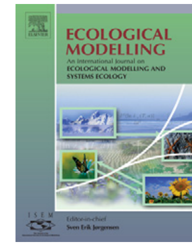
ECOLOGICAL MODELLING 202 (2007) 385–396

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/ecolmodel



A methodology for developing simulation models of complex systems

*Craig A. Aumann**

CW-405, Department of Biological Sciences, University of Alberta, Edmonton, Alta., Canada T6G 2E9

ARTICLE INFO

Article history:
Received 24 November 2005
Received in revised form
22 October 2006
Accepted 3 November 2006
Published on line 15 December 2006

Keywords:
Simulation
Individual-based model
Hierarchy theory
Assessment
Scientific explanation
Emergent property
Complexity

ABSTRACT

While the complexity of ecological simulation models has increased along with advances in computer hardware, recent reviews have concluded that such complex models have generally not fulfilled their potential for advancing ecological understanding. This state of affairs is partially the result of little attention being given to methodological issues for developing simulation models. This paper presents one methodology for developing models of complex systems. The first part of this methodology places the modeling process in the context of general research planning and thus emphasizes the process of synthesis—combining the constituent pieces of knowledge into a unified description of the entities and processes comprising the system to be modeled. The second methodological step involves operationalizing this synthetic description into a composition of smaller models specified over three scalar hierarchical levels. The hierarchical structure of the decomposition also structures the explanations given about system behaviors in that the “mechanism” for a behavior arises at the lower levels while its “purpose” is found at higher levels. The third part of the methodology involves using model assessment to explicitly establish the veracity of the links between the implemented model, the model design specification, and the synthesis. The veracity of these links are established by demonstrating: (i) that the model has been built correctly relative to the model design (verification), (ii) that the right model has been built as judged by the behaviors of the model components across the hierarchical levels (validation), and (iii) that the assessment criteria used are credible for judging model adequacy for the stated modeling objectives (critique). By creating confidence in the mapping between the model implementation and synthesis, the assessment process enables confidence in the explanations given for model behaviors and the theoretical understanding expressed in the synthesis. Alternatively, the failure of the model to satisfy assessment criteria may necessitate a reformulation of the understanding on which the model is based. Regardless of which possibility occurs, because the reasons for model behaviors can be mapped back to the synthesis, ecological understanding is advanced.

1. Introduction

The typical reasons for constructing models are that they enable predictions, act as guides for future experimentation, and aid knowledge synthesis (Oreskes, 2003). To achieve such goals, scientists make use of alternative idealizing assumptions that result in different kinds of models like statistical models that represent images of data, models representing

* Tel.: +1 780 438 6249.
E-mail address: caumann@permanentmail.com.

0304-3800/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2006.11.005

causal implications of our theoretical ideas (e.g., models of theoretical population or community ecology), or simulation models that are constructed to be analogous to some real system (Morton and Suárez, 2001). In representing systems in ecology, much debate has centered around whether the three goals of modeling are best achieved using simpler models containing only a few equations expressing a small number of general principles, or with complex simulation models that are more "structurally realistic" and contain more detailed representations of the basic processes and interactions determining the system's dynamics (Palladino, 1991; DeAngelis and Mooij, 2003). With the advances in the field of computer science and the increasing power of computing hardware, the number of such complex simulation models appearing in the ecological literature has certainly increased over the last 20 years. However, recent reviews suggest that simulation models have not fulfilled their expected potential for advancing ecological understanding (DeAngelis et al., 1994; Grimm, 1999).

This failure can certainly be added to the long list of reasons typically given for favoring simple over complex models: simple models can predict as well as complex models, are more transparent and easier to communicate, require less data for parameter estimation and model validation, may be less prone to error propagation, and since they embody only a few heuristic principles, may be more likely to lead to general causal understanding (Levins, 1966; Palladino, 1991; Pace, 2003; DeAngelis and Mooij, 2003; Van Nes and Scheffer, 2005). Thus, two questions must be asked: (1) Is it worthwhile to construct complex simulation models of real systems? and if so, (2) What methods must be followed so that the model can be used to make scientific inferences about the system represented?

The underlying debate between those favoring simple versus complex models is the result of a larger ongoing paradigm shift in ecology (Cooper, 2003). It is now realized that the preference for simple models expressing general heuristic principles is predicated on the metaphysical assumption that general regularities exist in nature (Palladino, 1991), but this assumption has been challenged and is currently regarded as an open question in ecology (e.g., Wang and Gutierrez, 1980; Simberloff, 1982; Strong, 1986; O'Hara, 2005). Since this question cannot be answered from first principles or empirically, one reason for attempting to build models that more accurately represent real ecological systems is to determine if the dynamics of such model systems can be reduced down to some set of simpler, general ecological principles (Judson, 1994). Another reason for favoring complex models is that ecologists today are much more likely to view the world as self-organizing where multiple causes occur among networks of processes leading to multiple possible explanations of system behavior. In such systems, it is thus difficult to speak of "the" cause of a particular system behavior (Pickett et al., 1994) and if the goal is to understand such systems, a misplaced emphasis on simple models will not lead to a deep understanding of how such complexity emerges (DeAngelis and Mooij, 2003). In addition, a number of management examples have shown that using models that do not adequately reflect the complexity of the management problem can have disastrous consequences (e.g., Walters and Maguire, 1996; Parker et al., 2002, 2003; Peterson et al., 2003). Further, even if simple models predict a few outputs as well or better than complex models, it is now realized that prediction alone is a poor criteria by which to judge theoretical adequacy since theoretically inaccurate models can make highly accurate predictions (e.g., the Ptolemaic system for predicting planetary motion Oreskes (2003)). In contrast, by representing a wider array of mechanisms complex models can be used to assess which factors most powerfully influence the state of the system—thereby providing greater guidance for observation and future experimentation (Oreskes and Belitz, 2001; Peck, 2004).

These are all strong arguments for the necessity of modeling ecological systems using complex models instead of attempting to simplify the model down to a few general principles a priori. Thus, the problem is not that complex models should be rejected, but that we must figure out how to meaningfully model this complexity. Thus, the focus of this paper is on proposing a methodology for modeling real ecological systems, showing how this methodology enables the resulting model to be used to make scientific explanations about the original system, and how the methodology structures these explanations. The presentation is focused around four questions that any modeler encounters: Is enough known about a system to produce a model that will satisfy the modeling objectives (Section 3)? How should the different parts of the system be represented (Section 4)? How closely must the behaviors of the model and source system agree (Section 5)? And finally, how can the model be used to increase understanding of the original system (Section 6)? As central as these questions are to the practice of simulation modeling in ecology, they are only now beginning to be more thoroughly explored (e.g., Grimm and Railsback, 2005). However, before presenting the methodology for answering these questions it is necessary to first introduce some terminology to clarify a number of modeling concepts.

2. Simulation framework overview

What makes models potentially useful is that they represent aspects of reality in a manner that is separate from both the theoretical and real worlds. By being separate from both these worlds, models can be manipulated and experimented with in ways that give them "a life of their own" (Cartwright, 1999; Morrison and Morgan, 1999; Winsberg, 2003; Magnani, 2004). The central assumption of this paper is that models and the process of modeling is best understood if considered as part of a larger simulation framework (see Fig. 1) that contains a model, a simulator, the source system, a set of experimental frames, and the relationships between these entities (see in particular Zeigler et al., 2000).

At a high level of abstraction, a *model* is just a state transition mechanism or function capable of generating output trajectories based on input trajectories. These trajectories are generated by a *simulator*—a computation system such as a micro-processor, the human mind or an abstract algorithm capable of executing the model instructions. The *source system* is the real or virtual system that is being modeled and is included in this framework so that comparisons can be made between the output of the model and source system. These outputs are compared in *experimental frames* which include the

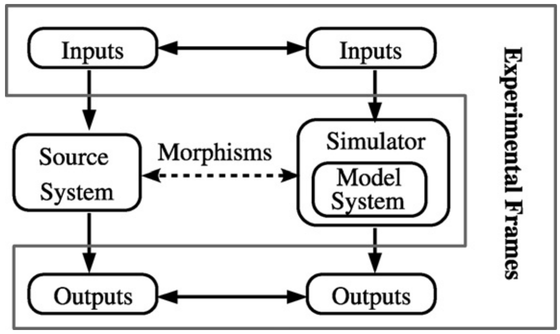


Fig. 1 – A simulation framework (Zeigler et al., 2000) consists of a source and model system that are related to each other via morphisms or mappings between the two systems, a simulator capable of executing the model instructions, experimental frames specifying the experimental conditions under which the inputs and outputs of both systems should be compared, and finally the inter-relationships between these entities.

specification of the conditions under which both the source system and model are to be observed or experimented with. For example, an experimental frame for evaluating a model of fire on a particular landscape might include conditions about vegetative states with associated probabilities of fire actually occurring. The central idea here is that under the "same" conditions, both systems should produce "similar" outputs. If they do, then the systems are analogous with respect to the conditions examined. The goal is to have the systems be analogous over a sufficiently wide range of conditions that encompass the modeling objectives. In this way, the experimental frames used are dependent on the modeling objectives.

Since modeling is severely constrained by complexity limitations, the essence of successful modeling is *valid simplification* or reducing the complexity of the system to enable the model to execute on a simulator while still ensuring that the model is valid within the experimental frames required by the modeling objectives. Any model component can be specified with varying degrees of detail and it is convenient to think of these different levels of detail according to a specification hierarchy (Klir, 1985; Zeigler et al., 2000). The highest level in this hierarchy contains the most detailed specifications and the lowest the most aggregated. At the lowest level, an *I/O Behavior* (i.e., Input/Output Behavior) specification is like a black-box in which inputs are mapped directly onto outputs. For example, if a fire initiation event occurs, the landscape always burns. In this case, there is no specification of mechanism beyond this one-to-one mapping. In a *I/O System* specification, a model maintains an internal state and inputs can change this internal state so that alternative outputs can be produced under the same inputs. For example, if the vegetative model contains more detailed information about vegetative state, then regardless of how many fire initiation events occur the landscape will not burn unless the vegetation is in some particular state. A *Coupled Component* model specification is a model composed of other I/O Behavior, I/O System, and possibly other coupled component models. For example, the fire model above may be combined with an I/O system model of vegetative succession. It is also possible to specify a derivability relation between higher and lower model specification levels, but the reverse is clearly not true since knowing a system's I/O Behavior provides little information about how such behavior was generated.

Since the simulation framework contains both source and model systems it is possible to specify a notion of similarity between these two systems using the concept of morphism. More formally, the two systems are said to be *morphic* at a given specification level if it is possible to establish a direct correspondence between the defining elements of each system at the same specification level within some experimental frame. The nature of this correspondence changes across specification levels. For an I/O Behavior specification, the morphism is simply the comparison of the inputs and outputs of both systems—if fire initiation events occur in both systems, do both systems burn? For an I/O System specification, the introduction of a state space requires the introduction of a mapping between the state spaces of both systems. This mapping is said to be *homomorphic* if there is a defined correspondence (but not necessarily an identity) between the states in both systems. Given that modeling involves simplification, the state-space in the source system will likely be larger than that in the model system and thus an identity is unlikely to hold. What is key, however, is that both systems progress through similar pathways to achieve similar model outputs. For example, following a fire event both systems transition to states in which fire initiation events do not lead to fire, and vegetative succession follows similar pathways. Finally, since Component Model specifications can involve all three model specification types, it must be ascertained not only that the output of the overall model is correct (similar to I/O Behavior), but also that the outputs are produced for the right reasons (i.e., that the homomorphisms between system components hold).

The final aspect of the simulation framework that needs to be discussed is how the veracity of the morphisms linking the model and source system are established. This is accomplished using the experimental frames. The set of experimental frames applied to a model are an operational formulation of the objectives motivating a modeling project. Returning to our simple fire example, if our objectives are to predict the range of vegetative patterns possible under a typical fire regime, then the role of the experimental frames are to specify the particular behaviors the model must display to meet these goals and to ensure that the model displays these behaviors. Such behaviors could include: the number of fires, their sizes, the risk of fire given particular vegetative states, and other criteria for evaluating the vegetative succession model. Practically, an experimental frame is a type of measurement or observer system consisting of a *generator* that generates the input to the system, an *acceptor* that monitors the experiment to ensure the desired experimental conditions are met, and a *transducer* that observes, analyzes and stores output (Zeigler et al., 2000). Thus, in the frame(s) for the fire model the generator produces the fire initiation events, the acceptor ensures that fires only occur when the vegetation has reached a particular state, while the transducer outputs statistics on the above assessment criteria.

If one has many different objectives in modeling a single system, many experimental frames may be formulated.

Alternatively, an experimental frame may apply across multiple systems if the same objectives are to be met. A frame is said to be *applicable* to a model if the experimental conditions required by the frame can be satisfied by the model while a model *accommodates* the frame if the frame is applicable to the model. Just as it is possible to specify a derivability relation for lower model specifications from higher specifications, it is also possible to specify a derivability relation for experimental frames. Frame A is said to be *derivable* from frame B if any model that accommodates frame B also accommodates frame A. Frame A is more *restrictive* than frame B because frame A allows a smaller set of experiments to be conducted and thus it is easier for the model to satisfy the fewer experimental conditions of the more restrictive frame than the greater number of experimental conditions of the less restrictive frame.

These background concepts and definitions will make it easier to present the methodology for modeling real systems for the purpose of increasing scientific understanding. Throughout the remainder of this paper, a single running example based on Aumann et al. (2006) will be used to illustrate the concepts presented. The background to this problem is that increased nutrient loading is increasing the extent and duration of hypoxia (patches of water with dissolved oxygen concentration <2 mg/L) in marine estuaries (Diaz and Rosenberg, 1995). Although adult blue crabs (*Callinectes sapidus*) prefer deeper waters, crab avoidance of hypoxia (e.g., Pihl et al., 1991; Das and Stickle, 1994) results in increased adult density in shallower waters in which juvenile crab density is highest. Increased adult density in the shallower waters could be leading to increased rates of crab cannibalism, thereby limiting the population. Alternatively, the decreases in the benthos caused by hypoxia (e.g., Buzzelli et al., 2002) could be leading to crab food limitation and thereby limiting the population. Because neither of these questions can be easily answered empirically, the two modeling objectives were: How do the extent and duration of transient hypoxic patches alter blue crab population dynamics and are these changes caused by increased crab cannibalism or food limitation? To determine if these questions could be answered with a model, it was necessary to first conduct a synthesis.

3. Specifying modeling objectives and performing knowledge synthesis

The synthesis process needs to be understood in the context of the five general processes of research planning (Ford, 2000). First, the questions to be addressed with the model must be defined. The second step, related to the first, involves defining why answering these questions is interesting from a larger ecological perspective and how the modeling effort could lead to more general ecological understanding. The third process requires ensuring that enough is known about the source system to actually construct a model to meet the modeling objectives. This must be determined by performing a knowledge synthesis. Determining how to apply modeling techniques to achieve the modeling objectives involves the fourth process—namely applying creativity to develop a model that can address the questions. Finally, the last process involves determining why it will be possible to use the model for its intended purposes before one actually invests the time implementing and assessing it.

Based on the questions to be answered in the example of blue crabs and hypoxia, it was realized early on that because key components of crab population dynamics are both size- and context-dependent (e.g., Jachowski, 1974; Pihl et al., 1991; Das and Stickle, 1993) a detailed representation of individuals would be needed to accurately characterize crab interactions with their habitats, prey, and each other. Such a representation is possible because blue crabs are one of the best studied of all marine organisms (Epifanio, 1995). For illustrative purposes, I will just focus on the synthesis relevant to determining which environmental variables need to be modeled. First, dissolved oxygen can be computed based on temperature, salinity, depth and total biomass (clams, etc.) (Mauersberger, 1983; Schnoor, 1996; Chapra, 1997). Second, current understanding of the crab life-cycle (van den Avyle, 1984), feeding (e.g., Eggleston, 1990; Mantelatto and Christofoletti, 2001), movement rates (e.g., Wolcott and Hines, 1990; Clark et al., 1999), and physiology (e.g., Houlihan et al., 1985; deFur, 1990; Guerin and Stickle, 1992; Booth and McMahon, 1992; McGaw and Reiber, 2000) all indicate that knowledge of salinity, temperature, and dissolved oxygen were sufficient for structuring the spatial distribution of crabs and predicting their physiology. Finally, clam growth rates (e.g., Nichols and Thompson, 1982), reproduction (e.g., Honkoop and Beukema, 1997) and mortality (e.g., Holland et al., 1987; Buzzelli et al., 2002) also indicated that salinity, temperature and dissolved oxygen are sufficient to model a clam's life-cycle. Thus, the conclusion is that the only environmental inputs needed for the model were salinity and temperature across time and space (with dissolved oxygen computed as described above). The other parts of the synthesis focused on crabs (life-cycle, feeding, movement, etc.) and clams (reproduction, growth rates, etc.) and required demonstrating that enough is known to arrive at specifications for all of these sub-models (see *Ecological Archives* M076-016-A1, Aumann et al., 2006, for all the details).

In general, synthesis proceeds in conjunction with the first two research planning stages and involves combining the constituent pieces of information about the system to be modeled into a unified description of the processes and inter-relations among the entities comprising the system (Pickett, 1999; Ford, 2000). Synthesis also involves conducting a conceptual and propositional analysis to define the axioms or propositions that are assumed to be true of the system, and specifying the data statements for how studies were conducted. In trying to determine whether enough is known about the source system to achieve the modeling objectives, a useful exercise is to consider how the set of possible achievable objectives would change if more or less were known about the source system. As in the pattern-oriented modeling approach (Wiegand et al., 2003; Grimm and Railsback, 2005; Grimm et al., 2005), the synthesis should specify plausible alternative hypotheses about system function. The adequacy of a synthesis for model construction can be judged by (Ford, 2000): (1) whether the individual propositions can be operationalized in a model and/or experimental frame, (2) whether concept definitions, and part and kind relationships are consistent throughout the theory network, model, and experimental frames, (3) whether ad hoc propositions are included to deal with special circumstances, (4) whether the fewest number of propositions are used to represent system behavior, and (5) if the explanation to be made with the synthesis encompasses the modeling objectives.

Many pitfalls can occur in conducting a synthesis (e.g., Balci and Nance, 1985; Balci, 1994). For example, in partially studied ecological systems (i.e, almost all cases) inherent biases in the personal standpoints of researchers can make it difficult to distinguish between known versus perceived causes or important vs unimportant processes. This is due to the inherent bias of explaining behavior based only on what is known about the system, instead of acknowledging this uncertainty. As a result, the problem boundary can be incorrectly specified and important unstudied components excluded from the problem formulation. If key knowledge gaps are identified, this may either necessitate a re-formulation of modeling objectives or require further empirical research to eliminate such gaps.

One of the greatest challenges in modeling is that each of the five general research processes needs to be done *before* the model is actually implemented. Too often, model implementation is attempted directly without any formal knowledge synthesis or even model design specification. This leads not only to a "build-and-fix" approach to model implementation (i.e., hacking), but also dramatically increases the risk of producing a model that bears no formal relationship (i.e., lacks explicit morphisms) to any well-defined scientific problem. Even if the model implementation is truly fantastic, the model results will be totally irrelevant. It cannot be over-emphasized: synthesis and problem formulation is one of the most important steps in the modeling process.

The deliverable at this stage of the modeling process is a document detailing the synthesis. In addition to facilitating communication between scientists and modelers, such a document greatly decreases the probability of problem miscommunication resulting in the wrong problem being solved. Ideally, those involved in model design should take the lead in writing this document since designing an effective model requires an understanding of both the biological nuances of the problem and simulation modeling.

4. Model design and implementation

Model design or specifying the structures to be modeled requires moving from lower to higher levels in the specification hierarchy. Based on (likely very incomplete) source system I/O Behavior coupled with the knowledge summarized in the synthesis about the way in which the internals of the system operate, the goal of model design is to come up with a data generating mechanism that is general enough to accommodate the experimental frames formulated based on the modeling objectives. Coming up with such a mechanism requires a process of decomposition. *Decomposition* is the process of transforming the description of the system given in the synthesis into a *composition* of descriptions of smaller systems so that global system properties arise from the interaction of more localized properties (Benjamin, 2000). For example, a single crab is made up of sub-models governing reproduction, feeding, movement, growth, energy balance, etc. The degree of coupling between these sub-models can be quite high, with movement and growth both highly dependent on energy balance, which in turn is dependent on the feeding rules. The model decomposition should respect the concept definitions, part and kind relationships in the synthesis, and should not introduce ad hoc postulates into the model formulation. Of course, there is no single "correct" model formulation (Oreskes and Belitz, 2001).

The suggestion that a system should be decomposed to generate its dynamics raises the old debate between reductionists and holists, with reductionists arguing that a system cannot be understood without decomposing it into parts and holists arguing that such a decomposed system cannot lead to the holistic behaviors of the original system. Although much has been written about the holism/reductionism debate in ecology (e.g., Levins and Lewontin, 1980; Shrader-Frechette, 1986), I focus on one possible solution to this problem that is useful for modeling complex systems. The key features of this approach are that complex systems are represented using at least three scalar hierarchical levels (e.g., O'Neill et al., 1986) where each level is defined by the existence of one or more emergent properties (Miller, 1978; Bergandi and Blandin, 1998; Bergandi, 2000). After defining these concepts in greater detail below, Section 4.2 describes the process necessary to represent a complex system using this framework.

4.1. *Scalar hierarchies and emergent properties*

In a *scalar hierarchy*, the objects at a given level contain, volumetrically and structurally, the objects of lower hierarchical levels while the given level is also part of higher hierarchical levels (Feibleman, 1954; Rowe, 1961; Bergandi and Blandin, 1998). Such a scalar hierarchy (Fig. 2) must be distinguished from the model specification hierarchy: model components within a scalar hierarchy can be specified at any level of the model specification hierarchy. For example, while an individual crab is a component model, its sub-models can consist of I/O Behavior, I/O System, and other Component models (Fig. 2).

While this scalar hierarchy will play a central role in how scientific explanations are given for model behaviors (Section 6), at this stage one can think of the highest level as providing the "purpose" for the system while the lowest level provides the "mechanism" (Feibleman, 1954). For example, in Fig. 2, one purpose of an individual crab's movement is to avoid hypoxic waters (focal + 1), but the way in which a crab avoids hypoxia depends on some internal rule-set (focal − 1) that biases its movement away from areas perceived as hypoxic.

Conceptually, an *emergent property* is something that is influenced by the properties and relations characterizing the lower levels, but is not strictly predictable from, nor reducible to, these lower-level entities (Damper, 2000; Rasmussen et al., 2001; Reuter et al., 2005). An emergent property represents something new at a given level that is not seen at the level below. For example, the total biomass of the crab population is not an emergent property since this aggregate property is simply the sum of the masses of each individual and the mass of the individual is just the sum of the masses its parts. However, the rate at which a crab grows, its inter-molt period or its spawning interval are new properties which do not exist for the crab's parts. While a crab's stomach and digestive

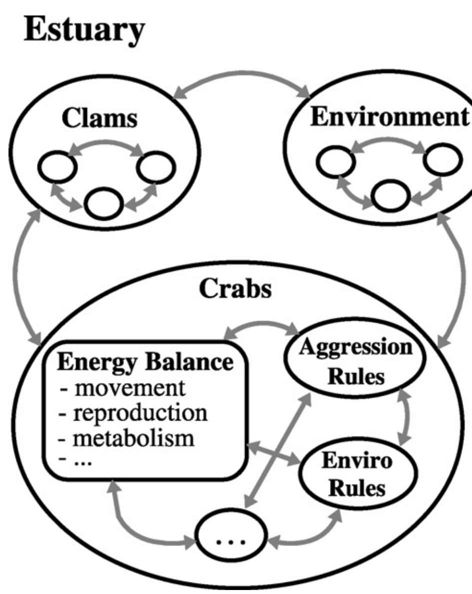


Fig. 2 – The scalar hierarchy used in the crab model consists of three levels: individual or focal, estuary or focal + 1, and the internals of the individual or focal − 1. In the figure, the three levels are only fully shown for crabs. The units within the focal level are composed of the units at the focal − 1 level. The coupling between units within a level is generally high, while the degree of coupling between levels is low. Each focal level is distinguished by the existence of one or more emergent properties.

processes contribute to such individual level properties, neither has the property of a spawning interval. Further, individual-level properties like growth rate are not solely predictable from just knowing a crab's lower-level metabolic rate or feeding rules. In addition, prey availability along the path traversed by the crab must be known. The key point is that both the higher and the lower levels are required for properties like crab growth rate to emerge since it is these higher level entities that impinge upon the individual crab as it moves throughout the estuary and the crab in turn responds via its lower level rules (e.g., for crab feeding/aggression, movement and internal energy balance). While the emergent property of growth rate is influenced by these lower level entities and processes it is not determined or reducible to them and thus the holistic argument fails. Similarly, basing the hierarchical levels on the existence of one or more emergent properties means that any claim of reductionism also fails since new properties do exist at the different levels that are not explicable in terms of lower level entities.

4.2. *Model decomposition and specifying an hierarchical design*

Having defined the nature of the scalar hierarchy based on emergent properties, the process of specifying a model design can now be discussed and involves assigning the entities or units used in the synthesis to the levels in the scalar hierarchical framework (Fig. 2). In particular (Bergandi and Blandin, 1998):

- Developing a representation of the focal or primary system level by proposing its spatio-temporal limits, describing its main phenomenological characteristics, and specifying the experimental frames operating at this level. In the crab-hypoxia example, the focal level was chosen to be that of the individual (Fig. 2). Individual crabs have the ability to move, respond to their environment, attack other crabs, feed on prey, etc. For both clams and the environmental variables, the focal level is a small region of space over which these variables are assumed homogeneous. This region of space keeps track of the age/size structure of clams and their abundance, and the dissolved oxygen, temperature, depth and salinity at that location. Assessment of the sub-models within each level is discussed in Section 5.
- Developing a representation of the sub-systems at the immediately lower level (focal − 1) of integration making up the focal level and specifying the relevant experimental frames. As discussed above, an individual crab consists of sub-models governing movement, reproduction, energy balance, aggression, etc. (Fig. 2). Similar models exist for clams while the way in which the environmental variables change over the small region of space is governed by the environmental inputs and equations governing dissolved oxygen.
- Developing a representation of the system from the immediately higher level (focal + 1) of integration which the focal level is part of and specifying its relevant experimental frames. In our example, we refer to this higher level as the estuary level as it is comprised of all individual crabs, the clams and the environmental variables over the entire 2D estuary (Fig. 2). It is at this level that one can talk about population averages for crabs and clams, and also environmental averages like the percent of the estuary hypoxic over the summer.
- Defining the processes and relationships linking the focal level to the focal + 1 and focal − 1 levels and determining which processes operating between the focal − 1 and focal levels are most dependent on the processes occurring between the focal + 1 and focal levels. Information needs to be passed down, up and within each of these levels. For example, as individual crabs move throughout the estuary, the environmental conditions encountered need to be passed to the crab's lower-level models. Alternatively, the location of individual crabs needs to be passed up to the estuary level so that local crab-crab interactions can occur.

Within and between each level, one has to decide on units and how these units inter-relate to each other. A *unit* is a collection of objects that are chosen by an observer in a way that allows them to be characterized as new relevant objects (Jax et al., 1998). To identify units in ecology, it is helpful to locate the proposed unit along four dimensions (Jax et al., 1998; Grimm, 1998). The first dimension specifies the degree to which the unit is defined topographically (due to the spatial relatedness of its elements) or functionally (due to shared interactions among its elements). The second dimensions relates to the degree of expected internal relationships within the unit, which should be high. The third dimension relates to the set of selected phenomena, elements or processes that must comprise the unit for it to exist. Finally, the fourth dimension relates to the resolution or degree of aggregation of these selected elements comprising the unit. Thus, an individual crab is a unit because of the tight spatial relatedness of its elements and the tight functional relationships between its elements. For the purposes of the model, the elements taken as necessary for a crab to exist include its reproductive, feeding and aggression sub-models, etc., along with attributes like its mass and size. Establishing the strongest possible morphism between the synthesis and model design requires ensuring that the relationships between the units of both are specified with a similar degree of resolution or abstraction.

Although synthesis precedes model design, in practice these steps are not disjoint activities, but part of an iterative cyclic process. A model design document detailing the decomposition is necessary for achieving a robust model and also for explicitly establishing the morphisms between the model design and synthesis. Together, the synthesis and design documents comprise one of the primary ways in which non-modelers will understand and thus critique the model.

In summary, the main challenge in model decomposition is arriving at a representation that uses the units described in the synthesis, is computationally efficient, and accommodates the specified experimental frames. In deciding on how detailed a representation is needed for the units, it is best to start with the simplest model necessary for the desired patterns to emerge. As discussed below, the assessment process will determine whether this representation is inadequate for the stated modeling objectives. Anyone who has done such work will agree that "*the hard part of building software [is] the specification, design, and testing of this conceptual construct, not the labor of representing it and testing the fidelity of the representation*" (emphasis in original Brooks Jr., 1995, p. 182). Achieving the above goals is not easy, but can be made easier by adopting good software design practices so that the final product is computationally efficient, modular, extensible, reliable, maintainable, testable, flexible, portable, reusable, and interoperable. Much research on software design has occurred in computer science in recent years and should be consulted for further guidance (e.g., Fishwick, 1995; Rumbaugh et al., 1999; Jacobson et al., 1999; Sommerville, 2001; Bruegge and Dutoit, 2004).

4.3. *Programming the model*

The model design document forms the basis for implementing the model in a computer language that can be executed by a simulator. Implementing the model involves the creation of another morphism linking the model design document to this implementation. Assessing the veracity of these morphisms is discussed in Section 5. Since the advantages/disadvantages of different programming languages (e.g., C++, JAVA, Visual Basic, etc.) are highly context and problem dependent, the issues to consider in choosing among them will not be considered here.

5. Model verification, validation, and critique

Unlike physical experiments which are grounded in the actual material of the system, simulation models can only be used to learn about the original system if the model stands in a certain relation of similarity or analogy to the represented system (Guala, 2002). It is not sufficient to merely claim that the morphisms connecting the two systems are adequate for the problem at hand. The purpose of model assessment is to explicitly demonstrate the adequacy and strengths of these morphisms. Metaphorically speaking, the aim of model assessment is to put the modeler, model and the knowledge underlying the model "on the hook" (Kleindorfer et al., 1998; Anderson and Bates, 2001). This can be facilitated by adhering to an explicit model assessment methodology and documenting the process so that the assessment can be evaluated by others, thereby building confidence in the model.

There are three stages to model assessment: model verification, model validation, and a critique of the way in which verification, validation, model design and the synthesis were carried out. This process is by far the most time consuming relative to the stages in the modeling process described above, easily consuming 50% or more of the total time required for the entire modeling process. Practically, it makes sense to work backwards from the model implementation and first assess the morphism linking the model implementation to the design document. *Model verification* is concerned with ensuring the model has been built right relative to the model design document (Balci, 1994; Zeigler et al., 2000). Verifying a model requires demonstrating that it has been implemented as described in the design document and is accomplished via a process of checking each model entity against its design specification.

In contrast, *model validation* is concerned with building the right model. This involves substantiating that the behavior of the model represents the behavior of the system with sufficient accuracy so that it is impossible to distinguish the behavior of both systems in the experimental frames (Zeigler et al., 2000; Wiegand et al., 2003). As others have noted (e.g., Oreskes and Belitz, 2001), validation is an unfortunate term to use in this context since the output of this process does not imply model assurance or "truth"; however, this terminology is in widespread use. Saying it is impossible to distinguish the two systems requires the notion of *replicative validity*, namely that for all experiments possible within the experimental frame, the behavior of the model and system agree within some specified tolerance at the I/O Behavior level (Zeigler et al., 2000). For models specified at the I/O System or Coupled Component levels, the introduction of a state-space necessitates the stronger notion of *structural validity* or that a model mimics in a step-by-step, component-by-component fashion the way the source system performs its transitions. Structural validity ensures that the model is generating the correct I/O Behavior for the right reasons, and not because incorrect behavior of one model component is compensated for by incorrect behaviors of other components. Methodologically, structural validity can be achieved by applying experimental frames to each of the model components within a given scalar level and across hierarchical levels to assess the larger scale consequence of these behaviors. By examining which experimental frames can be accommodated and which cannot reveals a great deal about the adequacy/inadequacy of the proposed model structure (Reynolds and Ford, 1999; Wiegand et al., 2003).

Returning to the crab-hypoxia example, our discussion of assessment presented here will focus primarily around crab physiology. The entire model assessment is available (see Appendix B, *Ecological Archives* M076-016-A2, Aumann et al., 2006) while other examples are described in Grimm et al. (2005). It is easiest to start the validation process at the lowest scalar hierarchical levels and then move upwards as one gains confidence in the lower-level model components. Thus, Frame A in Fig. 3 is concerned with validating the components of the model associated with crab physiology. Assessing a model component like its feeding algorithm would include examining the way in which the crab feeds, what size of clams it feeds on, and whether the feeding behavior is activated under appropriate input conditions. The rate of energy usage due to movement would be assessed using inputs of different rates of movement along with the outputs of amount of energy consumed, and by ensuring that both inputs and outputs agree with what is known empirically (e.g., Houlihan et al., 1985). Rates of energy usage from a crab's base metabolism would be assessed using a frame that specifies different environmental temperatures and by ensuring that the rates predicted agree with those observed experimentally (e.g., Booth and McMahon, 1992; McGaw and Reiber, 2000). Similarly for all the other components involved in crab physiology.

Moving up to Frame B (Fig. 3), the focus shifts to assessing the individual crab object. The particular input conditions used would include the conditions a typical crab is subject to (which should be close to what exists in an actual estuary as determined by the experimental frames for assessing the habitat and environmental conditions) and the crab unit would be assessed using criteria like: the time it takes to reach sexual maturity, the time between molts, its size at maturity, its distribution of movement rates, the time between reproductive events, etc. (e.g., van Engle, 1958; Leffler, 1972; Ryer et al., 1990; Fisher, 1999). Given that we have some confidence in the behavior of lower level components of the model, a failure to accommodate these crab-level criteria may indicate a failure in the specification of the lower level criteria, or a failure in the way these lower level components are interacting.

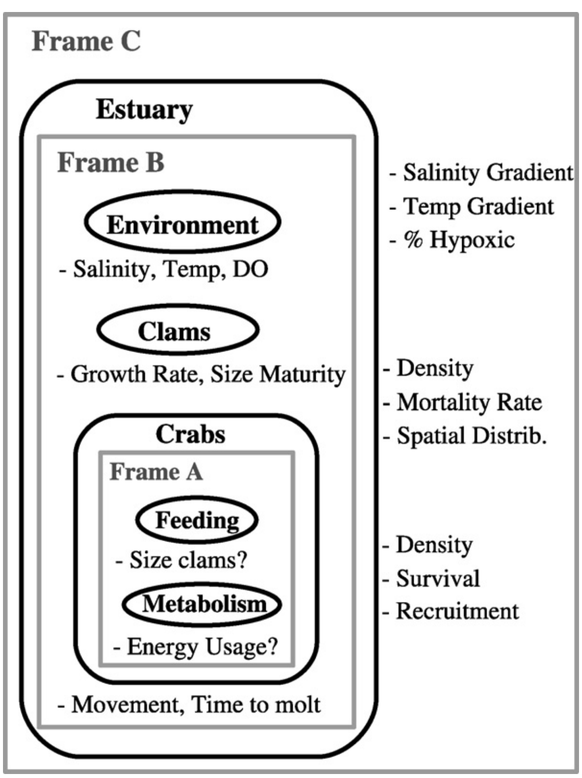


Fig. 3 – Representation of the experimental frames (gray lines) used to assess the model components (black lines) across the scalar hierarchical levels. The conditions in Frame C encompass those from Frame B which in turn encompasses those from Frame A. Thus, Frame A is derivable from Frames B or C. See Section 5 for a more detailed description of the frames.

Moving up further to Frame C (Fig. 3), the inputs to the frame would include estuary-level environmental variables and prey while the criteria used would include larger estuary-level criteria like: crab densities across instars, crab survival and mortality rates, spatial distribution, rates of recruitment, etc. (e.g., Orth and van Montfrans, 1987; van Montfrans et al., 1995; Rugolo et al., 1998; Eggleston, 1998; Clark et al., 1999). If all of these criteria are satisfied within the acceptable levels of agreement between the two systems, and if all of the lower level frames are also applicable, then we can have increased confidence that the observed high level behaviors are occurring for valid reasons. Further, applying experimental frames across the hierarchical levels of the model in this way reduces the propagation of uncertainty throughout the model since if sub-models are deemed to behave correctly this constrains the possible behaviors of higher level model components (DeAngelis and Mooij, 2003; Wiegand et al., 2004). Finally, starting with the simplest models believed necessary to satisfy the experimental frames avoids including detail that may not be required. The validation process will illuminate if particular factors or processes are not adequately represented while also enabling the evaluation of alternative representations.

One point that must be stressed is that if empirical data are available, their validity for judging model performance relative to the overall modeling objectives must be evaluated (Balci, 1997). Measurement error inherent in the data must also be considered (Wiegand et al., 2003), since it is erroneous to try and achieve a greater degree of agreement between the two systems than is possible given this error. It must also be stressed that formulating experimental frames to validate models is more general than simply comparing model output with empirical patterns. As Oreskes and Belitz (2001, p. 32) note "the capacity to mimic data is not evidence that you have captured underlying processes, and therefore not evidence of predictive capacity" since models that inaccurately capture the underlying causal processes can predict quite well, possibly due to trade-offs among model components. This is why the experimental frames need to include structural integrity criteria. Even in no data situations, such structural integrity criteria coupled with numerous "vague" patterns (patterns that singly seemingly contain little information) can be applied and may provide as much information as a single strong data pattern (Grimm et al., 2005, and Online Material). Experimental frames can also be used when attempting to establish the validity of a simpler model relative to a more complex model. In summary, possessing detailed empirical data is not a pre-requisite for formulating model validation criteria and the lack of empirical data is no justification to forgo this step of the assessment process.

Relative to model verification, performing model validation is a more open-ended process since it involves decisions about whether the behavior of a model is "close enough" to that demanded by the experimental frames. The outcome at any stage should not be considered as absolutely correct or incorrect, but instead as a degree of credibility. In addition to being aware of the accuracy of the data or patterns being used to judge model validity, a number of other suggestions may prove helpful (Balci, 1994; Kleijnen, 1995; Rykiel, 1996; Balci, 1997; Yilmaz and Balci, 1997). First, because multiple inputs and outputs need to be compared for both systems, multi-criteria evaluative techniques will be needed (e.g., Refsgaard and Storm, 1996; Reynolds and Ford, 1999; Komuro et al., 2006). Second, a simulation model must be judged relative to the overall modeling objectives (operationalized via the experimental frames) and these objectives must guide the degree of credibility required. Increased credibility will entail increased development costs, yet both more and less detailed models may be equally valid within a particular experimental frame and thus there may be no justifiable reason for such increased costs. Third, it must be realized that model validation is challenging and cannot be done by just anyone since devising test cases requires detailed knowledge of the problem, a through understanding the entire simulation model design, expertise in modeling, and experience with validation techniques. Fourth, model validity can only be claimed for the prescribed conditions under which the model is tested. Since it is not possible to completely test any simulation model, the test cases should be designed to cover potential future model uses consistent with the modeling objectives. How much testing is required and when to stop are determined by the model's desired domain of applicability—the larger this domain, the more testing required. Fifth, the validation process must be documented, so that both non-modelers and model developers can critique it. Finally, model validation must be conducted throughout the modeling process, since correcting errors and inadequacies from earlier modeling stages (e.g., synthesis and model design) becomes much more time-consuming and expensive if discovered in later modeling stages (e.g., after implementation).

The final stage of model assessment involves a larger critique of the processes of model verification, validation, design, and knowledge synthesis. This is necessary because model validity is a necessary but insufficient condition for establishing the credibility of simulation results (Balci, 1997). Model validity is only established relative to the study objectives as operationalized through the experimental frames. If these objectives are incorrectly specified and/or the model is incorrectly defined, the model may still be valid with respect to this incorrect specification even though the simulation results will not be credible. The aim of this larger critique is to detect such inadequacies. Judgments about whether a model has been adequately verified, validated or is credible involve increasing degrees of subjectivity. The way to counter this subjectivity is via the continuous application of just and effective criticism across all stages of the modeling process.

6. Using the model to make scientific inferences

One of the primary reasons for building complex simulation models of real systems is to use the model to construct scientific explanations (Oreskes and Belitz, 2001; DeAngelis and Mooij, 2003). By a *scientific explanation*, we mean the answer to a why-type question that specifies its topic of concern, defines its contrasting sets of alternatives, defines the explanatory relevance it requires, and is constructed based on causal and/or associated organizational processes that are consistent and general within the theory network (Ford, 2000). This section explains why models constructed using the above methodology can be used to make scientific explanations about the original system.

The reason models can answer such why-type questions depends on two key properties of the above methodology. The first is that a series of morphisms link the model implementation back to the theory network expressed in the synthesis document. The second is the process of model assessment that creates confidence in the strength of these morphisms. It is these two components that allow the modeler to explicitly demonstrate how the model is like or not like the source system and enables model behaviors to be mapped back to behaviors, properties and processes in the source system. In this way, the model becomes a surrogate for the source system. Thus, if a valid and credible model produces simulation results inconsistent with expectations we must either believe that the assessment process is inadequate, or we must modify our beliefs about how the system functions. Treating the model as an experimental system and subjecting it to additional simulation experiments or sensitivity analysis (Peck, 2004) can help to hone in on suspect postulates in the model design or synthesis.

The consequence of the above modeling methodology is that the scalar hierarchical decomposition used to structure model design also provides an hierarchical structure to any explanation. That these explanations are structured differently than in other scientific disciplines need not concern us, because as Shrader-Frechette (1986, p. 91) notes, "if biological phenomena are hierarchically organized, as indeed they seem to be, then perhaps they require special types of explanation". To understand the structure of the explanations given, recall that model components at the focal − 1 level are generally specification I/O Behavior models. These models do not explain, for example, why the crab's base metabolism changes with temperature, but merely describe such changes. This enables us to use such descriptive facts without seeking even lower-level explanations for such metabolic changes. If such mechanistic explanations were deemed necessary to the modeling objectives, then the model decomposition would have been altered to include even lower levels. Similarly, the focal + 1 level provides the context in which the focal level behaviors occur. Thus the question: "How does altering the degree and extent of hypoxia in the estuary alter crab population dynamics?" is answered by manipulating the environmental conditions at the estuary level and observing how such changes alter individual-level dynamics, such changes being dependent on the rules at the focal − 1 level.

For example, suppose the simulations show that increasing hypoxic extents (focal + 1) decreased population density (focal + 1). Further analysis showed that at least part of this decrease is because the per female rate of egg production decreased (focal) and that this reduction in reproductive output was caused by crabs spending more time in warmer waters which altered their energy balance (focal − 1) so less energy was available for egg production.

It is worth considering how the nature of this explanation would change if the model were only specified at a single level (e.g., the population level). First, any explanation offered solely based on the model would of necessity be much more limited, since the physiological processes and environmental conditions would have to be aggregated – somehow – into the single-level formulation. Second, aggregation would make it more difficult to assess the importance of these factors relative to other factors such as cannibalism and avoidance of hypoxia. In contrast, the hierarchical formulation of the simulation model can explicitly include each of these factors, thereby making it possible to test the strength and relevance of possible alternative explanations.

7. Concluding remarks

One of the reasons simulation models have not been more successful in advancing ecological understanding is that most have not adequately represented the complexity in the source systems (Grimm, 1999). The standpoint of this paper is that constructing models using the methodology described would go a long way towards remedying this situation. This methodology is focused around four main points. First, the process of building simulation models should be planned even more thoroughly than other research inquiries. One key step in this planning process is conducting a synthesis to establish what is known about the system and whether the stated objectives can be achieved with available knowledge. If the objectives cannot be achieved they must either be modified, or key knowledge gaps will first have to be filled via additional empirical inquiry. Second, the synthesis document forms the basis against which the model decomposition is done and is key to ensuring that similar units and part-whole relationships are used through all stages of the modeling process. The model design should be structured hierarchically with each level being defined by the existence of one or more emergent properties—lower level entities contained within higher level entities. This decomposition will also structure the scientific explanations made from the model. Third, the strength of the morphisms linking the model implementation to the model design and synthesis need to be explicitly established via a process of model verification, validation, and critique. This process is necessary to explicitly establish how the model is “like” and “unlike” the original system. Finally, the previous steps need to be documented to facilitate critique by the larger scientific community. The purpose of this critique is to challenge both the model and to stimulate novel empirical inquiry—thereby increasing the chances of advancing ecological understanding.

The scientific explanations constructed based on the model are highly dependent on each of these four points. Some may see this as a weakness, but it is a false expectation to think that any method of scientific inquiry (whether empirical or model based) will give objective knowledge of reality that is beyond the scope of criticism. Instead, our knowledge, and the models we build based on that knowledge, must be seen as provisional. This is a very different perspective from asserting that the principles embodied by the model are general and thus beyond the scope of criticism. In contrast, the perspective presented here is that scientific understanding will advance more quickly if such dependencies and uncertainties are openly acknowledged, as this helps to focus future inquiry. As new experimental results become available, either the theory network and model will have to be modified, or we will be justified in placing increased confidence in the theory. If modifications are required, then the resulting model structure (and thus new theory network) can again be subjected to novel simulation experiments. In this way, the theory network underlying a complex model can be subjected to much more rigorous examination than is possible with any simple model embodying general principles.

Acknowledgments

I would like to thank E. David Ford for sharing his insights into the scientific process over the years, and for his critical comments on an early draft. Comments from Mark Andruskiw, Jalene LaMontagne, Nicole McCutchen and two annonymous reviewers also improved this manuscript.

REFERENCES

Anderson, M.G., Bates, P.D., 2001. Hydrological science: model credibility and scientific understanding. In: Anderson, M.G., Bates, P.D (Eds.), Model Validation: Perspectives in Hydrological Science. John Wiley & Sons, New York, pp. 1–10 (Chapter 1).

Aumann, C.A., Eby, L.A., Fagan, W.F., 2006. How transient patches affect population dynamics: the case of hypoxia and blue crabs. Ecol. Monogr. 76 (3), 415–438.

Balci, O., 1994. Validation, verification, and testing techniques throughout the life cycle of a simulation study. Ann. Oper. Res. 53, 121–173.

Balci, O., 1997. Principles of simulation model validation, verification, and testing. Trans. Soc. Comput. Simul. Int. 14 (1), 3–12.

Balci, O., Nance, R.E., 1985. Formulated problem verification as an explicit requirement of model credibility. Simulation 45 (2), 76–86.

Benjamin, D.P., 2000. On the emergence of intelligent global behaviours from simple local actions. Int. J. Syst. Sci. 31 (7), 861–872.

Bergandi, D., 2000. Eco-cybernetics: the ecology and cybernetics of missing emergences. Kybernetes 29 (7/8), 928–942.

Bergandi, D., Blandin, P., 1998. Holism vs reductionism: do ecosystem ecology and landscape ecology clarify the debate? Acta Biotheor. 46, 185–206.

Booth, C.E., McMahon, B.R., 1992. Aerobic capacity of the blue crab, *Callinectes sapidus*. Physiol. Zool. 65 (6), 1074–1091.

Brooks, Jr., F.P., 1995. The Mythical Man-Month: Essays on Software Engineering, anniversary ed. Addison-Wesley, Ontario.

Bruegge, B., Dutoit, A.H., 2004. Object-Oriented Software Engineering Using UML, Patterns and Java. Prentice Hall, Toronto.
Buzzelli, C.P., Luettich Jr., R.A., Powers, S.P., Peterson, C.H., McNinch, J.E., Pinckey, J.L., Paerld, H.W., 2002. Estimating the spatial extent of bottom-water hypoxia and habitat degradation in a shallow estuary. Mar. Ecol. Progr. Ser. 230, 103–112.
Cartwright, N., 1999. The Dappled World: A Study of the Boundaries of Science. Cambridge University Press, Cambridge.
Chapra, S.C., 1997. Surface Water-Quality Modeling. McGraw-Hill, New York.
Clark, M.E., Wolcott, T.G., Wolcott, D.L., Hines, A.H., 1999. Foraging and agonistic activity co-occur in free-ranging blue crabs (*Callinectes sapidus*): observation of animals by ultrasonic telemetry. J. Exp. Mar. Biol. Ecol. 233, 143–160.
Cooper, G.J., 2003. The Science of the Struggle for Existence: On the Foundations of Ecology. Cambridge University Press, New York, NY.
Damper, R.I., 2000. Editorial for the special issue on 'Emergent Properties of Complex Systems': emergence and levels of abstraction. Int. J. Syst. Sci. 31 (7), 811–818.
Das, T., Stickle, W.B., 1993. Sensitivity of crabs *Callinectes sapidus* and *C. similes* and the gastropod *Stramonita haemastoma* to hypoxia and anoxia. Mar. Ecol. Progr. Ser. 98, 263–274.
Das, T., Stickle, W.B., 1994. Detection and avoidance of hypoxic water by juvenile *Callinectes sapidus* and *C. similes*. Mar. Biol. 120, 593–600.
DeAngelis, D.L., Mooij, W.M., 2003. In praise of mechanistically rich models. In: Canham, C.D., Cole, J.J., Lauenroth, W.K. (Eds.), Models in Ecosystem Science. Princeton University Press, Princeton, NJ, pp. 61–82 (Chapter 5).
DeAngelis, D.L., Rose, K.A., Huston, M.A., 1994. Individual-oriented approaches to modeling ecological populations and communities. In: Levin, S.A. (Ed.), Frontiers in Mathematical Biology. Springer, Berlin, pp. 290–410.
deFur, P.L., 1990. Respiration during ecdysis at low salinity in blue crabs, *Callinectes sapidus* Rathbun. Bull. Mar. Sci. 46 (1), 48–54.
Diaz, R.J., Rosenberg, R., 1995. Marine benthic hypoxia: a review of its ecological effects and the behavioral responses of benthic macrofauna. Oceanogr. Mar. Biology: Ann. Rev. 33, 245–303.
Eggleston, D.B., 1990. Foraging behavior of the blue crab, *Callinectes sapidus* on juvenile oysters, *Crassostrea virginica*: effects of prey density and size. Bull. Mar. Sci. 46 (1), 62–82.
Eggleston, D.B., November 1998. Population dynamics of the blue crab in North Carolina: statistical analyses of fisheries survey data. Tech. rep., North Carolina State University, Department of Marine, Earth and Atmospheric Sciences, Raleigh, NC, pp. 27695–28208.
Epifanio, C.E., 1995. Transport of blue crab (*Callinectes sapidus*) larvae in the waters off mid-Atlantic states. Bull. Mar. Sci. 57 (3), 713–725.
Feibleman, J.K., 1954. Theory of integrative levels. Br. J. Philos. Sci. 5 (17), 59–66.
Fisher, M.R., 1999. Effect of temperature and salinity on size at maturity of female blue crabs. Trans. Am. Fish. Soc. 128, 499–506.
Fishwick, P.A., 1995. Simulation Model Design and Execution: Building Digital Worlds. Prentice Hall, Englewood Cliffs, NJ.
Ford, E.D., 2000. Scientific Method for Ecological Research. Cambridge University Press, New York.
Grimm, V., 1998. To be, or to be essentially the same: the 'Self-Identity of Ecological Units'. Trends Ecol. Evol. 13 (8), 298–299.
Grimm, V., 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?. Ecol. Model. 115, 129–148.
Grimm, V., Railsback, S.F., 2005. Individual-Based Modeling and Ecology. Princeton University Press, Princeton.
Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H., Weiner, J., Wiegand, T., DeAngelis, D.L., November 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310 (11), 987–991.
Guala, F., 2002. Models, simulations and experiments. In: Magnani, L., Nersessian, N.J. (Eds.), Model-Based Reasoning: Science, Technology, Values. Kluwer Academic, New York, pp. 59–74.
Guerin, J.L., Stickle, W.B., 1992. Effects of salinity gradients on the tolerance and bioenergetics of juvenile blue crabs (*Callinectes sapidus*) from water of different environmental salinities. Mar. Biol. 114, 391–396.
Holland, A.F., Shaughnessy, A.T., Hiegel, M.H., 1987. Long-term variation in mesohaline Chesapeake Bay macrobenthos: spatial and temporal patterns. Estuaries 10 (3), 227–245.
Honkoop, P.J.C., Beukema, J.J., 1997. Loss of body mass in winter in three intertidal bivalve species: and experimental and observation study of the interacting effects between water temperature, feeding time and feeding behaviour. J. Exp. Mar. Biol. Ecol. 212, 277–297.
Houlihan, D.F., Govind, C.K., Haj, A.E., 1985. Energetics of swimming in *Callinectes sapidus* and walking in *Homarus americanus*. Comp. Biochem. Physiol. A 82 (2), 267–279.
Jachowski, R.L., 1974. Agonistic behavior of the blue crab, *Callinectes sapidus* Rathbun. Behaviour 50, 232–253.
Jacobson, I., Booch, G., Rumbaugh, J., 1999. The Unified Software Development Process. Addison-Wesley, Don Mills, Ontario.
Jax, K., Jones, C.G., Pickett, S.T., 1998. The self-identity of ecological units. Oikos 82, 253–264.
Judson, O.P., 1994. The rise of the individual-based model in ecology. Trends Ecol. Evol. 9 (1), 9–14.
Kleijnen, J.P., 1995. Verification and validation of simulation models. Eur. J. Oper. Res. 82, 145–162.
Kleindorfer, G.B., O'Neill, L., Ganeshan, R., 1998. Validation in simulation: various positions in the philosophy of science. Manage. Sci. 44 (8), 1087–1099.
Klir, G.J., 1985. Architecture of Systems Complexity. Saunders, New York.
Komuro, R., Ford, E., Reynolds, J.H., 2006. The use of multi-criteria assessment in developing a process model. Ecol. Model. 197 (3–4), 320–330.
Leffler, C.W., 1972. Some effects of temperature on the growth and metabolic rate of juvenile blue crabs *Callinectes sapidus*, in the laboratory. Mar. Biol. 14, 104–110.
Levins, R., 1966. The strategy of model building in population biology. Am. Sci. 54, 421–431.
Levins, R., Lewontin, R., 1980. Dialectics and reductionism in ecology. Synthese 43, 47–78.
Magnani, L., 2004. Conjectures and manipulations. Computational modeling and the extra-theoretical dimension of scientfic discovery. Minds Mach. 14, 507–537.
Mantelatto, F.L.M., Christofoletti, R.A., 2001. Natural feeding activity of the crab *Callinectes ornatus* (Portunidae) in Ubatuba Bay (Sao Paulo, Brazil): influence of season, sex, size and molt stage. Mar. Biol. 138, 585–594.
Mauersberger, P., 1983. General principles in deterministic water quality modeling. In: Orlob, G. (Ed.), Mathematical Modeling of Water Quality. John Wiley and Sons, New York, pp. 42–115 (Chapter 3).
McGaw, I.J., Reiber, C.L., 2000. Integrated physiological responses to feeding in the blue crab *Callinectes sapidus*. J. Exp. Biol. 203, 359–368.
Miller, J.G., 1978. Living Systems. McGraw-Hill, Toronto.

Morrison, M., Morgan, M.S., 1999. Models as mediating instruments. In: Morgan, M.S., Morrison, M. (Eds.), Models as Mediators. Cambridge University Press, pp. 10–37 (Chapter 2).
Morton, A., Suárez, M., 2001. Kinds of models. In: Anderson, M.G., Bates, P.D. (Eds.), Model Validation: Perspectives in Hydrological Science. John Wiley & Sons, New York, pp. 11–21 (Chapter 2).
Nichols, F.H., Thompson, J.K., 1982. Seasonal growth in the bivalve *Macoma balthica* near the southern limit of its range. Estuaries 5 (2), 110–120.
O'Hara, R.B., 2005. The anarchist's guide to ecological theory. Or we don't need no stinkin' laws. Oikos 110 (2), 390–393.
O'Neill, R.V., DeDangelis, D.L., Waide, J.B., Allen, T.F.H., 1986. A Hierarchical Concept of Ecosystems. Princeton University Press, Princeton, NJ.
Oreskes, N., 2003. The role of quantitative models in science. In: Canham, C.D., Cole, J.J., Lauenroth, W.K. (Eds.), Models in Ecosystem Science. Princeton University Press, Princeton, NJ, pp. 13–31 (Chapter 2).
Oreskes, N., Belitz, K., 2001. Philosophical issues in model assessment. In: Anderson, M.G., Bates, P.D. (Eds.), Model Validation: Perspectives in Hydrological Science. John Wiley & Sons, New York, pp. 23–41 (Chapter 3).
Orth, R.J., van Montfrans, J., 1987. Utilization of a seagrass meadow and tidal marsh creek by blue crabs *Callinectes sapidus*. I. Seasonal and annual variations in abundance with emphasis on post-settlement juveniles. Mar. Ecol. Progr. Ser. 41, 283–294.
Pace, M.L., 2003. The utility of simple models in ecosystem science. In: Canham, C.D., Cole, J.J., Lauenroth, W.K. (Eds.), Models in Ecosystem Science. Princeton University Press, Princeton, NJ, pp. 49–62 (Chapter 4).
Palladino, P., 1991. Defining ecology: ecological theories, mathematical models, and applied biology in the 1960s and 1970s. J. Hist. Biol. 24 (2), 223–243.
Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., Deadman, P., 2003. Multi-agent systems for the simulation of land-use and land-cover change: a review. Ann. Assoc. Am. Geogr. 93 (2), 314–337.
Parker, P., Letcher, R., Jakeman, A., Beck, M., Harris, G., Argent, R., Hare, M., 2002. Progress in integrated assessment and modeling. Environ. Model. Software 17, 209–217.
Peck, S.L., Oct 2004. Simulation as experiment: a philosophical reassessment for biological modeling. Trends Ecol. Evol. 19 (10), 530–534.
Peterson, G.D., Carpenter, S.R., Brock, W.A., 2003. Uncertainty and the management of multistate ecosystems: an apparently rational route to collapse. Ecology 84 (6), 1403–1411.
Pickett, S.T., Kolasa, J., Jones, C.G., 1994. Ecological Understanding. Academic Press, Toronto.
Pickett, S.T.A., 1999. The culture of synthesis: habits of mind in novel ecological integration. Oikos 87 (3), 479–487.
Pihl, L., Baden, S.P., Diaz, R.J., 1991. Effects of periodic hypoxia on distribution of demersal fish and crustaceans. Mar. Biol. 108, 349–360.
Rasmussen, S., Baas, N.A., Mayer, B., Nilsson, M., Olesen, M.W., 2001. Ansatz for dynamical hierarchies. Artif. Life 7 (4), 329–353.
Refsgaard, J.C., Storm, B., 1996. Construction, calibration and validation of hydrological models. In: Abbott, M.B., Refsgaard, J.C. (Eds.), Distributed Hydrological Modelling. Kluwer, Dordrecht, pp. 41–54.
Reuter, H., Hölker, F., Middelhoff, U., Jopp, F., Eschenbach, C., Breckling, B., 2005. The concepts of emergent and collective properties in individual-based models-summary and outlook of the Bornhöved case studies. Ecol. Model. 186, 489–501.
Reynolds, J.H., Ford, E.D., 1999. Multi-criteria assessment of ecological process models. Ecology 80 (2), 538–553.
Rowe, J.S., 1961. The level-of-integration concept and ecology. Ecology 42 (2), 420–427.
Rugolo, L.J., Knotts, K.S., Lange, A.M., Crecco, V.A., 1998. Stock assessment of Chesapeake Bay blue crab (*Callinectes sapidus* Rathbun). J. Shellfish Res. 17 (3), 906–930.
Rumbaugh, J., Jacobson, I., Booch, G., 1999. The Unified Modeling Language Reference Manual. Addison-Wesley, Reading, MA.
Ryer, C.H., van Montfrans, J., Orth, R.J., 1990. Utilization of a seagrass meadow and tidal marsh creek by blue crabs *Callinectes sapidus*. II: Spatial and temporal patterns of molting. Bull. Mar. Sci. 46 (1), 95–104.
Rykiel Jr., E.J., 1996. Testing ecological models: the meaning of validation. Ecol. Model. 90, 229–244.
Schnoor, J., 1996. Environmental Modeling: Fate and Transport of Polutants in Water, Air, and Soil. Wiley, New York.
Shrader-Frechette, K., 1986. Organismic biology and ecosystems ecology: description or explanation. In: Rescher, N. (Ed.), Current Issues in Teleology. University Press of America, Lanham, pp. 77–92.
Simberloff, D., 1982. A succession of paradigms in ecology: essentialism to materialism and probabilism. In: Saarinen, E. (Ed.), Conceptual Issues in Ecology. Reidel, Dordrecht, pp. 63–99.
Sommerville, I., 2001. Software Engineering, sixth ed. Addison-Wesley.
Strong, D.R., 1986. Population theory and understanding pest outbreaks. In: Kogan, M. (Ed.), Ecological Theory and Integrated Pest Management Practice. Wiley, Toronto, pp. 37–58 (Chapter 3).
van den Avyle, M.J., 1984. Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (South Atlantic)-blue crab. U.S. Fish and Wildlife Service, FWS/OBS-82/11. U.S. Army Corps of Engineers, TR EL-82-4.
van Engle, W.A., 1958. The blue crab and its fishery in Chesapeake Bay. Commerical Fish. Rev. 20 (6), 6–17.
van Montfrans, J., Epifanio, C.E., Knott, D.M., Lipcius, R.N., Mense, D.J., Metcalf, K.S., Olmi III, E.J., Orth, R.J., Posey, M.H., Wenner, E.L., West, T.L., 1995. Settlement of blue crab postlavae in western North Atlantic estuaries. Bull. Mar. Sci. 57 (3), 834–854.
Van Nes, E.H., Scheffer, M., 2005. A strategy to improve the contribution of complex simulation models to ecological theory. Ecol. Model. 185, 153–164.
Walters, C., Maguire, J.-J., 1996. Lessons for stock assessment from the northern cod collapse. Rev. Fish Biol. Fish. 6, 125–137.
Wang, Y.H., Gutierrez, A.P., 1980. An assessment of the use of stability analyses in population ecology. J. Anim. Ecol. 49 (2), 435–452.
Wiegand, T., Jeltsch, F., Hanski, I., Grimm, V., 2003. Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. Oikos 100, 209–222.
Wiegand, T., Revilla, E., Knauer, F., 2004. Dealing with uncertainty in spatially explicit population models. Biodivers. Conserv. 13, 53–78.
Winsberg, E., 2003. Simulated experiments: methodology for a virtual world. Philos. Sci. 70, 105–125.
Wolcott, T.G., Hines, A.H., 1990. Ultrasonic telemetry of small-scale movements and microhabitate selection by molting blue crabs (*Callinectes sapidus*). Bull. Mar. Sci. 46 (1), 83–94.
Yilmaz, L., Balci, O., 1997. Object-oriented simulation model verification and validation. In: Obaidat, M.S., Illgen, J. (Eds.), The Proceedings of the 1997 Summer Computer Simulation Conference, The Society for Computer Simulation International, Arlington, Virginia, pp. 835–840.
Zeigler, B.P., Praehoffer, H., Kim, T.G., 2000. Theory of Modeling and Simulation, second ed. Academic Press, New York.