Opinion TRENDS in Ecology and Evolution Vol.not known No.not known Month 0000

Rules of thumb for judging ecological theories

Lev R. Ginzburg and Christopher X.J. Jensen

Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794, USA

An impressive fit to historical data suggests to biologists that a given ecological model is highly valid. Models often achieve this fit at the expense of exaggerated complexity that is not justified by empirical evidence. Because overfitted theories complement the traditional assumption that ecology is 'messy', they generally remain unquestioned. Using predation theory as an example, we suggest that a fit-driven appraisal of model value is commonly misdirected; although fit to historical data can be important, the simplicity and generality of a theory – and thus its ecological value – are of comparable importance. In particular, we argue that theories whose complexity greatly exceeds the complexity of the problem that they address should be rejected. We suggest heuristics for distinguishing between valuable ecological theories and their overfitted brethren.

Convinced that we exist in a geocentric universe, Ptolemy sought to model quantitatively the motions of the planets. Viewed from Earth, the motions of other planets appear enigmatic – they display retrograde motion, slowly carving out self-intersecting paths in the night sky. Ptolemy 'explained' this motion by introducing epicycles. Imaginary points, rather than the planets themselves, orbited the Earth. These imaginary points were themselves the centers of other orbiting imaginary points. Eventually, once the requisite number of nested suborbits around imaginary points was introduced, the planets themselves follow a final orbit. This fitting process was not general; each planet had to be outfitted separately with its own unique set of epicycles. Although the process was not pretty, the product was: the beauty of Ptolemy's epicycle model was its impressive fit to observational data. More than 1500 years later, Newton's gravitational theory paired with the Copernican heliocentric view of the universe provided a far simpler and parsimonious theory of planetary motion (Table 1). Although Newton's theory failed at the time of its introduction to provide a fit equal to the accuracy of Ptolemy's, it was recognized instantly to be superior because of its rejection power. We view Ptolemy's epicycles as the best historical example of overfitting. Thus, Ptolemy epitomizes the standard but persistent error of theoreticians: their belief that explanatory power can be achieved by building complex models with many parameters (Box 1).

Nowhere is the erroneous Ptolemaic behavior of theorists better recognized than in financial mathematics, where models predicting market moves are abundant. Next to actual price data sequences, successful financial theorists always keep an equal number of 'random walks' simulating pseudo-prices that result from a sequence of random changes. Anyone with a theory on how to play the markets has to first check that it does not 'work' on the random walk sequence. Any market theory that can fit a random walk as well as it fits historical price data is discarded. If a theory passes the random walk test, the next test is for generality; overly specific theories (e.g. a theory that works for British pounds but not Swiss francs) are rejected because of their limited utility and reliability. A theory that is unjustifiably complex or inexplicably specific is too risky, and will not be entertained by successful investors. The selective forces in this field operating against overfitted price-fitting models are simply too strong for any to persist.

Unfortunately, ecological theories are not subject to comparably strong forces of selection. Although we might wish it otherwise, the weakness of selection on ecological theories is a reflection of the relative unimportance, in social terms, of the field. An engineering firm that builds a faulty bridge based on an overfitted model will be sued or fined out of existence; to date, we know of no ecological theorist whose similarly overfitted model has evoked comparable penalties. Because society demands little from theoretical ecology, one can have a successful lifetime career in the field without any of one's theories being put to the practical test of actual prediction; the weak test of retrofitting to past data is deemed as adequate validation of an ecological theory.

From describing to explaining

Models describe, theories explain. 'Describing' and 'explaining' are not discrete conditions, but extremes on

Box 1. The big picture

- Many ecological models designed to illuminate the mechanisms that drive predator–prey cycles are overparameterized.
- The strong fit of these models to ecological data inspires false confidence; as a result, the work of empirical ecologists might be led astray.
- With the intent of weeding out over-parameterized models, we suggest rules of thumb for comparing the complexity of proposed models with the complexity of the problem that they seek to address.

Corresponding author: Lev R. Ginzburg (lev@ramas.com).

http://tree.trends.com 0169-5347/$ - see front matter © 2003 Elsevier Ltd. All rights reserved. doi:10.1016/j.tree.2003.11.004

Table 1. Overfitting in astronomical and ecological models

	Total number of parameters[a]	Number of unsupported parameters[b]	Number of parameters describing the theory goal	'Degree of overfitting'[c]
Theories explaining the motions of other planets in our solar system				
Newton (1687)	5[d]	0	5[j]	0
Ptolemy (*ca.* 150)	~10[e]	~10	5[j,k]	+5
Theories explaining the trajectories of prey populations				
Akcakaya [4]	5[f]	1	5[l]	0
Turchin and Batzli [5]	7[g]	4	3[m]	+4
Hanski *et al.* [5]	9[h]	3	4[n]	+5
King and Schaffer [7]	11[i]	3	5[l]	+6

[a] Scaling is often performed on equations to reduce the number of parameters used. Scaling of abundances can potentially reduce the final number of parameters by eliminating one parameter per abundance variable. Therefore, to maintain a fair comparison of all models, we have listed the minimum number of post-scaling parameters for each model, irrespective of whether the author(s) employed scaling. We did not consider scaling of time (another means of reducing the number of parameters), because this procedure eliminates the potential to fit absolute periods.

[b] The number of parameters employed for which no empirically determined range exists.

[c] The difference between the number of parameters employed in the model and the number of parameters in the function that describes the goal of the theory.

[d] Mass of the planet (first), two coordinates for the initial position of the planet (second and third) and two components of the orbital velocity of the planet (fourth and fifth).

[e] Ptolemy employed 80 epicycles to fit the observed motions of the sun, moon, and five known planets. Each epicycle requires two parameters: a radius and a velocity. Our estimate of five epicycles (ten parameters) per planet is conservative.

[f] S of Equation 3; R_1, R_2, X^*, K, Y_d of Equation 5a/b; and N_{min} of Equation 8 in [4] (two parameters removed by abundance scaling).

[g] Table 3 of [5] (three parameters removed by abundance scaling).

[h] r, K, G, H, C, D, S, Q of Equation 1; e in Equation 2; and σ and σ_{obs} of [15] (two parameters removed by abundance scaling).

[i] Table 1 of [7] (three parameters removed by abundance scaling).

[j] Parameters describing an elliptical orbit: two-dimensional coordinates of each focus (first to fourth), length of string attached to foci (fifth).

[k] Based on Kepler laws, of which Ptolemy was unaware.

[l] Parameters describing actual predator–prey trajectories: prey period, relative amplitude, and cycle asymmetry (first to third); and predator amplitude and time lag (fourth and fifth).

[m] Parameters describing actual herbivore trajectory: herbivore period, relative amplitude, and cycle asymmetry (first to third).

[n] Parameters describing actual prey trajectory: prey period, relative amplitude, and cycle asymmetry (first to third); and geographical gradient in dynamics (fourth).

a continuum [1]. Theoreticans build relatively simple models – imperfect descriptions – to understand ecological phenomena, and to move close enough to 'explaining' to yield valuable insights into the mechanisms that drive natural systems. For the purpose of this paper, we use the terms 'model' and 'theory' rather interchangeably because we find meaning where description and explanation overlap.

To describe, a model need be manipulated only to fit historical data. We can argue about the relative fidelity of such descriptions, but they are difficult to falsify. If we really want to explain, our model ought to make predictions that can lead to empirical testing, predictions that have the potential to be falsified. For this reason, understanding and prediction cannot be decoupled. But: what about chaotic dynamics? Don't chaotic dynamics demonstrate that, even with a perfect understanding, we remain unable to make valid predictions? Although it is true that chaotic dynamics do not allow reliable long-term predictions, excellent short-term predictions can still be made. We equate short-term predictions with understanding. Many ecological theorists separate prediction from understanding. They contend that insights can be gleaned from the model that best fits a given dataset, regardless of whether the model makes valid predictions. This *ad hoc* approach improves the quality of the description at the expense of ultimate understanding.

When fitting is our sole criterion for model validation, testing predictions is problematic. Making reliable predictions in the face of significant variability and noise can be difficult, and we could find ourselves falsifying every theory based on the rather poor fit of its predictions. For this reason, we favor an approach that does not rely exclusively on fitting as a way of judging between theories. In particular, we value theoretical prohibitions, patterns that a given model predicts cannot appear. We are not as impressed when theoretical predictions agree with existing data as when theoretical prohibitions are absent from existing data.

There are always many theories that will explain a given observation. We prefer the simplest one, the most general one, and the one that allows for the least parameter fitting. A large number of unsupported or weakly supported parameters suggests potential overfitting and decreases our trust in the model. Additionally, we value single theories that provide general predictions over a set of different models for specific situations. In suggesting that we appraise models in the preceding manner, we are making an aesthetic argument: we seek the minimal description commensurate with the complexity of the problem.

Model selection methods such as likelihood ratio tests, Akaike's information criterion (AIC), and the Schwartz criterion (SC) provide a means of balancing the conflicting goals of simplicity and goodness of fit [2,3]. AIC and SC consider the likelihood of observed data given a particular model (i.e. goodness of fit), but also include an objective 'penalty' that increases with the number of parameters used (i.e. complexity). Models of varying complexity can then be ranked: simple models that provide substantial fit will be favoured over models that deliver only slightly better fit at the expense of increased complexity. Perhaps model selection criteria should be employed more widely, but we can see three reasons why even their prevalent use might not solve the problem of overfitting completely: (i) most datasets and modeling scenarios do not fit the assumptions of model selection criteria, which lose their ability to balance simplicity and goodness of fit objectively when assumptions are violated; (ii) many authors forward a single model and do not provide data that would enable even qualified readers to compare published models using AIC or SC; and (iii) most critically, these methods will always select one best model from the list of models considered, but the completeness of this list is always at question [3]. When authors do not provide a selection criterion that balances the twin goals of fit and simplicity, readers must turn to other means of assessing whether a theory is overfitted.

One measure of complexity is the number of parameters employed, which remains a crude but objective criterion. Although particular theories with seemingly 'simple' mechanisms can contain a large number of 'nuisance' parameters, we consider such theories complex because they suffer from the same overfitting problems as all highly parameterized models. There are several considerations to be made in judging whether a particular model is appropriately complex: (i) the total number of parameters, (ii) the number of parameters out of this total that are not supported strongly by data, and (iii) the number of parameters needed to represent the pattern of the observed data (the 'goal'). Larger overall numbers of parameters, even if they are supported by data, decrease our trust in the model. Box 2 provides an illustrative example of how a simple, valid methodology can be transformed into a complex, overparameterized model.

The use of model fitting to judge competing ecological hypotheses is valid when relevant datasets are abundant and high in quality. When data are limited, this form of adjudication lacks power, and can be derailed by overfitting. Under such circumstances, we look to alternative means of assessing hypotheses.

Box 2. Recent advances in overfitting

An example of an ecological model that is overfitted strongly is the Non-linear Time Series Modeling (NLTSM) of Peter Turchin [10]. Turchin fits a surface to the growth rate, $R_t = N_{t+1}/N_t$, as a function of two previous abundances, N_t and N_{t-1}, based on time-series data. The resulting function is then used to study the dynamic properties of the system. Used prudently (that is with few parameters, typically three), this is reliable methodology, and has been employed successfully by Turchin, Royama, Berryman, Stenseth, and others [11–14]. No mechanisms of delayed (N_{t-1}) and direct (N_t) density dependence are suggested by this approach, but it is a useful tool to analyze the time series. NLTSM takes this valid approach a step further. It allows for a total of ten parameters: the surface is assumed to be quadratic (six parameters), abundances are rescaled by power functions (two more parameters), and detrending is employed whenever needed, which means always (two more parameters).

Most prey abundance data series to which this model has been applied contain between 15 and 30 points. Fitting a ten-parameter function to such a short data series reliably is impossible; even a nonsensical series of comparable data points can be fit when there are so many parameters allowed. The way to check such a model for overfitting is to fit random walks on top of what is explained by simple three-parameter models to see if the move from three to ten parameters is justified. We suspect that the results of this experiment would reveal the overfitted nature of the NLTSM. This point is lost on unsuspecting readers, who must rely on the information provided, where the reported R^2 values (measuring the fit of the ten-parameter surface to the data points) look impressive.

An overfitting parable: examples from predation theory

Unfortunately, most ecologists are either unaware of the dangers of overfitting or are unable to spot Ptolemaic explanations when they appear in the ecological literature. Theoreticians (or modelers) are small minority of the overall ecology community, comprising ~6% of the total number of ecologists*. The specialized publications of this small community have limited impact on the larger field of ecology; to reach a broader audience, the theoretical ecologist must publish in mainstream ecology journals. When theoretical work is submitted to major journals that do not specialize in theory, a problematic dynamic emerges: because most ecologists are inexperienced mathematically and theoretically, they are ill equipped to judge the validity of theoretical publications. The small contingent of theoreticians becomes a 'priesthood' that, as the sole providers of theoretical guidance, can lead others astray through the use of rhetorical devices.

Overfitting is a potent rhetorical device that has been used extensively in a subfield with which we are intimately familiar: population ecology explaining predator–prey cycling. This subfield provides an illustrative example of the greater phenomenon of overfitting in ecology. We consider four models: Akcakaya [4], Hanski *et al.* [5], Turchin and Batzli [6], and King and Schaffer [7].

The goal of the theorist is to identify the mechanism(s) generating observed fluctuations of predator and/or prey populations. The four models considered were not all generated to address the same dataset, but they share a common basic approach. Data employed in these inquiries come in three forms: (i) time series data for predator and/or prey populations, (ii) field estimated ranges of parameters relevant to the proposed causative mechanisms, and (iii) auxiliary evidence for or against the presence of specific mechanisms in natural systems. Given time series data, one or more mechanisms are proposed to explain the observed pattern(s) and formulated as mathematical models. The parameters of these models are constrained to independently determined (i.e. field estimated) ranges whenever such data exist. We refer to these constrained parameters as 'supported'; parameters for which no observed ranges exist are 'unsupported'. Parameters are adjusted to produce the best fit to the observed data. When multiple models are considered, some measure of goodness of fit is applied to determine the relative success of each model at producing the observed dynamics.

Some authors [5–8] claim to include only those mechanisms supported by biological evidence (i.e. auxiliary 'data'). As a result, more complex theories (containing mechanisms for which data exist) are elevated, and less complex theories (containing mechanisms for which data are absent) are eliminated. This approach assumes that biologists have collected relevant data within the span of all possible mechanisms. We cannot imagine that such an assumption is ever valid, and therefore discourage the use of this approach.

Since the time of Lotka and Volterra, predation models have increased in complexity, much of which emerged from

---

* Holland, M.M. *et al.* (1992) Profiles of ecologists: results of a survey of the membership of the Ecological Society of America. *Ecol. Soc. Am.* 47

attempts to fit models to major aspects of the historical data. The fact that complex models employ supported parameters is often used to deflect accusations of overfitting; models constrained in their fitting of crucial parameters are viewed as 'less overfitted' than their unconstrained counterparts. This defense has to be considered skeptically. If there were just a few parameters, the claim would be credible. When there are many parameters, and ranges are very large and uncertain for a good percentage of them, the potential for overfitting exists even when supported parameters are utilized. The volume of the parameter space that is consistent with the observations declines rapidly with the number of parameters (Figure 1).

The number of parameters employed serves as an indirect measure of how much squeezing one has to do to fit the theoretical construct to the evidence; in other words, how much trial and error went into producing the model. Good theories work well from the first trial; they are simple, general and have few parameters. Subjectively, we forget all our trial and error and present a published model as if it was the first one that came logically to our mind. Because the answer we seek is known to us in advance (in the form of historical data), it is very hard not to subconsciously fit models to the facts *ad hoc*. Certainly, the relatively complex model structure that gives the best fit to historical data is usually developed with a good amount of trial and error. Thus, all claims of 'not fitting to the data' have to be considered cautiously. We always fit, directly or indirectly. It is for this reason that we have to develop more objective ways to judge the plausibility of a theory.

Kepler described the elliptical nature of planetary orbits. An ellipse can be fully described with five parameters (Table 1). Although Kepler was unaware of the gravitational theory that would eventually explain these orbits, his work suggested that a reasonable explanatory theory should have about five parameters. Although we do not know what drives population cycles, our awareness of their mathematical properties defines our 'goal': the exact number of parameters that should describe a given time series. The simplest non-mechanistic description of a cycle is a sine wave, which requires two parameters: period and amplitude. If we wish to pay attention to cycle asymmetry (e.g. slower up than down), we can describe this behavior with another parameter. If our theory takes latitudinal gradients into account, we can allow ourselves an additional parameter. If we have data for two interacting species (assuming that the period is the same), we need another amplitude for the second species and the 'phase shift' or time lag between the two waves. In so doing, we have counted a maximum of six potential 'goal' parameters. Table 1 enumerates the 'goal' parameters of each of the four models considered. Although none of the historical time series are characterized by all six potential parameters described above, all of them pull from this same menu of parameters.

Table 1 also compares the number of parameters employed by each model with the number of parameters that describe its 'goal'. Of the predator-prey models considered, the one by Akcakaya [4] is not only first historically, but also the most solid. The only weakly supported parameter employed is the size of refugia for hares. It is, however, a crucial parameter because the very presence of cycles is controlled by the size of refugia. In contrast with the other three models considered, the Akcakaya model [4] appears to be extremely reasonable. The other three models [5–7] are overparameterized by from four to six parameters. Additionally, all of them utilize at least three unsupported parameters. With so many unsupported parameters, these models ought to be able to fit every cycle.

When we view predator-prey theories such as most of those described above, our aesthetic sense tells us that there has to be something wrong with them. Passing our 'aesthetic criteria test' is not a guarantee of being correct; a diversity of theories might work aesthetically. However, aesthetic tests can serve as an ideological sieve, filtering out unviable theories so that differences between the remaining viable models can be settled by clear manipulative experiments.

Note that we include the generality of an explanation in the list of aesthetic criteria, but this is far from a universal view. Many people are much more comfortable thinking

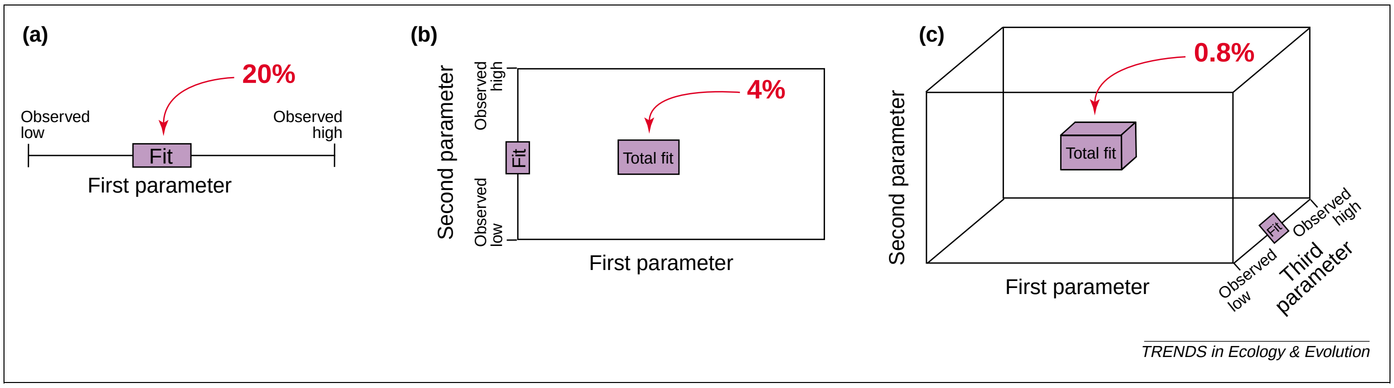

Figure 1. The significance of ranges in multiple-parameter space. Empirically based parameter ranges have been heralded as a means of creating 'realistic' models [7,15]. In reality, models that constrain parameter values within biologically observed ranges are fitted only slightly less than those that allow arbitrary parameter ranges. **(a)** A hypothetical range of observed values for a single parameter, 20% of which fits the target data of a one-parameter model. **(b)** A similar two-parameter model with empirically based ranges. The resulting parameter area that fits the target data represents only 4% of the total area of biologically realistic parameter combinations. **(c)** An analogous three-parameter model fits the target data only within 0.8% of the total parameter volume. As more parameters beyond three are added, the probability that a given set of parameters (randomly selected from known ranges) will fit the model approaches zero.

that every case has a different explanation. For these authors, even the cycle of the voles is fundamentally different from that of lemmings [6]. This view reminds us again of Ptolemaic epicycles, fitted separately and differently for the trajectory of every planet. It took over 1500 years to reject this erroneous view in astronomy. We hope that we can progress more rapidly in ecology.

Sensitivity analysis offers another means of eliminating overfitted theories

For those uncomfortable with the potential subjectivity of our aesthetic approach, another means of testing the viability of a complex model is to do an extensive sensitivity study. This would entail checking, at a minimum, the highs and lows of every parameter range. With one parameter, only two cases need to be checked, but with ten parameters, just the 'high' and 'low' check entails 2^{10} or over 1000 combinations. Although authors of complex models might claim that complete sensitivity analyses are too time-consuming to perform, we suspect that such analyses would discover the overfitted and thus unreliable character of the model. King and Schaffer do a good job of pairwise sensitivity analysis for some of the pairs of influential parameters [7]. This analysis suggests that most of the volume in the parameter space *does not* produce the desired result (Figure 1 describes this problem in visual terms). In other words, picking values randomly within observed ranges in most cases will not produce the desired ecological phenomenon. As a result, the theory does little to explain the actual dynamics of natural populations, and we are left with a model that is just 'consistent with' past data.

Box 3. Recent advances in rhetoric

The apparent 'objective' nature of theoretical work obscures the great potential for using rhetoric to support one's models. As biologists, we have come to expect data that are somewhat messy, and we question instinctively every result that seems 'too perfect'. Theorists are acutely aware of this concern, and sometimes use rhetorical devices to hide the overfitted nature of their models (see Rule of Thumb *#4* in Box 4 for more on this rhetorical device). Hanski and Turchin [15] add environmental noise and experimental error, producing a messier result that is more palatable to most biologists. Turchin and Batzli's [6] model produces a cycle period of 5.5 years, the 'best fit' to the true period of four years among the models reviewed in their paper. Curiously, the authors then admit that some values of an unsupported parameter (α) can produce the actual period of four years. As such, the reported value (5.5 years) seems strategically selected to avoid being unbelievably close to the target. Turchin [10] reports a cycle period for the same model in a later publication that is even further from the target (7.0 years). We doubt that any value for the period is outside the uncertainty range of this model.

How the parameters are presented can also impede our ability to detect overfitting. Whereas some authors provide clear lists of parameters employed [6,7], others do not. Turchin introduces the many parameters of his Non-linear Time Series Modeling (NLTSM) model (Box 2) sporadically among large tracts of discussion [10]; the parameters are never compiled into a single list, possibly leading some readers to believe that he is using a fairly simple model. Hanksi *et al.* present their model equations in a stripped-down format [5]. To count the actual number of parameters used, one has to track down the original model in other publications [15,16]. The claim that parameters are 'empirically based' (i.e. supported) can also be abused (Figure 1, main text). Turchin and Batzli [6] list 'median values' of parameters for which there are no known ranges. The correct term is 'point estimates'; the word "median" implies knowledge of the range.

Rules of thumb for judging ecological theories: recognizing rhetoric

Most ecologists do not read theoretical papers. Considering the amount of overparameterized theory in the literature, and the rhetoric that often surrounds this theory (Box 3), this might represent a fairly prudent general strategy for the nontheoretician. The problem is that specific theories do make sense sometimes, and this quality theoretical work can guide empirical research. For those nontheoretical ecologists who do wish to safely tap into the theoretical literature, we offer four rules of thumb for judging ecological theories in Box 4. In positing these, we are suggesting that ecologists adopt a more skeptical view of theoretical work. However, it is not our view that theoreticians intentionally mislead their fellow ecologists through the use of rhetoric. We suspect that theoretical ecologists simply suffer from excessive parental love. Because they fall in love with their own theories and treat them as their 'children', many theoreticians are blind to the overfitted nature of their models.

To be fair, we admit that we too might be guilty of rhetoric. In fact, one of the authors (L.R.G.) supports a competing view on population cycles, one based on the theory of maternal effects [9]. It captures the period and the asymmetry of prey cycles and employs two scaled parameters, one of which is unsupported. It can be falsified, because it predicts that population cycles with a period between two and six generations should never be found. These periods do appear to be missing from existing datasets. Because of this conflict of interest, we cannot claim impartiality with respect to the predation-based population cycle theories that we criticize here.

Box 4. Rules of thumb for judging ecological theories

1. Compare the number of parameters with the number of data points. When a model uses ten parameters to fit to a time series of 25 data points, chances are that it can fit almost any 25 data points.
2. Compare the complexity of the proposed model with the complexity of the phenomenon that it seeks to explain. Often, proposed models turn out to be dramatically more complex than the ecological problems that they seek to solve. If one can state the ecological phenomenon in fewer words than it takes to formulate the model, the theory is probably not useful.
3. Beware of meaningless caveats confessing oversimplification. Eager for their work to be embraced by ecologists, theoreticians like to conclude that their models are oversimplified. An already complex model that 'admits' that there are more mechanisms to be taken into account (read: more parameters) betrays a tendency towards further unjustified complexity.
4. Beware of being given what you expect. As ecologists, we have come to expect that our data will be 'messy', and many theoreticians will go out of their way to meet this expectation. One way to make the curves look 'less perfect' is to simply add environmental noise and observational error (each variance adding one more parameter). Suspect that rhetoric is at work when models that are fully capable of producing a perfect fit are tweaked to show a more palatable near-perfect fit.

Increasing selection against overfitted theories

Natural selection of ideas makes science different from astrology. The strength of selection depends on how important a particular field is. In unimportant fields, theories evolve neutrally, and thus flourish unchecked. As long as overfitted models flourish, most ecologists will remain justifiably skeptical of the theoretical minority. Two factors will increase the quality of ecological theories: (i) increased social demand for practical applications, and (ii) the related improvement in mathematical education of all ecologists. Waiting for these changes to occur is potentially dangerous. As we come to rely more heavily on predictive theories to instruct our conservation efforts, natural populations might suffer from the use of overfitted models. It is our hope that we will not have to 'learn the hard way'. If adopted, our suggested rules of thumb can begin the selective process now.

Acknowledgements

We thank Rudy Boonstra, Elizabeth Milliman and three anonymous referees for insightful comments and suggestions that greatly improved the quality of this article.

References

1 Colyvan, M. and Ginzburg, L.R. (2003) Laws of nature and laws of ecology. *Oikos* 101, 649–653
2 Forster, M. and Sober, E. (1994) How to tell when simpler, more unified, or less *ad hoc* theories will provide more accurate predictions. *Br. J. Philos. Sci.* 45, 1–35
3 Johnson, J.B. and S. Omland, K.S. (2004) Model selection in ecology and evolution. *Trends Ecol. Evol.* DOI:10.1016/j.tree.2003.10.013
4 Akcakaya, H.R. (1992) Population cycles of mammals – evidence for a ratio-dependent predation hypothesis. *Ecol. Monogr.* 62, 119–142
5 Hanski, I. *et al.* (2001) Small-rodent dynamics and predation. *Ecology* 82, 1505–1520
6 Turchin, P. and Batzli, G. (2001) Availability of food and population dynamics of arvicoline rodents. *Ecology* 82, 1521–1534
7 King, A.A. and Shaffer, W. (2001) The geometry of a population cycle: a mechanistic model of snowshoe hare demography. *Ecology* 82, 814–830
8 Turchin, P. *et al.* (2003) Dynamical effects of plant quality and parasitism on population cycles of larch budmoth. *Ecology* 84, 1207–1214
9 Ginzburg, L.R. and Colyvan, M. (2004) *Ecological Orbits: How Planets Move and Populations Grow*, Oxford University Press
10 Turchin, P. (2003) *Complex Population Dynamics: A Theoretical/Empirical Synthesis*, Princeton University Press
11 Turchin, P. and Taylor, A.D. (1992) Complex dynamics in ecological time series. *Ecology* 73, 289–305
12 Royama, T. (1992) Analytical Population Dynamics, Chapman & Hall
13 Berryman, A.A. (1999) *Principles of Population Dynamics and their Application*, Stanley Thornes Publishers
14 Stenseth, N.C. (1999) Population cycles in voles and lemmings: density dependence and phase dependence in a stochastic world. *Oikos* 87, 427–461
15 Turchin, P. and Hanski, I. (1997) An empirically based model for latitudinal gradients in vole population dynamics. *Am. Nat.* 149, 842–874
16 Hanski, I. *et al.* (1991) Specialist predators, generalist predators and the microtine rodent cycle. *J. Anim. Ecol.* 60, 353–367