

Int. J. General Systems, Vol. 0(0), pp. 1–22
Reprints available directly from the publisher
Photocopying permitted by license only

© 2000 OPA (Overseas Publishers Association) N.V.
Published by license under
the Gordon and Breach Science
Publishers imprint.
Printed in Malaysia.

MODELLING AS A DISCIPLINE

WILLIAM SILVERT

*Instituto de Investigação das Pescas e do Mar, Avenida de Brasília, s/n,
1449-006 Lisboa, Portugal*

(Received 24 March 2000; Revised ; In final form 26 April 2000)

Modelling is an essential and inseparable part of all scientific, and indeed all intellectual, activity. How then can we treat it as a separate discipline? The answer is that the professional modeller brings special skills and techniques to bear in order to produce results that are insightful, reliable, and useful. Many of these techniques can be taught formally, such as sophisticated statistical methods, computer simulation, systems identification, and sensitivity analysis. These are valuable tools, but they are not as important as the ability to understand the underlying dynamics of a complex system well enough to assess whether the assumptions of a model are correct and complete. Above all, the successful modeller must be able to recognise whether a model reflects reality, and to identify and deal with divergences between theory and data. Theories can be wrong or merely incomplete, and even “raw” data are just the outputs of experimental interpretations, *i.e.*, models. These points are illustrated with examples from the scientific literature, accompanied by horror stories of modelling projects gone awry.

Keywords: Models; modelling; ecology; simulation; validation

INTRODUCTION

This paper was originally developed as a talk to a group of young scientists in the fields of marine biology and oceanography. It has been expanded and modified for a broader audience, and although the examples are drawn from the ocean sciences, I hope that it will prove of interest to a much broader readership.

* E-mail: wsilvert@ipimar.pt.

WHAT IS MODELLING?

Modelling is one of the most universal activities. We can learn about the modelling process by studying its development in small children. A baby sees a pretty red disk on top of a stove and touches it. He learns that cherry red objects are painful. Later he sees a red coal that falls out of the fireplace and either applies the model, thereby saving himself from another painful burn, or tests the model, and thus validates it. Then his parents are going to a party and his mother comes in wearing a cherry red dress. She is baffled at his screams of panic—the model, validated or not, does not apply.

If modelling is so universal, what is special about modellers? Modelling is like breathing. Singers, other musicians, actors, etc., learn special breathing techniques, and modelling is much the same—everyone else does it, but we do it better.

One of the main roles of the professional modeller is to apply quantitative reasoning to observations about the world, in hopes of seeing aspects that may have escaped the notice of others. This, after all, is an extension of what all scientists do, namely looking at things in new ways in hopes of seeing what no one has noticed before. Galileo used new technology, namely the telescope, to discover spots on the sun and the phases of Venus. Kepler on the other hand used his mathematical skills to discover that the motions of the planets were best described not by cycles and epicycles based on circular orbits, but by ellipses.

An example of how even skilled scientists can overlook quantitative information arose when I first started working in marine ecology and was working on the dynamics of plankton blooms. A colleague of mine was explaining his data on the spring bloom, and gave me the standard picture—during the stormy and dark winter season nutrients accumulate in the upper part of the water column, but when spring comes there is enough sunlight to initiate the explosive growth of phytoplankton, referred to as a bloom. After a while the phytoplankton consume all of the nutrients, and the bloom collapses from nutrient depletion. The only problem with this classical picture was that when I actually started working with his data I found that the nutrient levels at the end of the spring bloom were still an order of magnitude higher than those required for algal growth. The standard theory was

so firmly implanted in his mind that he hadn't thought it necessary to look at the numbers, even at his own data!

SKILLS AND TECHNIQUES

There are many specific techniques that modellers use, which enable us to discover aspects of reality that may not be obvious to everyone. One is the use of sophisticated statistical methods that go well beyond the standard textbook procedures that most scientists learn. Unfortunately, they often do not go far enough. One of the most common failings of statistical models is that they are based on linear assumptions which often do not apply in the real world. For some strange reason it is considered reasonable to use linear methods for no other reason than that they appear more sophisticated than the cruder methods that exist for non-linear problems, but there is great reluctance to explore chaos or catastrophe theory, which are considered "controversial". Unfortunately, the use of sophisticated statistical techniques without fully understanding them can lead to ludicrous, and sometimes disastrous, results.

A simple example will suffice to show this point. I once went to a talk where data similar to those in Fig. 1 were displayed. To my great astonishment, the speaker referred to the linear relationship between x and y . When I protested that the data were most certainly not linear, I was condescendingly informed that the statistical test for significance of a linear regression clearly showed that there was a significant linear correlation. Of course this is a misunderstanding—the test shows only that there is a relation between x and y , it does not establish that the relationship is linear. This illustrates the fact that a little statistical knowledge can be very dangerous!

Another area where modellers tend to have esoteric specialised knowledge is in computer simulation. In fact, many people seem to think that all modellers do is write computer simulations! The essential point to remember about computers is GIGO, Garbage In, Garbage Out. As Lee (1973) once pointed out, "bigger computers simply permit bigger mistakes". This is not to say that computer simulation is wrong—I do lots of it myself in fact. The problem is that too much effort goes into the programming, and not enough into thinking about the assumptions and understanding the equations that are being programmed.

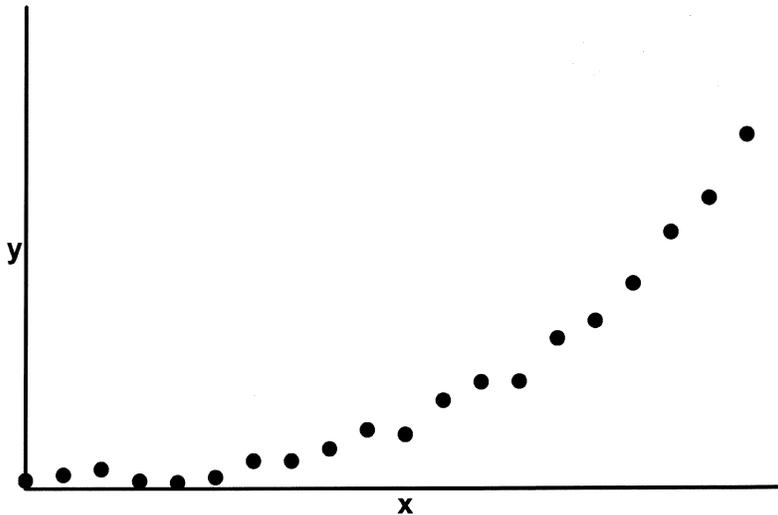


FIGURE 1 Hypothetical scatter plot of two variables. A figure like this was once described as a linear relationship because it met the test for significance of linear regression!

One of the biggest problems with computer simulation is proving to be its popularity. This has led to a rash of new simulation packages designed to make simulation easier, and more accessible to non-programmers. While it is hard to criticise the democratisation of one's profession, the fact is that most users of these programs are unfamiliar with their limitations and the assumptions implicit in their use. For example, the first of the programs that might be labelled as "simulation for the rest of us" was Stella. Stella was very popular, but at least in the earlier versions it imposed very strict limitations on the processes that could be represented. Time series inputs and non-linear feedback were difficult to incorporate, and required appreciable amounts of programming expertise. The current rage in marine ecology is ECOPATH, which is really just a set of protocols for putting together static nutrient budgets and cannot be used to resolve the kinds of questions about dynamic system responses that people are constantly throwing at it (a successor program called ECOSIM claims to address these concerns, but this has yet to be demonstrated). Certainly these programs are valuable tools and have been used for many good studies, but in the hands of people who do not understand them they can lead to nonsense.

The dangers of using scientific tools that one does not fully understand is not limited to theoretical simulation programs. The microscope provides a good analogy; in the hands of an expert it is one of the most valuable tools in science, but when used by someone who does not fully understand the importance of proper sterilisation techniques it can lead to totally meaningless results.

Nor should one think that these problems only arise in connection with specialised and sophisticated methods. Back in the early days of the personal computer revolution, with 8-bit machines that ran only BASIC, population biologists discovered that these were ideal tools for simulating Leslie matrix models for age-structured populations. The central idea of the Leslie matrix is that if the number of individuals of age a is represented by N_a , then the number of individuals of age $a + 1$ at the next time step is given by $N_{a+1} = S_{a,a+1}N_a$. The matrix $S_{i,j}$ is called the Leslie matrix, and the elements are zero unless $j = i + 1$ (I ignore the reproduction terms, which are not relevant to this discussion). Clearly a discrete matrix model like this is ideally suited to computer simulation, since the age structure can be stored in a vector (N_0, N_1, N_2, \dots) and the incrementation of the population vector from one time step to another can be carried out by a loop like

```
DO 100, I = 0 TO IMAX - 1
  N(I + 1) = S(I, I + 1) * N(I)
NEXT I
```

with some additional code to calculate the number newborns, $N(0)$. The problem is that this code generates nonsense! The reason is clear if one actually traces through the code; on the first iteration of the loop, the old value of $N(0)$ is used to calculate the new value of $N(1)$. The next iteration uses $N(1)$ to calculate the new value of $N(2)$, but since the value of $N(1)$ has already been updated, the calculation of $N(2)$ is incorrect, and the error gets worse with succeeding iterations. The solution is of course simply to reverse the order of computation by changing the first line to read

```
DO 100, I = IMAX - 1 TO 0
```

but I have to confess that I have lost track of the number of times I have had to explain this simple fact to colleagues who reluctantly sought help with their computer simulations.

Lest I seem overly critical about the tools in the modeller's kit, let me praise the systems identification approach, which I have written about in the past using the term "top-down modelling" (Silvert, 1981a; 1981b), but which might better be called "data-driven modelling". Both terms refer to the process of developing models that fit the data, rather than making a set of assumptions about the system and creating models on the basis of these assumptions, without regard to whether they fit the model or not.

No discussion of quantitative modelling techniques would be complete without reference to sensitivity analysis. Sensitivity analysis can be used not only to test the performance of models, but can also be a valuable guide to the development of models and can help identify what can safely be omitted. In some cases sensitivity analysis can even show the impossibility of developing a model that does what we want it to; Miller *et al.* (1973) found that a mosquito control model that they were working on was so sensitive to weather conditions during the larval stage that it was impossible to use it for practical predictions. Models are always a simplification of reality, and by carrying out sensitivity analysis at an early stage of the planning process it may be possible to avoid getting caught in a quagmire of detail that ultimately will not affect model performance. This can be as much a political as a scientific issue, since modellers are under constant pressure to include detailed information that is not really useful. After all, someone who has spent his career studying the reproductive behaviour of some organism is not likely to be happy to discover that it does not play a large role in the aspects of the ecosystem that are being modelled. It therefore falls on the shoulders of the modeller to maintain a clear vision of just how complex the model should be (see Silvert, 1996 for some ideas on how to do this), and to fight the inclusion of material that will make the development of the model harder without making its performance any better.

MODEL ASSUMPTIONS

More than anything else that a modeller learns, it is essential to understand the role that assumptions play in development of a model. Every equation, including those used for statistical analysis, is based on

certain underlying assumptions, and the primary role of the modeller should be to identify and understand these assumptions, and to keep them constantly in mind in order to see whether they are valid.

An interesting example of the role of assumptions in modelling is the book *The Limits to Growth* (Meadows *et al.*, 1972) which brought simulation modelling to public attention and inspired the present interest in modelling. The book deals with many different scenarios describing how mankind will deal with the issues of population growth, resource depletion, pollution, and so on, but there seems to be a curious assumption about how we will deal with this increasing stress and suffering—no matter how bad things get, we won't go to war about them. As the authors explain, “we have also ignored discontinuous events such as wars or epidemics” (Meadows *et al.*, 1972, p. 132), but this drastic assumption received little attention in the reviews of their study. A look at history shows that wars and epidemics are among the more continuous phenomena governing human existence, and under the extreme pressures described in *The Limits to Growth* they would be almost inevitable, and thus predictable. To be sure, the authors acknowledge that their models are optimistic estimates because they ignore these factors, but this sounds a bit like modelling the time it takes to drive across Russia non-stop, based on the assumption that one doesn't run out of gas!

The responsibility of a modeller to understand and criticise the assumptions of a model arise not only from a sense of professional obligation, but also from the circumstance that the modeller is usually the most mathematically sophisticated member of a research team and thus can bring quantitative judgement to bear even on problems where the other team members understand the workings of the system better. For example, I once reviewed a study of primary production in an estuary where one of the critical factors was the turbidity of the water—the more turbid the water, the less light reaches the phytoplankton, and thus the lower the primary production. The paper was based on a model in which primary production was inversely proportional to the average turbidity. However, this line of reasoning is misleading, since what matters most to the phytoplankton is the volume of water with sufficient light to grow, and thus the key variable is not turbidity, but the depth of the euphotic zone (the zone with sufficient light), and that is inversely related to turbidity. Consequently

the model should have used the average inverse turbidity, rather than the inverse of the average turbidity. Of course this sounds like a play on words, unless we consider the following extreme case; suppose that a dump truck empties a huge load of sand at one end of a clear lake. During the time that it takes the sand to settle, the average turbidity of the water is very high, since it is proportional to the total sediment load. However, since the sand is concentrated in only a small area, the volume of the euphotic zone, which depends on the average inverse turbidity, is decreased only slightly. In an actual estuary the situation is not this extreme, but since there can be localised areas where the sediment load is high (such as channels or regions which receive sewage wastes), the difference between average inverse turbidity and the inverse of the average turbidity can be considerable. These subtle points, which often occur only to those who are used to dealing with mathematical systems, can make the difference between a sloppy model and one which describes the system correctly.

These considerations do not only apply to complex scientific models, since, as pointed out above, modelling is a universal (if sometimes unconscious) human activity. The ongoing debate about intelligence and race offers a depressing example of the dangers of neglecting the assumptions behind a model. Numerous studies (which I will not dignify by citing) claim to show that the average intelligence of blacks is lower than that of whites, and these results are taken to mean that blacks are not qualified to be doctors, engineers, or football quarterbacks. The inference is made using an implicit model, and the “data” on intelligence are inputs to this model. But the model, although based on the perfectly reasonable assumption that intelligence is necessary to be a doctor, engineer, or football quarterback, it is clearly wrong, as can be seen by replacing the concept of intelligence with that of height. It is entirely possible that the average height of blacks is lower than that of whites, so an analogous model would say that blacks are not qualified for positions requiring height. Would this mean that the owners of major basketball teams would have to fire all of their black stars? Obviously not, and clearly this is only one example of the tragically common error of ascribing mean properties of groups to individuals—unfortunately this is one of the most common modelling errors encountered in the real world, as illustrated by many newspaper reports about how police forces treat minority groups.

DATA-DRIVEN MODELLING, OR SYSTEMS IDENTIFICATION

The usual approach to model development is to characterise the system, make some assumptions about how it works, translate these assumptions into equations, and start programming. This does not always work. If the model does not fit the data, it may well be that the assumptions are wrong. If so, what does one do next?

One approach is to ask what kinds of models could fit the data, and see whether any of these provide plausible alternatives. This approach, where one bases the model structure directly on the data, has been called “top-down modelling”, but since this can lead to confusion with top-down control, I will use the terms “data-driven modelling”, or “system identification” which is widely used in time-series analysis.

Path Analysis

To illustrate how analysis of the data can overthrow some of our assumptions in surprising ways, I’ll begin with a problem that I worked on during the 1970s when I was involved with fisheries management issues. I was trying to understand how fish landings were affected by both water temperature and fishing effort, and I decided to use path analysis to interpret the data (Silvert, 1981b; Silvert and Dickie, 1982). Path analysis is a quasi-statistical method for inferring causal relationships, and it is widely used in the social sciences (Li, 1975). It seemed reasonable to assume that both temperature and effort would influence landings (temperature as an environmental variable affecting recruitment and growth), but consistently I found that the landings affected the effort and not the other way around. This was baffling until I spoke to one of the scientists who collected the data, who laughed at my dilemma and explained that they never did collect any effort data, even though they were supposed to. Instead they simply looked at the landings and inferred effort from these data—if over half the fish in the hold was flounder, they listed it as a flounder trip. In other words, the effort “data” really did depend on the landings.

Although this was frustrating news at the time—I really wanted those effort data!—in retrospect this was very pleasing. It showed me

that path analysis was a useful tool for system identification, and it taught me that if a model doesn't make any sense, it may help to look at the experimental data with a critical eye and review the assumptions on which both the model and the interpretation of the data are based.

Sedimentation Problem

As an example of how modelling can serve as a tool to reveal aspects of nature that may not be directly evident in the data I would like to describe a frustrating experience in the analysis of sedimentological data (Silvert, 1981b). The experiment was in principle simple—a beaker of sediment was stirred and then allowed to settle out. Samples taken with a pipette at fixed depth over time showed a decreasing concentration of particulate matter, especially in the larger size range, due to the more rapid sinking of larger particles. Straightforward application of Stokes Law should explain the decrease in concentration over time. However, the concentration did not fall off as rapidly as it should, and there was a tail of fine particles that persisted for much longer than Stokes Law settling predicted. The only explanations that I could come up involved reintroduction of fine particles into the water column through microbial contamination, dust settling on the surface, or inadvertent mixing during the pipetting operation. The experimentalist with whom I was working angrily rejected these possibilities, which were seen as an attack on his laboratory technique, and said that if I couldn't model these data he would find someone else who could (a common response!). Some time later I asked whatever happened to that experiment, and he told me that the data had to be discarded because there proved to be microbes growing in the beakers. When I rashly pointed out that I had suggested this two years earlier, he got even angrier than he had the first time.

There are two morals to this unhappy story. The first is that a good model can reveal information that cannot easily be seen in the data (I might add that some really sophisticated scaling arguments went into this theoretical analysis). The second is that experimentalists don't always appreciate it when a modeller discovers something that they missed.

Toxicology Model

I once tried to model the toxicology of a motile alga (Vandermeulen *et al.*, 1983), and started off with the perfectly plausible assumption that this was a dose-response situation. In other words, the longer the algae were exposed to a contaminant, the greater the probability that they would exhibit a toxic response. No matter how I adjusted the parameters of the model, it produced unacceptable results. I finally decided to follow my own advice and use a data-driven approach. As soon as I did this, it became apparent that the toxic response was independent of dose—in other words, the probability that an alga would exhibit a toxic response during the next unit of time was completely independent of how long it had been exposed. The resulting model fit the data perfectly, although we still don't know why!

VALIDATION: CAN WE TRUST DATA?

It is widely believed that a model should fit the data, and if a model does not fit the data, the model is wrong. There are in fact two possibilities: the model may be wrong, or the data may be wrong (or both). Data are not reality, no matter what experimentalists may think. They are the outputs of experiments, and experiments are based on models. There is no reason to believe that a model devised by an experimentalist is better than one developed by a theorist. In any case, if one discards a model, it is essential to determine what went wrong—were the underlying assumptions of the model, which are usually specified by the experimentalists, wrong, or was the model incorrectly constructed? If the former, then both the model and the experiments need to be re-evaluated, and if the latter is true, then it should not be hard to fix.

Aliasing

Some of the most common problems in the design and analysis of experiments involve aliasing, which occurs when discrete sampling is used to measure continuous processes. Typically the measurements are interpolated with straight lines or higher order splines (de Boor, 1978;

a straight line can be thought of as a first-order spline), but if the system varies on scales small compared to the interval between measurements, serious errors can arise.

This is exemplified by studies of shellfish toxicity, where we use models to relate the presence of toxic algae in the water column to toxins in shellfish. The algae are usually sampled with relatively small bottles, and because of resource constraints it is difficult to maintain a high sampling frequency. This means that algae, which are very patchy and occur at varying depths in the water column, are sampled at a few discrete locations one or two times a week.

In a study of Paralytic Shellfish Poisoning (PSP) toxins in mussels (Silvert and Cembella, 1995), there were several apparent discrepancies between the model and the data, indicated by the arrows in Fig. 2. In some cases there was a high value of measured algal toxicity but no significant increase in shellfish toxicity, while in others the shellfish reached high toxin levels even though no toxicity was measured in the water column. If the model is fundamentally wrong, then the underlying assumption that mussels accumulate toxins from the algae they consume must be questioned.

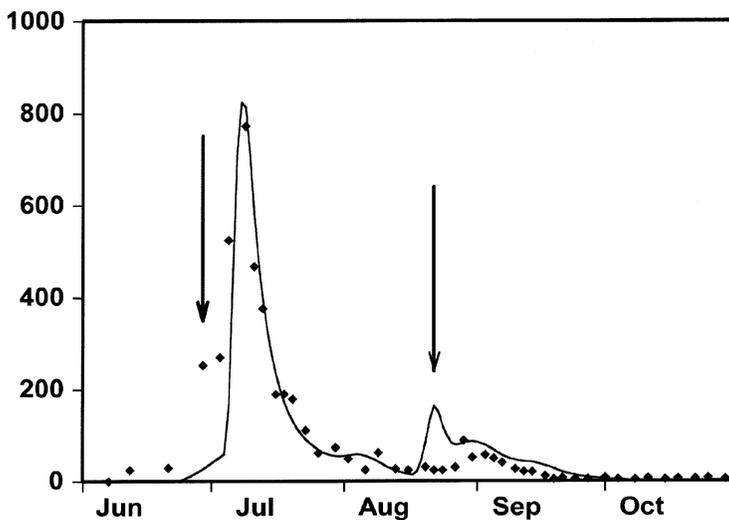


FIGURE 2 Data and simulated toxicity for PSP toxins in mussels, adapted from Silvert and Cembella (1995). The points represent experimental data, the line is the result of a simulation model.

This is clearly an issue that arises at the end of June in the simulation shown in Fig. 2, since the mussels already contain toxin levels over $200 \mu\text{g STX equivalent } (100 \text{ g})^{-1}$ while the simulation predicts much lower levels (these complicated units refer to the toxicity as compared to the “standard” STX toxin equivalent, measured in $\mu\text{g per } 100 \text{ g}$ of tissue). It is easy to see why the simulation model does not predict the observed toxicity—no toxic algae had yet been detected! Presumably toxic algae were present and had been consumed by the mussels, but they had been missed in the sampling procedure.

The opposite happens in August, when the spike in simulated toxicity is due to a single sample, but it was probably just a small patch of toxic algae that happened to be at the sampling location at the right time, and was not representative of the water column. In both cases it makes much more sense to ascribe the discrepancies to problems with the sampling procedure than to errors in the model, which is based on very simple and fundamental assumptions about how shellfish become toxic. It is however worth noting that problems like this, which are very common, usually prompt experimentalists to criticise the model, although I have never had any problem showing them that the errors are inherent in the model assumptions and have nothing to do with bad modelling procedures.

Replication

A major problem with validation in oceanography is that we seldom have an adequate set of replicate data to work with, and consequently there are often special situations which cause unexpected discrepancies. For example, none of the fisheries models in use at the time correctly predicted the dramatic decline in fish landings in the North Atlantic during the late 1930s and early 1940s. The explanation was of course the decrease in effort when fishing fleets were driven off the water by German U-boats or converted to military vessels. Surely no one would fault fisheries modellers for failing to take these factors into account!

I recently had an assuming experience trying to test a fisheries model that I helped develop in the mid-1980s (Silvert and Crawford, 1988). The model predicted that catches of small pelagic fishes in the coastal waters of Japan were at an unsustainable peak and would soon start to

fall. However, although the total catch has fallen to a third of its value ten years earlier, and the catch of the dominant species in this category, *Sardinops melanosticta* (sardine), has been falling, the combined catch of *Engraulis japonicus* (anchovy) and *Trachurus japonicus* (horse mackerel) suddenly started to climb and has been going in just the opposite direction from what the model predicts. Clearly this proves the model false—unless one looks more carefully at the data and considers an alternate hypothesis. This analysis is based on data from the Fisheries and Agriculture Organisation (FAO), obtained from statistics provided by national fisheries agencies, and the data “show” that before 1990 the People’s Republic of China was not catching any fish! During the past 10 years their catches of *Sardinops* and *Engraulis* have been increasing at a remarkable rate (Silvert, 1997), and by 1997 the Chinese were catching more fish in this area than all other countries combined, as illustrated by Fig. 3 which shows catches over time (the Chinese still do not report any catches of *Trachurus*).

It may be that the Chinese were catching fish before 1990 and simply not reporting their landings data to FAO, in which case the data cannot be used to falsify my model. Given the many wonderful ways in which the Chinese prepare fish it is hard to believe that they really were not catching any, at least in this area, before 1990, but this is not a question that can easily be resolved by simply looking at “scientific” data.

Similarly, the Silvert and Crawford (1988) paper fails to predict a recent surge in catches of *Trachurus symmetricus* off the coast of Peru and Chile, but this increase is associated with a huge expansion in the area fished, which now extends almost as far as New Zealand, and it is also hard to use this discrepancy to falsify the model. These are not special cases—it is rare for a model to be tested with data that are fully consistent with those on which the model is based.

False Implicit Assumptions

One of the most common experiments in our field is to measure the abundance of organisms by dragging a net through the water. Think of all the assumptions and modelling that this involves! Aside from the mathematical issue of calculating the volume of water filtered, there are more serious issues of net avoidance, extrusion of soft-bodied organisms through the net, and diversion of organisms pushed aside

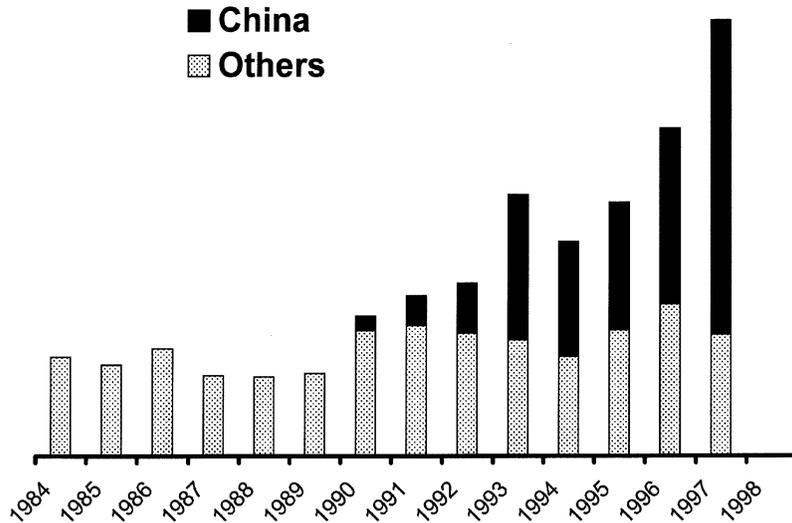


FIGURE 3 Fish catches in the coastal waters near Japan, showing reported landings by China and by other countries.

by the bow wave. Also, how does one deal with patchiness? At a more fundamental level, was the sampling carried out in the right part of the water column? Midwater trawls don't catch many crabs and flounder. Perhaps the most common experimental mistake is to assume that any effect that isn't measured can be ignored, even when this flies in the face of every rational argument. For example, since there is little quantitative data on net avoidance, it is commonly assumed that it can be ignored. This can lead to huge underestimates.

During the 1970s I carried out an informal project to see how much of the data I used was wrong. The criterion was not that I thought it was wrong, but that I could convince the people who collected the data that they were wrong. In numerically tabulated data the error rate consistently ran around 2%—perhaps now that data are logged automatically this has improved, but it shows that outliers are to be expected. More serious are gross errors due to carelessness in the calculations. I once received tables of plankton concentrations that were outrageously high. It turned out that the data were processed by someone who confused *mm* with *cm* and thus produced estimated displacement volumes that were too high by a factor of 1000. Perhaps the most disturbing aspect of this experience was the discovery that

most of my colleagues, as well as the financial officers whom I asked to suspend payment to the contractor until the data were corrected, felt that I was unreasonable to make a fuss about such a small mistake!

Consistently Wrong Data

There are some areas where I believe that the data are consistently wrong. I have been working for many years on the toxicology of shellfish, and consistently find myself unable to get correct results (*e.g.*, Silvert and Subba Rao, 1992). This could of course reflect my own inadequacies as a modeller, but the situation is so simple that I think there must be more to it than that.

Models of shellfish kinetics are pretty simple, although like most models they can be elaborated without limit. If the concentration of toxins in the water column is X , then the concentration C in the shellfish is given by the uptake-clearance equation

$$dC/dt = aX - bC,$$

where a represents uptake and assimilation, and b is a loss term. The problem I repeatedly find is that a can be obtained from experimental values, and it is almost always too low to generate correct curves of C over time. Even when b is set to zero, so that there is no detoxification at all, observed concentrations of toxins in shellfish can reach levels much higher than the model can generate. Since the main purpose of this kind of model is to warn us when toxin levels can reach dangerous levels, this is a serious problem.

My personal opinion is that shellfish can feed at much higher rates than have been observed in the laboratory, and can therefore ingest large quantities of toxin when the conditions are right. The experimentalists whose data I question do not agree with me, but I propose this example to illustrate the type of question that modellers have to be prepared to deal with.

HORROR STORIES

Although some of my own experiences have been frustrating, these problems are minor compared with some of more dramatic case his-

tories of modelling gone astray. Here are a few tales that should prove as instructive as they are terrifying.

Starved Amphipods

I'll start with a story that is funny in retrospect, although it was a disaster for scientist involved. We had a post-doctoral student in our laboratory who was studying the growth rates and feeding efficiency of amphipods (small marine invertebrates). This was a subject of considerable interest to me, so I asked if we could collaborate, but he informed me that this was a purely experimental study that did not require any modelling or other theoretical nonsense.

He continued for several months, feeding the bugs various quantities of food and measuring their growth rates. Finally, to wrap up the study, he decided to try starving them. No food. They continued to grow! Obviously the results of the experiment were suspect, to say the least.

He wasn't the first experimentalist to assume that the things he couldn't see (in this case diatoms growing on the sides of the Petri dishes) didn't exist, and he certainly wasn't the first to discover that there are some theoretical concepts that one cannot avoid, like conservation of energy. Unfortunately, he certainly will not be the last to make these mistakes.

Port Hacking Estuary

This may be the most expensive example of bad modelling that exists in marine ecology. The Port Hacking Estuary Project was a huge multi-year and multi-million dollar study carried out in Australia. The project was intended to produce an ecosystem model that would tie all the results together. They hired a young mathematician with what seemed to be a good background in biological modelling. But as time went by, the results of his model seemed to be getting more and more ridiculous, to the point where they decided that they should look over his shoulder.

It turns out that before being hired for this project he had been working on pharmacological models, which are linear donor-acceptor structures. He approached ecological modelling the same way. But

ecological models are not linear (not even the well-known Lotka–Volterra model), and their mathematical structure is totally different. The modelling aspect of the work could not be salvaged in the time frame remaining, and the entire project suffered as a result. As one of the participants put it, with admirable restraint and understatement, “It seemed that we may have erred in opting for linear, donor-dependent functions, thereby promoting disbelief and discontent among the participants; the published models generally contained complex non-linear functions . . .” (Cuff, 1983).

The Case of the Missing Menhaden

Models with spatial structure are particularly dangerous. A well-known theoretical ecologist landed a nice contract to study the potential impacts of a project in the Gulf of Mexico that would increase salinity over a small area, about two square kilometres. I was present when he spoke on his results, and was surprised to see that one of the major effects would be a severe reduction in the menhaden population. When I asked the mechanism for this, he explained that the increased salinity would cause a reduction in plankton, and the menhaden would starve. I protested that since menhaden are pelagic fish that move around a lot, depletion of their food resources over such a small area certainly could not cause starvation. Clearly the ability of fish to swim was left out of this model!

Foxes and Hares

This model is another example of many that founder on the ability of animals to move around. Several years ago I refereed a very sophisticated paper that used cellular automata to model the spatial distribution of animals, and a worked example was included that described the spatial distribution of foxes and hares. The model showed that if the distribution of hares was patchy, the foxes would starve. The way that this occurred was that the hares might end up in some cells of the grid, and foxes in the other parts of the grid would not find them. The grid spacing was about 100 m, and frankly I think that if you put a fox 100 m away from food, he will find it before he starves to death!

I don't mean to suggest by these two examples that migration and foraging are easy to model—far from it. But they are present in most ecosystems, and if you either ignore these effects or model them incorrectly, you will get nonsensical results.

Patchiness

Most organisms are distributed in patches. We don't always understand why this happens, although some theorists have suggested that uniform distributions of predators and prey are unstable and automatically form patches. But whether your models can predict patchiness or not, you have to recognise it as a reality that must be included in models.

About twenty years ago I was trying to put together large-scale ecosystem models of continental shelf ecosystems, and consulted a number of zooplanktologists and fisheries biologists about grazing rates (Silvert, 1988). The values were so low that I couldn't make any sense out of them—by my calculations, there wasn't enough energy flow to produce fish, even on the Grand Banks of Newfoundland. Then I remembered a lovely paper by some of my colleagues on how whales feed, in which they showed that if baleen whales consumed krill in the southern ocean by swimming at random, rather than exploiting patches, they would have to move at the speed of sound with their mouths wide open, 24 hours per day (Brodie *et al.*, 1978). So I went back to the experts to see what role patchiness played in their estimates, and I was informed that no one had ever measured this effect, so the values they had given me were based on laboratory experiments, typically ones where zooplankton were fed on uniform concentrations of algae in stirred tanks. Clearly this kind of experiment is not a good model for what happens in the wild!

The Complexity–Stability Debate

One of the most elaborate areas of investigation in theoretical ecology has been the debate about whether ecosystem complexity produces greater or lesser stability. Hundreds of papers have been written on this topic, and the most common approach has been to carry out numerical simulations of randomly connected ecosystem models. Much of the interest in the issue is driven by the surprising discovery that complex

systems prove to be less stable than simple ones in most cases, which is just the opposite of what many ecologists had long believed.

The models used were more sophisticated than the linear ones of the Port Hacking project, and consisted of generalised versions of the well-known Lotka–Volterra model, with terms of the form

$$dx_i/dt = \dots + a_{ij} x_i x_j + \dots,$$

where the x_i are the different populations and the a_{ij} are random interaction coefficients.

One interesting aspect of these simulation studies was the discovery that omnivory is destabilising. Almost any biologist will tell you that the opposite should be the case. If you ask him why he believes that, he will explain that since omnivores have a wider choice of prey, they will switch to the more abundant prey and ignore prey that are scarce, thereby giving endangered species a chance to recover.

Aha! This mechanism is not present in any of the computer simulations. Careful examination of the Lotka–Volterra interactions shown above shows that the amount of time that predators spend feeding on their prey is independent of the prey's abundance. This is why omnivory is destabilising—the omnivores relentlessly pursue even the scarcest prey, and they gain enough energy from the more abundant prey to keep going until the scarcer species have been driven extinct. In real systems predators can switch from one prey to another, and this is the crucial difference between theoretical instability and real stability (Silvert, 1983).

Of course it is not easy to model the response of an omnivore to changes in relative prey abundance. Most data on how animals choose their food (electivity indices) are based on time-averaged feeding rates, and are not very reliable. But once again, the dynamics of the ecosystem very often depend on something that is difficult to measure, but if we ignore the effect, the results of our models will probably turn out to be nonsense.

SUMMARY

Modelling is not an easy profession, and it is certainly not one that inspires much respect. There is a tendency to retreat into abstruse

mathematical formalism and focus on complex statistical techniques and computer programming, areas in which the properly trained modeller can excel without fear of criticism. Unfortunately, this does not automatically lead to good models.

Modellers have to meet their experimental colleagues on the ground. They have to understand what happens in the field, how experiments are conducted, and what the data mean. They have to put the priority on science, not on mathematics. Above all, models have to be based on correct assumptions about how reality works. Getting the mathematics right is important, definitely, but good mathematics will never salvage a model that is conceptually incorrect.

Acknowledgements

This paper was originally prepared as a keynote address for the Young Scientists Conference organised by ICES, the International Council for the Exploration of the Seas, in Denmark in November 1999. I wish to express my thanks to the organisers and participants for the opportunity to develop, present, and discuss my ideas.

Thanks is also due to my many colleagues who have shared both their ideas and their hard-earned experimental data with me. A theorist or modeller cannot work alone. Often however a modeller is viewed as a parasite who sucks out people's data and leaves them with no credit for their labours. If it were not for the trust and confidence of my co-workers I would have no story to tell.

References

- Brodie, P.F., Sameoto, D.D. and Sheldon, R.W. (1978), "Population densities of euphausiids off Nova Scotia as indicated by net samples, whale stomach contents, and sonar." *Limnological Oceanography*, **23**, 1264–1267.
- Cuff, W.R. (1983), "An evaluation of the Port Hacking Project from the Viewpoint of Applied Science Synthesis and modelling of Intermittent Estuaries." In: *Synthesis and Modelling of Intermittent Estuaries*, edited by W.R. Cuff and M. Tomczak Jr., Springer, Berlin, 273–292.
- Carl, de Boor (1978), *A Practical Guide to Splines*, Springer-Verlag, New York.
- Lee, D.B., Jr. (1973), "Requiem for large-scale models." *Journal of the American Institute of Planners*, **39**, 163–178.
- Li, C.C. (1975), *Path Analysis—a primer*, Boxwood Press, Pacific Grove, California.
- Meadows, Donella H., Meadows, Dennis L., Randers, Jørgen and Behrens III, William W. (1972), *The Limits to Growth*, Signet, New York.

- Miller, R.D., Weidhaas, D.E. and Hall, R.C. (1973), "Parameter sensitivity in insect population modeling," *Journal of Theoretical Biology*, **42**, 263–274.
- Silvert, William (1981a), "Principles of ecosystem modelling." In: *Analysis of Marine Ecosystems*, edited by A.R. Longhurst, Academic Press, London, 651–676.
- Silvert, William (1981b), "Top-down modelling in marine ecology." In: *Progress in Ecological Engineering and Management by Mathematical Modelling*, edited by D.M. Dubois, Editions Cebedoc, Liege, 259–270.
- Silvert, William (1982), "Top-down modeling of multispecies fisheries." In: *Multispecies Approaches to Fisheries Management Advice*, edited by M.C. Mercer, Canadian Special Publication in Fisheries and Aquatic Sciences, **59**, 24–27.
- Silvert, William and Dickie, Lloyd M. (1982), *Multispecies Approaches to Fisheries Management Advice*, edited by M.C. Mercer, Canadian Special Publication in Fisheries and Aquatic Sciences, **59**, 163–169.
- Silvert, William (1983), "Is dynamical systems theory the best way to understand ecosystem stability?" In: *Population Biology*, edited by H.I. Freedman and C. Strobeck, Springer-Verlag Lecture Notes in Mathematics, **52**, 366–371.
- Silvert, William (1985), "Model uniqueness." In: *Towards the Inclusion of Fishery Interactions in Management Advice*, edited by R. Mahon, Canadian Technical Report in Fisheries and Aquatic Sciences, 1347, 132–135.
- Silvert, William (1988), "Generic Models of Continental Shelf Ecosystems." In: *Eco-dynamics: Contributions to Theoretical Ecology*, edited by W. Wolff, C.-J. Soeder and F.R. Drepper, Springer-Verlag Lecture Notes in Physics, Berlin, 153–161.
- Silvert, W. and Crawford, R.J.M. (1988), "The Periodic Replacement of One Fish Stock by Another." In: *Long Term Changes in Marine Fish Populations*, edited by T. Wyatt and M.G. Larrañeta, Imprenta Real, Bayona, Spain, 161–180.
- Silvert, W. and Subba Rao, D.V. (1992), "Dynamic model of the flux of domoic acid, a neurotoxin, through a *Mytilus edulis* population." *Canadian Journal of Fisheries and Aquatic Sciences*, **49**, 400–405.
- Silvert, W.L. and Cembella, A.D. (1995), "Dynamic modelling of phycotoxin kinetics in the blue mussel, *Mytilus edulis*, with implications for other marine invertebrates." *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 521–531.
- Silvert, William (1996), "Complexity." *Journal Biol. Syst.*, **4**, 585–591.
- Silvert, William (1997), "The role of interactions in long-term population cycles." *Environmental Modeling and Assessment*, **2**, 49–54.
- Vandermeulen, J.H., Silvert, W. and Foda, A. (1983), "Sublethal hydrocarbon toxicity in the marine unicellular alga *Pavlova lutheri* Droop," *Aquatic Toxicology*, **4**, 31–49.



William Silvert is a theoretical ecologist living in Portugal. He received a Ph.D. in theoretical physics from Brown University in 1965 and taught physics in the United States and Canada for 10 years before becoming interested in marine science during the mid 1970s. In 1978 he joined the Marine Ecology Laboratory at the Bedford Institute of Oceanography in Nova Scotia, and he worked there as a research scientist in habitat ecology for over 20 years. Recently he moved to Portugal and is currently a visiting scientist at the Institute for Fisheries and Sea Research in Lisbon.