Environmental Modelling: Finding Simplicity in Complexity*. Edited by J. Wainwright and M. Mulligan
 ISBNs: 0-471-49617-0 (HB); 0-471-49618-9 (PB)

and real-world observations, or less objectively as the process insight gained from the model) with the least parameter complexity and the least model complexity.

Parsimony (using no more complex a model or representation of reality than is absolutely necessary) has been a guiding principle in scientific investigations since Aristotle who claimed: 'It is the mark of an instructed mind to rest satisfied with the degree of precision which the nature of the subject permits and not to seek an exactness where only an approximation of the truth is possible' though it was particularly strong in medieval times and was enunciated then by William of Ockham, in his famous 'razor' (Lark, 2001). Newton stated it as the first of his principles for fruitful scientific research in *Principia* as: 'We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances.'

Parsimony is a prerequisite for scientific explanation, not an indication that nature operates on the basis of parsimonious principles. It is an important principle in fields as far apart as taxonomy and biochemistry and is fundamental to likelihood and Bayesian approaches of statistical inference. In a modelling context, a parsimonious model is usually the one with the greatest explanation or predictive power and the least parameters or process complexity. It is a particularly important principle in modelling since our ability to model complexity is much greater than our ability to provide the data to parameterize, calibrate and validate those same models. Scientific explanations must be both relevant *and* testable. Unvalidated models are no better than untested hypotheses. If the application of the principle of parsimony facilitates validation, then it also facilitates utility of models.

### 1.1.3 The nature of modelling

Modelling is not an alternative to observation but, under certain circumstances, can be a powerful tool in understanding observations and in developing and testing theory. Observation will always be closer to truth and must remain the most important component of scientific investigation. Klemeš (1997: 48) describes the forces at work in putting the modelling 'cart' before the observation 'horse' as is sometimes apparent in modelling studies:

> *It is easier and more fun to play with a computer than to face the rigors of fieldwork especially hydrologic fieldwork, which is usually most intensive during the most adverse conditions. It is faster to get a result by modeling than through acquisition and analysis of more data, which suits managers and politicians as well as staff scientists and professors to whom it means more publications per unit time and thus an easier passage of the hurdles of annual evaluations and other paper-counting rituals. And it is more glamorous to polish mathematical equations (even bad ones) in the office than muddied boots (even good ones) in the field.*
>
> (Klemeš, 1997: 48)

A model is an abstraction of a real system, it is a simplification in which only those components which are seen to be significant to the problem at hand are represented in the model. In this, a model takes influence from aspects of the real system and aspects of the modeller's perception of the system and its importance to the problem at hand. Modelling supports in the conceptualization and exploration of the behaviour of objects or processes and their interaction as a means of better understanding these and generating hypotheses concerning them. Modelling also supports the development of (numerical) experiments in which hypotheses can be tested and outcomes predicted. In science, understanding is the goal and models serve as tools towards that end (Baker, 1998).

Cross and Moscardini (1985: 22) describe modelling as 'an art with a rational basis which requires the use of common sense at least as much as mathematical expertise'. Modelling is described as an art because it involves experience and intuition as well as the development of a set of (mathematical) skills. Cross and Moscardini argue that intuition and the resulting insight are the factors which distinguish good modellers from mediocre ones. Intuition (or imagination) cannot be taught and comes from the experience of designing, building and using models. Tackling some of the modelling problems presented on the website which complements this book will help in this.

### 1.1.4 Researching environmental systems

Modelling has grown significantly as a research activity since the 1950s, reflecting conceptual developments in the modelling techniques themselves, technological developments in computation, scientific developments in response to the increased need to study systems (especially environmental ones) in an integrated manner, and an increased demand for extrapolation (especially prediction) in space and time.

Modelling has become one of the most powerful tools in the workshop of environmental scientists who are charged with better understanding the interactions between the environment, ecosystems and the populations of humans and other animals. This understanding is increasingly important in environmental stewardship (monitoring and management) and the development of increasingly sustainable means of human dependency on environmental systems.

Environmental systems are, of course, the same systems as those studied by physicists, chemists and biologists but the level of abstraction of the environmental scientist is very different from many of these scientists. Whereas a physicist might study the behaviour of gases, liquids or solids under controlled conditions of temperature or pressure and a chemist might study the interaction of molecules in aqueous solution, a biologist must integrate what we know from these sciences to understand how a cell – or a plant or an animal – lives and functions. The environmental scientist or geographer or ecologist approaches their science at a much greater level of abstraction in which physical and chemical 'laws' provide the rule base for understanding the interaction between living organisms and their nonliving environments, the characteristics of each and the processes through which each functions.

Integrated environmental systems are different in many ways from the isolated objects of study in physics and chemistry though the integrated study of the environment cannot take place without the building blocks provided by research in physics and chemistry. The systems studied by environmental scientists are characteristically:

- *Large-scale, long-term.* Though the environmental scientist may only study a small time-scale and space-scale slice of the system, this slice invariably fits within the context of a system that has evolved over hundreds, thousands or millions of years and which will continue to evolve into the future. It is also a slice that takes in material and energy from a hierarchy of neighbours from the local, through regional, to global scale. It is this context which provides much of the complexity of environmental systems compared with the much more reductionist systems of the traditional 'hard' sciences. To the environmental scientist, models are a means of integrating across time and through space in order to understand how these contexts determine the nature and functioning of the system under study.
- *Multicomponent.* Environmental scientists rarely have the good fortune of studying a single component of their system in isolation. Most questions asked of environmental scientists require the understanding of interactions between multiple living (biotic) and nonliving (abiotic) systems and their interaction. Complexity increases greatly as the number of components increases, where their interactions are also taken into account. Since the human mind has some considerable difficulty in dealing with chains of causality with more than a few links, to an environmental scientist models are an important means of breaking systems into intellectually manageable components and combining them and making explicit the interactions between them.
- *Nonlaboratory controllable.* The luxury of controlled conditions under which to test the impact of individual forcing factors on the behaviour of the study system is very rarely available to environmental scientists. Very few environmental systems can be re-built in the laboratory (laboratory-based physical modelling) with an appropriate level of sophistication to adequately represent them. Taking the laboratory to the field (field-based physical modelling) is an alternative, as has been shown by the Free Atmosphere $CO_2$ Enrichment (FACE) experiments (Hall, 2001), BIOSPHERE 2 (Cohn, 2002) and a range of other environmental manipulation experiments. Field-based physical models are very limited in the degree of control available to the scientist because of the enormous expense associated with this. They are also very limited in the scale at which they can be applied, again because of expense and engineering limitations. So, the fact remains that, at the scale at which environmental scientists work, their systems remain effectively noncontrollable with only small components capable of undergoing controlled experiments. However, some do argue that the environment itself is one large laboratory, which is sustaining global-scale experiments through, for example, greenhouse gas emissions (Govindasamy *et al.*, 2003). These are not the kind of experiments that enable us to predict (since they are real-time) nor which help us, in the short term at least, to better interact with or manage the environment (notwithstanding the moral implications of this activity!). Models provide an inexpensive laboratory in which mathematical descriptions of systems and processes can be forced in a controlled way.
- *Multiscale, multidisciplinary.* Environmental systems are multiscale with environmental scientists needing to understand or experiment at scales from the atom through the molecule to the cell, organism or object, population of objects, community or landscape through to the ecosystem and beyond. This presence of multiple scales means that environmental scientists are rarely just environmental scientists, they may be physicists, chemists, physical chemists, engineers,

biologists, botanists, zoologists, anthropologists, population geographers, physical geographers, ecologists, social geographers, political scientists, lawyers, environmental economists or indeed environmental scientists in their training but who later apply themselves to environmental science. Environmental science is thus an interdisciplinary science which cuts across the traditional boundaries of academic research. Tackling contemporary environmental problems often involves large multidisciplinary (and often multinational) teams working together on different aspects of the system. Modelling provides an integrative framework in which these disparate disciplines can work on individual aspects of the research problem and supply a module for integration within the modelling framework. Disciplinary and national boundaries, research 'cultures' and research 'languages' are thus no barrier.
- *Multivariate, nonlinear and complex.* It goes without saying that integrated systems such as those handled by environmental scientists are multivariate and as a result the relationships between individual variables are often nonlinear and complex. Models provide a means of deconstructing the complexity of environmental systems and, through experimentation, of understanding the univariate contribution to multivariate complexity.

In addition to these properties of environmental systems, the rationale behind much research in environmental systems is often a practical or applied one such that research in environmental science also has to incorporate the following needs:

- *The need to look into the future*. Environmental research often involves extrapolation into the future in order to understand the impacts of some current state or process. Such prediction is difficult, not least because predictions can only be tested in real time. Models are very often used as a tool for integration of understanding over time and thus are well suited for prediction and postdiction. As with any means of predicting the future, the prediction is only as good as the information and understanding upon which it is based. While this understanding may be sufficient where one is working within process domains that have already been experienced during the period in which the understanding was developed, when future conditions cross a process domain, the reality may be quite different to the expectation.
- *The need to understand the impact of events that have not happened (yet)*. Environmental research very often concerns developing scenarios for change and understanding the impacts of these scenarios on systems upon which humans depend. These changes may be developmental such as the building of houses, industrial units, bridges, ports or golf courses and thus require environmental impact assessments (EIAs). Alternatively, they may be more abstract events such as climate change or land use and cover change (LUCC). In either case where models have been developed on the basis of process understanding or a knowledge of the response of similar systems to similar or analogous change, they are often used as a means of understanding the impact of expected events.
- *The need to understand the impacts* ***of*** *human behaviour.* With global human populations continuing to increase and *per capita* resource use high and increasing in the developed world and low but increasing in much of the developing world, the need to achieve renewable and nonrenewable resource use that can be sustained into the distant future becomes more and more pressing. Better understanding the impacts of human resource use (fishing, forestry, hunting, agriculture, mining) on the environment and its ability to sustain these resources is thus an increasing thrust of environmental research. Models, for many of the reasons outlined above, are often employed to investigate the enhancement and degradation of resources through human impact.
- *The need to understand the impacts* ***on*** *human behaviour.* With human population levels so high and concentrated and with *per capita* resource needs so high and sites of production so disparate from sites of consumption, human society is increasingly sensitive to environmental change. Where environmental change affects resource supply, resource demand or the ease and cost of resource transportation, the impact on human populations is likely to be high. Therefore understanding the nature of variation and change in environmental systems and the feedback of human impacts on the environment to human populations is increasingly important. Environmental science increasingly needs to be a supplier of reliable forecasts and understanding to the world of human health and welfare, development, politics and peacekeeping.

## 1.2 MODELS WITH A PURPOSE (THE PURPOSE OF MODELLING)

Modelling is thus the canvas of scientists on which they can develop and test ideas, put a number of ideas together and view the outcome, integrate and communicate those ideas to others. Models can play one or more of many roles, but they are usually developed with one or two roles specifically in mind. The type of model built will, in some way, restrict the uses to which the model may be put. The following seven headings outline the purposes to which models are usually put:

1. *As an aid to research.* Models are now a fairly commonplace component of research activities. Through their use in assisting understanding, in simulation, as a virtual laboratory, as an integrator across disciplines and as a product and means of communicating ideas, models are an aid to research activities. Models also facilitate observation and theory. For example, understanding the sensitivity of model output to parameter uncertainty can guide field data-collection activities. Moreover, models allow us to infer information about unmeasurable or expensively measured properties through modelling them from more readily measured variables that are related in some way to the variable of interest.
2. *As a tool for understanding.* Models are a tool for understanding because they allow (indeed, require) abstraction and formalization of the scientific concepts being developed, because they help tie together ideas that would otherwise remain separate and because they allow exploration of the outcomes of particular 'experiments' in the form of parameter changes. In building a model, one is usually forced to think very clearly about the system under investigation and in using a model one usually learns more about the system than was obvious during model construction.
3. *As a tool for simulation and prediction.* Though there are benefits to be gained from building models, their real value becomes apparent when they are extensively used for system simulation and/or prediction. Simulation with models allows one to integrate the effects of simple processes over complex spaces (or complex processes over simple spaces) and to cumulate the effects of those same processes (and their variation) over time. This integration and cumulation can lead to the prediction of system behaviour outside the time or space domain for which data are available. This integration and cumulation are of value in converting a knowledge or hypothesis of process into an understanding of the outcome of this process over time and space – something that is very difficult to pursue objectively without modelling. Models are thus extensively employed in extrapolation beyond measured times and spaces, whether that means prediction (forecasting) or postdiction (hindcasting) or near-term casting (nowcasting) as is common in meteorology and hydrology. Prediction using models is also a commonly used means of making better any data that we do have through the interpolation of data at points in which we have no samples, for example, through inverse distance weighting (IDW) or kriging techniques. Furthermore, an understanding of processes can help us to model high resolution data from lower resolution data as is common in climate model downscaling and weather generation (Bardossy, 1997; see also Chapters 2 and 19).
4. *As a virtual laboratory.* Models can also be rather inexpensive, low-hazard and space-saving laboratories in which a good understanding of processes can support model experiments. This approach can be particularly important where the building of hardware laboratories (or hardware models) would be too expensive, too hazardous or not possible (in the case of climate-system experiments, for example). Of course, the outcome of any model experiment is only as good as the understanding summarized within the model and thus using models as laboratories can be a risky business compared with hardware experimentation. The more physically based the model (i.e. the more based on proven physical principles), the better in this regard and indeed the most common applications of models as laboratories are in intensively studied fields in which the physics are fairly well understood such as computational fluid dynamics (CFD) or in areas where a laboratory could never be built to do the job (climate-system modelling, global vegetation modelling).
5. *As an integrator within and between disciplines.* As we will see in Chapters 14 and 15, models also have the ability to integrate the work of many research groups into a single working product which can summarize the understanding gained as well as, and sometimes much better than, traditional paper-based publications. Understanding environmental processes at the level of detail required to contribute to the management of changing environments requires intensive specialization by individual scientists at the same time as the need to approach environmental research in an increasingly multidisciplinary way. These two can be quite incompatible. Because of these two requirements, and because of funding pressures in this direction, scientific research is increasingly a collaborative process whereby large grants fund integrated analysis of a particular environmental issue by tens or hundreds of researchers from different scientific fields, departments, institutions, countries and continents working together and having to produce useful and consensus outcomes, sometimes after only three years. This approach of big science is particularly clear if we look at the scientific approach of the large UN conventions: climate change, biological diversity, desertification and is also evident

in the increasingly numerous authorship on individual scientific papers.

Where archaeologists work with hydrologists working with climate scientists working with ecologists and political scientists, the formal language of mathematics and the formal data and syntax requirements of models can provide a very useful language of communication. Where the research group can define a very clear picture of the system under study and its subcomponents, each contributing group can be held responsible for the production of algorithms and data for a number of subcomponents and a team of scientific integrators is charged with the task of making sure all the subcomponents or modules work together at the end of the day. A team of technical integrators is then charged with making this mathematical construct operate in software. In this way the knowledge and data gained by each group are tightly integrated where the worth of the sum becomes much more than the worth of its individual parts.

6. *As a research product*. Just as publications, websites, data sets, inventions and patents are valid research products, so are models, particularly when they can be used by others and thus either provide the basis for further research or act as a tool in practical environmental problem solving or consultancy. Equally, models can carry forward entrenched ideas and can set the agenda for future research, even if the models themselves have not been demonstrated to be sound. The power of models as research products can be seen in the wealth of low cost publicly available models especially on the online repositories of models held at the CAMASE Register of Agro-ecosystems Models (http://www.bib.wau.nl/camase/srch-cms.html) and the Register of Ecological Models (http://eco.wiz.uni-kassel.de/ecobas.html). Furthermore, a year-by-year analysis of the number of English language academic research papers using one prominent, publicly available hydrological model (TOPMODEL) indicates the amount of research that can stem from models. According to ISI, from 1991 to 2002 inclusive, some 143 scientific papers were published using TOPMODEL, amounting to more than 20 per year from 1997 to 2002 (this figure therefore does not include the rash of papers in conference proceedings and other nonpeer-reviewed publications). Models can also be very expensive 'inventions', marketed to specialist markets in government, consultancy and academia, sometimes paying all the research costs required to produce them, often paying part of the costs.

7. *As a means of communicating science and the results of science*. To write up science as an academic paper, in most cases, confines it to a small readership and to fewer still users. To add the science as a component to a working model can increase its use outside the research group that developed it. In this way, models can make science and the results of science more accessible both for research and for education. Models can be much more effective communicators of science because, unlike the written word, they can be interactive and their representation of results is very often graphical or moving graphical. If a picture saves a thousand words, then a movie may save a million and in this way very complex science can be hidden and its outcomes communicated easily (but see the discussion below on the disadvantages in this approach).

Points 1–4 above can be applied to most models while 5, 6 and 7 apply particularly to models that are interface-centred and focused on use by end users who are not the modellers themselves. In environmental science these types of model are usually applied within the context of policy and may be called 'policy models' that can be used by policy advisers during the decision-making or policy-making process. They thus support decisions and could also be called decision-support systems (DSS: see Chapters 14 and 15) and which perform the same task as do policy documents, which communicate research results. The hope is that policy models are capable of doing this better, particularly in the context of the reality of scientific uncertainty compared with the myth that policy can be confidently based on exact, hard science (Sarewitz, 1996). The worry is that the extra uncertainties involved in modelling processes (parameter uncertainty, model uncertainty, prediction uncertainty) mean that although models may be good communicators, what they have to communicate can be rather weak, and, worse still, these weaknesses may not be apparent to the user of the model output who may see them as prophecy, to the detriment of both science and policy. There is no clear boundary between policy models and nonpolicy (scientific models) but policy models or DSS in general tend to focus on 5, 6 and 7 much more than purely scientific models, which are usually only used by the model builder and few others. We will see later how the requirements of research and policy models differ.

### 1.2.1 Models with a poorly defined purpose

We have seen why modelling is important and how models may be used but before building or using a model we must clearly define its purpose. There are no generic models and models without a purpose are models that will find little use, or worse, if they do find use, they will often be inappropriate for the task at hand. In defining the purpose of a model, one must first clearly define the purpose of the research: what is the problem being addressed?; what are the processes involved?; who are the stakeholders?; what is the physical boundary of the problem and what flows cross it from outside of the modelling domain?; over which timescale should the problem be addressed?; and what are the appropriate levels of detail in time and space?

Here, the research question is the horse and the model the cart. The focus of all activities should be to answer the research question, not necessarily to parameterize or produce a better model. Once the research has been defined, one must ask whether modelling is the appropriate tool to use and, if so, what type of model. Then follows the process of abstracting reality and defining the model itself.

## 1.3 TYPES OF MODEL

Models are by no means a new tool in the scientists' toolbox. Environmental scientists have used spatial models of populations, environments, infrastructures, geologies and geographies in the form of maps and drawings for as long as science itself. Maps and drawings are abstractions of the *form* of nature in the same way that models are (usually) abstractions of the *process* of nature. Mathematics has its origins in the ancient Orient where it developed as a practical science to enable agriculture and agricultural engineering through the development of a usable calendar, a system of mensuration and tools for surveying and design. With the ancient Greeks, mathematics became more abstract and focused much more on deduction and reasoning. Mathematical models have been developed since the origin of mathematics, but there was a significant increase in modelling activity since the development of calculus by Newton and Leibniz working independently in the second half of the seventeenth century. Cross and Moscardini (1985) define three ages of modelling: (1) the 'Genius Age'; (2) the 'Transitional Age'; and (3) the 'Contemporary Age'. The 'Genius Age' followed the development of calculus and is characterized by the development of models of complex physical phenomena such as those of gravitation by Newton, of electromagnetic waves by Clerk Maxwell (of our own university) and of relativity by Einstein. Modelling in the 'Genius Age' was always limited by the need to produce analytical solutions to the set of equations developed. Cross and Moscardini's 'Transitional Age' was initiated by the availability of mechanical and then electromechanical aids to arithmetic but these devices were expensive, difficult to use and slow. The development of increasingly inexpensive computer power, a bank of numerical techniques that can yield accurate solutions to most equation sets and the softening of the human–computer communication barrier through the development of personal computers (PCs) and high-level programming languages have moved the 'Transitional Age' into the 'Contemporary Age'. Thus, modern numerical computer models can be seen rather simply as the continuation of the relationship between science and mathematics with the greater sophistication afforded by advances in the power of computer input, processing and display. No-one knows what will come next but we make some educated guesses in the last chapters of this book.

Models can be classified hierarchically. The two top-level model types are the *mathematical* models and the physical or *hardware* models (not to be confused with physically based, mathematical models). Hardware models are scaled-down versions of real-world situations and are used where mathematical models would be too complex, too uncertain or not possible because of lack of knowledge. Examples include laboratory channel flumes, wind tunnels, free atmosphere $CO_2$ enrichment apparatus, rhizotrons and lysimeters, and the BIOSPHERE 2 laboratory (Cohn, 2002). Many instruments are also hardware models which allow control of some environmental conditions and the measurement of the response of some system to these controls. The Parkinson leaf chamber which forms the basis of most leaf photosynthesis systems is a good example. The chamber is a small chamber (usually 5 $cm^3$) which is clamped onto a leaf *in vivo* and controls the input humidity and $CO_2$ concentration and measures the output of the same so that photosynthesis and transpiration can be measured.

Hardware models are usually small-scale compared with the systems which they simulate and their results may be prone to uncertainty resulting from scale effects which have to be balanced against the increased cost of larger-scale hardware models. Hardware models are also expensive. The 1.27 ha BIOSPHERE 2 (BIOSPHERE 1 is, apparently, the Earth) experiment in the Santa Catalina Hills, near Oracle, Arizona, USA, is now used to test how tropical forest, ocean and desert ecosystems might respond to rising $CO_2$ and climate change and cost some US$150 million (UK-based readers may want to visit the Eden project near St Austell, Cornwall, to get an impression of a similar physical model [www.edenproject.com]). Most physical models are considerably less ambitious and thus much cheaper but this does, nevertheless, give an indication of the kind of costs that are required to 'simulate' ecosystems in hardware. Hardware models give a degree of control on the systems under investigation, in the case of BIOSPHERE 2, $CO_2$, water and nutrient inputs can be controlled. There is, however, always the problem that the hardware representation of a process is only as good as our understanding of that process and our ability to replicate it: BIOSPHERE 2 cannot simulate storms with high winds or damaging downpours or the diversity of soil complexes that exist in the rainforest which it simulates. This may render the BIOSPHERE 2 rainforest response to climate change rather simplistic compared with the field reality. Furthermore, the structure of the hardware model may also interfere with natural processes in the environment, for example, the glass windows of BIOSPHERE 2 cut out some 50% of the incoming light. Because of the cost involved, it is also usually difficult to replicate hardware experiments: there is only one BIOSPHERE 2 and laboratories usually only have one wind tunnel or channel flume. Nevertheless hardware models couple the scientific rigour of observation with the controllability of mathematical modelling. For many applications, it is only logistic difficulties and cost which keep them as a relatively uncommon approach to modelling: physical models require a great deal more set-up and maintenance costs than software and data.

Mathematical models are much more common and represent states and rates of change according to formally expressed mathematical rules. Mathematical models can range from simple equations through to complex software codes applying many equations and rules over time and space discretizations. One can further define mathematical models into different types but most models are actually mixtures of many types or are transitional between types. One might separate mathematical models into empirical, conceptual or physically based:

- Empirical models describe observed behaviour between variables on the basis of observations alone and say nothing of process. They are usually the simplest mathematical function, which adequately fits the observed relationship between variables. No physical laws or assumptions about the relationships between variables are required. Empirical models have high predictive power but low explanatory depth, they are thus rather specific to the conditions under which data were collected and cannot be generalized easily for application to other conditions (i.e. other catchments, other forests, other latitudes).
- Conceptual models explain the same behaviour on the basis of preconceived notions of how the system works in addition to the parameter values, which describe the observed relationship between the variables. A conceptual model of hillslope hydrology may separately model processes of surface runoff, subsurface matrix quickflow and subsurface pipeflow (see Chapter 4). While all these models are empirical, their separation incorporates some process understanding. Conceptual models have slightly greater explanatory depth but are as nongeneralizable as the empirical models which make them up.
- Physically based models should be derived deductively from established physical principles and produce results that are consistent with observations (Beven, 2002) but in reality physically based models often do one of these but rarely both. In general use, there is a continuum of models that falls broadly under the heading of physically based, but that might include some level of empirical generalization in order to allow them to operate at an appropriate environmental scale, or to fill gaps where the physics is not known. *Process* models emphasize the importance of the processes transforming input to output. In some respects, this tradition may arise from the link between the study of process and form within the discipline of geomorphology. Similarly, ecologists often talk of *mechanistic* models. It may be that the choice of terminology (relating to physics or mechanics) represents a desire to suggest the science underlying the model has a sound basis. Physically based models tend to have good explanatory depth (i.e. it is possible to interrogate the model to find out exactly why, in process terms, an outcome is as it is). On the other hand, physically based models are characterized by low predictive power: they often do not agree with observations. This lack of agreement often demonstrates a poor understanding of the physics of the system. These models thus often need to be calibrated against observations (see Chapter 4, for example). One has to think carefully about the explanatory depth of a model that does not replicate the observed reality well (Beven, 2001). Where they are not highly calibrated to observed data *and* if they are appropriately and flexibly structured, physically based models offer a greater degree of generality than empirical models.

According to the level of process detail and understanding within the model, it may be termed black box or white box. In a black box model, only the input and output are known and no details on the processes which transform input to output are specified, and the transformation is simulated as a parameter or parameters defining the relationship of output to input. On the other hand, in a white box model, all elements of the physical processes transforming input to output are known and specified. There are very few systems for which white box models can be built and so most models, being a mixture of physically based and empirical approaches, fall in between white and black to form various shades of grey boxes. Empirical models are usually closer to the black box end of the spectrum while physically based models fall in between this and the white box, depending on their detail and the extent to which they are calibrated to observed data.

There are no universally accepted typologies of models and, given the diversity of approaches apparent in even a single model code in the multi-process models, which are increasingly common, there is little point in specifying one. Nevertheless, it is useful to understand the properties according to which models may be classified. We have already discussed the different types of model (empirical, conceptual, physically based). Models can be further subdivided according to how the equations are integrated (either analytically solving the model equations as differential equations or numerically solving them within a computer as difference equations, see this chapter). Further subdivision can take place according to the mathematics of the model, for example, whether the equations are deterministic, that is, a single set of inputs always produces one (and the same) output. In the alternative, stochastic approach, a single set of inputs can produce very different outputs according to some random processes within the model. Up to this point most models are still mixtures of many of these types, though two further properties are still to be specified. Models are of different spatial types. Lumped models simulate a (potentially) spatially heterogeneous environment as a single – lumped – value. Semi-distributed models may have multiple lumps representing clearly identifiable units such as catchments. Distributed models break space into discrete units, usually square cells (rasters) or triangular irregular networks (TINs, e.g. Goodrich *et al.*, 1991) or irregular objects. The spatial realm of a model may be one-dimensional, two-dimensional (sometimes within the context of a geographical information system or GIS) and, sometimes, three-dimensional. All models are lumped at the scale of the cell or triangle and because the sophistication of modelling techniques is way ahead of the sophistication of measurement techniques, data limitations mean that most distributed models use lumped data. Finally, one has to consider the manner in which the model handles time. Static models exclude time whereas dynamic ones include it explicitly. A summary of the potential means of classifying models is given below:

Conceptual type: empirical, conceptual, physically based or mixed
Integration type: analytical, numerical or mixed
Mathematical type: deterministic or stochastic or mixed
Spatial type: lumped, semi-distributed, distributed, GIS, 2D, 3D or mixed
Temporal type: static, dynamic or mixed.

## 1.4 MODEL STRUCTURE AND FORMULATION

### 1.4.1 Systems

Contemporary mathematical modelling owes much to systems thinking and systems analysis. A system is a set of inter-related components and the relationships between them. Systems analysis is the 'study of the composition and functioning of systems' (Huggett, 1980). In practice, systems analysis involves the breaking down or modularization of complexity into simple manageable subsystems connected by flows of causality, matter, energy or information. The purpose of systems analysis is to make complex systems more easily understood. Systems usually comprise of *compartments* or *stores* that represent quantities such as height (m), mass (kg), volume ($m^3$), temperature (°C), annual evaporation (mm) and which are added to or subtracted from by flows or fluxes such as height increment (m $a^{-1}$), weight gain (kg $a^{-1}$), volume increment ($m^3$ $a^{-1}$) and evaporation (mm $month^{-1}$). Further details on systems analysis can be found in Chorley and Kennedy (1971), Huggett (1980) and Hardisty *et al.* (1993).

In modelling a *variable* is a value that changes freely in time and space (a compartment or flow) and a state variable is one which represents a state (compartment). A *constant* is an entity that does not vary with the system under study, for example, acceleration due to gravity is a constant in most Earth-based environmental models (but not in geophysics models looking at gravitational anomalies, for example). A *parameter* is a value which is constant in the case concerned but may vary from case to case where a case can represent a different model run or different grid cells or objects within the same model.

### 1.4.2 Assumptions

In order to abstract a model from reality, a set of assumptions has to be made. Some of these assumptions will be wrong and known to be so but are necessary for the process of abstraction. The key to successful modelling is to know which assumptions are likely to be wrong and to ensure that they are not important for the purpose for which the model is intended. Further, one should only use the model for that purpose and ensure that no-one else uses the model for purposes which render incorrect assumptions significant or correct assumptions invalid. The value of a model depends totally on the validity and scope of these assumptions. These assumptions must be well understood and explicitly stated with reference to the conditions under which they are valid and, more importantly, the conditions under which they are invalidated. Abstraction should always be guided by the principle of parsimony. Perrin *et al*. (2001) indicate that simple models (with few optimized parameters) can achieve almost as high a level of performance as more complex models in terms of simulating their target variable. Although the addition of model parameters and more detailed process description may have benefits from a theoretical point of view, they are unlikely to add greatly to model predictive capability even if substantial data resources are available to keep the uncertainty in these parameters to a minimum (de Wit and Pebesma, 2001). The greater the number of parameters, the greater the likelihood that they will be cross-correlated and that each extra parameter will add relatively little to the explanation (van der Perk, 1997). Perrin *et al*. (2001) concur with Steefel and van Cappellen (1998) who indicate, for models with equal performance, that the best model is the simplest one. Simplicity must be strived for, but not at the cost of model performance. In this way building models is best achieved by starting with the simplest possible structure and gradually and accurately increasing the complexity as needed to improve model performance (see Nash and Sutcliffe, 1970). Figures 11.3–11.5 (in Chapter 11) indicate the manner in which model performance, model complexity, the costs (in time and resources) of model building and the uncertainty of results can be related.

### 1.4.3 Setting the boundaries

We have now defined the problem and agreed that modelling is part of the solution. Further, we know what the purpose of the model is and can thus define what kind of model is appropriate. The next task is to begin building the model. We must first set the boundaries in time and space to identify which times and spaces will be modelled and which must be supplied as data. We call these data the boundary conditions for data, representing processes outside the spatial domain of the model and the initial conditions for data, representing processes internal to the model spatial domain but external (before) the model temporal domain. Model results, from those of planetary rotation (Del Genio, 1996), through river flooding (e.g. Bates and Anderson, 1996) to moth distributions (Wilder, 2001) are usually sensitive to the initial and boundary conditions so these must be carefully specified. In the case of a general circulation model of the atmosphere (GCM), the boundary is usually the top of the atmosphere and thus one of the variables which must be supplied as a boundary condition, because it is not modelled within the GCM, is the incoming solar radiation flux. It is important to mention that boundary conditions may exist outside the conceptual space of a model even if they are inside the physical space of the same. For example, until recently, global vegetation cover was not simulated within GCMs and thus, despite the fact that it resides within the spatial domain of GCMs, it had to be specified as a boundary condition for all time steps (and also as an initial condition for the first time step). Nowadays many GCMs incorporate an interactive vegetation modelling scheme so this is no longer necessary. GCMs have to be supplied with an initial condition sea surface temperature (SST) field for each sea surface grid cell (GCMs are usually 3D raster distributed) for time zero after which the model will calculate the SST for each timestep.

Let us consider a simple model of soil erosion. The objective of the model is for us to understand more of the relative importance of climatic, landscape and plant factors on soil erosion. Specifically, we will simulate wash erosion ($E$, variable over space and time) which is simulated on the basis of runoff ($Q$, variable over space and time), slope gradient ($S$, variable over space constant over time), vegetation cover ($V$, state variable over space and time) and the erodibility ($K$, variable over space constant in time) of the soil and three parameters, $m$, $n$ and $i$. Wash erosion is soil erosion by water running across the land surface and, after Thornes (1990) the model can be expressed as:

$$E = kQ^m S^n \mathrm{e}^{-iV} \tag{1.1}$$

where

$E$ = erosion (mm month$^{-1}$)
$k$ = soil erodibility
$Q$ = overland flow (mm month$^{-1}$)
$m$ = flow power coefficient (1.66)
$S$ = tangent of slope (mm$^{-1}$)
$n$ = slope constant (2.0)
$V$ = vegetation cover (%)
$i$ = vegetation erosion exponential function (dimensionless).

In this way, soil erosion is a function of the erodibility of the soil (usually controlled by its organic matter content, structure, texture and moisture), the ‘stream power' of runoff ($Q^m$), the slope angle effect ($S^n$) and the protection afforded by vegetation cover ($e^{-0.07V}$). Let us say that it is a distributed model running over 25-metre raster cells and at a time resolution of one month for 50 years. Erosion and runoff are fluxes (flows), vegetation, slope gradient and erodibility are states (compartments). Since the model is of soil erosion, it does not simulate vegetation growth nor runoff generation, hence these are outside the boundary of the model and they must be specified for each timestep as a boundary condition. No initial conditions need be specified since the only state variable, vegetation cover, is already specified as a boundary condition. If we were to calculate the change in soil thickness ($Z$, state variable) according to erosion, then we must specify an initial condition for $Z$.

### 1.4.4 Conceptualizing the system

The understanding gained from being forced to rationalize one's conceptual view of a process or system and quantify the influence of each major factor is often the single most important benefit of building a mathematical model (Cross and Moscardini, 1985). If the model results match reality, then this conceptual view is supported and, if not, then the conceptualization may be, partly or wholly, at fault and should be reviewed. Different people will produce quite different conceptualizations of the same system depending upon their own background and experience: to a climate scientist a forest is a surface cover interacting with the atmosphere and land surface and affecting processes such as the $CO_2$ concentration of the atmosphere and the energy and water fluxes across the atmospheric boundary layer. To a forester a forest is a mass of wood-producing trees of different ages, sizes and monetary values. Cross and Moscardini (1985) specify five stages of mathematical modelling: problem identification, gestation, model building, simulation and pay-off. These stages are taken sequentially, although one may move back to earlier stages at any time as needed.

Problem identification is a fairly obvious first stage. If the problem is not properly identified, then it will not be possible to arrive at a solution through modelling or any other means. The gestation stage is an important, though often neglected, stage consisting of the gathering of background information, amassing an understanding of the system under investigation, separating relevant information from irrelevant and becoming familiar with the whole context of the problem. The two substages of gestation may be considered as modularization and reviewing the science. Modularization breaks the problem down into solvable chunks in the manner of systems analysis. A clear mental map of the processes involved will help very much in the separation of big complex processes into families of small and solvable self-contained modules. Once modularized, the existing science and understanding of each process must be reviewed, allowing abstraction of the relevant from the irrelevant. On departing from this stage the modeller should have a good conceptual understanding of the problem, its context and how it will be solved. This is where the process of abstraction is most important and the modeller's intuition as to what is and is not important will be most valuable. Having devised an acceptable conceptual framework and a suitable data subset, the formulation of the mathematical model is usually fairly straightforward. The process of model building incorporates a further three substages: developing the modules, testing the modules and verifying the modules. Developing the modules will often involve some re-use of existing models, some routine model development and some flashes of inspiration. It is important that, even at this early stage, the modules are tested so that the solution developed has a reasonable chance of producing consistent results and not, for example, negative masses, humidities or heights. In addition to defining the model itself, one will also have to, at this stage, give some thought to the method of solution of the model equations. Before detailing the numerical methods available for this, we will outline in some detail the practical aspects of putting a simple model together.

### 1.4.5 Model building

Model building may consist of stringing together sets of equations to which an analytical solution will be derived, but more likely these days it will involve compartmentalization of the problem and its specification as either compartments and flows within a graphical model building environment such as STELLA (http://hps-inc.com), VENSIM (http://www.vensim.com), PowerSim (http://www.powersim.com), ModelMaker (http://www.cherwell.com), SIMULINK, the graphical modelling environment of MATLAB or SIMILE (http://www.simulistics.com) or as routines and procedures in a high level computer programming language such as Basic, Fortran, Pascal, C++ or Java, or a custom modelling language such as PCRASTER (http://www.pcraster.nl).

Some modellers prefer to build their models graphically by adding compartments and flows, linking them with dependencies and entering the appropriate equations into the relevant compartments, flows or variables. Indeed, this approach is fairly close to systems analysis and the way that many non-programmers conceptualize their models. Others, usually those who are well initiated in the art of computer programming, prefer to construct their models in code, or indeed even in the pseudo-code which modern spreadsheet programs allow for the solution of equations as reviewed extensively by Hardisty *et al.* (1994). By way of an introduction to model building let us look back on the soil-erosion equation we introduced a few pages back and examine the manner in which this could be constructed in (a) a graphical model building program, in this case SIMILE which is produced by the Edinburgh-based simulistics.com and is currently distributed free of charge for educational use from http://www.simulistics.com; (b) a spreadsheet program, in this case Microsoft Excel, available from http://www.microsoft.com; and (c) a spatial model building language called PCRASTER produced by the University of Utrecht in The Netherlands and distributed free of charge for educational use from http://www.pcraster.nl. For simplicity we will keep the model simple and nonspatial in the SIMILE and Excel implementation and make it spatial in the PCRASTER implementation.

Though the graphical user interface (GUI) and syntax of these specific software tools will not be identical to any others that you may find, the basic principles of working with them will be similar and so this exercise should also prepare you for work using other tools. There is not space here to enter into the complexities of high-level programming syntax but suffice to say that, for a modeller, knowing a programming language – and they all share the same basic constructs – is a very useful but not indispensable skill to have mastered. Coding in a high-level language does allow more efficient models to be developed in terms of the time they take to run (excluding BASIC and Java from the above list because they interpret commands just like PCRASTER) and the memory they occupy. Coding also sidesteps the limitations of software tools by giving the programmer full, unlimited access to the computer's functions.

### 1.4.6 Modelling in graphical model-building tools: using SIMILE

SIMILE is a new addition to the suite of graphical systems-modelling tools and is attractive because of its low cost, comprehensive online documentation and powerful modelling capabilities. The construction of a model in SIMILE is achieved first through the specification of compartments and the flow variables between them and subsequently by the specification of the parameters that affect the flow variables and the direction of the influence they have. The SIMILE interface is as shown in Figure 1.1. The modelling canvas is initially blank and the modeller adds compartments, flows, parameters, variables and influences using the toolbar short cuts. These are then given a label and are populated with the relevant values or equations through the equation bar above the model canvas, which becomes active when a compartment, flow or variable is clicked with the computer mouse or other pointing device. The model shown on the canvas is our soil-erosion model but do not look too closely at this yet as we will now attempt to construct it.

The first step is to define the compartment or state variable. In this case of soil erosion, this could be soil thickness so we will label it as such. Note that the symbols that SIMILE produces are standard systems-analysis symbols. Now let's add the flow (Figure 1.2). This is, of course, soil erosion that is a flow out of the soil-thickness compartment and not into it because, for simplicity, we are only simulating erosion and not deposition. The flow is given the label **E** and an influence arrow must be drawn from **E** to the compartment which represents soil thickness since erosion affects soil thickness (note that we use a **bold** symbol to denote the parameter as used in the model formalization compared to the *italicized* symbol when discussing the model form in an equation). Note that until all influences and equations are fully specified, the flow and its associated arrows remain red. On full parameterization all model components will be black in colour. We cannot now specify the equation for erosion until we have fully specified all of the parameters and variables that influence it. We therefore add the variables: **Q**, **k**, **s**, **m**, **n**, **i** and **V** (Figure 1.3) and draw influence arrows from each of them to soil erosion (Figure 1.4) since they all affect soil erosion. We can now specify the parameter values of each of **Q**, **k**, **s**, **m**, **n**, **i** and **V** either as (fixed) parameters or as variables which are time dependent and either calculated in SIMILE or read from an input file or is the outcome of another equation or submodel in SIMILE. To keep things simple we will enter these values as constants, $\mathrm{Q} = 100$ (mm month$^{-1}$), $\mathrm{k} = 0.2$, $\mathrm{s} = 0.5$, $\mathrm{m} = 1.66$, $\mathrm{n} = 2.0$, $\mathrm{i} = -0.07$ and $\mathrm{V} = 30(\%)$. Finally, we can click on the flow, **E** and enter the equation (from Equation 1.1), which determines the value of **E** as a function of **Q**, **k**, **s**, **m**, **n**, **i**, **V** (see the equation bar in Figure 1.1). Note that in the SIMILE syntax we only enter the right-hand side of the equals sign in any equation and so Equation 1.1

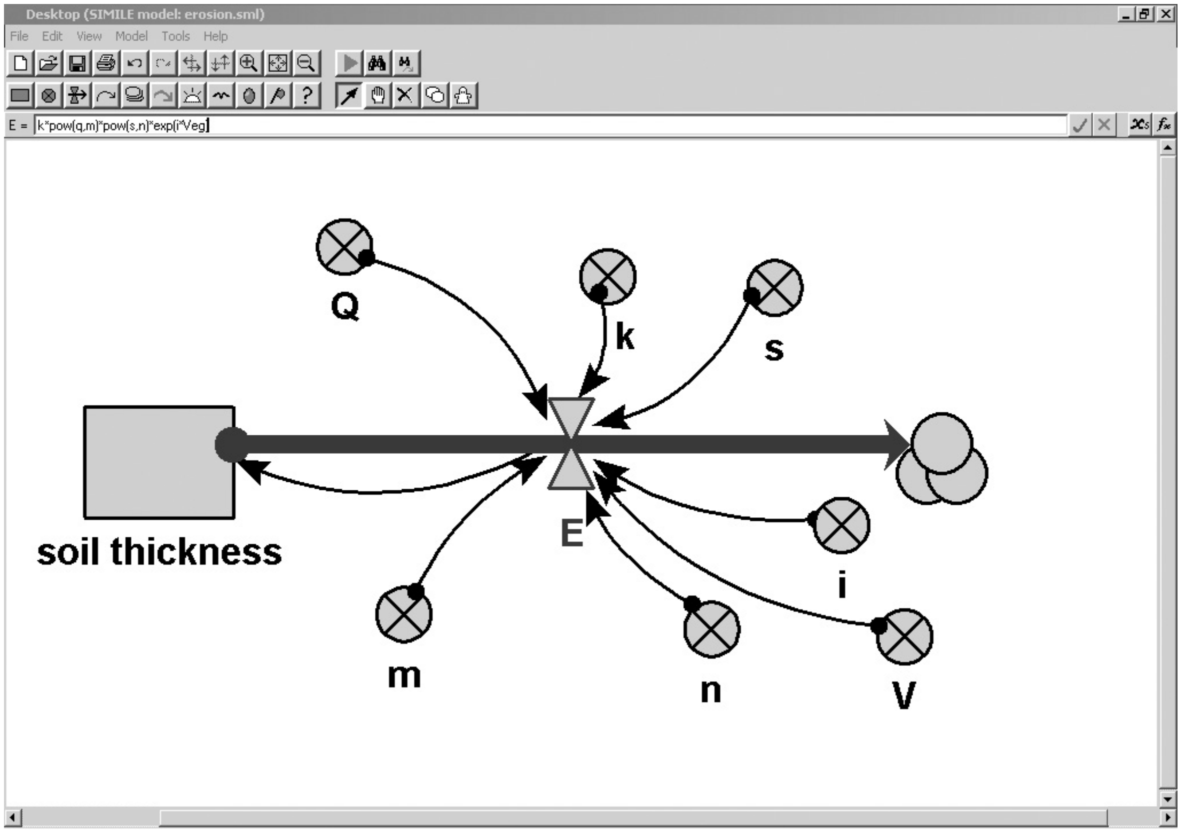



**Figure 1.1** Interface of the SIMILE modelling software

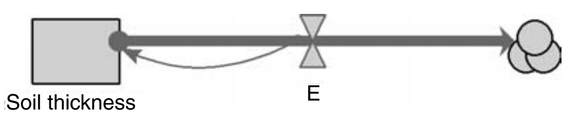



**Figure 1.2** Adding the flow component to the SIMILE model developed in the text

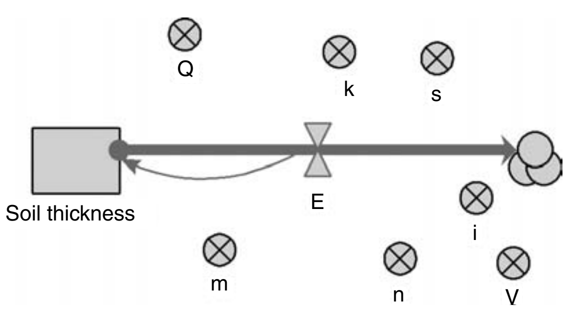



**Figure 1.3** The SIMILE erosion model with the variables (**Q**, **k**, **s**, **m**, **n**, **i**, **V**) added

becomes k*pow(Q,m)*pow(s,n)*exp(i*V) with pow(**x**) representing to the power of **x** and exp(**x**) representing **e** (the base of natural logarithms) to the power of **x**. We must also specify the initial condition for soil thickness, in mm because **E** is also in mm. Let's say the initial thickness is 10 m (10000 mm). All that remains is to build or compile the model allowing SIMILE to convert this graphic representation to a pseudo-code which is then interpreted into machine code at runtime.

To run the model we need to specify the number and frequency of timesteps in the runtime environment and then run the model. Output is logged for each timestep and the user can examine the model output variable by variable, graphically or as text. If checked and working, this model can be wrapped and used as a component of a larger, more complex model. Many of the graphical modelling tools also have helpers or wizards for performing model calibration, optimization and sensitivity analysis. The great advantage of graphical model-building tools is the rapidity with which they can be learned and the ease with which even very complex systems can be represented. Their disadvantages are that they are generally expensive and it can be rather difficult to do more advanced modelling since they are

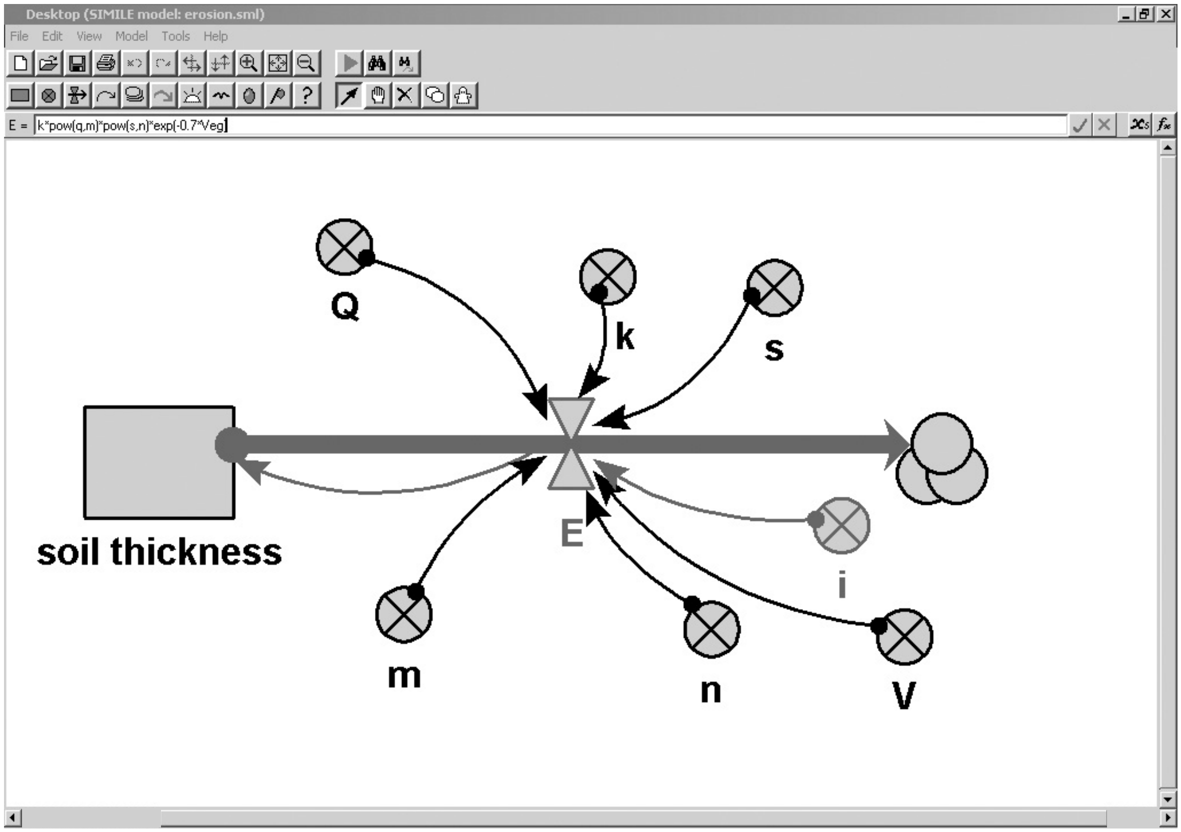



**Figure 1.4** Influences between the variables and the flow for the SIMILE soil-erosion model

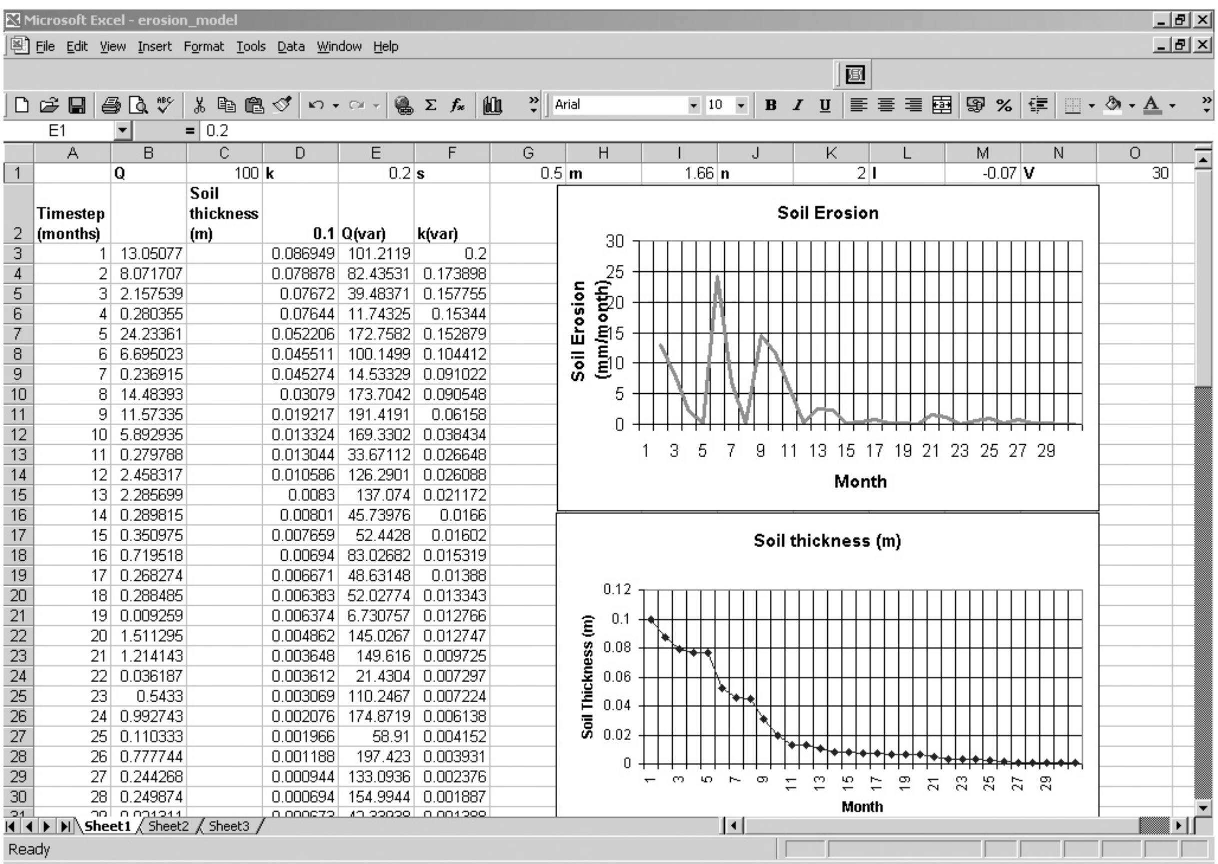
rather more constrained than modelling in code. If you want to look into it further, the online help and tutorial files of SIMILE provide much greater detail on the more powerful aspects of its functionality including its ability to deal with many (spatial) instances of a model representing for example, individual hillslopes or plants or catchments. You can download this erosion model from the companion website for this book.

### 1.4.7 Modelling in spreadsheets: using Excel

A very detailed training for modelling in Excel is given in Hardisty *et al*. (1994), so we only provide a simple example here. To build the same soil-erosion model in a spreadsheet such as Microsoft Excel requires the specification of the time interval and number of timesteps upfront since these will be entered directly into the first column of the spreadsheet. Open the spreadsheet and label column **A** as **Timestep** at position **A:2** in the spreadsheet (column A, row 2). Then populate the next 30 rows of column **A** with the numbers 1 to 30. This can be easily achieved by adding the number 1 to **A:3**. At position **A:4** type the following equation: **=A3+1** and press enter. The spreadsheet will now calculate the results of this equation which will appear as the number in **A:4**. We can now highlight cells **A:4** through to **A:32** with our pointing device (left click and drag down) and go to the menu item **Edit** and the **Fill** and then **Down** (or Ctrl-D) to copy this equation to all the cells in the highlighted range. Note that in each cell, the cell identifier to the left of the plus sign in the equation (which for the active cell can be viewed in the equation bar of Excel) is changed to represent the cell above it so that in **A:4** it is **A:3** but in **A:5** it is **A:4** and so on. This is relative referencing which is a very useful feature of Excel as far as modelling is concerned because it allows us to define variables, i.e. parameters which change with time and time, which are usually represented along the rows of the spreadsheet or even on another spreadsheet. Model instances representing different locations or objects are usually represented across the columns of the spreadsheet, as are the different compartment, flow and variable equations.

If we want to avoid relative referencing in order to specify a constant or parameter (rather than a variable), we can enter it into a particular cell and use absolute referencing to access it from elsewhere on the spreadsheet. In absolute referencing a $ must be placed in front of the identifier which we do not want to change when the equation is filled down or across. To specify a parameter which remains constant in time (down the spreadsheet) we would place the $ before the row identifier, for example, **=A$3+1**, to specify a parameter which remains constant across a number of model instances (across the spreadsheet) we place the $ before the column identifier, **=$A3+1**. To specify a constant that changes nowhere on the spreadsheet, then we use for example **=$A$3+1**. This form always refers to a single position on the spreadsheet, whenever the equation is filled to.

Going back to our soil-erosion model, we have a number of model parameters which will not change in time or space for this implementation and these can be sensibly placed in row 1, which we have left uncluttered for this purpose. We can label each parameter using the cell to the left of it and enter its value directly in the cell. So let's label **Q**, **k**, **s**, **m**, **n**, **i** and **V** and give them appropriate values, $Q = 100$ (mm month$^{-1}$), $k = 0.2$, $s = 0.5$, $m = 1.66$, $n = 2.0$, $i = -0.07$ and $V = 30(\%)$. See Figure 1.5 for the layout. Now that we have specified all of the parameters we can enter the soil erosion equation in cell **B:3** (we should place the label Soil Erosion (mm month$^{-1}$) in **B:2**). We will use absolute referencing to ensure that our parameters are properly referenced for all timesteps. In column **D** we can now specify soil thickness in metres so let us label **C:2** with 'Soil Thickness (m)' and specify an initial condition for soil thickness, of say 10 m, in cell **D:2**. In the syntax of Excel Equation 1.1 becomes =E$1*(C$1^I$1)*(G$1^K$1)*EXP(M$1*O$1) where E$1 holds the value of **k**, C$1 holds the value of **Q**, I$1 holds the value of **m**, G$1 holds the value of **s**, K$1 holds the value of **n**, M$1 holds the value of **i** and O$1 holds the value of **V**. In Excel a caret (^) represents 'to the power of' and EXP(**x**) represents **e** (the base of natural logarithms) to the power of **x**. On pressing enter, this equation will produce a value of




| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | | **Q** | 100 | **k** | 0.2 | **s** |
| 2 | **Timestep (months)** | | **Soil thickness (m)** | **0.1** | **Q(var)** | **k(var)** |
| 3 | 1 | 13.05077 | | 0.086949 | 101.2119 | 0.2 |
| 4 | 2 | 8.071707 | | 0.078878 | 82.43531 | 0.173898 |
| 5 | 3 | 2.157539 | | 0.07672 | 39.48371 | 0.157755 |
| 6 | 4 | 0.280355 | | 0.07644 | 11.74325 | 0.15344 |
| 7 | 5 | 24.23361 | | 0.052206 | 172.7582 | 0.152879 |
| 8 | 6 | 6.695023 | | 0.045511 | 100.1499 | 0.104412 |
| 9 | 7 | 0.236915 | | 0.045274 | 14.53329 | 0.091022 |
| 10 | 8 | 14.48393 | | 0.03079 | 173.7042 | 0.090548 |
| 11 | 9 | 11.57335 | | 0.019217 | 191.4191 | 0.06158 |
| 12 | 10 | 5.892935 | | 0.013324 | 169.3302 | 0.038434 |
| 13 | 11 | 0.279788 | | 0.013044 | 33.67112 | 0.026648 |
| 14 | 12 | 2.458317 | | 0.010586 | 126.2901 | 0.026088 |
| 15 | 13 | 2.285699 | | 0.0083 | 137.074 | 0.021172 |
| 16 | 14 | 0.289815 | | 0.00801 | 45.73976 | 0.0166 |
| 17 | 15 | 0.350975 | | 0.007659 | 52.4428 | 0.01602 |
| 18 | 16 | 0.719518 | | 0.00694 | 83.02682 | 0.015319 |
| 19 | 17 | 0.268274 | | 0.006671 | 48.63148 | 0.01388 |
| 20 | 18 | 0.288485 | | 0.006383 | 52.02774 | 0.013343 |
| 21 | 19 | 0.009259 | | 0.006374 | 6.730757 | 0.012766 |
| 22 | 20 | 1.511295 | | 0.004862 | 145.0267 | 0.012747 |
| 23 | 21 | 1.214143 | | 0.003648 | 149.616 | 0.009725 |
| 24 | 22 | 0.036187 | | 0.003612 | 21.4304 | 0.007297 |
| 25 | 23 | 0.5433 | | 0.003069 | 110.2467 | 0.007224 |
| 26 | 24 | 0.992743 | | 0.002076 | 174.8719 | 0.006138 |
| 27 | 25 | 0.110333 | | 0.001966 | 58.91 | 0.004152 |
| 28 | 26 | 0.777744 | | 0.001188 | 197.423 | 0.003931 |
| 29 | 27 | 0.244268 | | 0.000944 | 133.0936 | 0.002376 |
| 30 | 28 | 0.249874 | | 0.000694 | 154.9944 | 0.001887 |

| | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.5 | **m** | 1.66 | **n** | 2 | **i** | -0.07 | **V** | 30 |

**Figure 1.5** Layout of the soil-erosion model in Excel