



## **Chapter 8 Effect of complexity on ecosystem models**

### **Abstract**

Ecosystem rather than species management has become an explicit part of policies that feature in international treaties and national legislation. Many of the tools that will be needed to fulfil the requirements of these policies are still in an early stage of development. One such tool is trophic ecosystem modelling. These models have been put forward as an instrument that can aid system-level understanding and provide insight into the potential impacts of human activities. Despite this, there are many gaps in our knowledge of their strengths and weaknesses. In particular, little is known about the effect of the level of detail in a model on its performance. There has been some consideration of the effects of model formulation, as well as the effects of physical, biological and chemical scope of ecosystem models on their performance. A review of existing research indicates that there is a humped relationship between model detail and performance, and that there are some guiding principles to consider during model development. Our review also gives some insight into which model structures and assumptions are likely to aid understanding and management, and which may be unnecessary. Further, it provides some understanding of whether some models can capture properties of real systems that other models cannot. The main recommendation is that, the use of a single “ultimate” ecosystem model is ill advised, while the comparative and confirmatory use of multiple “minimum-realistic” models is strongly recommended.

### **Keywords**

biogeochemical, ecosystem, model, theoretical modelling, mass balance, complexity, model structure, model scope, model closure, multispecies

## 8.1 Introduction

During the last 30 years, every time there has been a major push to understand entire ecosystems, rather than isolated components, ecosystem models have become a popular tool (Watt 1975, Halfon 1979, Walters et al. 1997, Sainsbury et al. 2000). However, in the past complex general models have acquired a poor reputation, at least in some quarters, primarily because of two factors (Jørgensen et al. 1992). First, these models are often so large and complex that they are not necessarily cost efficient, with the majority of the modelling resources often spent in development and maintenance rather than on their application (Watt 1975). Second, complexity introduced for the sake of completeness accomplishes nothing if the resulting model is actually of poor quality (O'Neill 1975, Silvert 1981, DeCoursey 1992). While modern computing power makes ecosystem models attractive as computational restraints are lifted (Beck 1999), this does not solve the problems of uncertain model specification, parameterisation and system understanding, or the effects of model structure and detail on model performance (Silvert 1981, Jørgensen 1994). These areas of modelling still require much attention and the need becomes more urgent with increasing pressure on scientists and managers for “whole system” approaches, predictions and policies.

### *Ecosystem Models*

Terminology associated with ecosystem models is confusing in that ecosystem models can refer to everything from total system models to models that focus solely on fisheries (the “top end” of the web) or water quality (the “bottom end” of the web). Here we use the term ecosystem model to refer to total system models rather than those multispecies models tied to one end of the food web or the other, which will be identified as water quality and fisheries multispecies models.

Within the last few decades, increasing computer power and a shift in the focus

of scientific and political thought has lead to a rapid growth in the number of multispecies and ecosystem models in existence (Silvert et al. 1981, Breckling and Müller 1994). For marine systems, the majority of multispecies and ecosystem models (a) use time-dependent differential equations, (b) conserve some currency (either biomass or nutrients), (c) concentrate on one part of the trophic web (either the nutrients and plankton, or the fish and other harvested species) and (d) separate physical and biological components and handle them in different ways. It is not surprising that the majority of published and applied multispecies models tend to concentrate on one end of the food web or the other, as the two main fields to embrace and make principal use of such models are fisheries and water quality. In both fields, the history of model development can be “characterised by the growing intricacy of their internal structure” (Fransz 1991). There have been a number of attempts at producing trophic ecosystem models for the marine environment, including: the cove model of Patten et al. (1975); the multispecies model by Andersen and Ursin (1977); the Fjord model of Bax and Eliassen (1990); ECOPATH with ECOSIM (Christensen et al. 2000); ERSEM I (Baretta et al. 1995) and II (Baretta-Bekker and Baretta 1997); and IGBEM and BM2 (chapters 1 and 2). Nevertheless, compared with the widespread use of water quality and fisheries multispecies models, the use of ecosystem models remains limited. As a result, there is still a lot of scope for the development of a thorough understanding of the implications of model structure on performance for these kind of models.

The various types of multispecies and ecosystem models each have associated advantages and problems (Table 8.1), but there is a list of features and potential drawbacks common to them all. In general, such models improve our understanding of systems by reflecting the two-way nature of system dynamics. Human impact on one part of a system can spread to other parts of the system, but system feedbacks and interrelations can also mean that an impact can have unexpected consequences even for

**Table 8.1:** Summary of the main strengths and weaknesses associated with the main types of multispecies and ecosystem models.

Model Type	Examples	Potential Strengths	Potential Weaknesses	References
Pelagic Pooled Ecosystem Models (or Water Quality Models)	NPZ, NPZD, FINEST	<ul style="list-style-type: none"> <li>• Usually include environmental forcing</li> <li>• Some include environmental heterogeneity (though often only vertically)</li> <li>• Usually include constraints (via nutrient limitation)</li> <li>• Effects of formulation have received some attention</li> <li>• Some include size-age structure of all biological components (though may not always be justified)</li> </ul>	<ul style="list-style-type: none"> <li>• Ignore top-down effects from fish and fisheries (some even ignore zooplankton)</li> <li>• Can have a heavy reliance on source and sink terms</li> <li>• Tend to ignore benthic communities</li> <li>• May be steady-state or purely empirical</li> </ul>	Fransz et al. 1991, Murray and Parslow 1997, Hollowed et al. 2000, Tett and Wilson 2000
Multi-species Models	Technical Interaction Models (TIM), Statistical Assessment Models (SAM), Multispecies Production Models (MPM), MSVPA, MSFOR	<ul style="list-style-type: none"> <li>• Can usually account for age-size dependencies and temporal variation (except for MPMs)</li> <li>• Often include constraints (e.g. conservation of total biomass)</li> <li>• The effects of many underlying assumptions are well known</li> </ul>	<ul style="list-style-type: none"> <li>• Ignore bottom-up effects from lower trophic levels (e.g. primary producers)</li> <li>• Often ignore nutrient or environmental forcing</li> <li>• Some do not explicitly describe the trophic interactions (TIMs) or feedbacks (SAMs).</li> <li>• Often employ unsophisticated mortality terms</li> <li>• Can be sensitive to assumptions regarding recruitment</li> </ul>	Hollowed et al. 2000
Biomass Size Spectrum Models		<ul style="list-style-type: none"> <li>• Can incorporate the entire trophic web</li> <li>• Do not require extensive diet matrices as the only distinguishing characteristic used is size class (body weight)</li> <li>• Do not require extensive population or taxonomic data</li> </ul>	<ul style="list-style-type: none"> <li>• Assumptions with regard to feeding habits may not hold for benthic communities</li> <li>• Ignore influx due to reproduction and mortality other than fishing and predation</li> <li>• Mathematical properties of the underlying equations make them unsuitable for use over long time periods</li> <li>• Assume one preferred predator-prey weight ratio for the entire system</li> <li>• Often do not incorporate satiation or seasonality</li> </ul>	Silvert 1996b, Duplisea and Bravington 1999

**Table 8.1:** Continued

Model Type	Examples	Potential Strengths	Potential Weaknesses	References
Static Aggregate Models	ECOPATH, Skeub	<ul style="list-style-type: none"> <li>• Can make use of existing information and meta-analyses that would not be sufficient for intensive species based biogeochemical models</li> <li>• Trophic aggregation not consistent across the web (if aggregation is appropriate then this reduces the data needed for parameterisation)</li> <li>• Can incorporate the entire trophic web.</li> <li>• Constraints included (mass-balance, with the capability to build in biomass accumulation and depletion)</li> </ul>	<ul style="list-style-type: none"> <li>• Steady-state or instantaneous system state</li> <li>• Often use constant growth and production terms</li> <li>• Trophic aggregation not consistent across the web (this may affect results if aggregation is inappropriate)</li> <li>• Can be sensitive to diet-matrix used</li> </ul>	Christensen et al. 2000, Hollowed et al. 2000
Dynamic Aggregate Models	ECOSIM, DYNUMES, ECOSPACE	<ul style="list-style-type: none"> <li>• Trophic aggregation not consistent across the web (split-pool handling of groups is possible), which reduces data needs</li> <li>• Can incorporate the entire trophic web.</li> <li>• Forcing functions and trophic mediation functions can be applied</li> <li>• Use of predator-prey vulnerability parameters captures within pool heterogeneity with regard to predation (removes mass-action assumptions)</li> <li>• Constraints included (mainly via mass balance assumptions and predator-prey vulnerability)</li> </ul>	<ul style="list-style-type: none"> <li>• Seasonal behavioural changes difficult or impossible to incorporate</li> <li>• Can be sensitive to the value of the predator-prey vulnerability parameters</li> <li>• Trophic aggregation not consistent across the web (this may affect results if aggregation is inappropriate)</li> <li>• Often ignore environmental forcing.</li> <li>• Lack buffering against the effects of eutrophication</li> </ul>	Christensen et al. 2000, Hollowed et al. 2000, chapter 7

**Table 8.1:** Continued

Model Type	Examples	Potential Strengths	Potential Weaknesses	References
Biogeochemical Total System Model	ERSEM, IGBEM, BM2	<ul style="list-style-type: none"> <li>• Incorporate age-size dependencies for members of the higher trophic levels</li> <li>• Incorporate temporal variation and environmental forcing</li> <li>• Can incorporate the entire trophic web (in the form of functional groups)</li> <li>• Nutrient dynamics and recycling explicit.</li> <li>• Use of prey availability parameters captures within pool heterogeneity with regard to predation (removes mass-action assumptions)</li> <li>• Usually include spatial structure</li> <li>• Can have individual based components</li> <li>• Many constraints built in (e.g. nutrient limitation and predator satiation)</li> </ul>	<ul style="list-style-type: none"> <li>• Can be process and parameter intensive (therefore validation can be difficult and there can be large uncertainty, particularly in the benthic components)</li> <li>• Can be sensitive to assumptions regarding recruitment</li> <li>• Lack buffering against the effects of fishing</li> <li>• Lower trophic levels are represented by an aggregated pool per functional group, with no size or age structure (inappropriate if there are ontogenetic changes in habit or long generation spans)</li> </ul>	Baretta et al. 1995, Baretta-Bekker and Baretta 1997, Hollowed et al. 2000, this thesis

those groups directly affected by anthropogenic activities (Hollowed et al. 2000, Mace 2001, chapter 7). There are also a number of potential problems common to multispecies and ecosystem models: (1) they often require more information than single species models (which has associated costs); (2) they are more prone to suffer from issues about optimal complexity (what should be included in the implementation and what should be omitted); (3) operational objectives and monitoring indices can be hard to define for the real systems to be modelled, and if indices are to be used to summarise model output then the problem of defining appropriate measures also arises for the models; and (4) there are often alternative hypotheses about system structure and function (Silvert 1981, Jørgensen 1994, Mace 2001). These features are particularly important when dealing with ecosystem models. However, as multispecies and ecosystem models are the only models with the potential to answer the environmental questions that single species and pure hydrodynamic models cannot (Hollowed 2000, Mace 2001), the advantages of the intelligent and attentive application of such models can outweigh their potential pitfalls.

One of the main criticisms aimed at ecosystem models is that their potentially immense complexity can make predictions highly uncertain (Duplisea 2000). If the model output is to be used directly to determine management actions (as in a fisheries stock assessment model), such a characteristic is clearly undesirable (Butterworth 1989). In contrast, when such models are used as a guide to possible impacts and to explore implications of alternative broad policies, this property is no longer such a problem. This is particularly true if the robustness of the conclusions is tested against a range of models incorporating different structural and parametric assumptions, representing a range of plausible alternative hypotheses about how the particular system may work. Such a comparative and confirmatory approach has many advantages as it allows for the identification of effects and policies robust across levels of complexity,



uncertainty and underlying system and model assumptions (Reichert and Omlin 1997, Duplisea 2000, chapter 7). Another related use for more complex ecosystem models is as a test bed for simpler models that may be used in assessment of the system, or part of it. Applying simpler assessment models to “data” generated from ecosystem models is a useful way of checking the robustness of the assessment models, and of identifying the circumstances in which it may be appropriate to use them for more “tactical” management advice.

Regardless of their potential drawbacks, ecosystem models have the potential to identify issues and causes beyond the bounds possible in single species models, or even multispecies models in some cases. Management strategies implemented to achieve a certain goal may have the opposite effect if multispecies or ecosystem considerations are not included. For example, without multispecies models, the potential of a seal cull to have an effect opposite to the one intended would be overlooked (Punt and Leslie 1995, Yodzis 2001a). Beyond even these multispecies considerations, without the inclusion of the links between the upper and lower ends of the food web and the forces driving them, erroneous conclusions may be drawn about environmental and anthropogenic impacts, as alternative explanations and scenarios are overlooked (Steele 1998, Yodzis 2001b, chapter 7). For example, a decline in the biomass of a herbivorous fish may indicate overfishing, but it may equally indicate degradation of their main food reserves as a result of eutrophication (chapter 7).

Unfortunately, although the need to integrate comprehensive biological, physical and chemical models is recognised, reconciling and reducing the dimensions of complexity required in each of the areas is an enormous and often poorly understood task (Mace 2001, Nihoul and Delhez 1998). Much of the handling of model complexity has been dealt with by drawing on experience from prior modelling efforts (Murray and Parslow 1997). As a result, the systematic understanding of the effects of model

structure and detail on the performance of ecosystem models is still at an early stage.

There have not been many studies of the effect of model structure on marine ecosystem models. Unfortunately, this is a problem shared by the many branches of mathematical modelling (Brooks and Tobias 1996). Despite this, the study of model complexity in ecology can provide a sound point of departure for the study of the effects of the structure of ecosystem models on their performance. In turn, this can aid in understanding the mechanisms underlying the dynamics of real ecological systems.

## **8.2 A general history of the study of model complexity in ecology**

### *Model aggregation*

Optimal levels of model complexity, or appropriate degrees of trophic aggregation, remain a major problem in describing ecological systems. Thus, the study of the general properties of aggregation, and its limitations as a tool for use in model development and application, has received some attention in theoretical systems over the last 40 years. For instance, a number of researchers have considered the effects of aggregation from a theoretical standpoint and have produced some useful guidelines (Zeigler 1976, O'Neill 1979, O'Neill and Rust 1979, Cale and Odell 1980, Gardner et al. 1982, Cale et al. 1983, Iwasa et al. 1987, Bartell et al. 1988, Rastetter et al. 1992, chapter 5). The two most important rules that emerge to guide the use of trophic aggregation in model simplification are: (1) do not aggregate serially linked groups (predator and prey) (Gardner et al. 1982, chapter 5); and (2) do not aggregate species, age classes or functional groups with rate constants more than two- to three-fold different (Wiegert 1977, O'Neill and Rust 1979, Cale and Odell 1980, Gardner et al. 1982, chapter 5). Ignoring either of these guidelines is likely to result in a significant decline in model performance. Within the context of ecosystem models, this means that aggregating species to the level of functional groups is acceptable, but further

aggregation will lead to a model that performs poorly (chapter 5).

### *Construction rules*

Generally most studies considering the complexity of ecological models have created (subjective) rules to apply during model creation and simplification to ensure that the most efficient model is employed (e.g. Innis and Rextad 1983). The most thorough of these was performed by Halfon (1983a,b) who used Bosserman's (1980) complexity measure ( $\bar{c}$ ) to consider the effect of additional links between existing state variables (Halfon 1983a) and Hasse diagrams to investigate the structural properties of a number of different models (Halfon 1983b). Another notable, though more theoretical, approach was put forward by Iwasa et al. (1987). They gave formal mathematical rules that could be used to determine where aggregation of model variables was possible without loss of information. While these predominantly abstract theoretical and model development studies did present some examples with their discussions, they did not apply their methods to investigate the effects of complexity.

### *Formative studies of model complexity*

Investigations of the effect of model structure on performance have usually occurred in less theoretical settings. One of the best analyses of the issues of model complexity and aggregation, and its impacts on subsequent management performance, was undertaken in the area of single species fisheries assessment and management. Ludwig and Walters (1981, 1985) demonstrated that for estimating the true optimal fishing policy a small and highly aggregated model can perform better than a more complex and "realistic" one, with the same fundamental structure, even if the "data" were generated using the more complex model. This has been attributed to the parametric sensitivity of more detailed models, and its potential to propagate errors,

which may mask the underlying contrast in the data and thereby create problems during parameter estimation (Iwasa et al. 1987).

This approach (comparing simpler “assessment” models against more complex ones) has subsequently been widely adopted in evaluating fishery harvest strategies, including stock assessment methods, mainly for single species management. This approach has been called “operational management procedures” (Butterworth and Punt, 1999) or “management strategy evaluation” (Smith et al. 1999) and has recently been extended to include evaluations involving multi-species interactions (Sainsbury et al. 2000, Punt et al. in press).

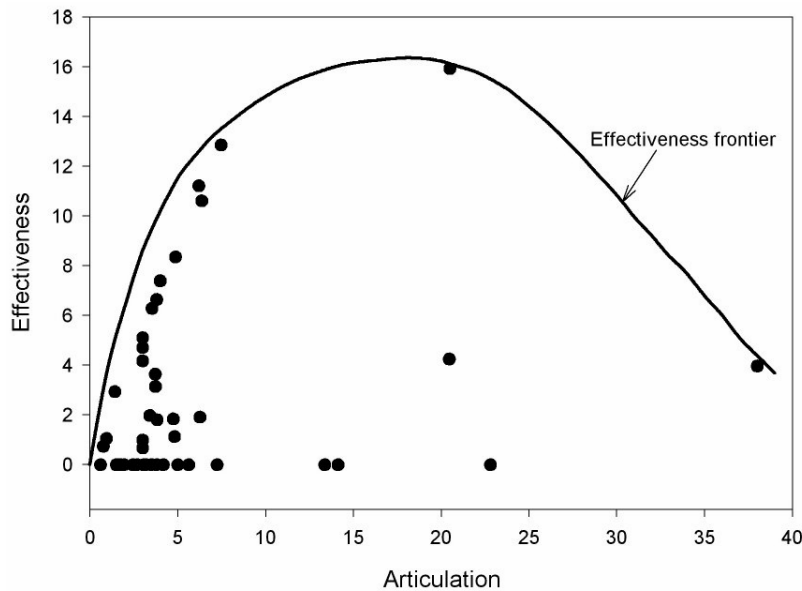
Within the more general ecological and ecosystem modelling literature, attempts to determine optimal complexity by comparing different models with each other or with data are scarce. Costanza and Maxwell (1994) began to span the divide between development and application when they examined the relationship between resolution and the predictive capacity of models – their discussion was based on the spatial dimension and was assumed to extend to the temporal dimension and number of components necessary. At the other extreme Kremer and Kremer (1982) and Hurtt and Armstrong (1996) both give examples of a search for optimal complexity through practical applications (they simplify or extend an existing model and discuss whether or not there is an improvement in performance). Overall, however, there are only a few attempts that take model performance into account while seriously considering the effects of model complexity on predictive models in a thorough or systematic way.

One of the earliest considerations of the structure of ecological models and its effect on model performance is the work by Wiegert (1975, 1977), who compared 5 models with differing trophic structure, levels of aggregation, spatial heterogeneity and formulation assumptions. The results suggested that the simplest of the models performed as well as the most complex, with a dip in performance for models with

intermediate levels of aggregation and simplification. This pattern of results was due to inadequate handling of spatial heterogeneity and time lags in the models with intermediate levels of complexity, whereas the simplest model did not require time delays and so that problem was avoided (Wiegert 1975). The modelling issues associated with varying degrees of model complexity in this study are strongly tied to the system being modelled and the modelling methods and assumptions employed. This may be why the relationship between complexity and performance observed by Wiegert (1975, 1977) is not reported in any other study considering the effects of model complexity.

The next major study to consider the effects of model complexity (Costanza and Sklar 1985) compared the predictive ability (judged against field data) of eighty-seven existing models of freshwater or shallow water bodies using three indices (articulation, accuracy and effectiveness). This work also found a non-linear relationship between model performance and complexity. The results were summarised in a plot showing that the relationship between effectiveness (explanatory power) and articulation (the amount of detail in conjunction with the physical and biological scope of the model) has a humped form (Figure 8.1). This finding supported anecdotal accounts from experienced modellers, particularly those that were involved in the International Biological Program of the early 1970s (e.g. Botkin 1977). One explanation for the shape of this relationship can be found in Jester et al. (1977). They suggested that at some point (as model complexity increases) the accuracy of the model reaches an asymptote, while the uncertainty continues to grow exponentially. Combining these into one plot of overall performance (incorporating accuracy and uncertainty) vs complexity gives the humped form found by Costanza and Sklar (1985).

**Figure 8.1:** Plot of articulation (measure of model complexity) against effectiveness (measure of model performance) for a number of existing aquatic models, redrawn from Figure 2 in Costanza and Sklar (1985).



*The implications of data with a low signal-to-noise ratio*

A different approach (a likelihood ratio test) was used by Yearsley and Lettenmaier (1987) to discriminate between three linear compartment models with varying levels of complexity (produced by aggregating compartments). The model comparison was made on the basis of synthetically generated data from a non-linear model of the global carbon cycle. They found that with realistic levels of measurement error even the most highly aggregated models are valid (a likelihood ratio test could not discriminate between even the most aggregated compartment model and the nonlinear model used to generate the “data”). Similar findings have also been reported for limnological models (van Tongeren 1995) and models of radionuclide transport in soils (Elert et al. 1999). These findings, that uncertainties in input data have a greater impact on the results than the model employed, are supported by the information theory concept of a model as a communication channel converting input data to output data. As

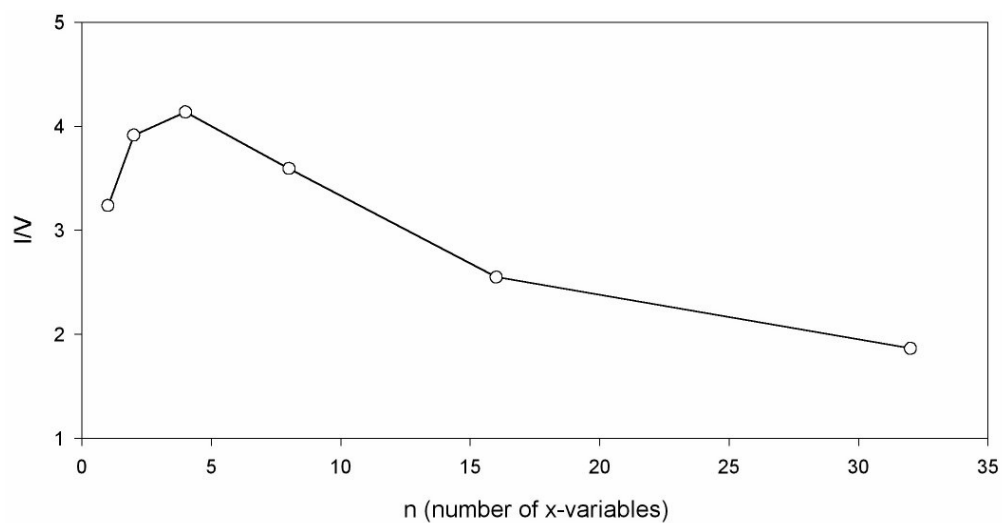
the signal-to-noise ratio of ecological and environmental data is typically low, there is no reason to expect that a large model spanning many noisy measurements will be any better than a model which deals with fewer, more precise, measurements (Silvert 1981).

### *Quantifying model complexity*

One of the biggest problems in considering the effect of model structure on performance is how to quantify the amount of detail in the model (Costanza and Sklar 1985). Håkanson (1995) considered predictive ecosystem models for lakes and coastal areas and tried to compare empirical and dynamic models by first considering the  $r^2$  values of a number of stepwise multiple regression models and then by examining the standard deviation of a number of additive and multiplicative models. Once again it was shown that there was a humped relationship between an indicator of performance and the number of variables included, though the peak of the plot was closer to the smaller model sizes in this case (Figure 8.2). Håkanson's (1995) work provides evidence (as does Costanza and Sklar (1985)) for the long held belief that predictive power increases quickly with the elaboration of simple models, but the trend is eventually reversed as accumulating errors and process and parameter uncertainty negate any potential benefits of increased detail. Extending this work, Håkanson (1997) defined a measure of predictive power (PP) of a model (based on the fit to multiple datasets) and used it to examine the predictive capabilities of models of the transport of radiocesium in lakes. The models considered included both empirical and dynamic models of differing size and complexity. This study also identified a humped relationship between predictive power and model size. More importantly, Håkanson (1997) gave another explanation for the decline in performance with increased detail. Accumulated uncertainty alone was not responsible for the drop. When it comes to using models in a prognostic sense, large models can be prescriptive rather than predictive (Håkanson 1997) - all the extra detail

can hardwire the responses rather than introduce flexibility. Further, Håkanson (1997) points out that the predictive power of a model is not determined by its strongest point, but by its weakest.

**Figure 8.2:** Plot of predictive power ( $I$  or  $r^2$ ) / accumulated error ( $V$  or  $SD$ ) against the number of x-variables ( $n$ ) for predictive lake models, redrawn from Figure 6a in Håkanson (1995).  $I$  is the information value,  $V$  is the accumulated uncertainty (relative standard deviation =  $SD/MV$ ),  $SD$  is the standard deviation and  $MV$  is the mean value.



### *The study of marine models*

Regarding models of marine systems in particular, several researchers have considered how the form of parts of models (such as grazing and mortality) affect their dynamics (Steele and Henderson 1992, Edwards and Brindley 1999, Murray and Parslow 1999b, Gao et al. 2000), but only a few have considered how the structure of entire models influences their behaviour (Yool 1998, Nihoul 1998, Nihoul and Djenidi 1998, Hoch and Garreau 1998, Tett and Wilson 2000, Murray 2001, this thesis). The work of Hoch and Garreau (1998) and Nihoul's body of research (Nihoul 1998a and 1998b, Nihoul and Delhex 1998, Nihoul and Djenidi 1998), though primarily concerned



with complexity from a hydrodynamic standpoint, should be of interest to all ecosystem modellers. Their findings highlight that not only are ecosystem models hampered by the same issues of internal complexity as any model, but they must also find a balance (or trade-off) between their various physical, chemical, geological and biological features (Hoch and Garreau 1998, Nihoul 1998a).

Within the field of single species fisheries modelling there is a well-established methodology for addressing the issue of model complexity. The approach involves simulation testing the performance of simpler models against a more complex one, using the latter to capture some of the possible complexities of the real world. Essentially, the “testbed” model is a simulation model incorporating complex processes, thought to occur in nature, which acts as an artificial world. This artificial world (and the “data” it provides) then becomes a useful testbed against which to judge the performance of simpler models that display a range of structural complexity and form. The more complex “testbed” model is referred to as a “deep” model, while the simpler operational models are called “shallow” models. Therefore the methodology is referred to as “deep-shallow model comparison”.

Given the signal-to-noise ratio of ecological data and what this may mean for model performance and interpretation, the use of the “deep-shallow” model comparison (as applied by Ludwig and Walters (1981, 1985) to fisheries science) is very attractive. The advantage of employing this methodology is that it allows the modeller to begin with a model that is known to work and then to “strip it back” to identify the level of detail that is effective and most efficient. Moreover, it provides a baseline of perfect knowledge (i.e. data generated by the deep model rather than gathered from the field) and therefore enables separation of the effects of model complexity due to model structure from those due to data uncertainty. Both sides of the problem must eventually be addressed, but keeping the two separate will clarify interpretation of any results.

Yool (1998), Murray (2001) and I (this thesis) apply the deep-shallow model methodology in a more general marine ecology setting. Yool (1998) decomposed the Fasham (1993) plankton ecosystem model and then reconstructed it stepwise in order to determine if a “best minimum model” existed. Murray (2001) compared a simplified version of the Port Phillip Bay Integrated Model with the original. Lastly, I compared a number of ecosystem models of varying degrees of detail and complexity. In addition I considered the effects of a range of simplifications or alternative assumptions on model performance. Each of these studies found that some degree of simplification (of structure, trophic coverage or physical scope) is possible without degrading the model, but simplifying too much leads to a substantial decline in model performance, particularly for the purposes of predicting the effects of changing conditions. While the study of the effects of the complexity of marine models as a whole is at an early stage, the work so far has provided some useful insights into two main areas of model construction and these are discussed in more detail below.

### **8.3 Model scope**

The scope of a model is largely defined by its spatial resolution, the temporal spacing of output (the model equivalent of sampling frequency in the field), the nutrients that are represented explicitly, and the biological web that the model describes. Along with the particulars of model formulation (discussed below), model scope is potentially one of the most important aspects of model complexity. Computational demands and the logistical requirements associated with collecting sufficient information to parameterise and validate ecosystem models are one of their biggest drawbacks (Silvert 1981, Sugihara et al. 1984, Lee and Fishwick 1998). Simplifying the biological or physical scope can make a model much easier to construct, summarise and interpret. The problem is that removing too much detail can leave a model that is simple

to parameterise and quick to run, but that has limited prognostic use (Tett and Wilson 2000, Murray 2001, chapters 4 and 5). The challenge is to define an optimum scope that minimises complexity, but which facilitates valid and robust predictions.

### *Trophic complexity*

The trophic complexity of a model is concerned with two aspects of model structure, the number of trophic levels explicitly represented in the model and how these levels are divided into species or functional groups (Murray and Parslow 1997, Pahl-Wostl 1997). It is understandable that multispecies models concerned primarily with water quality or harvested species may not include all trophic levels, but this assumes that groups at one end of the web don't significantly influence the behaviour at the other. In ecosystem models that may be used to consider many alternative anthropogenic scenarios, or the direct and indirect effects of processes such as fishing, the representation of a greater number of trophic levels is probably required. Unfortunately, the problem of where to draw the line remains an issue as every additional group included increases the complexity, modelling and data requirements. It may be desirable to truncate the web and omit the top predators, but the way this is done can strongly influence model behaviour (chapter 6). This aspect of the problem is discussed below in the context of model closure.

The number of species or functional groups (as defined by size classes, feeding linkages, shared predators and life history strategies) represented in a model needs careful consideration. Species level detail could mean the complete collapse of many biogeochemical ecosystem models and is beyond what is possible based on collected data from many systems. Nonetheless, without the flexibility inherent in a trophic web rather than a chain, realistic dynamics, especially under changing conditions, may be very difficult to capture (Baretta et al. 1995, Pahl-Wostl 1997). Thus, the systematic

consideration of the effects of trophic complexity on model behaviour is an important concern.

Randomly constructed food web models have been investigated by many researchers (Gardner and Ashby 1970, May 1973, Siljak 1974 and 1976, Waide and Webster 1976, Pimm and Lawton 1978). Within the context of ecosystem models however, consideration of realistically structured webs is more enlightening (Bosserman 1982). Consideration of changes in the behaviour of models with realistic web structure when the web is simplified by aggregation or omission of groups indicates that simplifying trophic structure too much is rarely beneficial (Christensen 1992, Sugihara et al. 1984, Optiz 1996, Pahl-Wostl 1997, Yool 1998, chapter 5). The simplified webs, especially those reduced to less than 25% of the size of the original model web, are not able to represent enough of the processes and interactions in the system to faithfully reproduce system dynamics, particularly when the strength of environmental or anthropogenic pressures change (chapter 5). This finding may also give some insight into the effect of a loss of biodiversity in real ecosystems. If real ecosystems respond to the loss of species or functional groups in the same way models do then it is likely that the impacts of a loss of biodiversity will be minimal if non-critical groups and interactions are lost and ecosystem conditions do not change. However, the magnitude of the impacts is much larger if conditions do change, as a result of environmental forcing (e.g. ENSO) or anthropogenic activities. This agrees with the ecological insurance hypothesis (Yachi and Loreau 1999), which proposes that diversity provides natural communities with a buffer against change.

As with the relationship between model performance and structural detail in general (Figure 8.1 or 8.2), there appears to be a nonlinear relationship between trophic complexity and behaviour. This relationship can be of two forms. It can be humped, like the overall relationship (Bosserman 1982, Gardner et al. 1982, chapter 5) or it can be in

the form of a threshold-triggered step-function (Tett and Wilson 2000, chapter 5). The second of these relationships is less common and is tied to groups with critical ecological roles, which must be explicitly represented to capture correct system dynamics.

Connectance (MacArthur 1955), the ratio of the number of strong:weak interactions in the web (McCann 2000) and redundant groups (Yachi and Loreau 1999) have all been proposed as explanations for the patterns of change in model dynamics with varying levels of trophic complexity. However, the work by Yool (1998), Edwards (2001) and the work presented in chapter 5 suggests that none of these alone can explain the patterns of performance observed and that the identity of the components and links included can be the most important determinants of performance. Thus, as with any other kind of model, it is far more important that crucial system characteristics are captured rather than rote application of a set of “modelling rules”. The findings given in chapter 5 also have important implications for the debate within ecology about whether increased diversity stabilises (MacArthur 1955) or destabilises (Gardner and Ashby 1970) an ecosystem. Many theories put forward in this debate assume a linear relationship between diversity and stability, but the relationship observed in the work on ecosystem models presented in chapter 5 is nonlinear. This suggests that there is no simple relationship between diversity and system behaviour and stability.

Simple rules may not always be guaranteed to work, but they may still prove to be useful. Within the context of ecosystem models and trophic complexity, the finding that over-simplifying is not usually advantageous is probably the strongest and most useful guideline for model development. One way in which this is exemplified is the effect of trophic aggregation on model behaviour. While aggregating species with rate constants that differ by less than threefold, which also have similar or common predators and prey, into functional groups is a successful modelling technique, pooling

functional groups is less successful than omitting the least important functional groups entirely (chapter 5). This stems directly from the guidelines given by the general papers on the effects of model aggregation mentioned above (Wiegert 1977, O'Neill and Rust 1979, Cale and Odell 1980, Gardner et al. 1982) and is shown explicitly in chapter 5. The other general guideline to come from studies of model complexity deals with the use of empirical submodels in the place of important, but poorly known, or omitted, processes, groups and linkages. This will be discussed in the section on model formulation below.

#### *Nutrients included explicitly in models*

Tett and Wilson (2000) advised that ecosystem models should be biogeochemical, as they must conserve one or more elements so that the potential growth of groups can be capped. The success of ECOSIM models (Walters 1998) suggests that conservation of biomass may also be sufficient.

In biogeochemical models, it is quite common to use the most limiting macronutrient (carbon, nitrogen or phosphorous) as the model currency and assume that the conversion of the other nutrients conforms to the Redfield ratio (Murray and Parslow 1997). However, this approach does not allow the model to adjust to spatial and temporal differences in nutrient availability (Baretta et al. 1995). This inability should not present a significant problem and will not (in general) lead to large model divergences if the system state is mesotrophic to eutrophic. This is because the other nutrients should remain in excess, even if not exactly in Redfield ratios. However, when a system is in an oligotrophic state the dependence of the model on a single macronutrient currency can be a problem (chapter 7). Under these conditions the identity of the limiting nutrient may well change and the preferential remineralisation of nitrogen and phosphorous will cause significant departures from the Redfield ratio, with

potentially catastrophic implications for production estimates and the dynamics of transmission up the food web. The successful application of models such as ECOPATH with ECOSIM to the open oceans (Christensen et al. 2000) suggests that models employing the conservation of biomass may not suffer as much from this problem. However, environmental influences are not usually present in such models, and if they are they are usually in the form of prescribed forcing functions (Hollowed et al. 2000). If environmental conditions became a more integrated part of these models, then the problems observed in biogeochemical models may appear in models like ECOSIM too, but this remains to be seen.

#### *Physical scope*

The last important aspect of model scope that can affect model behaviour is the physical scope of the model, and in particular the spatial resolution of the model. Many multispecies and ecosystem models (e.g. mass balance aggregate system models such as ECOSIM) do not include any explicit spatial representation. However, space is a vital system resource in its own right in many marine systems and as such the way in which it is represented can have a significant impact on model dynamics and predictions (Murray 2001, chapter 4). Many of the model stability issues identified in ecological and ecosystem models in the past (May 1974, Pimm 1982, Cohen and Newman 1988, Christensen et al. 2000) disappear with the introduction of explicit spatial (and thus environmental and/or ecological) heterogeneity (Johnson 1997, chapter 4). This assertion does not only apply to biogeochemical ecosystem models (like those employed in chapter 4), but extends to other types of multispecies and aggregate system models (like ECOSIM). There are many examples of these models being explicitly (through the development of ECOSPACE (Walters et al. 1999)) or implicitly (by separating individual model groups into inshore and offshore components (Christensen

pers. com.)) expanded to incorporate spatial partitioning. This is not to say that ecosystem models must be tied to general circulation models. Such a move would be computationally prohibitive and probably of little assistance. The box-model approach to transport processes is useful as it neglects small-scale gradients, but still allows for regional differences and spatial self-structuring which, in turn, lead to the formation of distinct communities and ecological zones (Baretta et al. 1995, chapters 1 – 3). Even when using box-models, large numbers of cells may not be necessary. For example, the work presented in chapter 4 shows that an 8-box model was a good compromise between the computational intensity associated with a 59-box version and the trophic self-simplification and degradation in performance associated with 3- and 1-box versions of the same model. This is another facet of model structure where intermediate complexity is best.

#### **8.4 Model Formulation**

The main concern of the majority of model studies considering the effect of model structure is model formulation not model scope. The way in which a model is implemented can have a large impact upon its performance and usefulness (Silvert 1981). The key areas of model formulation which have received some attention for marine system models concern process detail (particularly with regard to grazing functions and mortality terms) and the role of empirical submodels and forcing.

##### *Process detail*

Whereas physical oceanographers have a set of “basic hydrodynamic equations” there is no such set of equations in ecological modelling (Tett and Wilson 2000). The problem is compounded in ecosystem models, as the modeller must integrate a variety of processes and interactions with differing characteristic temporal and spatial scales



(Barthel and Goñi 1995). Consequently, ecosystem models run the very real risk of incorporating too much detail to be comprehensible, or over-compensating in the other direction and not including enough to be realistic or of any real use. Given this risk, and the increasing number of ecosystem models, there has been surprisingly little published on the effects of formulation detail on the dynamics of ecosystem models. Moreover, despite the contention that confirmatory and comparative model studies show the greatest promise for guiding management decisions, there have been few studies to compare different forms of ecosystem models (chapter 7).

Specific aspects of the effect of process detail will be discussed in the following sections, but there are some pertinent observations about overall detail that arise. The work by Håkanson (1997) and I (chapters 2 and 3) indicate that highly detailed, often physiologically based, process detail is not a pre-requisite for a successful ecosystem model. In chapters 2 and 3, I compared the outputs and predictions of two ecosystem models that covered the same web and processes, but with differing levels of process detail. The Integrated Generic Bay Ecosystem Model (IGBEM) is highly physiologically detailed, while Bay Model 2 (BM2) uses the same functional groups, but much simpler assimilative equations. It was found that, with regard to understanding system dynamics and qualitative responses to changing conditions, BM2 can represent systems as well as IGBEM and in some cases the performance of BM2 is better than that of IGBEM. In comparison with real bays, the predictions from BM2 are better than those from IGBEM in the areas of microfaunal dynamics, relative community composition, size structure, and sediment chemistry. IGBEM is better for biomass and consumption estimates of some groups, such as zooplankton, the ratio of chlorophyll *a* to dissolved inorganic nitrogen in the water column and for robust dynamics across a wide range of nutrient loadings. BM2 does have some weaknesses related to its simplified form: it is more sensitive to parameter values than IGBEM; the detritus

feeding benthic groups and microphytobenthos occasionally show almost exponential growth (suggesting the lack of a limiting factor for these groups, such as the availability of space); and the assumption that nutrients are always in Redfield ratios causes BM2 to perform poorly in oligotrophic conditions (where this assumption is often violated) (chapters 2 and 3). Despite these weaknesses, the overall performance of BM2 is as good or better than IGBEM and this shows that physiological detail is not necessarily required and that simpler formulations can work. This is a boon given that when using BM2 instead of IGBEM the number of parameters required drops by more than 50% and the computational demands also decline substantially (the run time drops by 30 – 60 %)(chapter 2). The amount of process detail required is only as great as that needed to successfully capture crucial system dynamics. The findings of Håkanson (1997) reinforce this assertion. In that study the simpler model had the most predictive power and the more complex the least.

The work of Tett and Wilson (2000) cautions against taking the process of simplification too far. They found that models that sacrifice large amounts of biogeochemical or ecological detail in favour of the other cannot adequately describe the dynamics of the plankton. A minimum level of both biogeochemical and ecological detail is required. The research of Murray and Parslow (1999b) and Murray (2000) arrives at a similar conclusion. They found that a much simpler model compared favorably with a more complex model of the same system, but with some caveats. This made the simplified model an excellent aid in the development of a more sophisticated model, but the inability of the simple model to capture certain dynamics in some conditions meant that it could not replace the sophisticated model. The extra detail of the larger model was required for fully informed system management and to allow scientists and managers to understand and consider a number of alternative scenarios. Håkanson (1997) considered the effects of complexity on performance by sequentially

increasing the detail included in the models considered, rather than systematically reducing detail as in this study. However, Håkanson's (1997) findings about the minimal realistic model also indicate that the "correct" level of process detail is that needed to successfully capture crucial system dynamics.

These studies on the "correct" level of process detail needed in ecosystem models may also be instructive when considering the dominant processes in real ecosystems. It is likely that the processes required in models to successfully capture crucial system dynamics are the dominant processes structuring systems. Furthermore, consideration of the dynamics displayed by ecosystem models may further the understanding of the behaviour, or potential behaviour, of real ecosystems. For example, even ecosystem models that do not incorporate detailed representations of physiological processes (e.g. BM2) can produce runs which show major transitions in system state (where some groups with high biomasses in one state decline to much lower levels in the other state and vice versa). In real systems the occurrence of such events are assumed to be linked to anthropogenic activities or changes in external forcing, but in ecosystem models they can occur even in the absence of these factors. This suggests that at least some of the transitions in system state observed in real systems may be emergent behaviour caused by internal system interactions or resource dynamics rather than the result of some external force.

Studies comparing the performance or predictions of different types of model are useful for judging how robust general findings are to the underlying assumptions of the models (chapter 7). However, they are also an excellent source of information on the effects of process detail on model dynamics. Duplisea and Bravington (1999) found that the results from a length-cohort model very similar to a MSFOR (multispecies forecast model) and a size-spectrum mass transfer model both lead to the same conclusions regarding fisheries management strategies. Thus, within the context of the system

dynamics related to the particular question of interest, the explicit process detail of the MSFOR did not confer any advantage over the far simpler size-spectrum model. In general, size-spectrum models are a successful methodology, at least for pelagic aquatic ecosystems (Silvert 1996b). However, in the context of the evaluation of management strategies Duplisea and Bravington (1999) recommend a few modifications, such as allowing some disaggregation into functional trophic-groups and including more realistic grazing terms. The popularity of ECOSIM suggests that it may also be a successful methodology. This is supported by the finding given in chapter 7 that, with a few exceptions, ECOSIM gave the same qualitative predictions as the biogeochemical ecosystem models IGBEM and BM2. The differences observed stem mostly from the lack of spatial detail in ECOSIM, or parts of the web that are poorly known. The main differences between the biogeochemical models and ECOSIM that are a direct result of model formulation are that the biogeochemical models are not as buffered against changes in fisheries as ECOSIM is, but they are more buffered against changes in nutrient loading. This is a reflection of the more realistic behaviour of the low to middle trophic groups in the biogeochemical models, whereas the higher trophic groups react more sensibly in ECOSIM (chapter 7). This is not surprising given their respective development histories and structure, but does caution against the assumption that a formulation that works at one level will work for every level.

### *Grazing terms*

General ecological research, as well as the results of more directed marine modelling, has shown that the form of grazing terms used can have important effects on overall model behaviour and predictions (May 1976, Hassell and Comins 1978, Begon and Mortimer 1986, Steele and Henderson 1992, Tett and Wilson 2000, Gao et al. 2000, chapter 6). Evaluation of the effects of the functional response used in plankton models,

built around relatively simple food chains, indicated that they do not have as great an impact as other parts of the model (Steele and Henderson 1992, Murray and Parslow 1999b). In contrast, consideration of the effect of the grazing terms used in a total system model, with a complex trophic web, (BM2) in chapter 6 indicates that they can have a substantial impact on model behaviour. In chapter 6 I concluded that, while there are biologically and mathematically sound arguments for the inclusion of sophisticated and dynamic functional responses, the extra parameterisation is not justified because the Holling “type II” response predicted the same general patterns of behaviour, and thus the same conclusions about system dynamics. Nevertheless, it is also stressed that very simple responses (such as the Holling “type I”) do not allow for realistic dynamics over the range of conditions of most interest in system management scenarios. Tett and Wilson (2000) reach a similar conclusion, whereas Gao et al. (2000) conclude that the optimal functional form will depend on the specific study and that more observations and understanding of real marine systems are required before the matter can be clarified. The latter may well be true, but for models incorporating a realistic food web, simpler grazing terms may suffice due to the many other stabilising features inherent in the web (Tett and Wilson 2000).

### *Model closure*

The other main aspect of model formulation that has received explicit attention is the form of model closure (Steele and Henderson 1992, Edwards and Brindley 1999, Murray and Parslow 1999b, chapter 6). Model closure refers to the form of the mortality term applied to the top most group(s) explicitly included in the model. Linear and quadratic mortality terms are the most common means of dealing with model closure and these reflect the cases when the effect of predators not included in the model are assumed to either be constant (linear mortality) or to change (quadratic mortality) with

the population of their prey (the top most modelled group(s)). The specific form used can have a substantial impact on model behaviour (Murray and Parslow 1999b).

Steady-state analysis of simple food chain plankton models indicates that model closure can be the most important determinant of model behaviour (Steele and Henderson 1992, Edwards and Brindley 1999, Murray and Parslow 1999b). In contrast, its effect on a total system model (with a complex trophic web) shows it is much less important than other aspects of model structure (chapter 6). However, even in the case of the total system model the need for the representation of higher predators (either explicitly or implicitly via quadratic mortality terms) is recognised. Unfortunately, conflicting conclusions regarding the dynamics of the highest predators (sharks, mammals and birds), when there are large changes in conditions and differential stability of the various forms of closure across a range of conditions, mean that further work on this topic is necessary (chapter 6). Nevertheless, it seems likely that, as a general guideline, the use of quadratic closure is acceptable regardless of the size of the implemented web (chapter 6).

#### *Forcing functions and empirical submodels*

The last facet of model formulation that has received some attention, is the value and usage of empirical formulations. Empirical formulations (or empirical models as they are also known) are functions that describe observed patterns or relationships in data, but without trying to capture real process dynamics. These formulations have both advantages and disadvantages. They can be developed without much understanding of the phenomenon of interest, they are simple (as they do not need to include complex causality) and they can be developed rapidly (DeCoursey 1992). All of these features make empirical models attractive, and within their range of applicability they can often provide better predictive power than dynamic models (Håkanson 1997). Their limiting

feature is that many domains of interest can be outside the range of applicability. This, along with the arbitrary nature of these models and the associated risk of adopting a misleading approach or false assumptions, can mean they are less than ideal (DeCoursey 1992).

Within the realm of marine ecosystem models, the choices are more complex than empirical vs. purely process models. The two approaches usually have different purposes and often complement rather than compete with each other (Håkanson 1995). Moreover, empirical models can play a role within dynamic process models. Given that the understanding of some components of marine ecosystems are poor (e.g. the processes and forces shaping the behaviour of the benthic infauna) and that a model is only as strong as its weakest part, the use of empirical submodels for the least known parts of the system is an attractive alternative. The inappropriate use of simple forcing functions can lead to very poor model performance, while the use of a structured empirical submodel can work very well (chapter 5). Such a submodel can compensate for not explicitly representing a poorly, or incompletely, known component, which has a potentially crucial role in the modelled system. For example, in the empirical model of nitrification-denitrification by Murray and Parslow (1999a) the amount of ammonia available for nitrification-denitrification was calculated using a temperature-dependent rate of breakdown of the form

$$R_X = \Phi \cdot X \quad (8.1)$$

where  $R_X$  is the ammonia released by remineralisation of component X,  $\Phi$  is the temperature-dependent rate of breakdown for the component X and X stands for labile detritus (DL), refractory detritus (DR) or dissolved organic nitrogen (DON) in the sediments. This available ammonia is then used in the following nitrification ( $S_{\text{NIT}}$ ) and denitrification equations ( $S_{\text{DENIT}}$ )

$$S_{\text{NIT}} = R_{\text{NET}} \cdot \theta_{\text{DMAX}} \cdot \max\left(0, 1 - \frac{R_{\text{NET}} \cdot \gamma_{\text{SED}}}{r_0}\right) \quad (8.2)$$

$$S_{\text{DENIT}} = S_{\text{NIT}} \cdot \min\left(1, \frac{R_{\text{NET}} \cdot \gamma_{\text{SED}}}{\theta_{\text{TD}}}\right) \quad (8.3)$$

where  $\theta_{\text{DMAX}}$  is the maximum rate of denitrification,  $\theta_{t0}$  is the temperature-dependent minimum rate of respiration that supports nitrification,  $\theta_{\text{TD}}$  is the peak of the nitrification-denitrification curve (as defined by Murray and Parslow 1999a),  $\gamma_{\text{sed}}$  the depth of the sediment layer considered in the model and  $R_{\text{NET}}$  is the total available ammonia, which is given by

$$R_{\text{NET}} = \max(0, R_{\text{DON}} + R_{\text{DL}} + R_{\text{DR}} - P_{\text{NH,MB}}) \quad (8.4)$$

where  $P_{\text{NH,MB}}$  is the amount of ammonia in the sediments taken up by microphytobenthos for growth. This empirical model performs well and is robust (Murray and Parslow 1999a), but it is not interactive, in the sense that it is not dependent on the activity of the benthic groups included in models with more complex trophic structure. Therefore, modifications were made to allow it to be interactive when it was included in the ecosystem model BM2. In BM2 the amount of ammonia available for nitrification and denitrification is determined by the activity of the attached bacterial populations and other sediment dwelling fauna and flora (chapter 2). Nitrification and denitrification is then completed using the empirical formulations of Murray and Parslow (equations 8.2 and 8.3 above). This empirically based bacteria-denitrification submodel is an improvement over the purely empirical model of Murray and Parslow (1999a). It was also a vast improvement over other process based attempts at modelling bacteria and denitrification (chapter 1), subjects that are still poorly known in many respects. Thus, the use of an empirical representation of an important process can prevent degradation of model performance. This is especially true if the causal mechanisms for the process are poorly known, or if explicit inclusion of the details of



the mechanism is beyond the scope of the model or the capability of the available data or computational resources.

## **8.5 Model performance under changing conditions**

Using ecosystem models to gain insight into a system and indicate (at least) qualitative trends associated with a change in “forcing” conditions is one of their most useful roles. It is also at this point that assumptions underlying the model formulation can have their greatest impacts. Thus, confirmatory comparison of models is strongly advocated. A comparison of three ecosystem models (ECOSIM, BM2 and IGBEM) in chapter 7 indicated that overall model structure and formulation can be robust (i.e. provide the same general predictions) under changing conditions, but still predict some potentially important differences in specific cases. For example, applying a fisheries management strategy developed purely to maximise economic gains to all three models produced predictions that coincided for the majority of the biological components in all the models and there was agreement between at least two of the models for all the components except detritus. The three models all gave different results for detritus (ECOSIM predicted no change, BM2 predicted a decline and IGBEM an increase). Given the role of detritus as a long-term storage of nutrients in enclosed bays like the one modelled, such a range of outcomes is a crucial result. This illustrates how conclusions drawn from different models can be very different for particular components of a system, even when the models generally agree overall.

The effects of implemented process detail and model scope are usually most apparent under changing conditions (such as changing nutrient loads or fishing pressures). A change in conditions or pressures on a system may be beyond the “range of applicability” of an empirical model or may expose a flaw in a chosen formulation (Murray 2001, chapter 5, chapter 6, chapter 7). I undertook many evaluations of different

aspects of model structure as part of an investigation of the effect of model structure on the behaviour of ecosystem models. It was repeatedly found that formulations of varying complexity can have very similar dynamics, or show only small divergences, under baseline conditions, but show much larger differences under altered nutrient loads or fishing pressures. For example, the problems associated with spatial resolution that are too restricted, or overly simplified trophic structures, or grazing and mortality terms that do not include some form of limitation all lead to poor performance under changing conditions. Thus, performance under changing conditions is an important measure of how robust model behaviour is to the level of complexity employed in a particular aspect of model structure or scope.

One of the clearest symptoms of this problem is model instability or manifestation of aberrant behaviours. This was one way to identify the potential weaknesses of the benthic deposit feeder group in BM2 (chapter 3). In certain circumstances the model allows this group to undergo almost exponential growth, an indication that this poorly known group has had a critical limiting factor omitted from the formulation during model development. While this highlights an area that needs more attention in the field, it is also an area that would have been missed if the model had been considered only under a restricted set of conditions.

Unfortunately, model failure need not be expressed in such an obvious way. A model that incorrectly specifies some process (like feeding or mortality) or has a scope that is too restricted may still appear to have “acceptable” behaviour (in relation to its state under current conditions), but the predicted behaviour may be incorrect (with regard to what would really occur) (chapters 4 – 6). For example, the lack of spatial structure in the 1 box model evaluated in Murray (2001) causes it to overestimate the nutrient loads that the Port Phillip Bay could tolerate (as predicted by the 59-box version) by 30%. This is a large problem if this version of the model is used to guide

management of nutrient loading and water quality in this bay.

## **8.6 Conclusions**

The multitude of links and processes that make up a real ecosystem mean that the ultimate effects of anthropogenic actions will probably be much wider than expected and may even lead to counterintuitive outcomes. Ecosystem models are a prime candidate as a tool to aid in the understanding of these potential outcomes. This does not mean they do not have potential drawbacks (often to do with their own size and complexity), but careful consideration of these problems and the intelligent application of the models (particularly in a confirmatory framework) can avoid or minimise many of these problems. As concerns and management at a system level become an increasing focus of many sectors of human society, ecosystem models can be a valuable tool in addressing such issues (Walters et al. 1997). However, greater understanding of the effects of model structure and scope on model performance are necessary. This is particularly the case if we are to avoid the situation where frustration resulting from poorly structured ecosystem models, or the inappropriate use of existing ecosystem models, leads us to reject the modelling approach altogether.

Studies of the effect of model structure that have already taken place indicate that there is a humped form to the relationship between model detail and performance (Costanza and Sklar 1985, Håkanson 1995, this thesis). Too much complexity leads to too much uncertainty and problems to do with interpretation of the model's dynamics and predictions, while too little detail results in models that cannot produce realistic behaviours. These studies have also provided a few important guidelines which can usefully extend or augment the "rules of thumb" proposed in earlier works on ecological models (Wiegert 1977, O'Neill and Rust 1979, Cale and Odell 1980, Innis and Rextad 1983, Halfon 1983 a and b, Gardner et al. 1982, Iwasa et al. 1987):

(1) Physiological detail is not always necessary. However, the use of explicit physiological detail can be important in certain circumstances (such as oligotrophic conditions when simple assumptions about nutrient uptake and the ratio of limiting nutrients are violated).

(2) If an important process or linkage (e.g. to an external web) is poorly known, or is not explicitly represented in the model, then an empirical representation should be included in its place. This can avoid introducing uncertainty without risking the degradation of performance associated with neglecting a crucial aspect of a system.

(3) Some level of spatial resolution is likely to be necessary for adequate performance of the model. A 1-box model is unlikely to be sufficient, as space is itself an important system resource. This is particularly true in systems where benthic groups are important. Moreover, there must be enough spatial resolution in the model to capture the major physical characteristics of the system. Trophic self-simplification of the web (the loss of one or more components from the web) is often a good indicator of an overly restricted spatial representation.

(4) Sampling frequency (the time period at which model output is recorded) can have a large impact on model interpretation and the predictions deduced from the output. At a system level, a 2-4 weekly scheme is an adequate compromise between excessive noise and a loss of information.

(5) The inclusion of a complete trophic web at the level of species is neither necessary nor desirable, but the way in which the web is handled is critical. The use of functional groups is a successful means of representing the system web realistically (particularly if some age or size structure is included for the highest groups). Aggregation beyond the level of functional groups is ill advised, and omission of the least important groups is a better strategy if further simplification is necessary. Moreover, simplifying a model web (which represents the food web of an

entire system aggregated to the level of functional groups) to less than 20-25% of its original size is rarely beneficial, as representing the distinctions between large and small, or mobile and sedentary, groups may be crucial.

(6) Quadratic closure of the top-most parts of the trophic web is a successful method of closing the web, regardless of its size. The explicit inclusion of the highest predators may only be necessary when they are of direct interest.

(7) The form of the grazing functions used must be given careful consideration so that they contain enough flexibility without introducing extraneous detail. Holling “type-I” responses are unlikely to be sufficient (especially under changing conditions), but the more complex Holling type functions may be acceptable. More sophisticated responses, incorporating more behavioural dynamics (e.g. balancing predator avoidance with the need to forage) may be required in some circumstances, but the value of their inclusion should be checked.

In summary, potentially the best rule of thumb to remember is, in the words of Albrecht (1992), “as simple and as highly aggregated as possible and as complex and disaggregated as necessary!”.

The topic of the effect of model complexity on model behaviour and performance is far from a closed chapter, especially with regard to ecosystem models. Within the context of trophic complexity and ECOSIM, Walters (pers. com.) stresses that “this [exploration of the effects] is a really crucial issue that has not yet been systematically explored by any of the science groups involved in ECOSIM modeling.” However, this point is not restricted to trophic complexity or ECOSIM, but pertains to many facets of model structure and the many methodologies and ecosystem models currently in use. The work covered in this review is a useful start, but it is still early days.