ECOGRAPHY

Forum

A parsimonious view of the parsimony principle in ecology and evolution

Marco Túlio P. Coelho, José Alexandre Diniz-Filho and Thiago F. Rangel

M. T. P. Coelho (http://orcid.org/0000-0002-7831-3053) ✉ (marcotpcoelho@gmail.com), J. A. Diniz-Filho and T. F. Rangel, Depto de Ecologia, Univ. Federal de Goiás, Goiânia, Goiás, Brazil. MTPC also at: Dept of Human Dimensions of Natural Resources, Colorado State Univ., Fort Collins, CO, USA.

Ecography
42: 968–976, 2019
doi: 10.1111/ecog.04228

Subject Editor: Pablo A. Marquet
Editor-in-Chief:
Jens-Christian Svenning
Accepted 19 November 2018

The idea that simplicity of explanation is important in science is as old as science itself. However, scientists often assume that parsimonious theories, hypothesis and models are more plausible than complex ones, forgetting that there is no empirical evidence to connect parsimony with credibility. The justification for the parsimony principle is strongly dependent on philosophical and statistical inference. Parsimony may have a true epistemic value in the evaluation of correlative and predictive models, as simpler models are less prone to overfitting. However, when natural mechanisms are explicitly modelled to represent the causes of biological phenomena, the application of the parsimony principle to judge the plausibility of mechanistic models would entail an unsupported belief that nature is simple. Here, we discuss the challenges we face in justifying, measuring, and assessing the trade-off between simplicity and complexity in ecological and evolutionary studies. We conclude that invoking the parsimony principle in ecology and evolution is particularly important in model-building programs in which models are viewed primarily as an operational tool to make predictions (an instrumentalist view) and in which data play a prominent role in deciding the structure of the model. However, theoretical advances in ecology and evolutionary biology may be derailed by the use of the parsimony principle to judge explanatory mechanistic models that are designed to understand complex natural phenomena. We advocate a parsimonious use of the parsimony principle.

Keywords: parsimony principle, statistical inference, Ockham's razor, correlative models, mechanistic models

Introduction

The idea that simple (parsimonious) theories are important in science is as old as science itself. It is difficult to disagree that simple theories, hypothesis and models are elegant, easy to understand and to remember. However, rooted in the parsimony principle, scientists often assume that simple theories, hypothesis and models are more plausible or reliable, although there is no empirical evidence to connect simplicity with credibility (Quine 1963). For example, in the quest to understand the causes of complex biodiversity patterns, such as spatial diversity gradients, ecologists and

NORDIC SOCIETY OIKOS

www.ecography.org

evolutionary biologists study how multiple processes, interacting over space and time, produce complex biological phenomena across all scales (Wiens and Donoghue 2004, Rangel et al. 2007, 2018, Warren et al. 2014, Rosindell et al. 2015, Connolly 2017, Coelho et al. 2018). However, even acknowledging that biodiversity patterns are immensely complex, and that disentangling their causes is an extremely challenging endeavor (Vellend 2010), simplicity has always had a special role in the judgment of ecological and evolutionary theories (Marquet et al. 2014), hypothesis (Hilborn and Stearns 1982) and models (Holling 1966, Evans et al. 2013, Houlahan et al. 2017). Yet, little discussion is found in ecology and evolution about how to justify, measure and trade-off simplicity.

The ecological and biogeographical phenomena (e.g. uneven distribution of species over space) do not provide universal opportunity for experimental research. Thus, ecologists and evolutionary biologists are usually limited to observational studies, in which curve-fitting techniques are employed to measure the statistical correspondence between observed biological patterns and suspected causal factors. In many cases, the main interest is to elect the most plausible model that fits to the data in hand, as the mathematical representation of a theory or hypothesis (Box 1). If goodness-of-fit was the only criterion to judge the plausibility of models, a complex model that best fits to the data would be assumed as the most plausible. However, parsimony is also viewed as valuable criterion to evaluate the plausibility of alternative models (Box 1). The main question that emerges from our scientific practice is whether/how parsimony can be used to understand natural processes. Although simple theories, hypothesis and models often strike us as beautiful and easier to understand (Quine 1963, Keuzenkamp et al. 2004), our psychological comfort when dealing with simpler explanations for complex phenomena is not a valid epistemic justification for its general use as a criterion to judge reliability (Pearl 1978, Sober 1981, 2015). For centuries scientists and philosophers have been struggling to find epistemic justifications for the parsimony principle, but no consensus has arisen yet. In this paper, we show that ecology and evolutionary biology employ different justifications for the role of parsimony in alternative inferential frameworks, such as frequentist, Bayesian, likelihood and multi-model inference. We also provide a historical perspective to the parsimony principle, discussing its role in predicting versus explaining nature, and its application to evaluate correlative and mechanistic models. We show that 1) there is an instrumental value in the use of parsimony to judge models that are designed for predictive purposes, but 2) the application of the parsimony principle to evaluate explanatory models is counterproductive, as it may prevent the elucidation of truly complex phenomena.

Box 1. The curve-fitting problem

Since Laplace and Gauss, curve-fitting is one of the most universal methodologies employed in scientific research (Mulaiktt 2001). Given the observations made during an experimental (or observational) study of the correspondence between a predictor (x) and a response variable (y), an infinite number of hypotheses (curves, Fig. 1) can be proposed to summarize and explain the relationship between the variables (Fig. 1). However, praxis is the choosing of only one hypothesis, based on an arbitrary balance between the criteria of goodness-of-fit and simplicity. The curve that best describes the data is always the one that comes closest to the largest number of observations, therefore the best-fit curve. Thus, if goodness-of-fit was the only criteria, one would have to choose the curve that connects all data points (dashed curve, Fig. 1). However, the geometric irregularity of the best-fitting curve indicates the mathematical complexity of the hypothesis, which is thus considered not parsimonious. Conversely, the simplest hypothesis is a straight line (red line, Fig. 1), which is far more regular but fails to account for all the available information about the relationship between the variables. Thus, the simplicity criterion is used as a counterweight to the goodness-of-fit criterion, and vice-versa. However, there are infinite alternatives between the simplest and best-fit curves (e.g. blue curve, Fig. 1). The use of simplicity as an additional criterion to select the most plausible hypothesis creates the philosophical problem of how to measure and justify the criterion. Because simplicity is so loosely defined and justified, the final choice is based on an arbitrary evaluation of how well the line fits the observation and how simple it personally feels to the scientist.

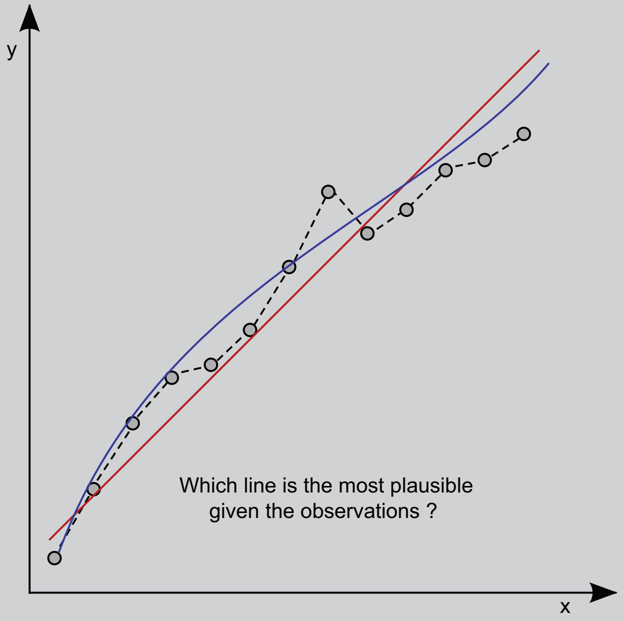

Figure 1. The curve-fitting problem. The choice of the most plausible hypothesis is based on an arbitrary balance between goodness-of-fit and simplicity criteria. The red curve is the simplest, indicating a linear relationship between the two variables, but also the worst-fit. The dashed line is the most complex, but also the best-fit. The blue line is just one of the infinite possible balances between the best-fit and simplest hypotheses.

Historical perspectives of the parsimony principle

Traditionally, most of the slogans used to discuss simplicity are found in William of Ockham's (1285–1348) writings such as 'It is futile to do with more what can be done with fewer', or 'plurality should not be posited without necessity' (Brown 1950, Sober 2015). Not surprisingly, the Ockham's razor criteria for judging competitive theories, hypothesis and models is one of the most popular names for the parsimony principle. However, the origins of the parsimony principle can be traced back at least to Aristotle, whom invoked the principle in many of his writings (Brown 1950, Keuzenkamp et al. 2004). Until the 20th century many philosophers and scientists believed that Ockham's razor was epistemically relevant because simplicity was assumed to be a true feature of nature (e.g. Descartes, Leibniz, Newton and Kant; Sober 2015). Some of them proposed theistic justifications for the parsimony principle (Descartes, Leibniz, Newton), arguing that the Abrahamic God had created a simple world, with simple rules (Sober 1981, 2015), and that understanding nature was equivalent of understanding the mind of God. Thus, the assumption of a simple nature would be a sufficient epistemic justification for the parsimony principle in science. However, other philosophers (e.g. Hume, Mill, Morgan) painted a different picture of parsimony. In their perspective, parsimony is viewed as necessary to make predictions about the world, without any assumption or assurance that nature is indeed simple (Sober 2015). Thus, parsimony would have an operational and methodological justification for scientific research, without any assertion of how nature works.

The rise of modern statistics in the 20th century transformed science into a data-driven investigation of natural phenomena, reducing the emphasis of narrative accounts and description/categorization of observables. However, modern statistics gave rise to many alternative inferential strategies, which differ not only methodologically, but also epistemologically. The epistemological distinction of alternative inferential strategies of modern statistics unchained the concept of plausibility from parsimony, therefore creating multiple directions in the search for a conceptual justification of the parsimony principle.

The parsimony principle in alternative frameworks of statistical inference

The most common use of Fisher's significance test is the assumption of a parsimonious null hypothesis as a premise (default position). In critical tests of this kind the null hypothesis invokes the least possible complexity by assuming the absence of relationship, association or effect among measured phenomena. Thus, the goal of the analysis is to estimate the degree of compatibility between the observed data and the scenario in which the null hypothesis is assumed to be true (Sokal and Rohlf 1981, Efron and Tibshirani 1994). The null hypothesis may only be rejected if the observed data is substantially (i.e. 'significantly') different from the expectation of the parsimonious null scenario. Because Fisher is also the proponent of the much flexible likelihood theory, it is arguable that he proposed the test of the null hypothesis for its mathematical convenience in analysis of experimental results, in which the experimental manipulation (i.e. treatment) is assumed to produce no effect. In addition, it is easier to derive the probability distribution of the clearly defined null hypothesis, as opposed to the infinite alternative hypotheses in which experimental treatment would promote some undefined degree of effect. Thus, the parsimony of the null hypothesis is a methodological strategy, rather than a statement about the reality of nature.

Parsimony also plays a prominent role in the comparison of models within the framework of model selection and multi-model inference (Burnham and Anderson 2001, 2002). Some information criteria (i.e. Akaike information criterion – AIC, corrected Akaike information criterion – AICc, kernel-based information criterion – KIC and empirical information criterion – EIC) are designed to asymptotically maximize the expected relative predictive accuracy of models (i.e. the ability to fit future data, Aho et al. 2014). Notably, the AIC have been widely used in ecological and evolutionary research in the last decade (MacNally et al. 2018). The simplicity criterion is mathematically relevant to access the relative predictive accuracy of models because very complex models tend to have poor predictive power (i.e. the curse of dimensionality; Burnham and Anderson 2001, Sober 2002, 2004a, b, Hastie et al. 2009, James et al. 2013). Because complex models tend to over-fit the data, therefore confounding noise with signal, AIC may be used to penalize model complexity to identify the best balance between goodness-of-fit and simplicity. Thus, parsimony has a justifiable epistemic value to gauge the relative predictive accuracy of a model. However, model selection does not downgrade complex models because they are less plausible than simpler models, but because they are often a poor choice for prediction of future data. In addition, model selection does not provide absolute estimates of predictive accuracy, but only a relative measure given the alternative models under consideration. Although model selection with AIC became a popular inferential strategy in ecology and evolutionary biology (MacNally et al. 2018), few models are designed for predictive purposes, as prediction is traditionally not a frequent goal in those sciences (Houlahan et al. 2017).

Some theoretical, simulation and applied studies have shown that AIC is not as parsimonious as once believed. Because AIC is designed to avoid underfitting errors (i.e. potential omissions of processes that do influence the phenomena under study), it becomes prone to overfitting errors (i.e. overrating models that include processes that do not influence the phenomena) (Taper 2004). However, there are more parsimonious alternative information criteria, which seek to avoid overfitting errors by accepting a higher rate of underfitting errors. For example, consistent AIC (CAIC), Bayesian information criterion (BIC), Hannan–Quinn information (HQC) and many others, can asymptotically identify the model that is closest to the true generating process, given a set of candidate models. As a result, this group of information criteria tend to select compact models, composed mostly of processes with strong influence in the phenomena under study, but potentially missing processes with minor effects.

In contrast with information theoretical approach, Bayesians aim to estimate the probability of different theories, given the available evidence and current knowledge (Ellison 2004), as they equate probability of a theory with its degree of plausibility. Thus, their concept of probability is a measure of credibility, certainty or belief (Horwich 1982, Hilborn and Mangel 1997, Hacking 2001, Ellison 2004). The Bayes theorem is used as a model for the learning process, indicating how scientists should update their knowledge in the light of new evidence. Thus, with the Bayes theorem one can identify the most probable theory, hypothesis or model, given the alternatives. However, parsimony does not play any role in traditional Bayesian estimation of the probability of a theory, as the simplicity of the theory does not increase its probability (Sober 2015).

Although parsimony is not a core principle of Bayesian inference, Jeffrey's postulate of simplicity suggests that when multiple alternative hypotheses are confronted, simple theories should be given higher prior probability (i.e. the degree of justified belief of a researcher in a hypothesis before new evidence of the studied phenomenon is uncovered) (Jeffrey 1965). However, Jeffrey does not provide a valid epistemic justification for his approach, just a heuristic methodological strategy to use the parsimony principle in Bayesian inference. Indeed, Jeffreys' framework is not commonly applied by Bayesians, and one should search for the most probable hypothesis based only on available knowledge and new evidence (Gelman et al. 2013, Sober 2015). However, there are multiple Bayesian techniques used for model selection that weight models by their goodness-of-fit and simplicity (Hooten and Hobbs 2015). Yet, those methods are seldom employed in ecology, and few discussions about their validity emerged in applied ecological problems. Some Bayesian techniques used in model selection justify the use of parsimony based on predictive accuracy, while others offer heuristic justification similar Jeffrey's postulate of simplicity (Hooten and Hobbs 2015, Sober 2015). Discussions of the correct use of Bayesian model selection emerged in social sciences where it is argued that performing Bayesian model selection is unimportant when the main task is to construct realistic models for complex systems that agree with both theory and data (Gelman and Rubin 1995), which seems to be what ecologists crave. Finally, ecologists frequently use AIC for model selection, instead of alternative Bayesian information criteria (MacNally et al. 2018), which makes the discussion about AIC even more relevant in ecology and evolution.

Likelihoodists and Bayesians share a similar perspective on how to deal with parsimony, but they disagree in the use of prior probabilities. While Bayesians attempt to carry knowledge over time, updating it as new information arises, likelihoodists see no value in the use of prior probabilities. For them, if the model with highest likelihood happens to be the simplest, then parsimony happens to mirror evidential support, but the likelihood of a hypothesis is still proportional to the probability of observing the data if that hypothesis was true (Sober 2004a, b).

Finally, as the parsimony principle started to be unchained from plausibility, it is no longer viewed as a pre-20th century magical argument used by classical philosophers and scientists. However, there is no universal justification for the parsimony principle. As different frameworks of statistical inference have different epistemic goals, the parsimony principle is invoked, or not, under different epistemic justification.

Simplicity and falsifiability in ecology and evolution

Karl Popper provided a different epistemic justification for parsimony by arguing that simple theories are more falsifiable (Popper 1959). According to him, because the goal of empirical science should be the refutation of theories, parsimonious theories are desirable because they are in principle easier to be refuted. For example, Popper's 'bold hypotheses' are parsimonious because they use few assumptions to make a myriad of predictions about natural phenomena, being easier to be refuted in case any of the predictions fail. However, according to Popper, if a theory is evaluated repeatedly over time and not refuted, it should not be considered more plausible, but only to have been more corroborated. Although most philosophers of science consider the concept of corroboration too obscure for a useful application in research (Godfrey-Smith 2003), it is frequently assumed that the most plausible theory are those that have been most corroborated (Sober 2004b).

One could argue that ecologists give great importance to parsimony when judging theories, hypothesis and models because of the Popperian argument that simple theories are easier to falsify. However, it is difficult to argue that ecologist and evolutionary biologists are strict Popperians, such that one single critical test that contradicts a model is sufficient to falsify its underlying theory (here not including 'naive falsifications'; Godfrey-Smith 2003). Let's take as an example the latitudinal gradient of species diversity, ecology oldest pattern (Hawkins 2001). It does not seem that we have falsified any hypothesis over time, which can be noticed by analyzing Pianka's (1966) seminal 'Latitudinal gradients in species diversity: a review of concepts' that still captures currently discussed hypotheses for latitudinal diversity gradients. Ecologists and evolutionary biologists work simultaneously with a plethora of competing theories, hypotheses and models, but rarely discard any, even when conflicting empirical evidence is found. Indeed, ecology and evolutionary biology seems to be best described by Lakatos' competition among multiple scientific programs (Godfrey-Smith 2003), in which scientific progress is associated with gradual improvements of theories that continuously explain the available empirical data. Also, as all scientists, we are deeply attached to our theories (Ginzburg and Jensen 2004). Thus, the importance that a scientist gives to a certain characteristic of a theory, hypothesis or model (e.g. specificity vs generality, simplicity vs complexity) is very much a matter of personal taste (Kuhn 1977).

The parsimony principle in predicting versus explaining nature

Does science truly explain how nature works, or does it identify patterns in natural phenomena? These are two opposite philosophical views about the role of science on understanding nature (Okasha 2002, Godfrey-Smith 2003). On the one hand, the scientific realism argues that the goal of science is to provide a description or explanation of the true underlying reality of nature. On the other hand, scientific instrumentalism holds that scientific theories and models are just instruments to assist researchers to predict natural phenomena. Thus, from an instrumentalist point of view, theories are used to provide predictions of observable data, and there is no contradiction if alternative theories are just as accurate. Conversely, under the realistic perspective, there is only one true theory, which perfectly describes the underlying mechanisms of nature.

From the realistic perspective, invoking parsimony to evaluate a theory is a metaphysical assumption that nature is indeed simple, like many pre 20th century philosophers and scientists defended. However, under an instrumentalist view of science, the Ockham's razor is a methodological criterion to judge among the infinite models that can fit the data (Box 1), as well as choosing the most reliable model to predict new data (higher predictive accuracy). These two different epistemic views of science, which are coupled with alternative uses and justifications of parsimony principle to evaluate theories, are present in discussions of alternative models to reconstruct phylogenetic history within evolutionary biology (Box 2). In fact, the core of that debate was the epistemic value of two alternative model building programs in ecology and evolution: models to describe nature (i.e. correlative) or models to explain nature (i.e. mechanistic) (Gotelli et al. 2009).

Correlative models are designed to probe nature and find functional relationships (patterns) among observed phenomena, but no causal story is built into the model. In fact, the data used to fit the model is more important than theory itself, as the statistical methods are used only in summarizing observed data, without explicitly including theories a priori (Gotelli et al. 2009, Pearl and Mackenzie 2018). Because the correlative model-building program is instrumentalist in its essence, the epistemic justification of parsimony is also conceptually tied to an instrumentalist view of science. Thus, parsimony should be invoked by ecologists and evolutionary biologists to judge competing correlative models because predictive accuracy is the main goal of such models. Since reliable predictions of highly complex phenomena and noisy data can be difficult to achieve (Burnham and Anderson 2001, Hastie et al. 2009), applying the parsimony principle to judge between the infinite alternative models reduces the probability of confounding noise with signal. Thus, the parsimony principle within the instrumentalist framework is justifiable only in predictive science, which is not the strong characteristic of ecology (Houlahan et al. 2017).

In contrast, mechanistic models have been used as vehicles for theoretical exploration, proposing causal explanation of a biological phenomenon (Peck 2004, Connolly et al. 2017, Rangel et al. 2018). Contrary to purely correlative models, the parameters of mechanistic models are designed with biological meaning, therefore contributing to the interpretation of the modelled phenomena (Gotelli et al. 2009). Therefore,

Box 2. The use of parsimony in evolutionary biology

Parsimony principle has been intensely discussed in evolutionary biology within the context of phylogenetic reconstruction (Sober 1989, 2015, Felsenstein 2004). A parsimonious model of trait evolution assumes that the fewest changes in a given trait has occurred over evolutionary history. Thus, one can apply the parsimony criterion as a realistic evolutionary process to estimate the most plausible phylogenetic tree from the observation of traits of a set of species (Felsenstein 2004). The puzzle that tormented evolutionary biologists was how to justify the use of the parsimony principle as an evolutionary model (Sober 1989, 2015). A realistic view of the parsimony principle was invoked by Camin and Sokal (1965) when affirming that the correctness of their method of phylogenetic reconstruction 'depends on the assumption that nature is, indeed, parsimonious'. Others provided a methodological justification for parsimony, arguing that parsimony could mirror likelihood (Cavalli-Sforza and Edwards 1967) which was later shown to occur only under specific assumptions (e.g. drift is the process governing trait evolution (Felsenstein 1973, 1978). Popperian arguments were also invoked by suggesting that the most parsimonious tree is easier falsifiable by data (Wiley 1975, 2011). The problem with the Popperian argument is that data used in phylogenetic reconstruction cannot, in any way, falsify the hypothesis of tree topology (Sober 2015). Later, an instrumentalist view of parsimony was clarified by suggesting that parsimony is not necessarily an assumption about how evolution occurs in nature, but just as a simple and useful methodology for phylogenetic reconstruction (Sober 1989), although not the only methodology available (Felsenstein 2004).

The real weakness of the use of parsimony in evolutionary biology is the impossibility of incorporating knowledge from different processes of evolution to reconstruct a phylogenetic tree (Yang and Rannala 2012). The sophisticated evolutionary models allowed by Bayesian and likelihood methods began to gradually replace parsimony, especially when molecular data is used in phylogenetic reconstruction. Parsimony remains a controversial issue in evolutionary biology (Yang and Rannala 2012), although it is still used in phylogenetic inference, especially when dealing with discrete characters (i.e. paleontology, behavioral sciences). Today, most phylogenies are reconstructed using statistical procedures that explicitly assume multiple models of evolutionary processes and make no use of parsimony (Felsenstein 2004, Sober 2015).

theory is more important than data for mechanistic modeling, since the final goal is to propose or evaluate the theory behind the ecological and evolutionary processes built into the model (Rangel et al. 2018). Philosophers and modelers argue that those models defy the restrictions found in common correlative models (Winsberg 1999, 2001), because they represent an attempt to realistically describe nature. Thus, invoking the parsimony principle to judge the plausibility a mechanistic model implies the belief that nature is simple, with few basic processes driving complex observed natural patterns.

Correlative and mechanistic models must not be compared using the parsimony criterion, as they emerge from alternative philosophical approaches, with different scientific goals (Winsberg 1999, 2001). Mechanistic models are in general more complex than correlative models by design, as they attempt to realistically incorporate and integrate complex causal processes. There are, of course, correlative models that are more complex than mechanistical models (e.g. geographically weighted regressions; Fotheringham et al. 2002), assuming that the number of adjustable parameters is a valid measure of complexity (Box 3).

One should not expect simplicity from mechanistic models that attempt to unify different areas of knowledge in ecology and evolution (Cabral et al. 2017), as ecological and evolutionary patterns at high organizational levels (e.g. latitudinal diversity gradients) are among the most complex natural phenomena. However, as Hilborn and Stearns (1982) affirmed over thirty years ago, we face a pathology of avoiding complex models, which promotes a major difficulty of publishing conclusions based on integrative mechanistic models (Evans et al. 2013). Neglecting complex models that try to understand and explain causal relationships among ecological and evolutionary processes prevents significant advances in our understanding of nature.

Because mechanistic models are attempts to explain nature as a vehicle for theoretical exploration, these models should be judged by their consistency and consilience (i.e. how well the built-in processes and assumptions of the model capture what is currently accepted by scientific community as valid explanation, or explanations with higher potential to integrate theories from different research fields), accuracy (i.e. how well the model reproduces empirical patterns) and precision (i.e. the amount of variation in model predictions among replicates of the model) (Gotelli et al. 2009). Conversely, evaluating mechanistic models by contrast against correlative models is analogous to comparing apples with oranges.

Notice that not employing the parsimony principle in the judgment of mechanistic models does not, in any way, neglect or denies the trade-offs between specificity and generality of model building strategies (Levins 1966, Grimm 2005). However, the traditional view that simple models are

Box 3. Measuring simplicity

Parsimony has a practical value in science but has little support for its epistemic value (Sober 2015), especially when complexity is expected in natural systems. The use of the parsimony principle as a valid criterion in scientific practice raises the problem of how to measure simplicity. The problem begins, first and simply, when using different communication formats. The degree of complexity used to describe reality through theories, models and hypothesis depends of language (Pearl 1978). Thus, our perception of complexity varies greatly according to the language used. Although it is assumed that models and hypotheses are described following a standardized scientific language, valid models and hypotheses can be proposed in a myriad of formats (e.g. verbal, deterministic equation, stochastic equation, computer simulation, etc).

Traditionally, the complexity of a model is measured using two criteria: 1) the number of adjustable parameters or, 2) the number of assumptions. However, even for very simple models these two criteria may lead to conflicting evaluations. Consider, for example, the linear (A) and parabolic (B) models:

$$(\boldsymbol{A})\; y = a + bx$$

$$(\boldsymbol{B})\; y = a + bx + cx^2$$

The linear relationship is a special case of the parabolic, where $c=0$. According to the criterion of simplicity based on the least number of parameters, the linear model is the simplest, as it has two adjustable parameters (a and b) while the parabolic model has three parameters (a, b and c). However, the fewer number of adjustable parameters in the linear model may also be considered as an assumption about a natural process (i.e. x^2 does not affect y, therefore $c=0$). Thus, one should deem the parabolic model as the simpler regarding the number of assumptions, as it has one less assumption than the linear model. Measuring complexity by counting the number of parameters gained much popularity in inferential criteria. However, model complexity in statistics depend on many factors such as parameter redundancy, parameter stability, random error structure of the model, linearity and nonlinearity of parameters and many others (Bozdogan 2000, Taper 2004).

Although it is common practice in ecology and evolution to evaluate models based on their simplicity, it is important to notice that it is difficult not only to justify the use of simplicity in model building and selection, but it is especially difficult to measure and compare the simplicity of models. While the parsimony of ecological theories is sometimes judged and compared without a clear epistemological justification, completely ignoring parsimony may lead to uninterpretable over-fit models, with hundreds of adjustable parameters (Phillips et al. 2006, Rangel and Loyola 2012).

the safest route towards a general theory is another common misconception, which has been identified as an obstacle for advance of ecological and evolutionary sciences (Evans et al. 2013). Both general and specific models may be complex or simple, as specificity and generality refer to the range of natural phenomena explained by the model, whereas complexity and simplicity refer to the number of assumptions and/or processes that are incorporated by the model. As Dayton (1973) affirm, simple models offer an easier shortcut for generalization, but if they are not based on a deep understanding of natural processes, then they are fated to produce the right predictions for the wrong reasons. Not rarely correlative models have higher predictive accuracy than mechanistic models, but they usually provide poor additional understanding of causal processes (Rangel and Loyola 2012).

Finally, when mechanistic models are used to perform predictions of natural patterns, then they may be compared against alternative models that attempt to predict the same patterns. Thus, parsimony reinserts itself into explanatory modelling as an instrumental criterion to identify competing models with poor predictive accuracy, even under the assumption that nature is complex. Greater evidential support for a model comes from the ability to predict new data than from the ability to match existing data. Echoing many before us, ecology and evolution should be more committed to predictions (Houlahan et al. 2017). However, one should always consider if the empirical patterns used to validate the model aren't too simple, because, if they are, then they could also be reproduced by some competing model, regardless of complexity. Yet, measuring the complexity or simplicity of models is conceptually and methodologically challenging (Box 3).

The parsimonious view of the parsimony principle

In this paper we discussed the challenge of justifying, measuring and trading-off simplicity/complexity. The parsimony principle can be dangerously used as a silver bullet to solve any argument when contrasting theories, hypothesis and models. Of course, we do not view the instrumental use of simple models as a problem. Indeed, they require less effort to fit and specially to understand. However, the usefulness of simplicity is not a valid justification to judge reliability. Ecologists would benefit from evidential statistics (Taper and Ponciano 2016) when evaluating their theories, models and hypothesis. Evidential statistics has in its core a measure of strength of evidence of a proposition about the studied phenomenon, and is explicit when not enough data is available to reach a conclusion (i.e. indeterminacy). Unfortunately, the framework of evidence statistics is new and has received little attention in ecology and evolution (Taper and Ponciano 2016).

Invoking the parsimony criterion in ecology and evolution is particularly important when building and comparing correlative models, in which theory plays only a modest role, whereas data is critical to decide the structure of the model. Thus, the parsimony principle tends to become even more important with the growing scientific use of big data, deep learning and artificial intelligence techniques, in which scientists use computing power to automate the building of predictive models, without emphasis in understanding or explaining natural phenomena. However, when natural mechanisms are explicitly modelled as representations of causal explanation of biological phenomena, the application of the parsimony principle to judge the plausibility of mechanistic models would imply a realistic belief in a simple nature. Thus, enforcing the parsimony principle in evaluating explanatory mechanistic models could halt the advance of ecology and evolutionary biology. We hope that ecologists and evolutionary biologists, in their roles as authors, reviewers and editors, use the parsimony principle parsimoniously.

Acknowledgements – We thank E. B. Pereira, J. F. M. Rodrigues, M. C. Gavin and P. Marquet for suggestions and constructive discussion during the development of the study. R. K. Colwell edited and substantially improved the abstract. Several ideas in this article were inspired by the graduate course in Philosophical Principles in Ecology and Evolution, taught by JAFD-F at Univ. Federal de Goiás.
Funding – MTPC is supported by a PhD scholarship provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, finance code 001). TFR and JAFD-F have been continuously supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grants PQ309550/2015-7 and PQ301799/2016-4). Our work was developed in the context of the National Inst. of Science and Technology (INCT) project in 'Ecology, Evolution and Biodiversity Conservation', supported by CNPq (grant 465610/2014-5) and FAPEG (grant 201810267000023).

References

Aho, K. et al. 2014. Model selection for ecologists: the worldview of AIC and BIC. – Ecology 95: 631–636.

Bozdogan, H. 2000. Akaike's information criterion and recent developments in information complexity. – J. Math. Psychol. 44: 62–91.

Brown, G. B. 1950. Science: its methods and its philosophy. – George Allen and Unwin.

Burnham, K. P. and Anderson, D. R. 2001. Kullback–Leibler information as a basis for strong inference in ecology. – Wildl. Res. 28: 111.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. – Springer.

Cabral, J. S. et al. 2017. Mechanistic simulation models in macroecology and biogeography: state-of-art and prospects. – Ecography 40: 267–280.

Camin, J. H. and Sokal, R. R. 1965. A method for deducing branching sequences in phylogeny. – Evolution 19: 311–326.

Cavalli-Sforza, L. L. and Edwards, A. W. F. 1967. Phylogenetic analysis: models and estimation procedures. – Evolution 21: 550–570.

Coelho, M. T. P. et al. 2018. Effects of neutrality and productivity on mammal richness and evolutionary history in Australia. – Ecography <https://doi.org/10.1111/ecog.03784>.

Connolly, S. R. et al. 2017. Process, mechanism, and modeling in macroecology. – Trends Ecol. Evol. 32: 835–844.

Dayton, P. K. 1973. Two cases of resource partitioning in an intertidal community: making the right prediction for the wrong reason. – Am. Nat. 107: 662–670.

Efron, B. and Tibshirani, R. J. 1994. An introduction to the bootstrap. – Chapman and Hall.

Ellison, A. M. 2004. Bayesian inference in ecology. – Ecol. Lett. 7: 509–520.

Evans, M. R. et al. 2013. Do simple models lead to generality in ecology? – Trends Ecol. Evol. 28: 578–583.

Felsenstein, J. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. – Syst. Zool. 22: 240–249.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. – Syst. Zool. 27: 401–410.

Felsenstein, J. 2004. Inferring phylogenies. – Sinauer Associates.

Fotheringham, A. S. et al. 2002 Geographically weighted regression: the analysis of spatially varying relationships. – Wiley.

Gelman, A. and Rubin, D. B. 1995. Avoiding model selection in Bayesian social research. – Sociol. Methodol. 25: 165–173.

Gelman, A. et al. 2013 Bayesian data analysis. – CRC Press.

Ginzburg, L. R. and Jensen, C. X. J. 2004. Rules of thumb for judging ecological theories. – Trends Ecol. Evol. 19: 121–126.

Godfrey-Smith, P. 2003. Theory and reality: an introduction to the philosophy of science. – Univ. of Chicago Press.

Gotelli, N. J. et al. 2009. Patterns and causes of species richness: a general simulation model for macroecology. – Ecol. Lett. 12: 873–886.

Grimm, V. 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. – Science 310: 987–991.

Hacking, I. 2001. An introduction to probability and inductive logic. – Cambridge Univ. Press.

Hastie, T. et al. 2009. The elements of statistical learning: data mining, inference and prediction. – Springer.

Hawkins, B. A. 2001. Ecology's oldest pattern? – Trends Ecol. Evol. 16: 470.

Hilborn, R. and Stearns, S. C. 1982. On inference in ecology and evolutionary biology: the problem of multiple causes. – Acta Biotheor. 31: 145–164.

Hilborn, R. and Mangel, M. 1997. The ecological detective: confronting models with data. – Princeton Univ. Press

Holling, C. S. 1966. The strategy of building models of complex ecological systems. – In: Watt, K. E. F. (ed.), Systems analysis in ecology. Academic Press, pp. 195–214.

Hooten, M. B. and Hobbs, N. T. 2015. A guide to Bayesian model selection for ecologists. – Ecol. Monogr. 85: 3–28.

Horwich, P. 1982. Probability and evidence. – Cambridge Univ. Press.

Houlahan, J. E. et al. 2017. The priority of prediction in ecological understanding. – Oikos 126: 1–7.

James, G. et al. 2013. An introduction to statistical learning: with applications in R. – Springer.

Jeffrey, R. 1965. The logic of decision, 2nd ed. – Univ. of Chicago Press.

Keuzenkamp, H. A. et al. 2004. The enigma of simplicity. – In: Zellner, A. et al. (eds), Simplicity, inference and modelling: keeping it sophisticatedly simple. Cambridge Univ. Press, pp. 1–11.

Kuhn, T. 1977. Objectivity, value judgment, and theory choice. – In: Kuhn, T. (ed.), The essential tension. Univ. of Chicago Press, pp. 320–339.

Levins, R. 1966. The strategy of model building in population biology. – Am. Sci. 54: 421–431.

MacNally, M. R. et al. 2018. Model selection using information criteria, but is the "best" model any good? – J. Appl. Ecol. 55: 1441–1444.

Marquet, P. A. et al. 2014. On theory in ecology. – Bioscience 64: 701–710.

Mulaiktt, S. A. 2001. The curve-fitting problem: an objectivist view. – Phil. Sci. 68: 218–241.

Okasha, S. 2002. Philosophy of science a very short introduction. – Oxford Univ. Press.

Pearl, J. 1978. On the connection between the complexity and credibility of inferred models. – Int. J. Gen. Syst. 4: 255–264.

Pearl, J. and Mackenzie, D. 2018. The book of why: the new science of cause and effect. – Basic books.

Peck, S. L. 2004. Simulation as experiment: a philosophical reassessment for biological modeling. – Trends Ecol. Evol. 19: 530–534.

Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Model. 190: 231–259.

Pianka, E. R. 1966. Latitudinal gradients in species diversity: a review of concepts. – Am. Nat. 100: 33–46.

Popper, K. R. 1959. The logic of scientific discovery. – Routledge Classics.

Quine, W. V. O. 1963. On simple theories of a complex world. – Synthese 15: 103–106.

Rangel, T. F. and Loyola, R. D. 2012. Labeling ecological niche models. – Nat. Conserv. 10: 119–126.

Rangel, T. F. L. V. B. et al. 2007. Species richness and evolutionary niche dynamics: a spatial pattern-oriented simulation experiment. – Am. Nat. 170: 602–616.

Rangel, T. F. et al. 2018. Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. – Science 361: eaar5452.

Rosindell, J. et al. 2015. Unifying ecology and macroevolution with individual-based theory. – Ecol. Lett. 18: 472–482.

Sober, E. 1981. The principle of parsimony. – Br. J. Phil. Sci. 32: 145–156.

Sober, E. 1989. Reconstructing the past: parsimony, evolution, and inference. – The MIT Press.

Sober, E. 2002. Instrumentalism, parsimony, and the Akaike framework. – Phil. Sci. 69: S112–S123.

Sober, E. 2004a. The contest between parsimony and likelihood. – Syst. Biol. 53: 644–653.

Sober, E. 2004b. What is the problem of simplicity? – In: Zellner, A. et al. (eds), Simplicity, inference and modelling: keeping it sophisticatedly simple. Cambridge Univ. Press, pp. 13–31.

Sober, E. 2015. Ockham's razors: a user's manual. – Cambridge Univ. Press.

Sokal, R. R. and Rohlf, F. J. 1981. Biometry: the principles and practice of statistics in biological research. – W. H. Freeman.

Taper, M. L. 2004. Model identification from many candidates. – In: Taper, M. L. and Lele, S. R. (eds), The nature of scientific evidence: statistical, philosophical and empirical considerations. Univ. of Chicago Press, pp. 448–524.

Taper, M. L. and Ponciano, J. M. 2016. Evidential statistics as a statistical modern synthesis to support 21st century science. – Popul. Ecol. 58: 9–29.

Vellend, M. 2010. Conceptual synthesis in community ecology. – Q. Rev. Biol. 85: 183–206.

Warren, D. L. et al. 2014. Mistaking geography for biology: inferring processes from species distributions. – Trends Ecol. Evol. 29: 572–580.

Wiens, J. J. and Donoghue, M. J. 2004. Historical biogeography, ecology and species richness. – Trends Ecol. Evol. 19: 639–644.

Wiley, E. O. 1975. Karl R. Popper, systematics, and classification: a reply to Walter Bock and other evolutionary taxonomists. – Syst. Zool. 24: 233–243.

Wiley, E. O. 2011. Phylogenetics – the theory and practice of phylogenetic systematics, 2nd ed. – Wiley.

Winsberg, E. 1999. Sanctioning models: the epistemology of simulation. – Sci. Context 12: 275–292.

Winsberg, E. 2001. Simulations, models, and theories: complex physical systems and their representations. – Phil. Sci. 68: S442–S454.

Yang, Z. and Rannala, B. 2012. Molecular phylogenetics: principles and practice. – Nat. Rev. Genet. 13: 303–314.