CellPress
REVIEWS

Opinion

Process, Mechanism, and Modeling in Macroecology

Sean R. Connolly,[1,*] Sally A. Keith,[2,3] Robert K. Colwell,[2,4,5] and Carsten Rahbek[2,6]

Macroecology has traditionally relied on descriptive characterization of large-scale ecological patterns to offer narrative explanations for the origin and maintenance of those patterns. Only recently have macroecologists begun to employ models termed 'process-based' and 'mechanistic', in contrast to other areas of ecology, where such models have a longer history. Here, we define and differentiate between process-based and mechanistic features of models, and we identify and discuss important advantages of working with models possessing such features. We describe some of the risks associated with process-based and mechanistic model-centered research programs, and we propose ways to mitigate these risks. Giving process-based and mechanistic models a more central role in research programs can reinvigorate macroecology by strengthening the link between theory and data.

Trends

Macroecologists are increasingly using process-based and mechanistic models.

Process-based models need not be reductionistic.

Such models can be used for theoretical, statistical, and virtual-worlds modelling.

These innovations allow stronger inferences about causes of ecological patterns.

The Evolution of Model-Based Ecology

The growth of mathematical modeling in ecology and the increasingly prevalent interpretation of data analysis as model fitting and evaluation (rather than falsifying specific hypotheses about nature) reflect a changing view of models that has paralleled the growth of model-based conceptualizations of science in philosophy (Box 1). From the middle to the end of the 20th century, ecology was dominated by a paradigm that reflected the influence of falsificationism [1], strong inference [2], and the use of threshold statistical significance as a decision rule [3]. However, ecologists have increasingly adopted approaches to data analysis that emphasize estimation of model parameter values and selection among alternative models [4,5]. This approach often eschews null hypothesis testing or strong inference, but, more fundamentally, it entails a reassessment of the roles, importance, and even nature of theories, models, and hypotheses in the scientific process [6]. This shift has occurred alongside several other trends in ecology. One such trend is the greater use of custom-built ecological models – commonly described as 'process-based' or 'mechanistic' – to fit to data, rather than models that test for associations between response and explanatory variables [7,8]. Simultaneously, improvements in computing technology have made fitting such models to data practicable [9]. Increasingly, ecologists have been called upon to supply knowledge for managing vulnerable ecosystems and exploited populations, which often requires estimates of effect sizes and associated confidence intervals to inform management decisions [10,11]. Not least of these trends has been the growth of macroecology.

Because macroecologists focus on large-scale patterns and processes, experimental manipulation is often impossible, limiting the extent to which macroecology could employ the methods of falsification and strong inference to test alternative explanations for phenomena. Instead, macroecologists have focused on the consistency (across taxa, ecosystems, and spatial and temporal scales) of macroecological patterns, such as Taylor's mean–variance scaling law, species–area relationships, and species-abundance distributions [12]. Recently,

[1]Marine Biology and Aquaculture, College of Science & Engineering, and ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Australia
[2]Center for Macroecology, Evolution, and Climate, National Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
[3]Lancaster Environment Centre, Lancaster University, Lancaster, UK
[4]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA
[5]University of Colorado Museum of Natural History, Boulder, CO, USA
[6]Department of Life Sciences, Imperial College London, Ascot SL5 7PY, UK

*Correspondence:
sean.connolly@jcu.edu.au
(S.R. Connolly).

Box 1. Theories of Scientific Explanation

For much of the first half of the 20th century, predominant theories of scientific explanation were 'syntactic', meaning that they treated theories as formal axiomatic systems, from which explanations or predictions of natural phenomena were deduced. The paradigmatic example of this approach is Hempel and Oppenheim's 'Covering Law Model' [76] (Figure IA). According to this model, a scientific explanation consists of a set of deterministic or statistical laws and a set of specific antecedent conditions that, if met, entail that a particular natural phenomenon either must follow, or follows with high probability (here termed the 'hypothesis'; formally the 'explanandum' in [76]). If empirical data conform to the hypothesis, then the explanation is supported; if they do not, then the explanation is either falsified or rendered unlikely to be true, implying that one of the putative laws or antecedent conditions is false. The key point is that the theory's axioms – its laws and antecedent conditions – are assumed to apply literally to the natural system itself, so an explanation is either true or false, and empirical data are used to assess the likelihood that the explanation is true (in the earlier positivist framework), or to show that the explanation is false or unlikely (in the later falsificationist framework).

In model-based conceptualizations of science, by contrast, hypotheses are derived from models that are known, *a priori*, to be idealized representations of nature (Figure IB). Model assumptions are akin to the axioms of the syntactic framework, but they are assumed to be strictly true only for the model itself, not the natural system(s) being modelled. Consequently, model-based explanations for natural phenomena are not strictly true or false in the same sense as in the syntactic framework. Rather, explanations about natural phenomena are evaluated based on the degree and nature of the (invariably imperfect) resemblance between the empirical data and the model on which the explanation is based [77]. Differences among proponents of model-based conceptualizations of science exist, for instance concerning the precise meaning of representation or explanation, and whether models are constituents of theory or mediate between theory and nature [78]. Nevertheless, this summary captures, to a first approximation, the key distinctions between the model-based view of science and the syntactic view that dominated philosophy of science in the first half of the 20th century [78].

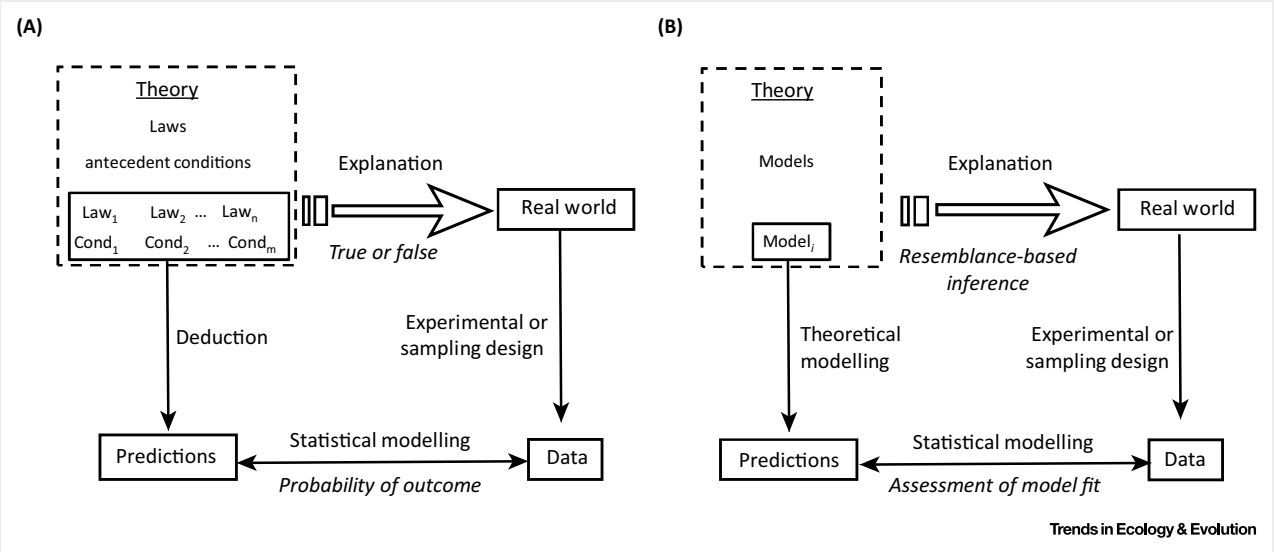

Figure I. Schematic Illustration of Two Different Models of Science. (A) Example of a 'syntactic' view, in which a theory is a formal axiomatic system, whose relevant laws and antecedent conditions are combined to derive predictions, and whose explanations are true or false based on the validity of its hypotheses. (B) Example of a 'model-based' view, according to which a theory is a constellation of models, which are used to derive predictions. According to this view, explanations are based on an assessment of the resemblance between the model and nature, rather than an assessment of whether the model is 'true' or 'false'. Panel (B) is adapted from Giere [6].

macroecological studies increasingly offer 'mechanistic' or 'process-based' explanations for macroecological patterns, using model-based methodological approaches. These include testing quantitative predictions about model parameter values [13], and model analyses that demonstrate how particular combinations of macroecological processes or mechanisms may give rise to well-documented macroecological patterns [14–16].

The growth of process-based and mechanistic models (hereafter PBMs and MMs, respectively) raises several questions. Firstly, what are the characteristics of a model that make it mechanistic or process-based (or not), as opposed to models that lack these characteristics (often termed 'descriptive', 'correlative', 'phenomenological', or 'purely statistical' models)? Secondly, what are the advantages of employing PBMs and MMs in macroecological research, and how do these advantages follow from the distinguishing features of PBMs and MMs? Finally,

does a PBM- or MM-centered research program entail any particular risks? In this contribution, we offer answers to these questions. Although we focus on macroecology, a field that has only recently seen widespread use of such models, much of what we say is relevant to ecology more broadly.

What Makes Models Mechanistic or Process-Based?

Recent reviews of PBMs and MMs in ecology consider these two types of models to be synonymous, either explicitly [7,8] or implicitly (for instance, by defining process-based models as incorporating mechanism or being reductionistic [17–20]). However, the use of the term 'mechanistic' by ecologists (including macroecologists) is somewhat inconsistent, and we know of no attempts to clarify what qualifies a model as 'mechanistic' or 'process-based' (other than simply that it includes mechanisms and processes, respectively). We therefore begin by proposing definitions for MMs and PBMs, in order to capture two distinct aspects of models that make it possible to employ them in a range of ways not possible with other models.

In philosophy of science, multiple definitions of 'mechanism' have been proposed, but these definitions share some key features: a phenomenon to be explained, and component parts whose activities and interactions produce the phenomenon [21,22]. In this spirit, we define a mechanistic model (MM) as 'a characterization of the state of a system as explicit functions of component parts and their associated actions and interactions'. Models that would fail to meet this definition include the standard statistical models usually employed by ecologists, such as regressions of species richness against environmental variables. However, our definition also excludes other models, such as the logistic model of population growth, for which changes in population size depend only on current population size, and no information about the states of the population's component individuals is needed. Nevertheless, in practice, we believe that most ecologists would consider fitting such a model to empirical data to be qualitatively different than, for instance, fitting a simple moving average to population size over time. To capture such models, we define process-based models (PBMs) as 'models that characterize changes in a system's state as explicit functions of the events that drive those state changes'. By this definition, the logistic model of population growth is a process-based model, because the state variable, population size, is modelled as an explicit function of the births and deaths that drive changes in population size. But neither a simple regression of species richness on, for example, temperature, nor a running-average model fitted to population size data, is a PBM, because their parameters or components do not correspond directly to events or entities in nature. We acknowledge that some ecologists use the term 'mechanistic model' more broadly than we have done here, and thus may prefer to consider these two groups to be (overlapping) subclasses of mechanistic models (e.g., 'component-based mechanistic' and 'process-based mechanistic'); conversely, others may consider an ecological model to be mechanistic only if it incorporates individual-level physiology or behavior (Box 2).

Although many mechanistic macroecological models are process-based, and vice versa, our definitions above imply that models may also belong to one class and not the other. Many niche models of species distributions, for example, are mechanistic but not process-based. Functional constraints on organisms' physiological tolerance may be derived from process-based studies, but if their implementation in species-distribution models is based on static functional relationships between performance and environment [23], then there is no explicit representation of process in the model. Conversely, island-biogeographic models that characterize the dynamics of species richness as a function of species-level immigration, speciation, and extinction rates are process-based (they explicitly include processes that drive changes in the state variable, richness), but not mechanistic, because the dynamics of species diversity depend only on the standing species richness (not on information about the variable states of particular species; e.g., [24]).

Box 2. What Is a Mechanistic Model?

The mechanism concept in philosophy of science, upon which we base our definition of a mechanistic model, is consistent with the concept's use in the macroecological literature on, for example, niche modeling: 'If a species' niche is to be modelled mechanistically . . . the organism must not enter the model as a point on a map but rather as a set of behavioral, morphological, and physiological traits' [23]. This implies that a mechanistic model must explicitly represent the mechanism's component parts (in this case, individual physiology) that produce the higher-level phenomenon (a species' distribution). Some ecologists have maintained that these component levels must include individual-level ecology, such as physiology or behavior [79]. However, this is not necessary to satisfy our definition. In macroecology, habitat patches and species are common components of mechanistic models. For instance, in neutral macroecological models, the component parts of the metacommunity are local communities, whose internal workings (community dynamics) and interactions (dispersal) are modelled explicitly. Such models have been explicitly termed mechanistic, even though they specifically ignore differences among individuals [80]. Similarly, in mechanistic models of species richness gradients, large-scale species richness patterns are produced by the modeling of variable and dynamic geographic range states of species, and species' effects on one another's colonization, extinction, or dispersal (i.e., their interactions) [56]. Because models with such group-level components (*sensu* [79]) offer the advantages summarized in this paper, we consider them to be mechanistic.

We also recognize that, even when they share a definition of mechanism, macroecologists may hold differing concepts of mechanistic model, arising from different views about how the mechanism concept should be employed in modeling. For instance, Tilman's consumer-resource competition model incorporates mechanism only implicitly: the Michaelis–Menton resource uptake function can be derived from a submodel of uptake by individual cells, even though individual cells' uptake dynamics are not explicitly modelled. Some ecologists consider such models to be mechanistic (e.g., [81,82]). Our definition requires explicit modeling of component states because we could not determine how to define MMs to include models with implicit mechanisms without allowing any model based on mechanistic reasoning (e.g., simple correlative statistical models), to be considered MMs as well. Nevertheless, macroecologists who prefer a broader definition of mechanistic model may prefer to consider our MMs and PBMs to be sub-classes of mechanistic model (component-based mechanistic models and process-based mechanistic models, for instance).

Advantages of Process-Based and Mechanistic Models

The preceding discussion identifies two distinct ways that PBMs and MMs can move beyond other models in how they represent nature, and clarifies that PBMs need not explicitly characterize entities or events at a lower or more fundamental level (such as explaining population growth in terms of the physiological states of individuals). What PBMs and MMs share is an explicit representation of causal structure. Indeed, past reviews have differentiated PBMs and MMs from other models because they incorporate 'causality', 'causal effects', 'causal mechanisms' or 'causal relationships' [7,8,17]. This structure means that PBMs and MMs have several important advantages as research tools. Specifically, they can be confronted with empirical data in more diverse and informative ways than models based purely on statistical associations between variables. Moreover, insights can be gained by using PBMs and MMs themselves as objects of study.

Richer Inferences from Statistical Modeling

Increases in computing power, coupled with the growth of methods for fitting purpose-built models to data, mean that many PBMs or MMs can be explicitly fitted to empirical data. For some models, it is possible to specify a goodness-of-fit function (e.g., a likelihood) explicitly, and use numerical methods to fit the model. In macroecology, such approaches have been important, for instance, in testing species-abundance models [14,25], in quantifying drivers of diversity–stability relationships at regional scales [26], and in inferring biogeographical dynamics in macroevolutionary time [27–31]. For models that are analytically intractable, stochastic simulation can be used to approximate expected distributions of observed (data) values whose correspondence with empirical data can then be quantified via summary statistics [32–35].

One major advantage of PBMs and MMs is that they have parameters with a direct biological interpretation: they refer to properties of entities that compose the system of interest, or to the rates at which particular processes occur. This correspondence makes those parameters

independently estimable, at least in principle, independent of the state of the system being modelled. For instance, one can estimate carrying capacity based on the resource requirements of individual organisms and knowledge of the distribution of limiting resources in a landscape [36]; one can measure the thermal reaction norms believed to determine a species' distributional limits without knowing that species' distributional limits [37]. Similarly, one can estimate parameters of many multispecies competition models from single-species and pairwise competition experiments (e.g., [38]). By contrast, the parameters of the statistical models traditionally used by ecologists are not measurable in such ways. In a regression model of site occupancy as a function of a set of environmental variables, for instance, the model parameters have no meaning independent of the calibrated relationship itself. Consequently, when such independent estimates are possible in practice, PBMs and MMs can be used to predict a system's state without directly fitting the model to observations of that state [36]. Conversely, a PBM or MM may be fitted to data obtained from observations of a system's state, and the resulting parameter estimates compared with those obtained by independent estimation [25]. Indeed, one major criticism of early neutral models of biodiversity was that estimates of speciation rates from fitted neutral models differed far too much from plausible values of those rates to be realistic [39]. Alternatively, these two approaches may be combined, with some parameter values independently estimated and others optimized to fit an observed macroecological pattern [40], or by independently estimating model parameter values and then using the results of these estimates as prior distributions when fitting a full, process-based model to the pattern under investigation [41].

The causal structure of process-based and mechanistic models usually implies greater capacity to make multiple predictions simultaneously than with other types of statistical models, which often specify relationships between a single response variable and one or more explanatory variables. Multiple predictions can be particularly useful when a model is formulated to explain a pattern that is already known, such as the hollow-curve shape of species-abundance distributions or the power-law shape of species-area curves. Because the parameters of PBMs or MMs tend to have biological meaning, they often yield predictions about additional macroecological patterns. Such predictions can be evaluated either by fitting a model to multiple patterns simultaneously, and asking whether the model's fit to all of the patterns is adequate [42], or by fitting a model to one or more patterns, and then using the estimated parameters to predict another pattern [34,43–45]. Such analyses are widely considered to be particularly strong tests of models [46], and they are less likely to be possible with models lacking explicit processes or mechanisms (see [47] for an exception).

Using Models as Objects of Investigation

Because PBMs and MMs explicitly represent a set of specific assumptions about the causal structure of a system, the models' behavior can be explored to determine the implications of a system being structured in a particular way. Indeed, historically, the most common use of PBMs and MMs in ecology has been for theoretical modeling: an analysis of the model itself, to understand how the model's assumptions influence the phenomenon or pattern that the model characterizes. Theoretical modeling is a deductive 'if–then' exercise: if aspects of the system's organization or the nature of the events is changed, how will the system respond? Such investigation may be undertaken using symbolic analysis, numerical analysis, or stochastic simulation, but in all cases the model itself is the object of study. In macroecology, symbolic analyses have, for example, been used to explicitly derive closed-form solutions for the species-abundance distribution [14,25], predict the trophic structure of island-biogeographic patterns [48,49], and derive expressions for species richness gradients in a bounded domain [50]. Numerical analyses have been used to determine the temporal evolution of species abundances in neutral communities [51], the distribution of species' range states under stochastic colonization and extinction dynamics [52], and food web dynamics under anthropogenic

environmental change [53,54]. For more heavily parameterized stochastic models, or models containing more complex functional forms, model predictions must be generated by stochastic simulation. In macroecology, such approaches have been employed to generate predicted species abundance distributions [55], species richness gradients [35,56–58], effects of geological history on island-biogeographic patterns [59], landscape patterns of species coexistence [60], and range shifts of species [61,62].

In addition to theoretical modeling, however, PBMs and MMs may also be used in conjunction with purely statistical models in various ways, to enrich the inferences that can be drawn from fits of such models to empirical data. For example, highly complex PBMs and MMs, which may be too heavily parameterized to be fitted to empirical data directly, can be used as virtual worlds to evaluate the robustness of inferences from simpler, more tractable models. Under this approach, virtual data, simulated from a PBM or MM, are analyzed with a simpler model (which may or may not be a PBM or MM). Then, inferences about what is giving rise to the pattern, based on the model used in the analysis, can be explicitly compared with the 'truth', since the processes generating the data are specified in the PBM or MM used to generate the virtual data. Virtual worlds simulation models have been used to test the performance of static and dynamic species distribution models [63,64], the extent to which effects of competition and environmental filtering are captured by null model tests of phylogenetic structure [65–67] (see [68] for a similar approach to earlier null model controversies), and the ability of functional diversity indices to detect convergence and divergence of traits [69].

In a conceptually different virtual-worlds application of PBMs and MMs, model analysis may suggest explanations for the existence of, or systematic variation in, a common macroecological pattern that is well-characterized by a conventional statistical model like a standard probability distribution or a simple regression, but the PBM or MM is not, itself, explicitly fitted to empirical data. For example, by modifying the importance of demographic stochasticity, environmental stochasticity, and density dependence in stochastic population models, the response of the power-law scaling exponent between the mean and variance of abundance to the nature or strength of such demographic processes has been explored [15,70]. Similarly, individual-based ecosystem models have been used to explain the slope of community size spectra [16,71,72]. Conversely, common predictions of models that share a particular assumption or set of assumptions may be identified, to evaluate the extent to which those assumptions approximate nature. A useful framework for this approach can be constructed by analogy to Imre Lakatos' philosophy of research programs (Box 3). To directly target the neutrality assumption in neutral theory of biodiversity, for example, common features of the abundance distributions produced by neutral models have been identified and compared with those exhibited by empirical data [73].

Risks of Process-Based and Mechanistic Models

The shift towards PBMs and MMs in macroecological research entails multiple, but manageable, risks. Firstly, moving away from methodologies focused on falsification of hypotheses increases the risk of confirmation bias. If the subjective disposition of researchers (or journals or reviewers) is towards studies demonstrating that models perform well, then researchers may be less inclined to actively interrogate models or model fit in ways that are likely to reveal weaknesses and failures. Such interrogations are essential to motivating theoretical progress in macroecology. A recognition that all models are idealizations and thus must fail to a greater or lesser degree, coupled with a culture and ethos (*sensu* [74]) that values rigorous and creative interrogation of model fit [46], can still encourage constructively skeptical approaches towards model testing and inference. Secondly, models that share particular assumptions of macroecological interest, but differ in auxiliary assumptions, often produce quite divergent patterns. Conversely, models that make markedly different assumptions can explain very similar

Box 3. Core and Auxiliary Assumptions in Model-Based Theory

In Lakatos' [83] philosophy of science, a research program consists of a 'hard core' of fundamental assumptions with a 'protective belt' of auxiliary hypotheses (Figure IA). The fundamental assumptions define the research program: if one or more of these is wrong, the research program is undermined. However, fundamental assumptions cannot be tested in isolation, because specific hypotheses hold only if the auxiliary hypotheses, as well as the fundamental assumptions, are correct. These auxiliary hypotheses protect the 'hard core' of fundamental assumptions if data appear that contradict predictions from the theory.

The hard core/protective belt dichotomy has been criticized as a model for research programs in general because ideas may move in and out of the hard core as a research program evolves [84]. Nevertheless, an analogous distinction between core and auxiliary model assumptions illuminates why tests of a single model rarely resolve the underlying questions that those tests aim to answer. Neutral theory of biodiversity, for example, seeks to explain patterns of abundance and biodiversity as the consequence of demographic stochasticity [85]. Thus, neutral models share the 'core' assumptions of ecological and demographic equivalence of individuals, regardless of species, so that all differences in abundance patterns between species arise only from chance variation in births, deaths, speciation and dispersal events (Figure IB). However, to predict those patterns of abundance and diversity, other assumptions are required, concerning mode of speciation, the nature of competition, and the rules governing dispersal in the landscape or seascape (Figure IB). Because none of these assumptions is core to neutral theory, neutral models differ from one another in these assumptions (e.g., [85–88]). Consequently, when neutral models fail to produce an empirical pattern (e.g., realistic species lifespans, or patterns of community similarity), new models are produced that seek to capture these patterns without violating the core neutrality assumptions (e.g., [39,55,87]).

Note that the concepts of 'core' and 'auxiliary' in Dynamic Energy Budget (DEB) theory [89] closely resemble the Lakatosian concepts employed here, since auxiliary parameters connect the core theory (which concerns unmeasurable quantities) with observable quantities that can be compared with theoretical predictions. Discrepancies between DEB models and data have been addressed by modifying elements of auxiliary rather than core theory [90,91], in a fashion similar to that described here.

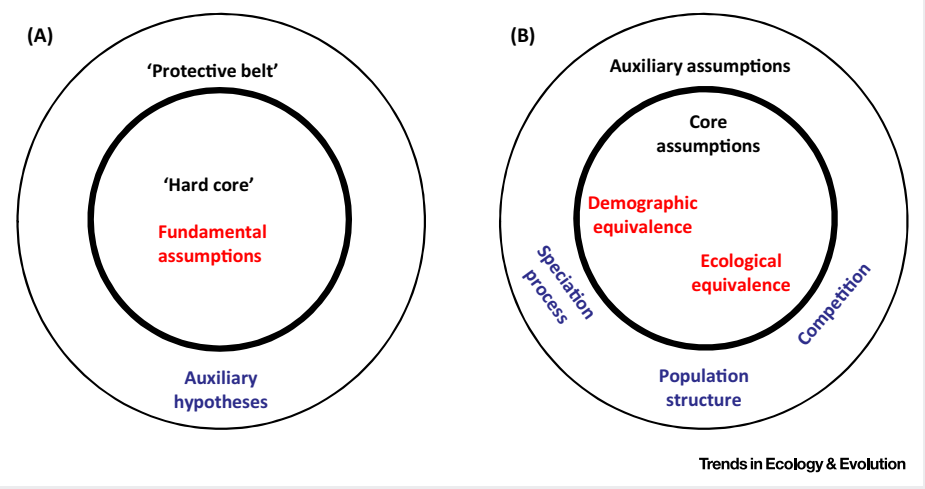

Figure I. Core and Auxiliary Model Assumptions. (A) Schematic representation of Lakatos' conception of research programs. (B) An analogous representation of neutral models of biodiversity. Neutral theory of biodiversity postulates that patterns of abundance and biodiversity can be explained without reference to demographic or niche differences between species. For this reason, neutral models share the 'core' assumptions of demographic and ecological equivalence, but differ in the additional, 'auxiliary' assumptions required to make predictions about species richness and species abundances.

patterns. This indeterminacy highlights the importance of employing multiple models, and testing models against multiple predictions, in model-based research programs.

More practically, PBMs and MMs tend to be more complicated than the statistical models more commonly employed by ecologists, increasing the risk of modeling errors. This risk is arguably greater in macroecology than many other fields, because advanced mathematics, statistics,

and programming are rarely viewed as core curricula for aspiring macroecologists. A greater emphasis on adequate training in the areas of mathematics most often used for macroecological models (for instance, differential equations, linear algebra, and stochastic processes), alongside requirements for greater transparency in scientific publication (such as a requirement to make relevant computer code publicly available) should help to reduce the frequency of errors and increase the likelihood of identifying and correcting errors that do occur.

Concluding Remarks

PBMs and MMs are in the process of dramatically changing the field of macroecology, away from a discipline that was once defined in terms of an analytical approach based on pattern description [75], towards one that employs PBMs and MMs to draw inferences about the processes and mechanisms that underlie macroecological patterns (see Outstanding Questions). The potential to infer the values of process parameters from macroecological data permits more stringent testing of models by producing estimates of parameter values or by predicting patterns in different places, at different scales, or along different dimensions (e.g., time), whose plausibility can then be examined. PBMs and MMs can also be used as virtual worlds, generating simulated data that can help us understand how well simpler models can yield inferences about the underlying drivers of macroecological patterns. Finally, because they can be used in theoretical analysis as well as directly confronted with data, PBMs and MMs have the potential to advance macroecology by strengthening the relationship between theory and data.

Outstanding Questions

What formal or informal criteria should be adopted to mitigate the risk of confirmation bias in model-based macroecology?

How much can PBMs and MMs enrich predictive macroecology (beyond species distribution modelling), by making more explicit the causal structure underpinning changes in macroecological patterns under novel conditions?

When fitting models to data by stochastic simulation, how sensitive are the conclusions drawn to the particular choices of summary statistics used to compare models with data?

How reliable are inferences about the robustness of purely statistical models that lack explicit process or mechanism, when they are drawn from tests based on virtual-worlds modeling?

Can we use families of PBMs and MMs to target core, rather than auxiliary, assumptions of macroecological theories besides neutral theory?

Can PBMs and MMs be used synergistically in ways that are more powerful and illuminating than when one or the other type is used in isolation?

Acknowledgments

We thank W. Connolly, D. Gravel, M. Kearney, T. Northfield, and an anonymous reviewer for helpful comments and suggestions on this manuscript, and S. Glennan for drawing our attention to some very recent philosophical work. This work was supported by the Australian Research Council (CE140100020), the Danish National Research Foundation (DNRF96), and VILLUM FONDEN (10114).

Supplemental Information

Supplemental information associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tree.2017.08.011.

References

1. Popper, K.R. (1962) *Conjectures and Refutations; The Growth of Scientific Knowledge,* Basic Books
2. Platt, J.R. (1964) Strong inference. *Science* 146, 347–353
3. Neyman, J. and Pearson, E.S. (1933) On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 231, 289–337
4. Burnham, K.P. and Anderson, D.R. (2014) *P* values are only an index to evidence: 20th-vs. 21st-century statistical science. *Ecology* 95, 627–630
5. Hooten, M.B. and Hobbs, N.T. (2015) A guide to Bayesian model selection for ecologists. *Ecol. Monogr.* 85, 3–28
6. Giere, R.N. (1991) *Understanding Scientific Reasoning.* (3rd edn.), Holt, Rinehart, and Winston
7. Cuddington, K. *et al.* (2013) Process-based models are required to manage ecological systems in a changing world. *Ecosphere* 4, 20
8. Cabral, J.S. *et al.* (2017) Mechanistic simulation models in macroecology and biogeography: state-of-art and prospects. *Ecography* 40, 267–280
9. MacNeil, M.A. and Connolly, S.R. (2015) Multi-scale patterns and processes in reef fish abundance. In *Ecology of Fishes on Coral Reefs* (Mora, C., ed.), pp. 116–124, Cambridge University Press
10. Stillman, R.A. *et al.* (2015) Making predictions in a changing world: the benefits of individual-based ecology. *BioScience* 65, 140–150
11. Petchey, O.L. *et al.* (2015) The ecological forecast horizon, and examples of its uses and determinants. *Ecol. Lett.* 18, 597–611
12. Gaston, K.J. and Blackburn, T.M. (1999) A critique for macroecology. *Oikos* 84, 353–368
13. Enquist, B.J. *et al.* (1998) Allometric scaling of plant energetics and population density. *Nature* 395, 163–165
14. Saether, B.E. *et al.* (2013) Species diversity and community similarity in fluctuating environments: parametric approaches using species abundance distributions. *J. Anim. Ecol.* 82, 721–738
15. Engen, S. *et al.* (2008) A general model for analyzing Taylor's spatial scaling laws. *Ecology* 89, 2612–2622
16. Blanchard, J.L. *et al.* (2017) From bacteria to whales: using functional size spectra to model marine ecosystems. *Trends Ecol. Evol.* 32, 174–186
17. Dormann, C.F. *et al.* (2012) Correlation and process in species distribution models: bridging a dichotomy. *J. Biogeogr.* 39, 2119–2131
18. Borregaard, M.K. *et al.* (2017) Oceanic island biogeography through the lens of the general dynamic model: assessment and prospect Oceanic island biogeography. *Biol. Rev.* 92, 830–853
19. Fordham, D.A. *et al.* (2014) Better forecasts of range dynamics using genetic data. *Trends Ecol. Evol.* 29, 436–443
20. Leroux, S.J. *et al.* (2013) Mechanistic models for the spatial spread of species under climate change. *Ecol. Appl.* 23, 815–828

21. Craver, C. and Tabery, J. (2017) Mechanisms in science. In *The Stanford Encyclopedia of Philosophy* (Zalta, E.N., ed.), Metaphysics Research Lab, Stanford University
22. Glennan, S. (2017) *The New Mechanical Philosophy,* Oxford University Press
23. Kearney, M. and Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecol. Lett.* 12, 334–350
24. Triantis, K.A. *et al.* (2015) Diversity regulation at macro-scales: species richness on oceanic archipelagos. *Glob. Ecol. Biogeogr.* 24, 594–605
25. Etienne, R.S. (2005) A new sampling formula for neutral biodiversity. *Ecol. Lett.* 8, 253–260
26. Thibaut, L.M. *et al.* (2012) Diversity and stability of herbivorous fishes on coral reefs. *Ecology* 93, 891–901
27. Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* 63, 951–970
28. Loader, S.P. *et al.* (2014) Persistence and stability of Eastern Afromontane forests: evidence from brevicipitid frogs. *J. Biogeogr.* 41, 1781–1792
29. de Bruyn, M. *et al.* (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Syst. Biol.* 63, 879–901
30. Goldberg, E.E. *et al.* (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst. Biol.* 60, 451–465
31. Valente, L.M. *et al.* (2015) Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. *Ecol. Lett.* 18, 844–852
32. Hartig, F. *et al.* (2011) Statistical inference for stochastic simulation models – theory and application. *Ecol. Lett.* 14, 816–827
33. Sukumaran, J. *et al.* (2016) Machine learning biogeographic processes from biotic patterns: a new trait-dependent dispersal and diversification model with model choice by simulation-trained discriminant analysis. *Syst. Biol.* 65, 525–545
34. May, F. *et al.* (2016) Do abundance distributions and species aggregation correctly predict macroecological biodiversity patterns in tropical forests? *Glob. Ecol. Biogeogr.* 25, 575–585
35. Colwell, R.K. and Rangel, T.F. (2010) A stochastic, evolutionary model for range shifts and richness on tropical elevational gradients under quaternary glacial cycles. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 3695–3707
36. Zurell, D. *et al.* (2015) Individual-based modelling of resource competition to predict density-dependent population dynamics: a case study with white storks. *Oikos* 124, 319–330
37. Kearney, M. (2012) Metabolic theory, life history and the distribution of a terrestrial ectotherm. *Funct. Ecol.* 26, 167–179
38. Vandermeer, J.H. (1969) Competitive structure of communities – an experimental approach with protozoa. *Ecology* 50, 362–371
39. Rosindell, J. *et al.* (2010) Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol. Lett.* 13, 716–727
40. Gutiérrez, A.G. *et al.* (2016) Using a dynamic forest model to predict tree species distributions. *Glob. Ecol. Biogeogr.* 25, 347–358
41. Hartig, F. *et al.* (2012) Connecting dynamic vegetation models to data – an inverse perspective. *J. Biogeogr.* 39, 2240–2252
42. Etienne, R.S. (2007) A neutral sampling formula for multiple samples and an 'exact' test of neutrality. *Ecol. Lett.* 10, 608–618
43. Bode, M. *et al.* (2012) Species differences drive nonneutral structure in Pleistocene coral communities. *Am. Nat.* 180, 577–588
44. McGill, B.J. *et al.* (2005) Community inertia of quaternary small mammal assemblages in North America. *Proc. Natl. Acad. Sci. U. S. A.* 102, 16701–16706
45. Ricklefs, R.E. (2006) The unified neutral theory of biodiversity: do the numbers add up? *Ecology* 87, 1424–1431
46. McGill, B. (2003) Strong and weak tests of macroecological theory. *Oikos* 102, 679–685
47. Connolly, S.R. *et al.* (2017) A unified model explains commonness and rarity on coral reefs. *Ecol. Lett.* 20, 477–486
48. Gravel, D. *et al.* (2011) Trophic theory of island biogeography. *Ecol. Lett.* 14, 1010–1016
49. Cazelles, K. *et al.* (2016) On the integration of biotic interaction and environmental constraints at the biogeographical scale. *Ecography* 39, 921–931
50. Connolly, S. (2005) Process-based models of species distributions and the mid-domain effect. *Am. Nat.* 166, 1–11
51. McKane, A.J. *et al.* (2004) Analytic solution of Hubbell's model of local community dynamics. *Theor. Popul. Biol.* 65, 67–73
52. Connolly, S.R. (2009) Macroecological theory and the analysis of species richness gradients. In *Marine Macroecology* (Witman, J. D. and Roy, K., eds), pp. 279–309, University of Chicago Press
53. Binzer, A. *et al.* (2016) Interactive effects of warming, eutrophication and size structure: impacts on biodiversity and food-web structure. *Glob. Change Biol.* 22, 220–227
54. Blanchard, J.L. *et al.* (2014) Evaluating targets and trade-offs among fisheries and conservation objectives using a multispecies size spectrum model. *J. Appl. Ecol.* 51, 612–622
55. Kalyuzhny, M. *et al.* (2015) A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecol. Lett.* 18, 572–580
56. Keith, S.A. and Connolly, S.R. (2013) Effects of diversity-dependent colonization-extinction dynamics on the mid-domain effect. *Glob. Ecol. Biogeogr.* 22, 773–783
57. Tittensor, D.P. and Worm, B. (2016) A neutral-metabolic theory of latitudinal biodiversity. *Glob. Ecol. Biogeogr.* 25, 630–641
58. Bertuzzo, E. *et al.* (2016) Geomorphic controls on elevational gradients of species richness. *Proc. Natl. Acad. Sci. U. S. A.* 113, 1737–1742
59. Borregaard, M.K. *et al.* (2016) The general dynamic model: towards a unified theory of island biogeography? *Glob. Ecol. Biogeogr.* 25, 805–816
60. Boeye, J. *et al.* (2014) Habitat structure mediates spatial segregation and therefore coexistence. *Landsc. Ecol.* 29, 593–604
61. Bocedi, G. *et al.* (2014) RangeShifter: a platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes. *Methods Ecol. Evol.* 5, 388–396
62. Kubisch, A. *et al.* (2014) Where am I and why? Synthesizing range biology and the eco-evolutionary dynamics of dispersal. *Oikos* 123, 5–22
63. Zurell, D. *et al.* (2016) Benchmarking novel approaches for modelling species range dynamics. *Glob. Change Biol.* 22, 2651–2664
64. Pagel, J. and Schurr, F.M. (2012) Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Glob. Ecol. Biogeogr.* 21, 293–304
65. Freilich, M.A. and Connolly, S.R. (2015) Phylogenetic community structure when competition and environmental filtering determine abundances. *Glob. Ecol. Biogeogr.* 24, 1390–1400
66. Münkemüller, T. *et al.* (2012) From diversity indices to community assembly processes: a test with simulated data. *Ecography* 35, 468–480
67. Münkemüller, T. and Gallien, L. (2015) VirtualCom: a simulation model for eco-evolutionary community assembly and invasion. *Methods Ecol. Evol.* 6, 735–743
68. Colwell, R.K. and Winkler, D.W. *et al.* (1984) A null model for null models in biogeography. In *Ecological Communities: Conceptual Issues and the Evidence* (Strong, D.R., ed.), pp. 344–359, Princeton University Press
69. Botta-Dukát, Z. and Czúcz, B. (2016) Testing the ability of functional diversity indices to detect trait convergence and divergence using individual-based simulation. *Methods Ecol. Evol.* 7, 114–126
70. Kilpatrick, A.M. and Ives, A.R. (2003) Species interactions can explain Taylor's power law for ecological time series. *Nature* 422, 65–68
71. Maury, O. and Poggiale, J.-C. (2013) From individuals to populations to communities: a dynamic energy budget model of marine ecosystem size-spectrum including life history diversity. *J. Theor. Biol.* 324, 52–71
72. Hartvig, M. *et al.* (2011) Food web framework for size-structured populations. *J. Theor. Biol.* 272, 113–122

73. Connolly, S.R. *et al.* (2014) Commonness and rarity in the marine biosphere. *Proc. Natl. Acad. Sci. U. S. A.* 111, 8524–8529
74. Polanyi, M. (1964) *Science, Faith and Society,* University of Chicago Press
75. Brown, J.H. (1995) *Macroecology,* University of Chicago Press
76. Hempel, C.G. and Oppenheim, P. (1948) Studies in the logic of explanation. *Philos. Sci.* 15, 135–175
77. Giere, R.N. (1988) *Explaining Science: A Cognitive Approach,* University of Chicago Press
78. Frigg, R. and Hartmann, S. (2017) Models in science. In *The Stanford Encyclopedia of Philosophy* (Zalta, E.N., ed.), Metaphysics Research Lab, Stanford University
79. Paslaru, V. (2017) Mechanisms in ecology. In *The Rutledge Handbook of Mechanisms and Mechanical Philosophy* (Glennan, S. and Illari, P., eds), pp. 348–361, Taylor and Francis
80. Tucker, C.M. *et al.* (2016) Differentiating between niche and neutral assembly in metacommunities using null models of β-diversity. *Oikos* 125, 778–789
81. Schoener, T.W. (1986) Mechanistic approaches to community ecology – a new reductionism. *Am. Zool.* 26, 81–106
82. Turchin, P. (1999) Population regulation: a synthetic view. *Oikos* 84, 153–159
83. Lakatos, I. (1978) *The Methodology of Scientific Research Programmes,* Cambridge University Press
84. Godfrey-Smith, P. (2003) *Theory and Reality: An Introduction to the Philosophy of Science,* University of Chicago Press
85. Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography,* Princeton University Press
86. Etienne, R. and Haegeman, B. (2011) The neutral theory of biodiversity with random fission speciation. *Theor. Ecol.* 4, 87–109
87. Volkov, I. *et al.* (2007) Patterns of relative species abundance in rainforests and coral reefs. *Nature* 450, 45–49
88. Rosindell, J. *et al.* (2008) A coalescence approach to spatial neutral ecology. *Ecol. Inform.* 3, 259–271
89. Kooijman, S.A.L.M. *et al.* (2008) From food-dependent statistics to metabolic parameters, a practical guide to the use of dynamic energy budget theory. *Biol. Rev.* 83, 533–552
90. Schwarzkopf, L. *et al.* (2016) One lump or two? Explaining a major latitudinal transition in reproductive allocation in a viviparous lizard. *Funct. Ecol.* 30, 1373–1383
91. Kooijman, S.A.L.M. (2014) Metabolic acceleration in animal ontogeny: an evolutionary perspective. *J. Sea Res.* 94, 128–137