



Getz, W. M., Marshall, C. R., Carlson, C. J., Giuggioli, L., Ryan, S. J., Romañach, S. S., ... O'Sullivan, D. (2018). Making ecological models adequate. *Ecology Letters*, 21(2), 153-166. <https://doi.org/10.1111/ele.12893>

Peer reviewed version

Link to published version (if available):  
[10.1111/ele.12893](https://doi.org/10.1111/ele.12893)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12893> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms>

# Making ecological models adequate

Wayne M. Getz<sup>1,7</sup>, Charles R. Marshall<sup>2</sup>, Colin J. Carlson<sup>1</sup>, Luca Giuggioli<sup>3</sup>,  
Sadie J. Ryan<sup>4,7</sup>, Stephanie S. Romañach<sup>5</sup>, Carl Boettiger<sup>1</sup>, Samuel D. Chamberlain<sup>1</sup>,  
Laurel Larsen<sup>6</sup>, Paolo D'Odorico<sup>1</sup>, David O'Sullivan<sup>6</sup>

Confidential October 9, 2017

<sup>1</sup> Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720

<sup>2</sup> Museum of Paleontology and Department Integrative Biology, University of California, Berkeley, CA 94720

<sup>3</sup> Bristol Centre for Complexity Sciences, Department of Engineering Mathematics, and School of Biological Sciences, University of Bristol, Bristol, UK

<sup>4</sup> Department of Geography, and Emerging Pathogens Institute, University of Florida, Gainesville, FL 32601

<sup>5</sup> Wetland and Aquatic Research Center, U.S. Geological Survey, Fort Lauderdale FL 33314

<sup>6</sup> Department of Geography, University of California, Berkeley, CA 94720

<sup>7</sup> Schools of Mathematical Sciences (Getz) and Life Sciences (Ryan) University of KwaZulu, Natal, South Africa

**Key Words:** Appropriate Complexity Modeling, Environmental management models, Hierarchical modeling, Coarse graining, Ecosystems restoration models, Extinction risk assessment, Disease modeling

**Authorship Statement** All authors participated in the production of this manuscript, beginning with their participation in a workshop held at the University of California, Berkeley, in February 2017. The workshop was organized by WMG with input from CRM and LL. All authors contributed ideas and text during the workshop, and helped with the revision and editing of several drafts, coordinated by WMG with the help of CRM. WMG created Figures 1 and 2, while Figure 3 was a group effort. All authors agreed to its submission to Ecology Letters.

**Data Accessibility:** Not applicable

**Manuscript Statistics**

Abstract: 192 words

Main text: 5925 words approximately (in latex rendered pdf file)

References: 145

Figures: 3

Tables: 0

Boxes: 2

## **Abstract**

Critical evaluation of the adequacy of ecological models is urgently needed to enhance their utility in developing theory and enabling environmental managers and policymakers to make informed decisions. Poorly supported management can have detrimental, costly or irreversible impacts on the environment and society. Here we examine common issues in ecological modeling and suggest criteria for improving modeling frameworks. An appropriate level of process description is crucial to constructing the best possible model, given the available data and understanding of ecological structures. Model details unsupported by data typically lead to over parameterization and poor model performance. Conversely, a lack of mechanistic details may limit a model's ability to predict ecological systems' responses to management. Ecological studies that employ models should follow a set of model adequacy assessment protocols that include: asking a series of critical questions regarding state and control variable selection, the determinacy of data, and that sensitivity and validity of analyses. We also need to improve model elaboration, refinement, and coarse graining procedures to better understand the relevancy and adequacy of our models and the role they play in advancing theory, improving hind and forecasting, and enabling problem solving and management.

# Introduction

Environmental management relies on both data and models, even when the data are few and the models simple. In situations requiring rapid response, decisions typically need to be made in the face of inadequate data and models. Even when not pressed for time, we still need to better understand the balance between collecting more data and formulating better, typically more sophisticated, models. This conundrum evokes the philosophical dichotomy between the Kuhnian and Galisonian views of scientific revolutions [1]: developing new ideas or enhancing our investigative tools. Is it better to reformulate our models using novel paradigms, or at least improve existing paradigms—or should we concentrate on obtaining higher quality data? What is the optimal trade-off between improving models versus collecting more data? If we chose model improvement, how do we do this? How do we determine how complex to make our model (see Larsen et al. [2] for a suggested protocol)? We feel that as a community we spend too little time addressing these questions. Here we call upon the ecological modeling community to pay more attention to the trade-off between model refinement and new data collection, and to assessing the adequacy of our models and data. These are daunting tasks. While the model-data-adequacy-evaluation road can be a multi-decade activity, requiring considerable commitment, it is worthy because the stakes are high. If we get things right, we may be able to save species, control pandemics, and effectively mitigate ecological disasters. If we get things wrong, disasters are likely to ensue.

## Ecological Models

Ecological models—whether formulated at the population, community, or ecosystems levels, or focusing purely on biomass, nutrient or chemical flows—are typically highly caricatured representations designed to address specific questions or consider narrowly defined, and thus tractable, viewpoints. These questions and viewpoints may concern understanding and demonstrating how particular components of ecological function, formulating new research hypotheses, or predicting the response of selected ecosystem components to a variety of driver scenarios. Depending on the questions at hand, the first task when developing a specific ecological model is to decide on the appropriate spatial, temporal, and organizational level of detail (Fig. 1) and whether the model is to be an Eulerian systems-based or a Lagrangian individual-based formulation [3, 4].

Until the 1980s, ecological modeling, at both the population and ecosystems levels, followed several relatively simple Eulerian—i.e., population systems description—threads (*inter alia* Lotka-Volterra-Gause population competition and predation models [5, 6], Kermack and McKendrick epidemiological models [7], Nicholson-Bailey host-parasite models [8], linear systems compartmental models [9], and Leslie matrix age-structured models [10]—see Fig. 2). The most basic of these models could generally be explored using

analytical methods. Over time these models became increasingly elaborate, as additional state variables and processes were included. Computational methods also began to be increasingly used to understand the behavior of these elaborated models. More complicated models, however, were not necessarily better [11,12]. Greater detail can lead to a loss of comprehension regarding a system's dynamic response to management (as discussed in [2]), or may result in the inclusion of more parameters than can be estimated from the available data. In addition, the inclusion of structures not critical to the question at hand may be counterproductive. For example, if we include age structure rather than movement-type in a model of an epidemiological system, where the most mobile individuals spread disease, then age-specific models could turn out to confuse rather than clarify our understanding if mobility is unrelated to age.

Critical environmental management or intervention decisions often rely on predictive models. A distinctive feature of the modeling activity in these cases is the urgency and impacts of the decisions they need to inform. Thus, model adequacy is sometimes evaluated on the basis of how quickly the models can be developed, modified, and used, as well as the accuracy of the forecasts they provide with respect to the critical decisions they are expected to support. In this context, we note the enormous difficulties in evaluating the accuracy of ecosystem model forecasts [13,14]. Population growth and mobility, climate change, land-use change, and resource extraction, have led to increased need for rapid decision-making. Some of these environmental decisions include control of invasive species [15,16], habitat conservation and extinction prevention of endangered species [17], harvest management (e.g., fisheries and forests, [18]), and vector control in zoonotic disease management [19]. Thus it is important that our predictive models are up to the task set for them.

## Assessing Adequacy

The ability of a model to meet given ecosystem management goals or needs depends on the appropriateness and quality of the model and the resolution of the available data. The degree to which the models and data we use are viewed as being adequate to meet the task at hand is context dependent, particularly with regard to the urgency with which decisions need to be made. For example, a moderately effective action to curb the growth of an epidemic *at its onset* may be far more valuable in terms of lives saved than a highly effective decision that extinguishes an epidemic, but is only implemented *after* the epidemic has reached its peak.

The process of evaluating model adequacy begins with an assessment of the appropriate hierarchical level of description for our model based on its spatio-temporal and organizational scale (Figure 1; also see Box 1). Identifying the characteristic scales is an essential task. Thinking that a finer-scale representation in a model is bound to lead to an improved description of a phenomenon under study is a trap that must

be avoided. Knowledge of an underlying process does not guarantee insights into the creation of larger scale patterns that depend on that process. For example, an individual's endocrine system regulates its metabolism that provides the energy necessary for its movement. Other than noting that sufficient energy must be available for individuals to draw upon while executing maneuvers, explanations of how coordinated maneuvers in flocking birds or schooling fish emerge lie elsewhere: viz., in sensory biases, cognitive abilities, and the motor response of individual animals to their external environment [20]. Seeking to explain group movement patterns through specific endocrine variables (e.g., hormone levels) is to completely miss the key processes at play, as well as the temporal, spatial and organization scales at which these processes are most appropriately described. Thus, for coordinated group movement, the key processes are cognitive and motor, and the appropriate temporal, spatial and organizations scales are sub-seconds to seconds, sub-meters to meters, and pairs of individuals to groups of neighboring individuals.

Similar considerations hold in selecting the fundamental unit of description of a phenomenon. For example, if we have a question about maintaining the diversity of biological communities, we may consider the fundamental entities in our model to be species, and hence formulate mathematical expressions describing the mechanisms involved in maintaining the integrity of multi-species communities over time and space. On the other hand, if we have a question relating to the preservation of a single species, then we may consider the fundamental entities in our model to be individuals of the species in question, and formulate mathematical expressions relating to the demography of remnant populations of that species. In the latter case we might go even further, and refine our notions of individuals by parsing them into age and sex categories to obtain a more detailed demographic model that can take into account the number of breeding females in the population, with at least one suitable male (e.g., preferably not a close relative) present to ensure successful reproduction.

The importance and visibility of ecological models has dramatically increased with the advent of rapid global change, predominantly climate and land use change. While some models may perform well (see criteria under Assessing Model Adequacy in Box 1) under a particular set of boundary conditions, they may not perform well under novel boundary conditions (e.g., temperature or precipitation levels outside ranges experienced by the modeled system). Thus, a model that might meet all the requirements laid out in Box 1 might not lead to accurate or useful forecasts of future community structure or environmental states. One way of testing the forecasting capacity of a model is to assess its accuracy via hindcasting [21]; i.e., determining the extent to which the model can account for data from past ecosystem or environmental configurations. In this context, we note that sometimes more robust forecasting can be achieved by deemphasizing the role of complex structures. This phenomenon has been demonstrated in the context of fisheries production, where models without age-structure provided better fits than age-structured models

to noise-corrupted simulation data that has been generated by the age-structured models themselves [22]. This result suggests that statistically-based, process-free analyses, as discussed below, may sometimes provide better predictions than models incorporating descriptions of the processes that were responsible for the phenomena being measured.

Adoption of statistically-based models comes with similar challenges to determining the adequacy of models. Species distribution models, for example, are used to extrapolate relationships between species occurrence and climate conditions through space and time. Depending on the methodology employed, results and therefore utility can vary widely [23]. Modelers typically view models as adequate when they are able to accurately classify species presences; however, in some management contexts, classifying absences accurately can be just as, if not more, important. Understanding the context in which these types of models will be applied is critical, because they are increasingly used in managing natural resources under the threat of rapid global change.

## Model Improvement

Models can be improved either through elaboration or refinement. Elaboration (see Box 1) involves adding new terms or making existing terms more elaborate while remaining at the same hierarchical level of description that we refer to as the model focal level (Fig. 3). Refinement involves unpacking system components at the focal level into smaller constituent parts (i.e., by adding processes at hierarchical levels below the focal level), typically followed by evaluation of the model at the original focal level using coarse graining—a process used to extract the essential features at the refined level to return, hopefully, to a more nuanced description at the focal level [24, 25] (Fig. 3). Numerous examples of model elaboration and refinement exist in the literature, several threads of which are illustrated in Fig. 3 using color themes, which have been organized in a 3D depiction of a modeling typology space spanned by *structural-/process-/utility-complexity* space (Box 2).

## A typology of ecological models

The typology of ecological models in the context of their complexity varies both within and across fields that embrace elements of ecological knowledge (from autecology through evolutionary and ecosystems ecology to socioecology and anthropogenic impacts on Earth systems) (Box 2). Rigorous measures of model complexity that can be used to organize this typology do not exist, and complexity itself can refer to different aspects of models. Here we frame model typology in terms of *structural*, *process*, and *utility* complexity (Figure 2). For our purposes, as elaborated in Box 2, *structural complexity* refers the level of detail represented by the independent (typically time and space, and sometimes trait measures, such as sex

or size) and dependent (state/component) variables used in the model. This detail is represented by the names (variables and scripting indices) and operators (difference, differential, partial differential, integral, delay), most often found on the left-hand-side, but also on the right-hand-side of the mathematical equations that describe how the variables change over time and space, and is extendible to individual- and agent-based models (IBMs and ABMs) [26–28]. *Process complexity*, as elaborated in Box 2, is represented by the level of detail contained in the functions that describe rates of change of the dependent variables over the independent variables, or other transformational changes in these variables such as jumps in value that occur at discrete values of the independent variables. *Utility complexity* pertains to the purpose of the model: whether it is to be used purely to gain understanding, purely for prediction, or both, particularly in the context of management. Philosophical treatise have been written on the role of mathematical models in epistemology [29–31]. In Box 2 we confine ourselves to a discussion of utility complexity in the context of space spanned by *model-complexity* and *standard-of-fidelity* axes, as described in [11].

## Elaboration and refinement

Model elaboration and refinement can take many forms, and comprehensive review is well beyond the scope of this paper. Here we simply give an example of each.

In our first example, we consider the prey-predator model  $dx/dt = rx - bxy$ ,  $dy/dt = -dx + axy$ , which has exponential population growth,  $dx/dt = rx$ , at its core. In this formulation the predator-prey interactions are characterized by oscillations that peculiarly depend on initial conditions (i.e. the equilibrium solution is neutrally stable so that solutions starting near this solution neither approach nor more away from it) [31]. However, if we elaborate our core growth model so that it has the logistic form,  $dx/dt = rx(1 - x/K)$ , then the resulting prey-predator model, with a prey equation of the form  $dx/dt = rx(1 - x/K) - bxy$ , has stability conditions that are now independent of initial conditions (i.e. either asymptotically stable or unstable), a much more robust situation because small perturbations to parameter values do not now change the stability characteristics of the equilibrium. Now, elaborating the logistically anchored prey-predator system by replacing the terms in  $xy$  in the model with  $xy/(1 + cx)$  leads to more realistic populations dynamics [31].

In our second example, we consider the fact that any ecosystem, with examples provided in Box 2, can undergo a process of refinement into its constituent communities, into the populations that make up those communities, into the trait groups within populations (e.g., age, sex, disease state), or even to the point of identifying unique individuals using appropriate measures (e.g., size, genetic identity, history of exposure to pathogens, hormonal state). We can refine individuals further, as illustrated in Figure 1, into organ systems, types of cells, macromolecules and so on [32]. In the context of disease modeling a further refinement involves representing individuals by a system of differential equations that describe



the dynamics of each individual’s immunological system in terms of cell types and messenger molecules (e.g., B-cells, T-cells, antibodies, cytokines, etc. [33]); or the model can be elaborated, as discussed on Box 2 to incorporate individual behavioral and stochastic population elements.

## Coarse Graining

A coarse-grained representation is a means to reduce and simplify detailed information into lower resolution models that seek to retain the key features of the system of interest. It allows one to connect a microscopic to a coarser level of description, often called mesoscale representation, and to distinguish it from the emerging larger scale patterns. The goal of coarse graining is to seek a simplified representation of the interaction among the constituent entities of a system. In the context of this paper, we see coarse graining as a necessary step to assess whether a refinement, where detail is added below the focal level, aids in model performance at the focal level (Figure 3).

Coarse graining is hardly discussed in the ecological literature, while refining system models by moving to an individual-based or agent-based approach has become quite common [34]. The concept of coarse graining emerged in the early part of last century [35] and became an integral part of the statistical mechanical theory of many-body systems [36]. Beyond its original purpose of explaining macroscopic irreversibility from underlying microscopic physical processes, coarse graining techniques and ideas are now routinely used in chemistry and biophysics [24, 37, 38].

Various tools to reduce spatio-temporal complexity in ecological models have been developed in the past and include modified mean-field equations [39], pair approximations [40], moment methods [41], and transition-scale theory [42]. These techniques rely on assumptions such as the nature of the interaction being short range, the lower moments of the distributions being sufficient to describe the process, or the process being Markovian. But when aggregating over processes that play out at different scales, these assumptions may lead to errors [43, 44], particularly when nonlinearities are involved [45]. Theoretically well-grounded techniques have been developed in physics, chemistry and biophysics, however, that do not make these assumptions. These approaches use projection operators [46–48], path-ensemble averaging [49], and the metadynamics method [50], among others [51].

Given the large number of degrees of freedom in ecological systems, it is surprising that course-graining techniques have not been employed more routinely in environmental models. This may be due, however, to the difficulty of characterizing individual-level interactions in ecological systems compared with physical systems. In the latter, the interactions between the constituent entities are derived from well-known physical laws, whereas in the former interactions are generally known only phenomenologically. Thus, in biological systems the approach is unconstrained by rigorous theory.

Despite these differences considerable opportunities exist for adapting coarse-graining techniques to

ecological problems. The principles underlining coarse-grained approximations are, in fact, already being utilized in ecological models even without the sophisticated computational and theoretical machineries routinely employed in the study of macromolecular or chemical processes [24, 37, 38]. The crux of these techniques consists of reducing the number of degrees of freedom of the original system by incorporating the average effect of the eliminated degrees of freedom into the lower-resolution model. In doing so, although temporal and spatial non-local dependence, not present explicitly in the finer description, may appear in the coarse-grained model [52], the theoretical output can be rendered at scales comparable or matching the resolution of the observations.

As a concrete example we consider a specific coarse-graining approach, namely the smoothing of the dynamics by describing an ecological system at a mesoscale level of description as a way to bridge the gap between micro and macro scales. The dynamics in question is that of territorial pattern formation in scent-marking animals, which has seen a resurgence of modeling interest since the early mathematical models of the 1990s [53, 54]. Recasting the phenomenon as a collective movement process [55, 56] has been the key to advancing our understanding of the link between animal avoidance interaction and the emergent segregation patterns of the entire population. However, establishing the relationship between the individual-level description and the segregation patterns has only been possible with a coarse-grained representation, where active scented territories are viewed as the fundamental unit of interaction [57]. With this mesoscale approach one deals with a distribution of territories, each represented by a disk of varying radius, rather than with the distribution of the animals per se. These territories, defined by the locations of active scent, are thus much bigger than the individuals generating them. This “extended particle” mesoscale description of the territories allows territorial patterns to be classified as being in one of three discrete states—formally analogous to solid, liquid and gaseous states of matter—and to characterize the spatial arrangement in terms of the likelihood of finding a neighboring territory as a function of an effective interaction. This procedure allows one to link the dynamics of the animals to that of the boundaries and goes beyond earlier phenomenological descriptions of their dynamics [58].

## Deemphasizing Model Structure

Above we have implicitly assumed that our models are process based, where elaboration, refinement, and coarse graining all represent different ways of improving model performance. However, sometimes the mathematical specification of a model can lead implicitly to the adoption of stronger assumptions than intended: for instance, theoretical considerations and empirical data may indicate that a key variable increases at an unspecified rate (e.g., population growth might increase as an unknown function of the current density, or the market value of a resource as an unknown function of the supply), while the model might rely on a particular functional form for this relationship (such as a logistic equation). Two ways

of dealing with this over-specification are to run the model with all possible parametric formulations consistent with the general relationship, or to simply work with a lower-level description. However, the former approach is typically intractable, while the latter may be simply too blunt a tool to be useful.

Non-parametric models, as have been used for example to study factors influencing plant reproduction rates [59], can often provide an alternative approach, allowing a researcher to statistically estimate or simulate data from a general class of models, such as a Gaussian Process (in the case of continuous data, such as spatial distributions, [60]) or Dirichlet Process (for discrete data, such as in the genetic sequence substitution models, [61]). For instance, by appropriate choice of the kernel function, a Gaussian Process can be used to indicate that a model has as a positive first and second derivative without assuming any particular parameterization for such a function.

Like coarse-grained models, non-parametric descriptions provide a way to simplify complex phenomena and focus on the most salient features. While non-parametric descriptions do not require the detailed mechanistic descriptions used in coarse graining, the price paid for this convenience is a hearty appetite for data—without a wealth of data to constrain them, these approaches are unable to transcend our initial uncertainty about the processes involved. Additionally, it is better to have a non-parametric model that honestly reflects this uncertainty when a mechanistically-derived parametric model is not available rather than a parametric model that consequently underestimates the level of uncertainty in the process. For example, such non-parametric models have been shown to allow fisheries managers to avoid tipping points, where model choice based on information criteria in parametric models would crash the fish stock [62].

The process-based approach can, in fact, be by-passed altogether. Machine learning methods (e.g., artificial neural networks or ANNs for short, support vector machines, regression trees), for example, provide an alternative statistical approach to prediction when a process-level understanding may not be available, or desirable [63]. These techniques have been applied frequently within the biogeosciences to spatially upscale carbon, water, and energy fluxes to continental [64–67] and global scales [68–71], primarily using a combination of remotely-sensed and ground-based climate data. These upscaled products are also useful for evaluating flux outputs from Earth system models where ground-based measurements are sparse [72]. Machine learning techniques demonstrate strong predictive capabilities across time and space [68]; however, this comes at the cost of large data requirements and, of course, does not typically lead to advances in process-level understanding (but see [63]).

Despite potentially limited process-level insights, ensembles of ANNs have been used to identify dominant drivers of changes in greenhouse gas fluxes from eddy covariance time series [73], where data constraints are not a concern. Because ANNs can identify dominant drivers from data, they may be useful as an inductive approach when elaborating model processes as described in Fig. 3, similar to information theory-based approaches (e.g., mutual information and transfer entropy) outlined in Larsen et al. [2]

where Shannon entropy-based metrics are used to identify key state variables and interactions [74–76].

## Modeling Under Time Pressure

The challenge of assessing model and data adequacy is particularly taxing when time is limited. Below we provide examples to illustrate the practical challenges when swift action is required in the contexts of ecosystem restoration, extinction risk assessment, or zoonotic disease containment.

### Ecosystem restoration

A series of ecosystem restoration planning exercises are being undertaken for the Florida Everglades. Restoration is urgent as sea-levels rise and the populations of critically endangered species fall. Reversing decades of human-caused damage in the Everglades will cost of billions of dollars, so the application of models is well-justified [77]. Below we outline the challenges that managers face when trying to develop restoration measures; but, unfortunately, we have no silver bullet for maximizing performance when under intense time pressure.

To plan for restoration on this scale, decision makers must integrate ecological information, such as output from ecological models, with social and economic considerations such as flood control for adjacent residential areas. Although ecological modeling is only one piece in a much larger puzzle, it is important to get it right given that the aim of Everglades restoration is to attend equally to human and wildlife needs. To assess the appropriate model focal level (Fig. 3) decision makers work with stakeholders to determine focal species and targeted habitat features [78]. This in turn provides a guide to the state and control variables needed to ensure the ecological models developed are adequate for the task of forecasting the potential impacts of proposed restoration plans on focal species and habitat features (cf. Box 1). Restoration planning exercises typically proceed on accelerated time lines, leaving modelers with limited time to assess the adequacy of models during each phase of restoration. Further, modelers are often asked to apply current models in new ways, such as expanding the geographic scope or including new habitats, with limited time allowed for model validation before the implementation of conservation measures needed to rescue species on the verge of extinction (e.g., the federally endangered Cape Sable Seaside Sparrow). If current models are not appropriately elaborated and then validated, mistakes arising from model forecasts can be extremely costly in terms of dollars and lost species.

Hydrological processes are key drivers of the Everglades ecosystem, and thus of primary concern when constructing or refining appropriate models [79]. In rapid planning exercises, some modelers have employed niche or species distribution modeling to examine how well hydrologic measures explain species occurrence [80]. These modelers must determine whether hydrologic metrics alone can serve to describe

species distributions. If not, given that collecting more data is an untenable option when decisions must be made quickly, then model elaboration or refinement becomes the primary option (Fig. 3), particularly through adding additional predictor variables. For each proposed restoration plan, hydrologic modelers develop forecasting models of expected hydrologic patterns for the ecosystem; however, forecasts of other variables (e.g., land cover) may not be available [81]. Given the long time span of restoration implementation, modelers are constantly switching between restoration planning exercises (to develop new models and methods) and forecasting exercises to help better inform subsequent planning [77].

## **Extinction risk assessment**

In the field of extinction-risk assessment, models are typically used to predict and validate estimates of extinction magnitudes, especially under various climate change or human land use change scenarios. For example, early modeling work at the global scale applied the species area relationship (SAR) to predict extinction risk based on habitat loss from climate change [82]. While the application of the SAR to the sixth mass extinction was an important development in global change biology, the method itself has come under significant criticism, with increasingly complex alternatives being developed to keep pace with macroecological theory (e.g., [83, 84]). Models based on the SAR have become more accurate, particularly as they accommodate the fact that macroecological patterns change across scales (e.g., the asymptotically declining slope of the SAR with increasing geographic scale [85]) and incorporate more complex, synergistic drivers of extinction such as habitat fragmentation [86]. To make models more accurate across spatial scales, more sophisticated data than basic range loss will be required. For terrestrial vertebrates and some plant groups, gathering additional data is likely to be much easier than for the majority of other species on Earth. For example, rough estimates for all terrestrial invertebrates suggest that 98% of extinctions on Earth are currently going unrecorded [87]. To make the modeling process even more challenging, SAR and range-based models are typically based on average habitat and climate conditions, so they will need to incorporate the extinction pressures generated by predicted increases in the magnitude of weather extremes.

The role human socio-demographic change in extinction risk assessment [88, 89] includes the impacts of human population growth, increasing wealth disparity, and land use and land cover change, including habitat fragmentation. The latter comprises land conversion for food production, urbanization and sprawl processes, as well as through the creation of protected areas for the restoration, recovery, and protection of ecosystems. Changes to the social component of social-ecological systems have a large, and often hard to quantify, impact on extinction risk.

## Zoonotic disease containment

Although modeling disease outbreaks stretches back more than a century [7], only recently have ecologists begun to incorporate significant ecological and environmental components into modeling the spread of disease, particularly zoonotic outbreaks. The development of modeling and research frameworks like One Health [90], EcoHealth [91], eco-epidemiology [92], etc., have bridged the divide and significantly increased collaboration among wildlife and livestock ecologists, veterinarians, clinicians, and the broader public health community. As new participants in pandemic preparedness and response, ecologists are frequently charged with the development of models that appropriately forecast (or retrospectively explain) the environmental drivers of zoonotic spillover, vector-borne disease emergence and spread, the rise of antibiotic resistance outside human populations, and the emergence of entirely new diseases [93]. Ecological modelers face some of the most severe operational challenges during the early days of a pandemic: information is generally severely limited, public and agency pressure to produce forecasts is intense and time-sensitive, and the stakes (measurable in human morbidity and mortality) are unusually high by ecological standards.

When zoonotic disease establishes itself in a human population, it then comes under the purview of epidemiologists and other health care practitioners. These zoonoses typically still maintain critical non-human ecological components (e.g., ebola viral disease), which require ecologists to consider when forecasting pandemic spread, and estimating the final boundaries of endemic transmission. The 2015 pandemic of mosquito (*Aedes*)-borne Zika virus in the Americas particularly highlights the challenges of modeling in this context. During the early days of the epidemic, the majority of predictions were made based on the dynamics of other *Aedes*-transmitted diseases, in particular two other flaviviruses (Dengue and Chikungunya) [94]. Forecasts based on this knowledge of *Aedes*-borne pathogens suggested a high potential for Zika to spread rapidly in North America [95], causing widespread alarm and preemptive vector control efforts (with unintended negative consequences for bees) especially in the southeast United States. Continued uncertainty about the relative risk of autochthonous Zika transmission poses an ongoing challenge for policymakers facing the task of allocating funds and resources for Zika prevention, including in states that have yet to record local transmission events. Of the 5,158 cases of Zika the Centers for Disease Control (CDC) has reported so far in the United States, only 222 cases appear to have been local to the U.S. so far (216 in Florida, and 6 in Texas), but widespread alarm and calls to prepare for more severe outbreaks persist.

The products that managers of disease emergence need from modelers tend to take the form of assessments of the impact available interventions will have. Rapidly constructed predictive disease models, however, may not include appropriate control variables (Box 1). For example, in the case of vector-borne diseases such as Zika, the only point of intervention may be vector control, so it is critical that this aspect

be included in models for evaluating intervention strategies. Further, models that produce numerical estimates of people affected at the end of an epidemic may be useful tools for agencies to anticipate palliative care needs, or to re-allocate funding to vector control, but they may miss the mark in pointing to regions where intervention needs to occur when spatial variation predominates. Conversely, top-down phenomenological distributional models can add geographic resolution using only limited data, but they lack the ability to predict case burdens, and can be overly sensitive to small differences in the model [96]. Tools to integrate these methods are lacking, as are tools to resolve consensus between conflicting forecasts. Further, the adequacy of ecological forecasts must match the specific challenges policymakers face on the short, high-pressure timescale of pandemic preparedness. In short, disease modelers may need to pay more attention to both the data adequacy assessment loop and model elaboration and refinement loop in Fig. 3 to better deal with crises such as the recent Ebola and Zika outbreaks.

## Discussion

The central issue we address here is where on the spectrum of model complexity should a particular model be to best meet the purpose for which that model was built. At one end of this spectrum are models built to address “super wicked problems” [97, 98] of how to rein in the impacts of climate and land-use change on ecosystems. At the other end of this spectrum are models built, following the principle of Occam’s razor, to obtain maximum explanatory power per free parameter [11, 12]. In addition, because the human mind is limited in its ability to grasp the full complexity of many problems [12], understanding can be greatly enhanced by partitioning complex problems into smaller, well-bounded, parts, followed by assembly of the resulting pieces into a larger more coherent total. Ignoring the typically ad-hoc assembly process, defining each of the parts still represents a considerable challenge. For example, although we can use relatively simple harvesting models to shed light on complex Ecopath with Ecosim multispecies fisheries management models (see Fig. 2 and Box 2), it is still difficult to devise management strategies under stable ocean conditions [99], let alone under various global change scenarios. Similarly, although we can use relatively simple epidemiological models to understand the basic dynamics of epidemics [100], all zoonotic outbreaks have important stochastic and individual movement components [101] that are often linked to environmental phenomena such as the El-Niño Southern Oscillation [102]. These complications still represent a considerable challenge to modelers and managers [103]. Thus, for example, in the recent Ebola outbreak in West Africa, a lack of knowledge regarding the role of increased urbanization, changes in population size, and movement of individuals thwarted efforts to predict the course of the epidemic with the needed confidence [101].

Our primary goal here has been to draw attention to the general need for evaluation of the adequacy

of the models, regardless of whether they are being used to generate knowledge or manage systems. In addition, it is important to bear in mind that high fidelity models (Box 2) are not always superior to low fidelity models. High fidelity models, for example, will be inadequate if they omit the instruments (i.e., control variables) needed to address intended management questions (Box 1), while low fidelity models may have considerable strategic value when comparing alternative management options, even if the predictive capabilities of the models are low; e.g., when comparing various land acquisition options for conserving biological diversity. Thus, for example, simple models of complex systems have proven to be important tools for formulating policy when implemented in adaptive management settings [104].

In the future, with an increasing availability of data, the question of how best to refine our models will take on additional relevancy. Model refinement is a challenging process, as some of us have described in the implementation of an Appropriate Complexity Method (ACME) in modeling floodplain structure [2], where the environmental elements remain relatively simple compared to such ecosystem elements as the behavior of organisms. In situations where we are hampered by a lack of data required to support needed refinement, we can still gain considerable power by asking simpler questions that can be answered with the data at hand or by data that can be obtained quickly enough to enable the generation of useful forecasts.

In applied situations when rapid action is needed, for example, on how to manage a pandemic outbreak, models may be inadequate because key drivers are missing, or their importance under appreciated. In some cases a model might already be known to be inadequate prior to the crisis, but in other cases the inadequacy might only become apparent in retrospect, as was the case with the recent Ebola outbreak. In the future, we need to be better prepared to meet such crises. Part of this preparation includes having improved procedures for evaluating the adequacy of our models, including those for the rational management of our natural resources, the prudent management of endangered species, and the outbreak containment of pandemic diseases.



## Acknowledgements

This paper is the product of a DIMACS Mathematics of Planet Earth Program workshop on Appropriate Complexity Modeling of the Impacts of Global Change on Ecosystems, supported under National Science Foundation grant 1246305 to Rutgers University. LG acknowledges support of EPSRC grant EP/I013717/1 and SJR of NSF DEB EEID 1518681 and NSF DEB RAPID 1641145. SSR was supported by the U.S. Geological Survey's Greater Everglades Priority Ecosystem Science. Use of trade, product, or firm names in this paper does not imply endorsement by the US Government.

## References

- [1] Dyson F.J. 2012 Is science mostly driven by ideas or by tools? Science **338**, 1426–1427.
- [2] Larsen L.G., Eppinga M.B., Passalacqua P., Getz W.M., Rose K.A., Liang M. 2016 Appropriate complexity landscape modeling. Earth-Science Reviews **160**, 111–130.
- [3] Vincenot C.E., Giannino F., Rietkerk M., Moriya K., Mazzoleni S. 2011 Theoretical considerations on the combined use of system dynamics and individual-based modeling in ecology. Ecological Modelling **222**, 210–218.
- [4] Grimm V., Railsback S.F. 2005 Individual-based modeling and ecology, volume 2005. BioOne.
- [5] Kingsland S. 1982 The refractory model: The logistic curve and the history of population ecology. The Quarterly Review of Biology **57**, 29–52.
- [6] Wangersky P.J. 1978 Lotka-volterra population models. Annual Review of Ecology and Systematics **9**, 189–218.
- [7] Hethcote H.W. 2000 The mathematics of infectious diseases. SIAM review **42**, 599–653.
- [8] Hassell M. 2000 Host–parasitoid population dynamics. Journal of Animal Ecology **69**, 543–566.
- [9] Mulholland R.J., Keener M.S. 1974 Analysis of linear compartment models for ecosystems. Journal of theoretical biology **44**, 105–116.
- [10] Caswell H. 2001 Matrix population models. Wiley Online Library.
- [11] Evans M.R., Grimm V., Johst K., Knuuttila T., De Langhe R., Lessells C.M., Merz M., O’Malley M.A., Orzack S.H., Weisberg M., et al. 2013 Do simple models lead to generality in ecology? Trends in ecology & evolution **28**, 578–583.
- [12] Feldman J. 2003 The simplicity principle in human concept learning. Current Directions in Psychological Science **12**, 227–232.
- [13] Evans M.R. 2012 Modelling ecological systems in a changing world. Phil. Trans. R. Soc. B **367**, 181–190.
- [14] Augusiak J., Van den Brink P.J., Grimm V. 2014 Merging validation and evaluation of ecological models to ‘evaludation’: a review of terminology and a practical approach. Ecological Modelling **280**, 117–128.
- [15] Pyšek P., Richardson D.M. 2010 Invasive species, environmental change and management, and health. Annual Review of Environment and Resources **35**, 25–55.

- [16] Epanchin-Niell R.S., Hufford M.B., Aslan C.E., Sexton J.P., Port J.D., Waring T.M. 2010 Controlling invasive species in complex social landscapes. Frontiers in Ecology and the Environment **8**, 210–216.
- [17] Schmolke A., Thorbek P., DeAngelis D.L., Grimm V. 2010 Ecological models supporting environmental decision making: a strategy for the future. Trends in ecology & evolution **25**, 479–486.
- [18] Getz W.M., Haight R.G. 1989 Population harvesting: demographic models of fish, forest, and animal resources, volume 27. princeton university press.
- [19] Kilpatrick A.M., Randolph S.E. 2012 Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. The Lancet **380**, 1946–1955.
- [20] Giuggioli L., McKetterick T.J., Holderied M. 2015 Delayed response and biosonar perception explain movement coordination in trawling bats. PLoS computational biology **11**, e1004089.
- [21] Davis E.B., McGuire J.L., Orcutt J.D. 2014 Ecological niche models of mammalian glacial refugia show consistent bias. Ecography **37**, 1133–1138.
- [22] Ludwig D., Walters C.J. 1985 Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences **42**, 1066–1072.
- [23] Watling J.I., Bucklin D.N., Speroterra C., Brandt L.A., Mazzotti F.J., Romañach S.S. 2013 Validating predictions from climate envelope models. PloS one **8**, e63600.
- [24] Binder K., Paul W., Santos S., Suter U. 2004 Coarse-graining techniques. In Simulation Methods for Polymers, (eds. M. Kotelyanskii, D.N. Theodorou), pp. 491–510. M. Dekker, New York.
- [25] Coveney P.V., Fowler P.W. 2005 Modelling biological complexity: a physical scientist’s perspective. Journal of The Royal Society Interface **2**, 267–280.
- [26] Getz W.M. 2013 Computational population biology: linking the inner and outer worlds of organisms. Israel Journal of Ecology & Evolution **59**, 2–16.
- [27] Macy M.W., Willer R. 2002 From factors to factors: computational sociology and agent-based modeling. Annual review of sociology **28**, 143–166.
- [28] Grimm V., Revilla E., Berger U., Jeltsch F., Mooij W.M., Railsback S.F., Thulke H.H., Weiner J., Wiegand T., DeAngelis D.L. 2005 Pattern-oriented modeling of agent-based complex systems: lessons from ecology. science **310**, 987–991.
- [29] Frigg R., Hartmann S. 2017 Models in science. In The Stanford Encyclopedia of Philosophy, (ed. E.N. Zalta). Metaphysics Research Lab, Stanford University, spring 2017 edition.

- [30] Rodgers J.L. 2010 The epistemology of mathematical and statistical modeling: a quiet methodological revolution. American Psychologist **65**, 1.
- [31] Getz W.M. 1998 An introspection on the art of modeling in population ecology. BioScience **48**, 540–552.
- [32] O’Neill R.V. 1986 A hierarchical concept of ecosystems, volume 23. Princeton University Press.
- [33] Abbas A.K., Lichtman A.H., Pillai S. 2014 Basic immunology: functions and disorders of the immune system. Elsevier Health Sciences.
- [34] Grimm V., Berger U., Bastiansen F., Eliassen S., Ginot V., Giske J., Goss-Custard J., Grand T., Heinz S.K., Huse G., et al. 2006 A standard protocol for describing individual-based and agent-based models. Ecological modelling **198**, 115–126.
- [35] Casimir H.B.G. 1945 On onsager’s principle of microscopic reversibility. Reviews of Modern Physics **17**, 343.
- [36] Landau L.D., Lifshitz E.M. 1980 Statistical physics, vol. 5. Elsevier, 3rd edition.
- [37] Kamerlin S.C., Vicatos S., Dryga A., Warshel A. 2011 Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. Annual review of physical chemistry **62**, 41–64.
- [38] Saunders M.G., Voth G.A. 2013 Coarse-graining methods for computational biology. Annual review of biophysics **42**, 73–93.
- [39] Pascual M., Roy M., Laneri K. 2011 Simple models for complex systems: exploiting the relationship between local and global densities. Theoretical Ecology **4**, 211–222.
- [40] Sato K., Iwasa Y. 2000 Pair approximations for lattice-based ecological models. In The geometry of ecological interactions: simplifying spatial complexity, (eds. U. Dieckmann, R.Law, J. Metz), pp. 341–358. Cambridge University Press.
- [41] Bolker B.M., Pacala S.W., Levin S.A. 2000 Moment methods for ecological processes in continuous space. In The geometry of ecological interactions: simplifying spatial complexity, (eds. U. Dieckmann, R.Law, J. Metz), pp. 388–411. Cambridge University Press.
- [42] Chesson P. 2012 Scale transition theory: its aims, motivations and predictions. Ecological Complexity **10**, 52–68.
- [43] Gardner R., Cale W., O’Neill R. 1982 Robust analysis of aggregation error. Ecology **63**, 1771–1779.

- [44] Rastetter E.B., King A.W., Cosby B.J., Hornberger G.M., O’Neill R.V., Hobbie J.E. 1992 Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecological Applications **2**, 55–70.
- [45] Iwasa Y., Levin S.A., Andreasen V. 1989 Aggregation in model ecosystems ii. approximate aggregation. Mathematical Medicine and Biology: A Journal of the IMA **6**, 1–23.
- [46] Mori H. 1965 Transport, collective motion, and brownian motion. Progress of theoretical physics **33**, 423–455.
- [47] Kenkre V. 1977 The generalized master equation and its applications. In Statistical Mechanics and statistical methods in theory and application, (ed. U. Landman), pp. 441–461. Plenum, New York.
- [48] Grabert H. 1982 Projection operator techniques in nonequilibrium statistical mechanics. Springer.
- [49] Calderon C.P., Janosi L., Kosztin I. 2009 Using stochastic models calibrated from nanosecond nonequilibrium simulations to approximate mesoscale information. The Journal of chemical physics **130**, 144908.
- [50] Ensing B., De Vivo M., Liu Z., Moore P., Klein M.L. 2006 Metadynamics as a tool for exploring free energy landscapes of chemical reactions. Accounts of chemical research **39**, 73–81.
- [51] Peter C., Kremer K. 2009 Multiscale simulation of soft matter systems—from the atomistic to the coarse-grained level and back. Soft Matter **5**, 4357–4366.
- [52] Giuggioli L., Sevilla F.J., Kenkre V. 2009 A generalized master equation approach to modelling anomalous transport in animal movement. Journal of Physics A: Mathematical and Theoretical **42**, 434004.
- [53] Lewis M., Murray J. 1993 Modelling territoriality and wolf-deer interactions. Nature **366**, 738–740.
- [54] Moorcroft P., Lewis M. 2006 Mechanistic Home Range Analysis. Monographs in Population Biology. Princeton University Press.
- [55] Giuggioli L., Potts J.R., Harris S. 2011 Animal interactions and the emergence of territoriality. PLoS computational biology **7**, e1002008.
- [56] Giuggioli L., Potts J.R., Rubenstein D.I., Levin S.A. 2013 Stigmergy, collective actions, and animal social spacing. Proceedings of the National Academy of Sciences **110**, 16904–16909.
- [57] Heiblum Robles A., Giuggioli L. 2017 Phase transitions in stigmergic territorial systems. preprint .

- [58] Potts J.R., Harris S., Giuggioli L. 2013 Quantifying behavioral changes in territorial animals caused by sudden population declines. The American Naturalist **182**, E73–E82.
- [59] McCune B. 2006 Non-parametric habitat models with automatic interactions. Journal of Vegetation Science **17**, 819–830.
- [60] Rasmussen C.E., Williams C.K.I. 2006 Gaussian Processes for Machine Learning. MIT Press, Boston.
- [61] Huelsenbeck J.P., Jain S., Frost S.W.D., Pond S.L.K. 2006 A Dirichlet process model for detecting positive selection in protein-coding DNA sequences. Proceedings of the National Academy of Sciences **103**, 6263–6268.
- [62] Boettiger C., Mangel M., Munch S. 2015 Avoiding tipping points in fisheries management through Gaussian process dynamic programming. Proceedings of the Royal Society B: Biological Sciences **282**, 20141631–20141631.
- [63] Breiman L., et al. 2001 Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical science **16**, 199–231.
- [64] Papale D., Black T.A., Carvalhais N., Cescatti A., Chen J., Jung M., Kiely G., Lasslop G., Mahecha M.D., Margolis H., et al. 2015 Effect of spatial sampling from european flux towers for estimating carbon and water fluxes with artificial neural networks. Journal of Geophysical Research: Biogeosciences **120**, 1941–1957.
- [65] Papale D., Valentini R. 2003 A new assessment of european forests carbon exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biology **9**, 525–535.
- [66] Xiao J., Zhuang Q., Baldocchi D.D., Law B.E., Richardson A.D., Chen J., Oren R., Starr G., Noormets A., Ma S., et al. 2008 Estimation of net ecosystem carbon exchange for the conterminous united states by combining modis and ameriflux data. Agricultural and Forest Meteorology **148**, 1827–1847.
- [67] Yang F., Ichii K., White M.A., Hashimoto H., Michaelis A.R., Votava P., Zhu A.X., Huete A., Running S.W., Nemani R.R. 2007 Developing a continental-scale measure of gross primary production by combining modis and ameriflux data through support vector machine approach. Remote Sensing of Environment **110**, 109–122.
- [68] Alemohammad S.H., Fang B., Konings A.G., Green J.K., Kolassa J., Prigent C., Aires F., Miralles D., Gentile P. 2016 Water, energy, and carbon with artificial neural networks (wecann):

A statistically-based estimate of global surface turbulent fluxes using solar-induced fluorescence. Biogeosciences Discussions .

- [69] Jung M., Reichstein M., Margolis H.A., Cescatti A., Richardson A.D., Arain M.A., Arneth A., Bernhofer C., Bonal D., Chen J., et al. 2011 Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. Journal of Geophysical Research: Biogeosciences **116**.
- [70] Jung M., Reichstein M., Bondeau A. 2009 Towards global empirical upscaling of fluxnet eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. Biogeosciences **6**, 2001–2013.
- [71] Tramontana G., Jung M., Schwalm C.R., Ichii K., Camps-Valls G., Ráduly B., Reichstein M., Arain M.A., Cescatti A., Kiely G., et al. 2016 Predicting carbon dioxide and energy fluxes across global fluxnet sites with regression algorithms. Biogeosciences **13**, 4291–4313.
- [72] Bonan G.B., Lawrence P.J., Oleson K.W., Levis S., Jung M., Reichstein M., Lawrence D.M., Swenson S.C. 2011 Improving canopy processes in the community land model version 4 (clm4) using global flux fields empirically inferred from fluxnet data. Journal of Geophysical Research: Biogeosciences **116**.
- [73] Albert L.P., Keenan T.F., Burns S.P., Huxman T.E., Monson R.K. 2017 Climate controls over ecosystem metabolism: insights from a fifteen-year inductive artificial neural network synthesis for a subalpine forest. Oecologia pp. 1–17.
- [74] Ruddell B.L., Yu R., Kang M., Childers D.L. 2016 Seasonally varied controls of climate and phenophase on terrestrial carbon dynamics: modeling eco-climate system state using dynamical process networks. Landscape ecology **31**, 165–180.
- [75] Sturtevant C., Ruddell B.L., Knox S.H., Verfaillie J., Matthes J.H., Oikawa P.Y., Baldocchi D. 2016 Identifying scale-emergent, nonlinear, asynchronous processes of wetland methane exchange. Journal of Geophysical Research: Biogeosciences .
- [76] Larsen L., Harvey J. Disrupted carbon cycling in restored and unrestored urban streams: Critical timescales and controls. Limnology and Oceanography .
- [77] National Research Council. 2008 2008 Progress toward restoring the Everglades: The second biennial review. Committee on Independent Scientific Review of Everglades Restoration Progress (CISRERP), chapter 3, p. 71–107. National Academies Press.

- [78] Romañach S.S., McKelvy M., Conzelmann C., Suir K. 2014 A visualization tool to support decision making in environmental and biological planning. Environmental Modelling & Software **62**, 221–229.
- [79] DeAngelis D.L., Gross L.J., Huston M.A., Wolff W.F., Fleming D.M., Comiskey E.J., Sylvester S.M. 1998 Landscape modeling for everglades ecosystem restoration. Ecosystems **1**, 64–75.
- [80] Beerens J.M., Noonburg E.G., Gawlik D.E. 2015 Linking dynamic habitat selection with wading bird foraging distributions across resource gradients. PloS one **10**, e0128182.
- [81] Darby P.C., DeAngelis D.L., Romañach S.S., Suir K., Bridevaux J. 2015 Modeling apple snail population dynamics on the everglades landscape. Landscape ecology **30**, 1497–1510.
- [82] Thomas C.D., Cameron A., Green R.E., Bakkenes M., Beaumont L.J., Collingham Y.C., Erasmus B.F., De Siqueira M.F., Grainger A., Hannah L., et al. 2004 Extinction risk from climate change. Nature **427**, 145–148.
- [83] Harte J., Kitzes J. 2012 The use and misuse of species-area relationships in predicting climate-driven extinction. In Saving a Million Species, pp. 73–86. Springer.
- [84] He F., Hubbell S.P. 2011 Species-area relationships always overestimate extinction rates from habitat loss. Nature **473**, 368–371.
- [85] Harte J., Smith A.B., Storch D. 2009 Biodiversity scales from plots to biomes with a universal species–area curve. Ecology letters **12**, 789–797.
- [86] Rybicki J., Hanski I. 2013 Species–area relationships and extinctions caused by habitat loss and fragmentation. Ecology letters **16**, 27–38.
- [87] Régnier C., Achaz G., Lambert A., Cowie R.H., Bouchet P., Fontaine B. 2015 Mass extinction in poorly known taxa. Proceedings of the National Academy of Sciences **112**, 7761–7766.
- [88] Brook B.W., Sodhi N.S., Bradshaw C.J. 2008 Synergies among extinction drivers under global change. Trends in ecology & evolution **23**, 453–460.
- [89] Bartlett L.J., Newbold T., Purves D.W., Tittensor D.P., Harfoot M.B. 2016 Synergistic impacts of habitat loss and fragmentation on model ecosystems. In Proc. R. Soc. B, volume 283, p. 20161027. The Royal Society.
- [90] Blackburn J.K., Krcalík I.T., Fair J.M. 2015 Applying science: Opportunities to inform disease management policy with cooperative research within a one health framework. Frontiers in public health **3**.



- [91] Wilcox B.A., Aguirre A.A., Daszak P., Horwitz P., Martens P., Parkes M., Patz J.A., Waltner-Toews D. 2004 Ecohealth: a transdisciplinary imperative for a sustainable future. EcoHealth **1**, 3–5.
- [92] Susser M., Susser E. 1996 Choosing a future for epidemiology: II. from black box to chinese boxes and eco-epidemiology. American journal of public health **86**, 674–677.
- [93] Daszak P., Cunningham A., Hyatt A. 2001 Anthropogenic environmental change and the emergence of infectious diseases in wildlife. Acta tropica **78**, 103–116.
- [94] Bogoch I.I., Brady O.J., Kraemer M., German M., Creatore M.I., Kulkarni M.A., Brownstein J.S., Mekaru S.R., Hay S.I., Groot E., et al. 2016 Anticipating the international spread of zika virus from brazil. Lancet (London, England) **387**, 335–336.
- [95] Mordecai E., Cohen J., Evans M.V., Gudapati P., Johnson L.R., Miazgowicz K., Murdock C.C., Rohr J.R., Ryan S.J., Savage V., et al. 2016 Temperature determines zika, dengue and chikungunya transmission potential in the americas. bioRxiv p. 063735.
- [96] Carlson C.J., Dougherty E.R., Getz W. 2016 An ecological assessment of the pandemic threat of zika virus. PLoS Negl Trop Dis **10**, e0004968.
- [97] Lazarus R.J. 2008 Super wicked problems and climate change: Restraining the present to liberate the future. Cornell L. Rev. **94**, 1153.
- [98] Levin K., Cashore B., Bernstein S., Auld G. 2012 Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change. Policy sciences **45**, 123–152.
- [99] Plagányi É.E., Punt A.E., Hillary R., Morello E.B., Thébaud O., Hutton T., Pillans R.D., Thorson J.T., Fulton E.A., Smith A.D., et al. 2014 Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. Fish and Fisheries **15**, 1–22.
- [100] Smith D.L., Battle K.E., Hay S.I., Barker C.M., Scott T.W., McKenzie F.E. 2012 Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. PLoS pathog **8**, e1002588.
- [101] Getz W.M., Gonzalez J.P., Salter R., Bangura J., Carlson C., Coomber M., Dougherty E., Kargbo D., Wolfe N.D., Wauquier N. 2015 Tactics and strategies for managing ebola outbreaks and the salience of immunization. Computational and mathematical methods in medicine **2015**.
- [102] Pascual M., Rodó X., Ellner S.P., Colwell R., Bouma M.J. 2000 Cholera dynamics and el nino-southern oscillation. Science **289**, 1766–1769.

- [103] Kramer A.M., Pulliam J.T., Alexander L.W., Park A.W., Rohani P., Drake J.M. 2016 Spatial spread of the west africa ebola epidemic. Royal Society Open Science **3**, 160294.
- [104] Williams B.K. 2011 Adaptive management of natural resources—framework and issues. Journal of environmental management **92**, 1346–1353.
- [105] Getz W.M., Muellerklein O.C., Salter R.M., Carlson C.J., Lyons A.J., Seidel D.P. 2017 A web app for population viability and harvesting analyses. Natural Resource Modeling **30**, e12120.
- [106] Dushoff J., Huang W., Castillo-Chavez C. 1998 Backwards bifurcations and catastrophe in simple models of fatal diseases. Journal of mathematical biology **36**, 227–248.
- [107] Ward B.A., Friedrichs M.A., Anderson T.R., Oschlies A. 2010 Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models. Journal of Marine Systems **81**, 34–43.
- [108] Lloyd-Smith J.O., Schreiber S.J., Kopp P.E., Getz W.M. 2005 Superspreading and the effect of individual variation on disease emergence. Nature **438**, 355–359.
- [109] Coyte K.Z., Schluter J., Foster K.R. 2015 The ecology of the microbiome: networks, competition, and stability. Science **350**, 663–666.
- [110] Getz W.M. 2011 Biomass transformation webs provide a unified approach to consumer–resource modelling. Ecology letters **14**, 113–124.
- [111] Getz W.M., Salter R., Lyons A.J., Sippl-Swezey N. 2015 Panmictic and clonal evolution on a single patchy resource produces polymorphic foraging guilds. PloS one **10**, e0133732.
- [112] Getz W.M., Salter R., Seidel D.P., Hooft P. 2016 Sympatric speciation in structureless environments. BMC evolutionary biology **16**, 50.
- [113] Spiegel O., Leu S.T., Bull C.M., Sih A. 2017 What’s your move? movement as a link between personality and spatial dynamics in animal populations. Ecology Letters **20**, 3–18.
- [114] Gilpin M. 2012 Metapopulation dynamics: empirical and theoretical investigations. Academic Press.
- [115] Phillips S.J., Anderson R.P., Schapire R.E. 2006 Maximum entropy modeling of species geographic distributions. Ecological modelling **190**, 231–259.
- [116] Carlson C.J., Burgio K.R., Dougherty E.R., Phillips A.J., Bueno V.M., Clements C.F., Castaldo G., Dallas T.A., Cizauskas C.A., Cumming G.S., et al. 2017 Parasite biodiversity faces extinction and redistribution in a changing climate. Science Advances **3**, e1602422.

- [117] Minasny B., McBratney A.B., Salvador-Blanes S. 2008 Quantitative models for pedogenesis—a review. Geoderma **144**, 140–157.
- [118] Ellner S.P., Rees M. 2006 Integral projection models for species with complex demography. The American Naturalist **167**, 410–428.
- [119] Rees M., Childs D.Z., Ellner S.P. 2014 Building integral projection models: a user’s guide. Journal of Animal Ecology **83**, 528–545.
- [120] Merow C., Latimer A.M., Wilson A.M., McMahon S.M., Rebelo A.G., Silander J.A. 2014 On using integral projection models to generate demographically driven predictions of species’ distributions: development and validation using sparse data. Ecography **37**, 1167–1183.
- [121] Dunne J.A., Williams R.J., Martinez N.D. 2002 Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecology letters **5**, 558–567.
- [122] Leibold M.A., Holyoak M., Mouquet N., Amarasekare P., Chase J.M., Hoopes M.F., Holt R.D., Shurin J.B., Law R., Tilman D., et al. 2004 The metacommunity concept: a framework for multi-scale community ecology. Ecology letters **7**, 601–613.
- [123] Coll M., Bundy A., Shannon L.J. 2009 Ecosystem modelling using the ecopath with ecosim approach. In Computers in fisheries research, pp. 225–291. Springer.
- [124] Colléter M., Valls A., Guitton J., Gascuel D., Pauly D., Christensen V. 2015 Global overview of the applications of the ecopath with ecosim modeling approach using the ecobase models repository. Ecological Modelling **302**, 42–53.
- [125] Heymans J.J., Coll M., Link J.S., Mackinson S., Steenbeek J., Walters C., Christensen V. 2016 Best practice in ecopath with ecosim food-web models for ecosystem-based management. Ecological modelling **331**, 173–184.
- [126] Getz W.M. 1993 Metaphysiological and evolutionary dynamics of populations exploiting constant and interactive resources: R—k selection revisited. Evolutionary Ecology **7**, 287–305.
- [127] Gillespie D.T. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of computational physics **22**, 403–434.
- [128] Gillespie D.T. 1977 Exact stochastic simulation of coupled chemical reactions. The journal of physical chemistry **81**, 2340–2361.
- [129] Vestergaard C.L., Géniois M. 2015 Temporal gillespie algorithm: Fast simulation of contagion processes on time-varying networks. PLoS computational biology **11**, e1004579.

- [130] Levin S.A. 1992 The problem of pattern and scale in ecology: the robert h. macarthur award lecture. Ecology **73**, 1943–1967.
- [131] Slingo J., Bates K., Nikiforakis N., Piggott M., Roberts M., Shaffrey L., Stevens I., Vidale P.L., Weller H. 2009 Developing the next-generation climate system models: challenges and achievements. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **367**, 815–831.
- [132] Dada J.O., Mendes P. 2011 Multi-scale modelling and simulation in systems biology. Integrative Biology **3**, 86–96.
- [133] Seidl R., Rammer W., Scheller R.M., Spies T.A. 2012 An individual-based process model to simulate landscape-scale forest ecosystem dynamics. Ecological Modelling **231**, 87–100.
- [134] Wieder W.R., Allison S.D., Davidson E.A., Georgiou K., Hararuk O., He Y., Hopkins F., Luo Y., Smith M.J., Sulman B., et al. 2015 Explicitly representing soil microbial processes in earth system models. Global Biogeochemical Cycles **29**, 1782–1800.
- [135] Luo Y., Ahlström A., Allison S.D., Batjes N.H., Brovkin V., Carvalhais N., Chappell A., Ciais P., Davidson E.A., Finzi A., et al. 2016 Toward more realistic projections of soil carbon dynamics by earth system models. Global Biogeochemical Cycles **30**, 40–56.
- [136] Miller J.H., Page S.E. 2009 Complex adaptive systems: An introduction to computational models of social life. Princeton university press.
- [137] Holland J.H. 2006 Studying complex adaptive systems. Journal of Systems Science and Complexity **19**, 1–8.
- [138] Clark C. 1990 Mathematical bioeconomics .
- [139] Getz W.M., Francis R.C., Swartzman G.L. 1987 On managing variable marine fisheries. Canadian Journal of Fisheries and Aquatic Sciences **44**, 1370–1375.
- [140] Hilborn R., Mangel M. 1997 The ecological detective: confronting models with data, volume 28. Princeton University Press.
- [141] Johnson J.B., Omland K.S. 2004 Model selection in ecology and evolution. Trends in ecology & evolution **19**, 101–108.
- [142] Beaumont L.J., Hughes L., Pitman A. 2008 Why is the choice of future climate scenarios for species distribution modelling important? Ecology letters **11**, 1135–1146.

- [143] Beissinger S.R., Westphal M.I. 1998 On the use of demographic models of population viability in endangered species management. The Journal of wildlife management pp. 821–841.
- [144] Getz W.M., Muellerklein O.C., Salter R.M., Carlson C.J., Lyons A.J., Seidel D. 2017 A web app for population viability and harvesting analyses. Natural Resource Modeling **30**.

## Box 1: Assessing Model and Data Adequacy, Improving Models

For purposes of clarity, we have organized key questions that should to be asked in assessing the adequacy of ecosystem models and the available data (but see Larsen et al. [2] for a considerably more detailed discussion in the context of landscape level geophysical models that incorporate ecological components).

### ASSESSING MODEL ADEQUACY

**State Variables:** *Does the model include all the state variables needed to accomplish the task at hand?* Beyond characterizing necessary physical (e.g., concentrations of focal chemical entities: gases, nutrients, etc.) and biological elements (e.g., species), are structures such as age, class (e.g., size, disease, condition) sex, and space included at the appropriate levels of detail?

**Control Variables:** *Does the model incorporate all the instruments needed to evaluate the system's response to interventions?* For example, if animal populations are to be manipulated through harvesting, stocking or translocations among subpopulations, then such variables need to be incorporated in ways that reflect the impact of associated management actions on population numbers [105].

**Sensitivity:** *Is the behavior or predications of the model sensitive to perturbations in parameter values and, if so, does this sensitivity likely reflect a real or spurious behavior?* For example, backward bifurcations are quite common in the stability analysis of disease models. These bifurcations, if proven to be real, imply that the disease can persist once established even if changes in transmission effectively reduce  $R_0$  below one [106].

### ASSESSING DATA ADEQUACY

**Determinacy:** *Are data sufficient to reliably estimate model parameters?* The values of as many parameters as possible should be pinned down using independent sets of data (e.g., birth and mortality rates from life tables for different species in ecosystems models) before using optimization procedures to estimate remaining parameter values. If the model has too many free parameters it is possible for several parameters to settle anywhere across a wide range of values while providing the same best fits to data using cutting edge estimation procedures [107].

**Validation:** *Beyond the data used to estimate model parameters are there sufficient additional data to carry out robust assessments of the reliability of model performance using cross-validation procedures, particularly in predicting patterns not used to fit the data?* For example, if a disease model is fitted to incidence data, does the model also correctly predict offspring distributions when such data are available [108].

### IMPROVING MODELS

**Elaboration:** *Do horizontally elaborated versions of these models provide better performance or new insights into behavior and patterns beyond those produced by the core models?.* Here elaboration typically means either adding terms that represent an additional process (e.g., adding migration into a demographic model) or making expressions more elaborate by replacing process-description functions with more general functions that include the original functions as special cases (Fig. 3).

**Refining:** *Does the replacement of a phenomenological construct with a module that explicitly computes the phenomenon of interest as the emergent behavior of a system described at a lower hierarchical level provide new insights into observed behaviors and patterns, as well as a richer set of behaviors and patterns?. A case in point is replacing the carrying capacity parameter  $K$  in the logistic model with a module that explicitly computes how individuals encounter one another and resolve conflicts that arising from competitive interactions [109]*

**Extracting Essentials–Coarse Graining:** *In systems we have been investigating using vertical elaboration (i.e. below the focal level), can we extract or interpret behaviors seen below the focal level in ways that add insight or predictive power at the focal level? (see coarse graining path in Fig. 3). For example, one can model the interaction of consumers and resources interacting on a landscape using a system of differential equations [110]. One can then model the same system at the refined level of an individual-based approach [26]. In this latter case, one can show that if individuals have different propensities to move, to avoid competitors, and to plan moves ahead, then several different movement behavioral types emerge [111, 112]. Thus the consumers can then organized into several syndromic movement groups [113] that can be modeled using a more nuanced system of differential equations than the original by now taking into account this new group structure.*

## Box 2: Complexity Typology of Ecological Models

### STRUCTURAL COMPLEXITY

**Space/time:** Spatial complexity ranges from no spatial structure, through metapopulation structure [114] to variation across a full Euclidean coordinate space. Temporal structure can either be static (e.g., species distribution [115,116] models) or dynamic (typically modeled by systems of difference or differential equations, or as individual-based processes).

**Environment:** The environment can be implicitly included in model parameter values (e.g., carrying-capacity, time varying growth rates) or explicitly included in terms of abiotic factors that drive or influence seasonal or long-term trends, and include biotic feedbacks as well (e.g., soil carbon levels [117], as in Earth Systems Models—ESMs: see Fig. 2).

**Traits:** Traits can be implicitly reflected in parameter values (e.g. Lotka-Volterra competition models) or explicit incorporated as discrete or continuous distributions (e.g., integral project models—IPMs—are distributional generalizations of discrete matrix age or other trait structure models [118–120]). The most comprehensive way to model multi-trait structure is to use an individual-based modeling (IBM) approach [4, 34].

**Hierarchy:** Beyond linear compartmental modeling [9], hierarchical structure is often incorporated into dynamic models using Lotka-Volterra type foodweb formulations [121]. These formulations can be extended to metapopulations [122], as in the Ecopath with Ecosim (EwE) implementation formulations of fisheries systems [123–125] (Fig. 2).

### PROCESS COMPLEXITY

**Transformations:** Transformation process functions can either be [31]: *empirically determined* by fitting a curve to data; *phenomenologically inferred*, as in the inclusion of a carrying capacity parameter in the logistic-growth equation; or *mechanistically derived*, as in unpacking the intrinsic growth-rate and carrying-capacity parameters in the logistic equation, using a metaphysiological model approach [126]. The latter underpins the biomass transformation web (BTW) formulation [110], which is a mechanistic alternative to LV foodwebs, and is directly transferable to an IBM trait-level implementation [26] (Fig. 2).

**Environment:** All functions in models can be extended to include stochastic elements that either implicitly or explicitly incorporate environmental fluctuations through additive or multiplicative noise, represented by samplings from statistical distributions. They may also involve scheduling events using randomly distributed occurrence times [127–129].

**Scales:** Multiscale models are challenging to implement from both modeling and computational points of view [26, 122, 130–132], but notable examples exist (Fig. 2): individual tree growth models have been scaled up to simulate landscape-level ecosystem dynamics [133]; Michaelis-Menton soil microbial process models have been incorporated into earth system models (ESMs) to link the micro scales in the soil with landscape macro scales to make multi-decadal climate change projections [134, 135]; and BTW processes (mentioned above) have been studied at evolutionary time scales revealing processes that lead to the emergence of foraging



guilds [111, 112]. The study of systems that include evolutionary process—so-called complex adaptive systems (CAS)—has become a research field in its own right [136, 137].

## UTILITY COMPLEXITY

**Explanatory:** Simple explanatory models, such as the logistic growth model, are referred to as minimal and strategic [11]. A mechanistic complexification of the phenomenological logistic-growth model is embodied in the metaphysiological-growth model (Fig. 2), which *explains* how carrying capacity emerges as a function of resource density in terms of four parameters: a maximum per-capita resource-extraction rate, an extraction-rate efficiency value, an interference-competition value, and a metabolic-conversion-efficiency value [126]. This model is synthetic because it provides testable predictions of how the environmental carrying-capacity will change in response to changes in the values of these parameters.

**Management:** Management models can be *strategic*, such as the logistic used to develop a bioeconomic population-harvesting theory [138]. They can also be tactical, such as age-structured population models used to set quotas in fisheries management [18]. These models have been extended to stochastic [139] and multispecies fisheries setting using the Ecopath with Ecosim (EwE) software platform [99] (Fig. 2).

**Fidelity:** Assessing the fidelity of models is a field of statistical research that involves uncertainty in both model structure and the data used to fit the models [21, 140–142]. Strategic models of low fidelity can be used to make *comparative statements* regarding our understanding of the effects of different management actions on particular systems. If we want to make *absolute statements*, however, then we need levels of fidelity that produce predications that have the desired levels of confidence. These require tactical models, such as Leslie matrix formulations extended to include density-dependent demography [18] (Fig. 2), and can be used in a population viability analysis (PVA) framework to assess the risk of a particular species going extinct by some specified future date. [143, 144].

## Figures

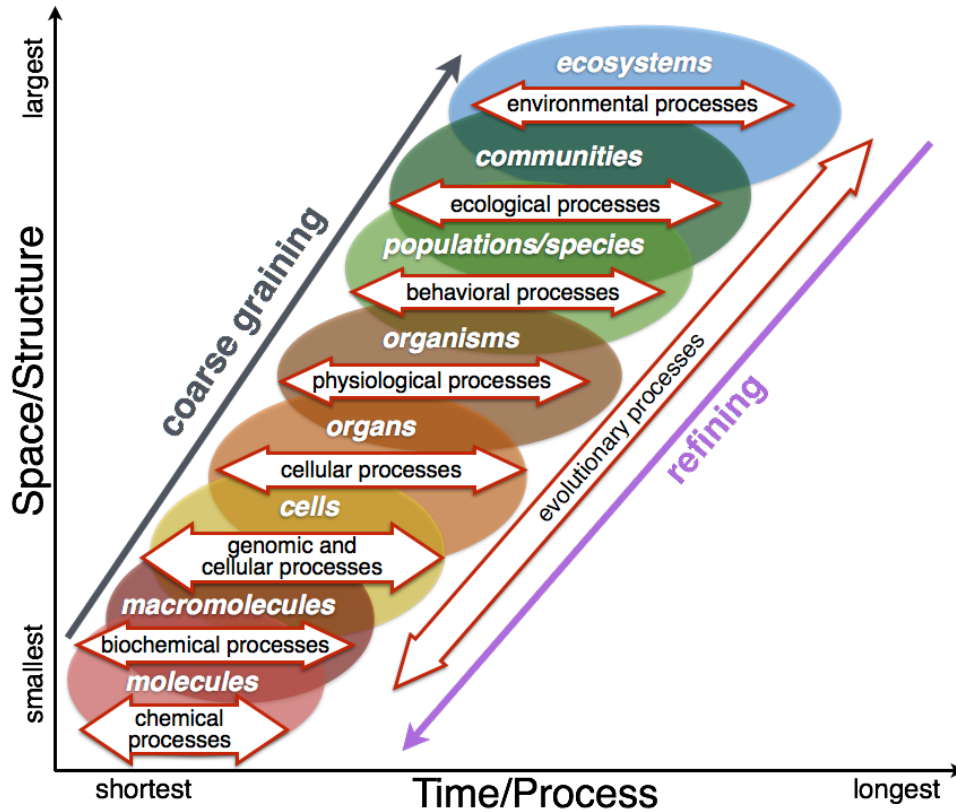


Figure 1: Hierarchical Organization. A hierarchical view of an ecosystem’s organizational scale is envisioned here as being composed of constituent communities, which in turn consist of species assemblages, populations of organisms, and so on all the way down to the constituent molecules of cells. The organization depicted here is neither unique nor complete, but includes a number of typical focal levels at which biological systems are described and modeled. The axes indicate a notion of increasing time and spatial scales at increasingly higher levels of organization. The challenge in developing adequate models may involve an “appropriate complexity approach” [2] to identifying an appropriate focal level for the question at hand and then alternately refining the focal level by moving to a lower level and then coarse-graining the model (as defined in Box 1) back to the focal level, in an approach to finding satisfactory answers to environmental problems using a procedure depicted in Figure 3.

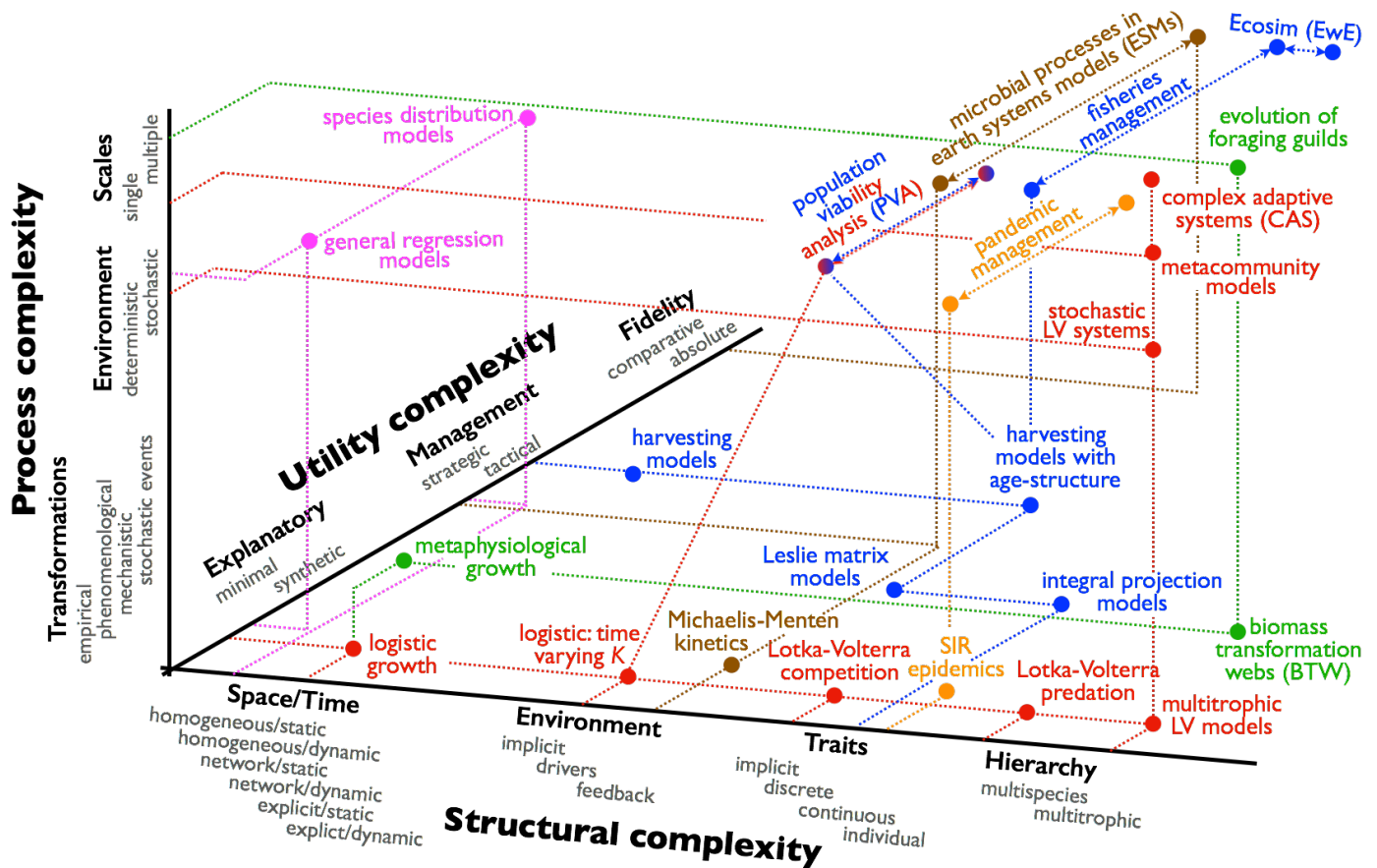


Figure 2: Model Typology. A way of categorizing models is to locate them in a *structure-/process-/utility-complexity* space. Although measures do not exist to ordinate this space, segments of this space can include a notion of increasing complexity with respect to the named primary categories along each axis: descriptions within each segment are indicated by the smaller text underlying each category name, staggered left to right to indicated increasing levels complexity. See Box 2 for further discussion of typologies and models: colors denote related modeling threads. Placement of models in this space is approximate, to facilitate labeling. Dots connected by arrows imply the labels above pertain to models that span the indicated range. PVA analysis is labeled using both blue and red text because the field depends critically on both threads.

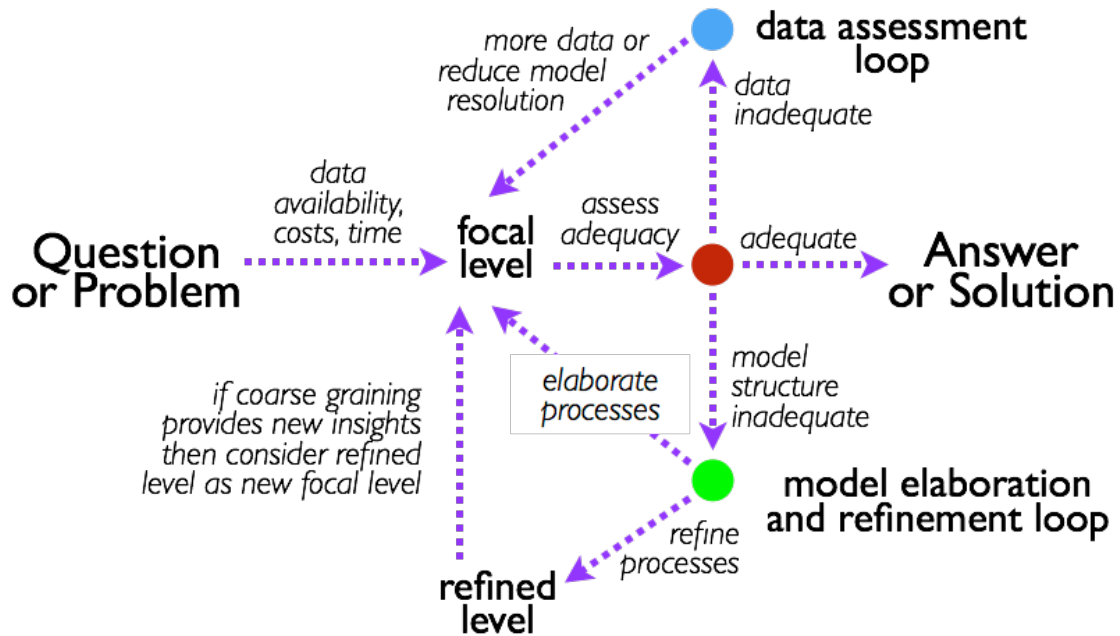


Figure 3: Assessing Model Adequacy. The process of assessing model adequacy to provide an answer to a question or a solution to a problem begins at the focal level at which the model was initially formulated (Fig. 1). Adequacy is assessed (Box 1) through evaluation of the relative benefits of new data and modification of the model structure (red circle). The decision might be to collect more data (blue circle) or elaborate the model (green circle, Box 1) or both. The process is iterative and may lead to an increase or decrease in the structural complexity of the model.