The Ecological Forecast Horizon, and examples of its uses and determinants

Petchey, Owen L.[1,2] owen.petchey@ieu.uzh.ch
Pontarp, Mikael[1,10] mikael.pontarp@ieu.uzh.ch
Massie, Thomas M.[1] thomas.massie@ieu.uzh.ch
Kéfi, Sonia[3] sonia.kefi@univ-montp2.fr
Ozgul, Arpat[1] arpat.ozgul@ieu.uzh.ch
Weilenmann, Maja[1] maja.weilenmann@ieu.uzh.ch
Palamara, Gian Marco[1] gianmarco.palamara@ieu.uzh.ch
Altermatt, Florian[1,2] florian.altermatt@ieu.uzh.ch
Matthews, Blake[4] blake.matthews@eawag.ch
Levine, Jonathan M.[5] jonathan.levine@usys.ethz.ch
Childs, Dylan Z.[6] d.childs@sheffield.ac.uk
McGill, Brian J.[7] brian.mcgill@maine.edu
Schaepman, Michael E.[8] michael.schaepman@geo.uzh.ch
Schmid, Bernhard[1] bernhard.schmid@ieu.uzh.ch
Spaak, Piet [2,5] piet.spaak@eawag.ch
Beckerman, Andrew P.[6] a.beckerman@sheffield.ac.uk
Pennekamp, Frank[1] frank.pennekamp@ieu.uzh.ch
Pearse, Ian S.[9] ianspearse@gmail.com

[1] University of Zurich, Institute of Evolutionary Biology and Environmental Studies, Zurich, Switzerland
[2] Eawag: Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Dübendorf, Switzerland
[3] Université de Montpellier II, CNRS, Institut des Sciences de l'Evolution, Montpellier, France
[4] Eawag: Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Center for Ecology, Evolution, and Biogeochemistry, Kastanienbaum, Luzern, Switzerland
[5] ETH Zurich, Institute of Integrative Biology, Zurich, Switzerland
[6] Sheffield University, Animal and Plant Sciences, Sheffield, South Yorkshire, United Kingdom
[7] University of Maine, School of Biology and Ecology and Mitchel Center for Sustainability Solutions, Orono, Maine, United States
[8] University of Zurich, Department of Geography, Remote Sensing Laboratories, Zurich, Switzerland
[9] The Illinois Natural History Survey, Champaign, Illinois, United States
[10] Department of Ecology and Environmental Science, Umeå University, Sweden

Statement of authorship: conception: OLP; methodology / computation / analysis: OLP, MP, G-MP, IP, BM, FA, FP; writing: all co-authors.

Corresponding author: Owen L Petchey, Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Tel. +41 44 635 47 70, No fax number, Email owen.petchey@ieu.uzh.ch

1 Abstract

Forecasts of how ecological systems respond to environmental change are increasingly important. Sufficiently inaccurate forecasts will be of little use, however. For example, weather forecasts are for about one week into the future; after that they are too unreliable to be useful (i.e., the forecast horizon is about one week). There is a general absence of knowledge about how far into the future (or other dimensions, e.g., space, temperature, phylogenetic distance) useful ecological forecasts can be made, in part due to lack of appreciation of the value of ecological forecast horizons. The ecological forecast horizon is the distance into the future (or other dimension) for which useful forecasts can be made. Five case studies illustrate the influence of various sources of uncertainty (e.g., parameter uncertainty, environmental and demographic stochasticity, evolution), level of ecological organisation (e.g., population or community), organismal properties (e.g., body size or number of trophic links) on temporal, spatial and phylogenetic forecast horizons. We propose that the ecological forecast horizon is a flexible and powerful tool for researching and communicating ecological predictability, and for motivating and guiding agenda setting for ecological forecasting research and development.

2 Introduction

Forecasts are statements about what the future is likely to hold in store (Coreau *et al.* 2009) and as such, they are an essential basis for all kinds of decisions, including economic, political and personal ones. In ecological systems examples of forecasts include species distributions (e.g., Guisan & Thuiller 2005; Araújo & New 2007), functional diversity (e.g., Kooistra *et al.* 2008; Schimel *et al.* 2013), phenology (e.g., Cannell & Smith 1983; Diez *et al.* 2012; Garonna *et al.* 2014), population size (e.g., Ward *et al.* 2014), species invasions (e.g., Levine & Antonio 2003), agricultural yield (e.g., Cane *et al.* 1994), pollinator performance (e.g., Corbet *et al.* 1995), extinction risk (e.g., Gotelli & Ellison 2006a), fishery dynamics (e.g., Hare *et al.* 2010), water quality (e.g., Komatsu *et al.* 2007), forest carbon dynamics (e.g., Gao *et al.* 2011), ecosystem services (e.g., Homolová *et al.* 2013), disease dynamics (e.g., Ollerenshaw & Smith 1969; Hijmans *et al.* 2000), and ecological interactions (e.g., Pearse & Altermatt 2013).

Although ecological forecasting has occurred in ecological research for decades, current and expected environmental changes are motivating ever increasing interest in ecological forecasting. There is a pressing need to deliver information about the likely future state of populations, communities, and ecosystems, in order to better inform conservation, management, and adaptation strategies (Clark *et al.* 2001; Sutherland *et al.* 2006; Tallis & Kareiva 2006; Evans 2012; Mouquet *et al.* 2012; Purves *et al.* 2013). Also, because accelerated environmental change may prevent equilibration of species ranges to new environmental conditions, historical information on range-environmental correlations becomes less useful for predicting species distribution (Schimel *et al.* 2013). Consequently, timely as well as high quality information will

fundamentally drive the predictive capabilities of forecasting systems (Dowd 2007; Laurent *et al.* 2014; Niu *et al.* 2014). Furthermore, accurate forecasting (i.e., correct prediction) is sometimes regarded as the hallmark of a successful science (Evans *et al.* 2012), and as such can be a powerful driver of advances in knowledge about how ecological systems work (Coreau *et al.* 2009).

This study rests on the premises that accurate ecological forecasts are valuable, but our knowledge about ecological forecasting is relatively sparse, contradictory, and disconnected. Ecologists need to know what properties and components of ecological systems are forecastable, and the uncertainties associated with these forecasts (Clark *et al.* 2001; Godfray & May 2014). A systematic understanding of forecast performance in relation to different types of modelling practices and sources of uncertainty can guide ecology to become an even more predictive science.

First we review opinion and evidence about the predictability of ecological systems, concluding that important, large, and exciting advances remain. We propose that these advances are constrained by lack of generally applicable and intuitive tools for assessing ecological predictability. We then introduce such a tool: the ecological forecast horizon, and suggest that it could be a hub for research about ecological predictability, as well as a tool for intuitively communicating the same. We provide case study illustrations of how various sources of uncertainty (e.g., imperfect or incomplete knowledge about parameter values, demographic stochasticity, evolution) and organismal characteristics influence forecast horizons, and discuss challenges and research priorities associated with the use of ecological forecast horizons. As such, this article aims to initialise and motivate further agenda setting for forecasting research in ecology.

2.1 Existing knowledge about ecological predictability

Recent reviews and commentaries provide encouraging views of the possibility to make useful ecological forecasts (Sutherland 2006; Purves & Pacala 2008; Evans *et al.* 2013; Purves *et al.* 2013). The argument goes that advances in data collection and handling, coupled with new methods for using that data to reduce uncertainty, will enable process-based models that provide useful predictions. Forecasts of influenza dynamics provide support for this standpoint. Despite the non-linearity and intrinsically chaotic nature of infectious disease dynamics, timing of the peak of a disease outbreak could be predicted up to seven weeks in advance (Shaman & Karspeck 2012). Models of population (e.g., Brook *et al.* 2000), community (e.g., Wollrab *et al.* 2012; Hudson & Reuman 2013), and ecosystem (e.g., Harfoot *et al.* 2014; Seferian *et al.* 2014) dynamics also suggest that forecasting ecological dynamics via process based models is possible. Emerging technologies and methods, combining advanced approaches of hind-, now- and forecasting mechanisms (Dobrowski & Thorne 2011; Stigall 2012) and a more timely

assessment of ecosystem states (Asner 2009; Loarie *et al.* 2009) provide data rapidly enough to parameterize land-atmosphere interaction models.

Less encouraging viewpoints exist. Beckage *et al.* (2011) argue that ecological systems have low intrinsic predictability because a species' niche is difficult to specify, because ecological systems are complex, and because novel system states can be created (e.g., by ecological engineering). Coreau *et al.* (2009) give a somewhat similar list of difficulties. These features should make ecological systems 'computationally irreducible', such that there is no substitute to observing the real thing. Furthermore, evolution may be an intrinsically chaotic process, thus limiting long-term predictability of ecological systems (Doebeli & Ispolatov 2014). If so, ecological responses to anthropogenic climate change are likely to be intrinsically unpredictable. Indeed, population dynamics of a laboratory-based aquatic community were predictable only to 15–30 days due to chaotic dynamics, and useful predictions thereafter could be "fundamentally" impossible (Benincà *et al.* 2008). Indeed, the theoretical discovery of chaos led to pessimism about forecasting. Even completely deterministic systems could have very limited forecast horizons due to sensitivity of initial conditions and our inability to precisely measure initial conditions (something that certainly holds in ecological systems). Chaos also magnifies non-modelled processes (e.g., stochasticity) (Ellner & Turchin 1995). Although there is debate about whether single species populations show chaotic dynamics, there is a general understanding that the higher dimensional a system is, the greater the likelihood that it is chaotic (Turchin 2003) and ecological systems are nothing if not high dimensional.

Other evidence comes from theoretical and empirical studies about interspecific effects. For instance, Yodzis (1988) studied whether effects of changes in abundance of one species on another (i.e., a press perturbation) were directional determined, i.e., whether the direction (increase or decrease) of an effect can be reliably predicted. He defined a prediction (e.g., algal biomass increases due to the addition of fish) as being directionally determined when its sign was consistent in at least 95% of cases. Each case was created by randomly drawing parameter values, specifically interaction strengths among the species in a simple food web model, from a uniform distribution with order of magnitude range. Yodzis found that over half of the net effects of press perturbations were directionally undetermined. That is, if uncertainty in interaction strength spans an order of magnitude, predictions of press perturbations will more often than not be unreliable, even in terms of the direction of the effect. Yodzis' findings paint a depressing picture of predicting ecological dynamics. Uncertainty in parameter values (specifically interaction strengths) interacts with complexity (specifically the presence of indirect effects) to make "implementing conservation and management strategies difficult because the effects of a species loss or an environmental perturbation become difficult to predict a priori" (quote from Wootton 2002).

Recent extensions and explanations of Yodzis' findings provide reasons for optimism and pessimism about prediction in ecology (Novak *et al.* 2011). First, some effects of press perturbations are determined (Dambacher *et al.* 2002; Aufderheide *et al.* 2013), though these effects reduce in number with increases in ecological complexity (species richness and connectance of a food web) (Dambacher *et al.* 2003; Novak *et al.* 2011). Some empirical studies suggest complexity begets predictability (McGrady-Steed & Harris 1997; Berlow *et al.* 2009) while others do not (France & Duffy 2006). Second, it seems that interaction strengths can be estimated with sufficient accuracy to provide determinacy, though the demands on accuracy increase as the complexity of the ecological system increases (Novak *et al.* 2011; Carrara *et al.* 2015). Third, some experimental studies show that models can predict dynamics (Vandermeer 1969; Wootton 2002, 2004). Fourth, much remains poorly understood regarding predicting effects of ecological systems to environmental change, such that great advances remain to be made. Fifth, prediction at the community and ecosystem level may still be possible even if predictions at population level are not.

Some final evidence suggests that more mechanistic models often make worse predictions of population dynamic than simple "model-free" or "statistical" forecasts. For example, simple state-space reconstructions based on relatively little observed data outperform more complex mechanistic models (though see Hartig & Dormann 2013; Perretti *et al.* 2013a, 2013b) and still can distinguish causality from correlation (Sugihara *et al.* 2012). Similarly, a comparison of population dynamic time series forecasting models of natural animal population data reported that the most accurate model was the one that used the most recent observation as the forecast (Ward *et al.* 2014). Also see the spatial example provided by Bahn & McGill (2007)

Whether contradictions exist among these different views about ecological predictability is unclear. Reductions in uncertainty will increase predictability, but little is known about how computationally irreducible are real ecological communities, or about whether different state variables (e.g., population size versus ecosystem processes) will have different predictability, or about the predictability of effects of different types of environmental change (though see Fussmann *et al.* 2014; Gilbert *et al.* 2014). Indeed, a review of marine ecosystem models revealed the assumptions are mostly left implicit and uncertainties often not considered (Gregr & Chan 2014). Ecologists must systematically and thoroughly address these challenges (Clark *et al.* 2001), though they might lack the tools needed to do so. We believe that a standard, flexible, quantitative, intuitive and policy-relevant method for assessing how well ecological models can forecast, such as the ecological forecast horizon, will greatly aid research and communication.

2.2 The ecological forecast horizon

The prediction / forecast horizon as a concept goes back at least to Lorenz (1965), who wrote about how the ability to predict the weather is closely related to "the amount of time in advance for which the prediction is made". Thus a forecast horizon is how far into the future (or dimensions other than time, e.g., space, phylogeny, environment) sufficiently good predictions can be made. A common reflection of the forecast horizon concept is the observation that weather forecasts are usually only made up to a specific time period into the future. After that specific period, predictions are not good enough to be useful. However, the notion of a dynamically changing forecast horizon is important: over the past decades, the forecast horizon of 'weather' has increased. Key improvements of the forecast horizon were achieved through external effects (e.g., increase in computational power) as well as optimizing the forecast system (internal: ensemble forecasting, data assimilation, Kalman filtering, etc.).

Quantifying a forecast horizon requires a measure of how good is a forecast (we term this the *forecast proficiency*) and a *forecast proficiency threshold* above which predictions are good enough, and below which forecasts are not good enough (below we deal with how the threshold can be set). The forecast horizon is the time at which average forecast proficiency drops below this threshold (figure 1). A far forecast horizon indicates greater ability to predict (high realised predictability), a close one a weaker ability to predict (low realised predictability).

In practice, there will usually be multiple possible forecasts (even given the same model if parameter values are uncertain), each with a particular forecast proficiency. This will result in a distribution of forecast proficiencies that then creates a distribution of forecast horizons (figure 1). Integrating information about the distribution of forecast proficiencies into analyses is important and, at least in the following case studies, was relatively straightforward.

The forecast horizon is a measure of predictability / forecastability that focuses on how far into the future forecasts are good enough. It is relatively straightforward to focus on the flip side of the coin, however, by setting a forecast horizon threshold and measuring the forecast proficiency at that horizon (one might term this forecast proficiency at desired horizon).

3 Case studies

We provide five case studies. Two involve analyses of models, three of empirical data. Three involve temporal forecast horizons (how far into the future can useful forecasts be made), one spatial forecast horizons (how far away in space can useful forecasts be made), and one phylogenetic forecast horizon (how far across a phylogeny can useful forecasts be made). The temporal case studies include analysis of a simple model, a more complex model, and a complex empirical food web, and among them illustrate how various sources of uncertainty can impact forecast horizons. Finally, the case studies illustrate different types of predictive models (from process-based to statistical).

3.1 Chaos and demographic stochasticity

Using a model, we can produce a time series of some variable that we can assume is the truth. We can also produce a time series that we can assume is a forecast. If the model used to make the forecast is different (e.g., in initial conditions, structure or parameter values) from the one used to make the truth, the true time series and the forecast time series can differ. This difference is the forecast proficiency of the predictive model, and could be any of many quantitative measures of difference (see later). Here, we use the correlation coefficient for a window of the time series. Moving this window provides measures of forecast proficiency as a function of how far into the future the forecast is made. Note that this approach will result in perfect correspondence of the truth and prediction at all time horizons for a fully deterministic model and no uncertainty in parameter values or initial conditions.

We illustrate this approach with the Ricker model in the chaotic dynamic regime, as this is a simple model that can produce non-trivial behaviour. We examined the effects on forecast horizon of uncertainty in the predictive model in the value of the intrinsic growth rate (r) and the initial population size (N_0). We also examined the effects of the presence or absence of demographic stochasticity in the model used to make the true time series. For each level of uncertainty in r and N_0 we drew a random value of r and N_0, simulated dynamics, and calculated the forecast proficiency and forecast horizon of population dynamics. We then calculated average forecast proficiency and average of the forecast horizon across simulations.

Forecast proficiency started high (correlation between true and predicted population size is close to 1) and by 20 generations dropped to near zero (figure 2). This is consistent with the chaotic nature of the modelled dynamics (see Box 1). Higher uncertainty in the growth rate r or initial population N_0 size results in earlier drop in forecast proficiency, compared to when there is low uncertainty. The presence of demographic stochasticity causes generally earlier drops in forecast proficiency.

Effects of uncertainty in r and N_0 interact (figure 3). For example, high uncertainty in r results in close forecast horizons regardless of uncertainty in N_0, while lower uncertainty in r allows lower uncertainty in N_0 to give farther forecast horizons. Demographic stochasticity in the true dynamics makes for a very close forecast horizon, consistent with the chaotic nature of the dynamics.

3.2 Level of organisation, evolution, and environmental uncertainty

We applied the same general approach as above to a model of a competitive community including evolutionary change, similar to that in Ripa *et al.* (2009). Briefly, each competing species has a trait value that determines its resource use requirements. Ecological dynamics result from resource depletion and therefore competition among the species, while evolutionary dynamics result from changes in trait values of a species (e.g., body size and resource uptake characteristics). The model also included environmental variability, implemented as random

variation in the resource distribution. As before, we evaluated the forecast proficiency by measuring the correlation between true and also predicted dynamics in a window along the time series for two variables: the abundance of one of the species and the total biomass of all species. We manipulated whether evolution operated in the model that was used to produce the true data, and also the amount of uncertainty about the nature of the environmental variability.

In the absence of evolution, forecast horizons for species abundance and total community biomass were very similar (figure 4). In the presence of evolution, forecast horizons were consistently farther for total community biomass. This may result from density compensation among the competing species, enhanced by supply of diversity by evolution, creating more predictable dynamics of total community biomass (e.g., Yachi & Loreau 1999). Unsurprisingly, forecast horizons are closer when there is greater uncertainty about future environmental conditions.

3.3 Dynamics of an aquatic food web

A phytoplankton community isolated from the Baltic Seas was kept in a laboratory mesocosm for more than eight years, and nutrients and abundance of organisms in ten functional groups were sampled 690 times (Benincà *et al.* 2008). This long ecological time series exhibited characteristics consistent with a chaotic system. A neural network model of the community displayed high predictability (0.70 to 0.90; measured as r-squared between observed and predicted data) in the short term only.

We extended the published study by examining variation in ecological forecast horizon among the ten functional groups and two nutrients. Forecast horizons were calculated by fitting a curve to the forecast proficiency (measured by r-squared)–forecast time relationships in Figure 2 of Benincà *et al.* (2008), and estimating the time at which forecast proficiency dropped below an arbitrarily determined forecast proficiency threshold of 0.6. Size ranges represented by organisms in each taxonomic group were gathered from literature and online sources.

Forecast horizons exhibited a triangular relationship with organism size, with only low forecasts horizons for smaller organisms, and a wide range of forecast horizons for larger organisms (Figure 5a). Forecast horizon was somewhat shorter for taxa with greater number of trophic links to other organisms (Figure 5b). Linear models with variance constant or a power function of log organism size and number of trophic links, with or without the interaction, had a minimum p-value of 0.05 for the association of forecast horizon with number of trophic links.

Generally longer generation times of larger organisms may partially explain this (albeit non significant) result, though generally lower population sizes should increase the importance of demographic stochasticity, making for nearer forecast horizons. Hence, we do not feel confident, based on verbal arguments, about making a hypothesis about the expected relationship between body size and forecast horizon. The trend towards nearer prediction horizon for organisms with greater number of trophic links may reflect the negative effects of

complexity on predictability (Dambacher *et al.* 2003; Novak *et al.* 2011) perhaps being related to processes linking complexity and stability (e.g., McCann 2000; May 2001).

3.4 Spatial forecast horizons and statistical models

Forecasting horizons can be made in space (maximum distance predicted to acceptable proficiency) and when the predictive model is a statistical one rather than a process-based model. One well known macroecological pattern is usable as a statistical model to assess spatial forecast horizons: the decay of similarity with distance curve (Nekola & White 1999; Nekola & McGill 2014). Here the statistical predictive model is simply using spatial autocorrelation to predict that a neighbouring community will have the same set of species (or same relative abundances) as the observed community. Similarity becomes a measure of prediction efficiency. If Sørensen similarity is used, it gives a measure of the percentage of species correctly predicted to be present. A decay of similarity curve shows some measure of community similarity between pairs of communities on the y-axis plotted against the distance apart the communities are found (figure 6). Repeated over many pairs of communities at different distances and with an exponential decay curve fit to the data this shows the expected or average similarity (which can also be treated as a measure of forecasting efficiency giving the % of species correctly predicted in a community) as a function of distance. Given any threshold level of similarity desired, one can quickly read off the distance at which this similarity can be (Figure 6). Spatial forecast horizons could also readily be applied to species distribution models (e.g., Pottier *et al.* 2014).

In this spatial case study the model is not process based, but rather a statistical model assuming autocorrelation. Currently this statistical model is a better predictor than a wide variety of commonly used covariates such as climate and other species (Bahn & McGill 2007). If an abundance-based similarity metric such as Bray-Curtis is used this becomes a prediction of not just species composition but relative abundance. Nekola and McGill (Nekola & McGill 2014) suggest that we should also plot decay of similarity curves for single species, in which case spatial autocorrelation becomes a predictor of species presence/absence or abundance and the individual decay of similarity curve allows the determination of forecasting thresholds for statistical predictions of abundance. There exist diverse and very well developed methods for statistical forecasting of time series, such as autoregressive models, that are used in business and economic forecasting, for example. Some of these methods are already used in ecology (Wootton 2004), but whether more can be usefully borrowed from fields with well developed statistical forecasting methods, such as economics.

3.5 Phylogenetic forecast horizons

A phylogenetic forecast horizon concerns how far across a phylogeny can useful forecasts be made. To illustrate a phylogenetic forecast horizon, we analysed a previously published study of

native Lepidoptera-plant interactions in Central Europe (Pearse & Altermatt 2013). We constructed a host-use model (a binomial GLM), in which the inclusion of a host plant in the diet of a herbivore was a function of the herbivore's host breadth and the phylogenetic distance of that plant from another known host. We then used this model to predict the inclusion of plants introduced into Central Europe into the diet breadth of herbivores. To construct a forecast horizon in phylogenetic distance, we divided the novel (prediction) dataset of known novel Lepidoptera-plant interactions and predictions into 12 phylogenetic distance slices (12 was large enough to construct the forecast proficiency versus phylogenetic distance curve, but not so many to have too little data in each slice). We then calculated the area under the ROC curve (AUC, the measure of forecast proficiency) within each phylogenetic distance slice.

AUC related linearly and positively to phylogenetic distance, with higher forecast proficiency at farther phylogenetic distances (i.e., between plant families), and lower forecast proficiencies at smaller phylogenetic distances (figure 7). Reducing the amount of data used to parameterise the forecasting model indicates that increased information allows better predictions of host use over plant phylogeny.

Interesting, this phylogenetic forecast increases its predictability with increasing distance, whereas forecasts over time typically decrease in predictability with increasing time. Because many herbivorous insects consume a set of plants delimited at roughly the family-level, the forecast horizon for the prediction of a novel plant-herbivore interaction might be set at the family level, where predictions at a lower and higher taxonomic level are less inaccurate (e.g., Pearse & Altermatt 2013). Conversely, when considering the over-dispersion of plant communities, co-occurence was unlikely among very close relatives (congeners), but this trend did not hold at higher taxonomic levels (Cavender-Bares *et al.* 2006), suggesting that the forecast horizon for co-occurence might be at the genus-level, where predictions at higher levels of taxonomy will be inaccurate. Cleary more research is required to better document and understand phylogenetic forecast horizons.

4 Discussion

4.1 What makes a forecast useful?

Generally speaking, a useful forecast will be about an important variable and be sufficiently accurate and precise. This raises at least three requirements: 1) a decision about the important variables to be predicted; 2) a measure of how closely a forecast is required to match the truth, i.e., a specific measure of forecast proficiency; and 3) a threshold forecast proficiency that defines "good enough". We consider each in turn.

Which variables are important to predict is difficult to answer generally. Species abundances and distributions would be the answer according to one textbook definition of

ecology (Begon *et al.* 1990). The sub-disciplines of ecology would have logical preferences: connectance in food web ecology (Petchey *et al.* 2010), species richness in community ecology (Algar *et al.* 2009), timing of infectious disease outbreaks in disease ecology (Shaman & Karspeck 2012), biomass or carbon in a system (Harfoot *et al.* 2014) and so on. Taking a more stakeholder-oriented approach, ecological forecasts and their horizons would be a service / product provided, and important variables should be specified by stakeholders during dialogue before predictive models are employed.

How to measure how closely a forecast matches the truth? When the forecast variable is continuous, a number of calculations on the residuals ϵ_i (predicted minus actual or $\hat{y}_i - y_i$) are useful, such as mean error (bias), mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), variance explained (R^2), and correlation between predicted and observed (see Glossary for details). Choices for binary variables (e.g., presence or absence, extinction or not) include the point-biserial correlation, statistics of the confusion matrix, and area under a receiver operating characteristic (ROC) curve. These vary in meaning, advantages, and disadvantages, and need to be carefully matched to purpose. For example, RMSE gives absolute error in units of the original variable while R^2 gives relative error on a scale of 0–1 and in proportion to the total variability in the value being predicted; AUC can be misleading because the range from predicting at random to predicting perfectly is 0.5–1 (rather than the 0–1 of R^2), which can lead people to interpret AUC scores as better than they are and there is little intuition of what counts as a good AUC score (Bahn & McGill 2013). In situations when predicting patterns (e.g., whether dynamics are cyclic or not) is more important than exact values (Levins 1966), "pattern-oriented modelling / prediction" and associated methods for comparing predictions with data could be used (Grimm & Railsback 2012). Finally, in many predictive situations, a key issue is to ensure that the data testing the predictions are independent of the data used to calibrate the model (Bahn & McGill 2007). A distinct advantage of focusing on forecasting and forecast horizons is that this level of rigor is automatically achieved.

For less applied research about ecological predictability an arbitrary forecast proficiency threshold is sufficient (e.g., the case studies above), or one could use a threshold based on the average performance of a simple statistical model. For more stakeholder-oriented services, stakeholders should be asked about how proficient is proficient enough. This may require translating the measures of forecast proficiency that are less accessible to stakeholders (e.g., r-squared, RMSE, AUC) into more intuitive ones: e.g., prediction is within an order of magnitude, is within a factor of two, or has the correct sign.

4.2 Uses of ecological forecast horizons

Ecological forecast horizons can be a general tool for assessing how well ecological variables / systems can be predicted. They are general in the sense that they can be applied in any situation where the value of a variable is predicted and there is knowledge about the known

or assumed true value of that variable. That is, they convert the output of any predictive model and any measure of forecast proficiency into a common currency: distance (be this distance in time, space, or environmental conditions). As such, ecological forecast horizons could be a powerful and flexible tool for answering questions about what in ecological is predictable, what methods offer greatest predictive power, and how forecasting is changing through time (Simmons & Hollingsworth 2002).

4.2.1 Model validation

Model validation is "the process of determining the degree to which a model and its associated data provide an accurate representation of the real world from the perspective of the intended uses of the model" (e.g., also see Chivers *et al.* 2014; quoted in Corley *et al.* 2014). When we know the truth, for example because we are predicting events that have already happened (e.g., retrodiction, postdiction, hindcasting), we can calculate prediction horizons and use these for model validation. When observations of the truth are unavailable, simulation-based forecast horizons can inform about what aspects of predictive models contribute to, or detract from predictability (for example, the first two case studies above). Obviously it will be wise to make such studies with models that are thought or known to be reasonable representation of the real system, based on knowledge of the biology of organisms and processes they are involved in. Nevertheless, model verification and validation is relatively rare for ecological models (e.g., less than half of the disease models reported in Corley et al. (2014) had experienced any model validation). Researchers and stakeholders should develop clear guidelines for verification and validation of ecological / environmental forecasting models and decide if accreditation is desirable.

In some research fields, model verification (did we build the model correctly) and validation (did we build the correct model) are extremely important, and necessary for formal accreditation and use of models (for further information see Corley *et al.* 2014). Ensemble of models (Araújo & New 2007) and use agreement (or lack of) among them could also be used as a measure of forecast proficiency. While all models make similar forecasts we might be more confident they are correct, until some distance into the future when their forecasts diverge.

4.2.2 Time, space, phylogeny, and other dimensions

The five case studies involved forecasting in time, space, and phylogeny. An ecological forecast horizon could also be used to estimate and convey predictability in environmental conditions (e.g., that species abundances can be usefully forecast for up to 5°C of warming, but not farther), ecological complexity (e.g., single species data can be used to usefully forecast in communities with up to 6 species, but not beyond), and changes in community structure (Gotelli & Ellison 2006b). Similarly, when the traits that define an organism's ecological niche are known, a forecast horizon may be defined along the axis of trait distance (Gravel *et al.* 2013).

We have concerned ourselves so far with forecasting in single dimensions. Nevertheless, forecasts simultaneously across time, environmental conditions, ecological complexity, space, phylogeny or other dimensions are likely to be quite useful.

4.2.3 Public, stakeholder and policy engagement

Harwood & Stokes (2003) proposed that ecologists face a dilemma: present persuasive simplified forecasts that pay little attention to uncertainty, or emphasise uncertainties. They go on to suggest that ecologists improve how they communicate uncertainty: "*ecologists must develop rigorous methods for evaluating these uncertainties*" (also see, e.g., Spiegelhalter *et al.* 2011; Raftery 2014).

Ecological forecast horizons could be an excellent tool for communicating predictability, as they are intuitive and the concept is already in common usage. One could argue they are more intuitive than other measures of predictability / uncertainty only because they hide details, such as the forecast proficiency measure. This seems to be only part of the reason, however, as one could hide details in an obscure and non-intuitive quantity. Perhaps another reason is because the quantity being communicated is a time (or distance in space, phylogeny, or environmental conditions). Another reasons for assisting in communicating ecological predictability is people's existing familiarity with the concept, e.g., from weather forecasting. The ease of communicating the results of quite complex research about predictability is illustrated by Shaman & Karspeck (2012) and Seferian *et al.* (2014), though one should not ignore the need to estimate, appreciate, and communicate uncertainty in forecast horizons (figure 1, and, for example, vertical error bars in figures 3, 4, & 5).

4.3 Advancing ecological predictability research

Constantly improving forecast horizons in weather forecasting are a good example of how research steered better forecasting and how emphasis on forecasting drove research (Simmons & Hollingsworth 2002). Effective forecast horizons have gone from 2–3 days to 5–7 days over the last 50 years (exact horizons depend on thresholds chosen). For the most part improved weather forecast horizons have been a result of 1) clear focus on achieving such improvements, 2) addition of subtle processes to improve the governing equations, 3) better computing power allowing models of larger areas, smaller grid cells, and more layers of the atmosphere, and 4) vastly improved measurement of the current weather conditions. Ecology will likely improve through analogous activities and practices.

4.3.1 Focusing on improving ecological forecasting

Below we list activities and practices that could increase focus on improving ecological forecasting, such as improved monitoring of ecological forecasting capabilities, development of an ecological forecasting toolbox, discussion about how to deal with uncertainties, and forecasting competitions.

- A systematic analysis of ecological forecast horizons in existing studies with appropriate data would be a worthwhile starting point to provide a baseline against which to assess improvements in ecological forecasting capabilities as well as being useful in providing information about correlates of ecological forecast horizons (e.g., figure 5).
- Ecologists could aim for a catalogue of forecasts that lists important ecological variables and their ecological forecast horizons (perhaps similar to the proposal for essential biodiversity variables Pereira *et al.* 2013). Producing this will require thorough and systematic investigations about the limits of ecological predictability. What is forecastable far into the future, what is forecastable only in the short term? Which parameters and initial conditions are more important than others, in their effects on predictability.
- Learning from the past and hindcasting has the potential to inform successful forecasting strategies. Past forecasting efforts can be objectively confronted with observations hence informing which predictions were met and the overall proficiency achieved. Both successful and failed predictions will be informative to tackle the sources of inaccurate predictions and forecast horizons are a tool that can be used to decide which refinements are vital to include in analogy with weather forecasting.
- Careful consideration is required about whether to organise research by sources of uncertainty (e.g, parameter uncertainty, model structure uncertainties, inherent stochasticity, and uncertainty in initial condition) or by effects of ecological and evolutionary processes and variables (e.g., this paper). Particularly profitable may be a combination of both, e.g., understanding the effects of processes via their effects on uncertainties.
- Making connections with the numerous dynamics systems theory tools that address predictability (Boffetta *et al.* 2002) is important. Box 1 shows how forecast horizon is related to the Lyapunov exponent of a time series. Investigating the functional importance for analysing ecological data of other methods from dynamical systems theory (e.g., Salvino *et al.* 1995; Bailey 1996; Aurell *et al.* 1997; Ziehmann *et al.* 2000; Garland *et al.* 2014) should be a research priority and will require close communication between the disciplines.
- Providing a standardized toolbox of methods for estimating and analysing ecological predictability (including via forecast horizons) applicable across the diversity of ecology study and data types (e.g., experimental, observational, replicated, unreplicated) would likely be quite useful, and we are working towards developing one. Those interested in contributing should write to the corresponding author or visit the corresponding github repository (github.com/opetchey/ecopredtools).
- Methods for dealing with a situation in which forecast of multiple variables are important should be developed. One could produce a multivariate measure of forecast proficiency, resulting in one forecast horizon for all variables. Alternatively, one could calculate a forecast horizon for each variable, perhaps using variable specific-measures of forecast

proficiency and forecast proficiency thresholds. The resulting set of forecast horizons could be presented individually, or combined into a single forecast horizon, depending on specific use cases.

- Given our acknowledged poor ability to forecast annual climate, even next year, ecological systems strongly controlled by environmental stochasticity will almost certainly show very short prediction horizons. This challenge could potentially be overcome by instead predicting a moving average of system dynamics, allowing one to evaluate longer-term trends despite shorter-term uncertainty. This would be akin to predicting climate as opposed to weather. In addition, the possibility of non-monotonic declines in forecast proficiency with forecast distance deserves further attention.
- Following the example of other fields with strong interests in accurate predictions such as economics, prediction competitions could advance methods and foster interest from non-ecologists with forecasting skills. They can provide platforms where predictions are confronted with observations on a regular basis stimulating improvements. Being based on common datasets they also allow direct comparisons of different methods in terms of forecasting proficiency. For instance, tests of ensembles of models (including process based and statistical ones) compared to predictions of single methods are easily possible as suggested in the model validation section. Such competitions are currently used in economics and also common for improving machine learning algorithms and approaches (e.g., www.kaggle.com).

4.3.2 Improving knowledge of the governing equations

The core equations governing weather forecasting are well understood (e.g., Shuman 1978). The governing equations for ecological systems include equations linking demographic rates with environmental constraints, organismal traits and dispersal abilities, and feeding rates to resource abundances, to name only a few. The previously mentioned optimism about the potential for process-based models for forecasting relies on continued efforts to better document these and other equations governing ecological dynamics: fundamental research is necessary for improved forecasting (Courchamp *et al.* 2015). Such research should, however, be explicitly combined with research about the impacts of the additional knowledge on predictive ability.

As stated previously, mechanistic models are often outperformed by simpler "model free" (Perretti *et al.* 2013a) or statistical models (see case study predictability across space; Ward *et al.* 2014) (though see Courchamp *et al.* 2015). They hence provide a baseline of minimum forecasting proficiency, on which process-based models should be judged. This could ensure that research into the governing equations is directed towards maximising increases in predictive ability.

4.3.3 Infrastructure improvements

Ecological forecast horizons will likely also improve if we continue to model larger spatial extents (making the systems modelled more closed), with finer grain sizes and with more attention to modelling multiple vertical layers (e.g., below ground processes). Predictions will likely improve as we continue to gather data with better spatial coverage and finer resolution and longer temporal extent data about the current and past conditions of variables of interest.

Large-scale integrated investment in infrastructure for predicting ecological and ecosystem states should therefore be considered. For example, ecologists, ecosystem scientists, and organisations such as the IPBES should consider aiming to develop forecasting infrastructure on the scale of the UK Meteorological Office (1,800 people employed at 60 globally distributed locations, processing over 10 million weather observations a day using an advanced atmospheric model running on a high performance supercomputer, creating 3,000 tailored forecasts and briefings a day [UK Met Office web site]).

As demonstrated above, the forecast horizon in part depends on the quality and comparability of data used to inform the predictive model. Compared to, for example, meteorology, data acquisition in the field of ecology is often less standardized across different research groups and geographic/temporal dimensions. Meteorology has used standardized tools to measure model-relevant variables, such as temperature or humidity, since the mid-19th century, such that standard weather stations based on the Stevenson screen (Stevenson 1864) have been contributing comparable data across the globe for more than a century. In ecology, even basic data (e.g., on following population abundances across different types of organisms) are acquired very differently across time and research groups, or are based on initiatives of individual researchers and then often lack spatial replication. Many "good" examples of time series of ecological data were actually collected without an ecologists' initiative (e.g., records of the number of Canada lynx and snowshoe hare pelts traded by Hudson's Bay, fisheries data, etc. which were collected mostly with an economic perspective in mind). Setting priority on which variables and parameters to measure (and how to do so in a standardized way), and following explicit information standards (e.g., Darwin Core, www.tdwg.org) and ontologies may thus be of high urgency in ecology. Efforts to make such data readily accessible (Kattge *et al.* 2011; Hudson *et al.* 2014; Salguero-Gómez *et al.* 2014) in a consistent and freely available form should be redoubled (e.g., meteorological data are not only collected in a standardized way, but also made available by National Meteorological Offices) (Costello *et al.* 2013).

4.3.4 General challenges and open questions

- Close collaboration with stakeholders is now desirable, to discover which types of stakeholders can benefit from knowing what kinds of forecast horizons. Scientific stakeholders, for example scientists that use a prediction as an input to a further model, may wish to know the forecast horizon and its consequences for predictability of their model.

Scientific organisations such as IPBES may prefer to deal with forecast horizons. Other stakeholders may require other products; understanding stakeholder diversity is key to communicating uncertainty and predictability (Raftery 2014).

- What causes observed patterns of predictability? Do they result from ecological systems being computationally irreducible (i.e., intrinsically unpredictable) such that even the best possible parameter estimates and knowledge of initial conditions cannot provide useful forecasts? Or are ecological systems intrinsically predictable, such that feeding more and more data into models will yield increases in predictability?
- Evolutionary change is increasingly recognized as an important driver of ecological dynamics, but very little is known about how evolution might affect forecast horizons. Existing work suggests that evolution could either increase or decrease the predictability of ecological dynamics. On the one hand, incorporating the potential for evolution into simple predator-prey models might substantially increase our ability to explain ecological dynamics through time (Yoshida *et al.* 2003; Hairston *et al.* 2005; Becks *et al.* 2010; Ellner *et al.* 2011; Matthews *et al.* 2011; Fischer *et al.* 2014) and might help explore how evolution could affect transitions between different dynamic states (Ellner & Turchin 1995; Fussmann *et al.* 2000). On the other hand, evolutionary trajectories that are strongly influenced by ecological dynamics causing frequency-dependent selection might lead to more unpredictable evolutionary dynamics in the long term (Doebeli & Ispolatov 2014). We need research directly addressing this uncertainty. Much less is known about how such eco-evolutionary dynamics might affect the predictability of population, community, and ecosystem level responses to environmental change (but see Vincenzi 2014).
- What are the effects of human behaviour on predictability, and how can social systems be coupled with ecological ones in predictive models (Palmer & Smith 2014)? Ecological systems include humans, such that forecasting models will need to include their actions (Palmer & Smith 2014). Scenarios coupled with quantitative models have been and may remain particularly important here (e.g., Cork *et al.* 2006)
- Recent theoretical and empirical studies emphasise predicting regime shifts in ecological systems (Takimoto 2009; Drake & Griffen 2010). Imminent changes at the population- or community-level are often preceded by internal processes such as the 'critical slowing down' in the case of population extinctions. These processes can be inferred in advance from early warning signs — in the form of generic statistical signatures — occurring after the onset of environmental perturbation and before the critical system transition. The forecast horizon of such signals remains relatively unexplored.

5 Conclusions

We believe we have shown only a fraction of the potential of forecast horizons in ecological research. They are a general and intuitive tool and have potential to guide future research agendas to improve predictability not only by stimulating scientists to make quantitative predictions, but also to actively confront these predictions with observed dynamics. Forecast horizons provide baselines about how well we can predict specific dynamics of interest, and when and why accurate predictions succeed or fail. Given these properties, we believe that the forecast horizon can be an important tool in making the science of ecology even more predictive. Nevertheless, research should also aim for complementary and perhaps even better tools for advancing and organising predictability research in ecology.

6 Acknowledgements

Brian McGill wrote a series of four articles on the dynamicecology.wordpress.com website that, with comments from readers, proved very valuable; these are recommended reading. Elisa Benincà provided the data from figure 2 of Benincà *et al.* (2008). Contributions of OLP, MS, and BS are supported by the University of Zurich Research Priority Program on '*Global Change and Biodiversity*' (URPP GCB). Contributions of OLP are supported by SNF Project 31003A_137921 *Predicting community responses to environmental change* and two Tansley Working Groups: *FunKey Traits: Using Key Functional Traits to develop ecosystem service indicators*, and *Advancing the ecological foundations of sustainability science*. Contributions of FA are supported by the Swiss National Science Foundation Grant No PP00P3_150698. Contributions of MP are supported by the Swedish Research Council.

7 Glossary

Prediction. Two types of predictions can be distinguished.

Explanatory predictions are formulations of what should be expected if the general hypotheses of a theory or model are correct. Their aim is to assist with testing a model or a theory. They can be rejected or not, which validates or not the hypothesis (Popper 2002). *Anticipatory predictions* are formulations of a possible future assuming that the current interactions and processes will hold in the future (i.e. that the hypothesis of the models/theories are validated and can be extended in to the future). Their aim is to give a statement about what the future will be.

Projection A statement about the future based on extrapolating models to domains for which there are no data (Coreau *et al.* 2009).

Forecast The best projection of the future from a model or expert.

Scenarios Alternative futures based on consistent sets of assumptions, interactions and driving forces (Bennett *et al.* 2003); provide a set of plausible pathways to the futures, rather than predicting what the future will actually be (Coreau *et al.* 2009).

Accuracy The difference between observed and predicted value. High accuracy implies good prediction and low accuracy poor prediction. Accuracy is an important component of forecast proficiency (see below).

Precision The amount of uncertainty in predictions. Precise predictions will have low uncertainty (i.e., be closely grouped around the mean prediction). Imprecise predictions will have high uncertainty. Unlike accuracy, very high precision may indicate a poor predictive model that might result, for example, from failing to include a stochastic process. Low precision is also a sign of a poor predictive model. Hence, best is if a predictive model produces a prediction that has the same uncertainty as the real system being modelled.

Uncertainty. Regan *et al.* (2002) give two classes of uncertainty: epistemic and linguistic. Epistemic uncertainty is lack of knowledge in the state of a system, for example in parameter values, processes operating, representation of processes, system components, and inherent randomness (also see Clark *et al.* 2001). See Gregr & Chan (Gregr & Chan 2014) for discussion of the relationship between modelling assumptions and uncertainties.

Intrinsic and realised predictability Beckage *et al.* (2011) recognise two types of predictability: the intrinsic predictability of a system, and the realised predictability achieved by a particular model of the system. The intrinsic predictability of a system is the predictability of the best possible model of that system, i.e., it is the greatest achievable predictability. Low realised predictability and high intrinsic predictability implies problems with the predictive model, such as uncertainty in parameter values. High predictability requires an intrinsically predictable system, and low uncertainty about the processes governing the system. A fully deterministic system has perfect intrinsic predictability, since perfect knowledge of parameters and initial conditions results in perfect predictions. A fully deterministic system may, however, be computationally irreducible.

Forecast proficiency A measure of how useful is a forecast, usually some function of accuracy and or precision. We first thought to use instead the term *forecast skill*, which comes from meteorology and there usually refers to a specific measure of accuracy, mean square error, and has already been used in environmental science to assess forecasts of marine net primary production (Seferian *et al.* 2014). Forecast skill is, however, often used to mean one measure, mean square error, and we do not wish to be so specific. We propose that in ecology, the term *forecast proficiency* be general, such that any measure of accuracy or match in precision can be a measure of forecast proficiency. Thus, a model with high accuracy and appropriate precision will have high forecast proficiency. Very high precision or very low precision may both be inappropriate and contribute to lower forecast proficiency.

Measures of forecast proficiency for continuous variables include mean error or bias=$E(\epsilon_i)=1/n\ \Sigma\ \epsilon_i$, which gives a measure of whether predictions are consistently wrong in one direction. Mean squared error is given by MSE=$E(\epsilon_i^{\ 2})=1/n\ \Sigma\ \epsilon_i^{\ 2}$. Taking the square root gives

root mean squared error, RMSE=$\sqrt{MSE}$ and is in the units of the original variable. Another common measure is variance explained, R^2=1-SSE/SST=1-$\Sigma\epsilon_i^2/\Sigma\ y_i^2$=1-MSE/VAR($y_i$). A relative of RMSE that is robust to outliers is Mean Absolute Error MAE=1/n $\Sigma|\epsilon_i|$. The correlation between predicted and observed, r=cor($\hat{y}_i, y_i$), is sometimes used but is a weaker assessment since predictions that are biased or not falling on a 1-to-1 line in a predicted vs. observed plot can still have a perfect correlation of one (r=1). MSE has the useful property of combining accuracy and precision.

For binary variables, the choices are less obvious. The observed values can be coded as zero and one and the predicted values kept as a probability between 0–1 and the Pearson correlation can be calculated. This is called the point-biserial correlation and is an easily understood metric but the values will be lower than correlation of continuous variables. Alternatively, a confusion matrix can be calculated. A confusion matrix is a 2x2 table giving counts of true and false positives and true and false negatives. One entry, the true positives, is a measure of accuracy, but a number of other values can be calculated from the confusion matrix that correct for an uneven ratio of positives to negatives. One commonly used metric in scenarios requiring thresholding is the AUC or Area Under the Curve (the curve being a receiver operator curve; though see the caveats in the main text).

Forecast horizon The distance in time, space, or environmental parameters at which forecast proficiency falls below the forecast proficiency threshold. Forecast horizon is closely related to concepts such as mean and maximal forecast time (e.g., Salvino *et al.* 1995).

Forecast proficiency threshold The value of forecast proficiency above which forecasts are useful, and below which forecasts are not useful.

Retrodiction / postdiction / hindcasting Each relates to the practice of testing the predictions of models / theories against observations already in existence at the time when the predictions were made. While care is required to understand how the existing observation might have influenced the predictions, prediction horizons can be calculated, and provide an indication about prediction into the future.

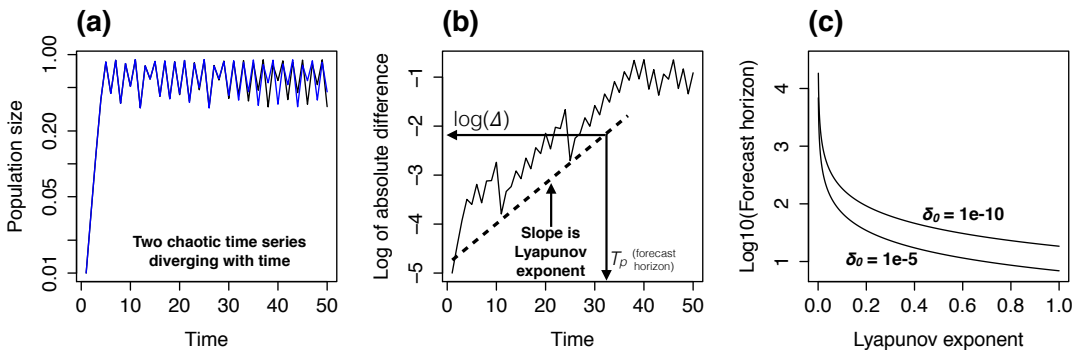
8 Box 1. Lyapunov Exponents and the ecological forecast horizon

Dynamical systems theory concerns, in part, the predictability of dynamics (e.g., Boffetta *et al.* 2002). In particular, the Lyapunov exponent (LE) is closely related to intrinsic predictability of a deterministic system. The LE is a measure of the rate of separation of close trajectories (box figure a). For example, consider the logistic map $x_{t+1} = rx_t(1 - x_t)$, where x_t is population size at time t and r is the growth rate. Let initial size of one replicate population be x_o, and $x_0' = x_0 + \delta_0$ is the starting size of another population. The difference in size of the two populations initially is δ_0, and the difference at time t is δ_t (box figure b). How δ_t changes through time is characterised by the LE (λ), according to the equation $\delta_t = \delta_0 e^{\lambda t}$. Thus, when $\lambda > 0$ the intial difference grows exponentially, whereas if $\lambda < 0$ the difference shrinks exponentially.

In order to translate the LE into a forecast horizon, we must know two things: 1) the amount of uncertainty in initial conditions (δ_0); 2) the required precision of the prediction Δ (i.e., the forecast proficiency threshold). The forecast horizon is given by the heuristic equation

$$T_p \sim \frac{1}{\lambda} \ln\left(\frac{\Delta}{\delta_0}\right) \quad \text{(equation 1)}$$

The forecast horizon T_p (otherwise known as the predictability time) is the time at which a small error in the initial condition becomes large enough to preclude a useful forecast. T_p is determined by the inverse of the LE, while it has weak dependence on δ_0 and Δ (box figure c). Negative LE result in an infinite forecast horizon. In case the system is multidimensional (e.g. a multispecies community) there is a LE for every dimension and predictability is determined by the largest LE of the system.



Box figure: (a) Two population dynamic time series originated by two nearby initial conditions ($x_0 = 0.01$, $x'_0 = x_0 + \delta_0$, with $\delta_0 = 10^{-5}$) using the Logistic map with growth rate = 3.6. (b) Growth of the logarithm of the difference of the two times series in panel (a). (c) Relationship between forecast horizon (Tp) and the Lyapunov exponent predicted by equation 1, for two sizes of δ_0.

9 Bibliography

1.
Algar, A.C., Kharouba, H.M., Young, E.R. & Kerr, J.T. (2009). Predicting the future of species diversity: Macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography (Cop)*, 32, 22–33.

2.
Araújo, M.B. & New, M. (2007). Ensemble forecasting of species distributions. *Trends Ecol Evol*, 22, 42–47.

3.
Asner, G.P. (2009). Tropical forest carbon assessment: Integrating satellite and airborne mapping approaches. *Environ Res Lett*, 4, 034009.

4.
Aufderheide, H., Rudolf, L., Gross, T. & Lafferty, K.D. (2013). How to predict community responses to perturbations in the face of imperfect knowledge and network complexity. *Proc R Soc B Biol Sci*, 280, 20132355.

5.
Aurell, E., Boffetta, G., Crisanti, A., Paladin, G. & Vulpiani, A. (1997). Predictability in the large: An extension of the concept of Lyapunov exponent. *J Phys A Math Gen*, 30, 1–26.

6.
Bahn, V. & McGill, B.J. (2007). Can niche-based distribution models outperform spatial interpolation? *Glob Ecol Biogeogr*, 16, 733–742.

7.
Bahn, V. & McGill, B.J. (2013). Testing the predictive performance of distribution models. *Oikos*, 122, 321–331.

8.
Bailey, B. (1996). Local Lyapunov Exponents: Predictability depends on where you are. In: *Nonlinear Dyn Econ Proc Tenth Int Symp Econ Theory Econom* (eds. Barnett, W.A., Kirman, A.P. & Salmon, M.). Cambridge University Press, pp. 345–360.

9.
Beckage, B., Gross, L.J. & Kauffman, S. (2011). The limits to prediction in ecological systems. *Ecosphere*, 2, art125.

10.
Becks, L., Ellner, S.P., Jones, L.E. & Hairston, N.G. (2010). Reduction of adaptive genetic diversity radically alters eco-evolutionary community dynamics. *Ecol Lett*, 13, 989–997.

11.
Begon, M., Harper, J.L. & Townsend, C.R. (1990). *Ecology: Individuals, Populations and Communities*. BLACKWELL PUBLISHING.

12.
Benincà, E., Huisman, J., Heerkloss, R., Jöhnk, K.D., Branco, P., Van Nes, E.H., *et al.* (2008). Chaos in a long-term experiment with a plankton community. *Nature*, 451, 822–825.

13.
Bennett, E.M., Carpenter, S.R., Peterson, G.D., Cumming, G.S., Zurek, M. & Pingali, P. (2003). Why global scenarios need ecology. *Front Ecol Environ*, 1, 322–329.

14.
Berlow, E.L., Dunne, J., Martinez, N.D., Stark, P.B., Williams, R.J. & Brose, U. (2009). Simple prediction of interaction strengths in complex food webs. *Proc Natl Acad Sci*, 106, 187–191.

15.
Boffetta, G., Cencini, M., Falcioni, M. & Vulpiani, A. (2002). Predictability: A way to characterize complexity. *Phys Rep*, 356, 367–474.

16.
Brook, B.W., O'Grady, J.J., Chapman, P., Burgman, M., Akçakaya, H.R. & Frankham, R. (2000). Predictive accuracy of population viability analysis in conservation biology. *Nature*, 404, 385–387.

17.
Cane, M.A., Eshel, G. & Buckland, R.W. (1994). Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea surface temperature. *Nature*, 370, 204–205.

18.
Cannell, M.G.R. & Smith, R.I. (1983). Thermal time, chill days and prediction of budburst in Picea Sitchensis. *J Appl Ecol*, 20, 951–963.

19.
Carrara, F., Giometto, A., Seymour, M., Rinaldo, A. & Altermatt, F. (2015). Experimental evidence for strong stabilizing forces at high functional diversity of microbial communities. *Ecology*, in press.

20.
Cavender-Bares, J., Keen, A. & Miles, B. (2006). Phylogenetic structure of floridian plant communities depends on taxonomic and spatial scale. *Ecology*, 87, S109–S122.

21.
Chivers, C., Leung, B. & Yan, N.D. (2014). Validation and calibration of probabilistic predictions in ecology. *Methods Ecol Evol*, 5, 1023–1032.

22.
Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J., *et al.* (2001). Ecological forecasts: An emerging imperative. *Science*, 293, 657–660.

23.

Corbet, S.A., Saville, N.M., Fussell, M., Prŷs-Jones, O.E. & Unwin, D.M. (1995). The competition box : A graphical aid to forecasting pollinator performance. *J Appl Ecol*, 32, 707–719.

24.

Coreau, A., Pinay, G., Thompson, J.D., Cheptou, P.-O. & Mermet, L. (2009). The rise of research on futures in ecology: Rebalancing scenarios and predictions. *Ecol Lett*, 12, 1277–1286.

25.

Cork, S.J., Peterson, G.D., Bennett, E.M., Petschel-Held, G. & Zurek, M. (2006). Synthesis of the storylines. *Ecol Soc*, 11, 11.

26.

Corley, C.D., Pullum, L.L., Hartley, D.M., Benedum, C., Noonan, C., Rabinowitz, P.M., *et al.* (2014). Disease prediction models and operational readiness. *PLoS One*, 9, e91989.

27.

Costello, M.J., Michener, W.K., Gahegan, M., Zhang, Z.-Q. & Bourne, P.E. (2013). Biodiversity data should be published, cited, and peer reviewed. *Trends Ecol Evol*, 28, 454–461.

28.

Courchamp, F., Dunne, J. a., Le Maho, Y., May, R.M., Thébaud, C. & Hochberg, M.E. (2015). Fundamental ecology is fundamental. *Trends Ecol Evol*, 30, 9–16.

29.

Dambacher, J., Li, H. & Rossignol, P. (2002). Relevance of community structure in assessing indeterminacy of ecological predictions. *Ecology*, 83, 1372–1385.

30.

Dambacher, J., Li, H. & Rossignol, P. (2003). Qualitative predictions in model ecosystems. *Ecol Modell*, 161, 79–93.

31.

Diez, J.M., Ibáñez, I., Miller-Rushing, A.J., Mazer, S.J., Crimmins, T.M., Crimmins, M., *et al.* (2012). Forecasting phenology: From species variability to community patterns. *Ecol Lett*, 15, 545–553.

32.

Dobrowski, S. & Thorne, J. (2011). Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecol Monogr*, 81, 241–257.

33.

Doebeli, M. & Ispolatov, I. (2014). Chaos and unpredictability in evolution. *Evolution (N Y)*, 68, 1365–1373.

34.

Dowd, M. (2007). Bayesian statistical data assimilation for ecosystem models using Markov Chain Monte Carlo. *J Mar Syst*, 68, 439–456.

35.

Drake, J.M. & Griffen, B.D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, 467, 456–459.

36.

Ellner, S. & Turchin, P. (1995). Chaos in a noisy world: New methods and evidence from time-series analysis. *Am Nat*, 145, 343–375.

37.

Ellner, S.P., Geber, M. & Hairston, N.G. (2011). Does rapid evolution matter? Measuring the rate of contemporary evolution and its impacts on ecological dynamics. *Ecol Lett*, 14, 603–614.

38.

Evans, M.R. (2012). Modelling ecological systems in a changing world. *Philos Trans R Soc Lond B Biol Sci*, 367, 181–190.

39.

Evans, M.R., Bithell, M., Cornell, S.J., Dall, S., Diaz, S., Emmott, S., *et al.* (2013). Predictive systems ecology. *Proc R Soc B Biol Sci*, 280, 20131452.

40.

Evans, M.R., Norris, K.J. & Benton, T.G. (2012). Predictive ecology: Systems approaches. *Philos Trans R Soc Lond B Biol Sci*, 367, 163–169.

41.

Fischer, B.B., Kwiatkowski, M., Ackermann, M., Krismer, J., Roffler, S., Suter, M.J.F., *et al.* (2014). Phenotypic plasticity influences the eco-evolutionary dynamics of a predator-prey system. *Ecology*, 95, 3080–3092.

42.

France, K.E. & Duffy, J.E. (2006). Diversity and dispersal interactively affect predictability of ecosystem function. *Nature*, 441, 1139–1143.

43.

Fussmann, G., Ellner, S., Shertzer, K. & Hairston, J.N. (2000). Crossing the Hopf bifurcation in a live predator-prey system. *Science*, 290, 1358–1360.

44.

Fussmann, K.E., Schwarzmüller, F., Brose, U., Jousset, A. & Rall, B.C. (2014). Ecological stability in response to warming. *Nat Clim Chang*, 4, 206–210.

45.

Gao, C., Wand, H., Weng, E., Lakshmivarahan, S., Zhang, Y. & Luo, Y. (2011). Assimilation of multiple data sets with the ensemble Kalman filter to improve forecasts of forest carbon dynamics. *Ecol Appl*, 21, 1461–1473.

46.

Garland, J., James, R. & Bradley, E. (2014). Model-free quantification of time-series predictability. *Phys Rev E*, 90, 052910.

47.

Garonna, I., de Jong, R., de Wit, A.J.W., Mücher, C., Schmid, B. & Schaepman, M.E. (2014). Strong contribution of autumn phenology to changes in satellite-derived growing season length estimates across Europe (1982-2011). *Glob Chang Biol*, 20, 3457–3470.

48.

Gilbert, B., Tunney, T.D., McCann, K.S., DeLong, J.P., Vasseur, D., Savage, V., *et al.* (2014). A bioenergetic framework for the temperature dependence of trophic interactions. *Ecol Lett*, 17, 902–914.

49.

Godfray, H.C.J. & May, R.M. (2014). Open questions: Are the dynamics of ecological communities predictable? *BMC Biol*, 12, 22.

50.

Gotelli, N.J. & Ellison, A.M. (2006a). Forecasting extinction risk with nonstationary matrix models. *Ecol Appl*, 16, 51–61.

51.

Gotelli, N.J. & Ellison, A.M. (2006b). Food-web models predict species abundances in response to habitat change. *PLoS Biol*, 4, e324.

52.

Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). Inferring food web structure from predator-prey body size relationships. *Methods Ecol Evol*, 4, 1083–1090.

53.

Gregr, E.J. & Chan, K.M.A. (2014). Leaps of Faith: How Implicit Assumptions Compromise the Utility of Ecosystem Models for Decision-making. *Bioscience*, 65, 43–54.

54.

Grimm, V. & Railsback, S.F. (2012). Pattern-oriented modelling: A “multi-scope” for predictive systems ecology. *Philos Trans R Soc Lond B Biol Sci*, 367, 298–310.

55.

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecol Lett*, 8, 993–1009.

56.

Hairston, N.G., Ellner, S.P., Geber, M., Yoshida, T. & Fox, J. (2005). Rapid evolution and the convergence of ecological and evolutionary time. *Ecol Lett*, 8, 1114–1127.

57.

Hare, J.A., Alexander, M.A., Fogarty, M.J., Williams, E.H. & Scott, J.D. (2010). Forecasting the dynamics of a coastal fishery species using a coupled climate – population model. *Ecol Appl*, 20, 452–464.

58.

Harfoot, M.B.J., Newbold, T., Tittensor, D.P., Emmott, S., Hutton, J., Lyutsarev, V., *et al.* (2014). Emergent global patterns of ecosystem structure and function from a mechanistic general ecosystem model. *PLoS Biol*, 12, e1001841.

59.

Hartig, F. & Dormann, C.F. (2013). Does model-free forecasting really outperform the true model? *Proc Natl Acad Sci*, 110, E3975.

60.

Harwood, J. & Stokes, K. (2003). Coping with uncertainty in ecological advice: Lessons from fisheries. *Trends Ecol Evol*, 18, 617–622.

61.

Hijmans, R., Forbes, G. & Walker, T. (2000). Estimating the global severity of potato late blight with GIS-linked disease forecast models. *Plant Pathol*, 49, 697–705.

62.

Homolová, L., Schaepman, M., Lamarque, P., Clevers, J.G.P.W., de Bello, F., Thuiller, W., *et al.* (2013). Comparison of remote sensing and plant trait-based modelling to predict ecosystem services in subalpine grasslands. *Ecosphere*, 5, 1–29.

63.

Hudson, L. & Reuman, D. (2013). A cure for the plague of parameters: Constraining models of complex population dynamics with allometries. *Proc R Soc B Biol Sci*, 280, 20131901.

64.

Hudson, L.N., Newbold, T., Contu, S., Hill, S.L., Lysenko, I., De Palma, A., *et al.* (2014). The PREDICTS database: A global database of how local terrestrial biodiversity responds to human impacts. *Ecol Evol*, 4, 4701–4735.

65.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., *et al.* (2011). TRY - a global database of plant traits. *Glob Chang Biol*, 17, 2905–2935.

66.

Komatsu, E., Fukushima, T. & Harasawa, H. (2007). A modeling approach to forecast the effect of long-term climate change on lake water quality. *Ecol Modell*, 209, 351–366.

67.

Kooistra, L., Wamelink, W., Schaepman-Strub, G., Schaepman, M., van Dobben, H., Aduaka, U., *et al.* (2008). Assessing and predicting biodiversity in a floodplain ecosystem: Assimilation of net primary production derived from imaging spectrometer data into a dynamic vegetation model. *Remote Sens Environ*, 112, 2118–2130.

68.

Laurent, V.C.E., Schaepman, M.E., Verhoef, W., Weyermann, J. & Chávez, R.O. (2014). Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2 top-of-atmosphere radiance image. *Remote Sens Environ*, 140, 318–329.

69.

Levine, J.M. & Antonio, C.M.D. (2003). Forecasting biological invasions with international trade. *Conserv Biol*, 17, 322–326.

70.

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

71.

Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009). The velocity of climate change. *Nature*, 462, 1052–1055.

72.

Lorenz, E. (1965). A study of the predictability of a 28-variable atmospheric model. *Tellus*, 17, 321–333.

73.

Matthews, B., Narwani, A., Hausch, S., Nonaka, E., Peter, H., Yamamichi, M., *et al.* (2011). Toward an integration of evolutionary biology and ecosystem science. *Ecol Lett*, 14, 690–701.

74.

May, R.M. (2001). *Stability and Complexity in Model Ecosystems*. Princeton University Press.

75.

McCann, K.S. (2000). The diversity-stability debate. *Nature*, 405, 228–233.

76.

McGrady-Steed, J. & Harris, P. (1997). Biodiversity regulates ecosystem predictability. *Nature*, 390, 162–165.

77.

Mouquet, N., Lagadeuc, Y. & Loreau, M. (2012). Ecologie prédictive and changement planétaire. *Les Cah Prospect - Rep*, 9–44.

78.

Nekola, J. & White, P. (1999). The distance decay of similarity in biogeography and ecology. *J Biogeogr*, 867–878.

79.
Nekola, J.C. & McGill, B.J. (2014). Scale dependency in the functional form of the distance decay relationship. *Ecography (Cop)*, 37, 309–320.

80.
Niu, S., Luo, Y., Dietze, M., Keenan, T. & Shi, Z. (2014). The role of data assimilation in predictive ecology. *Ecosphere*, 5, 1–16.

81.
Novak, M., Wootton, J.T., Doak, D.F., Emmerson, M., Estes, J.A. & Tinker, M.T. (2011). Predicting community responses to perturbations in the face of imperfect knowledge and network complexity. *Ecology*, 92, 836–846.

82.
Ollerenshaw, C.B. & Smith, L.P. (1969). Meteorological factors and forecasts of helminthic disease. *Adv Parasitol*, 7, 283–323.

83.
Palmer, P.I. & Smith, M.J. (2014). Model human adaptation to climate change. *Nature*, 512, 365–366.

84.
Pearse, I.S. & Altermatt, F. (2013). Predicting novel trophic interactions in a non-native world. *Ecol Lett*, 16, 1088–1094.

85.
Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., *et al.* (2013). Essential biodiversity variables. *Science*, 339, 277–278.

86.
Perretti, C.T., Munch, S.B. & Sugihara, G. (2013a). Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. *Proc Natl Acad Sci*, 110, 5253–5257.

87.
Perretti, C.T., Munch, S.B. & Sugihara, G. (2013b). Reply to Hartig and Dormann : The true model myth. *Proc Natl Acad Sci*, 110, 3976–3977.

88.
Petchey, O., Brose, U. & Rall, B. (2010). Predicting the effects of temperature on food web connectance. *Philos Trans R Soc Lond B Biol Sci*, 365, 2081–2091.

89.
Popper, K. (2002). *The logic of scientific discovery*. 2nd edn. Routledge Taylor & Francis Group.

90.

Pottier, J., Malenovský, Z., Psomas, A., Homolová, L., Schaepman, M.E., Choler, P., *et al.* (2014). Modelling plant species distribution in alpine grasslands using airborne imaging spectroscopy. *Biol Lett*, 10, 20140347.

91.

Purves, D. & Pacala, S. (2008). Predictive models of forest dynamics. *Science*, 320, 1452–1453.

92.

Purves, D., Scharlemann, J., Harfoot, M., Newbold, T., Tittensor, D., Hutton, J., *et al.* (2013). Ecosystems: Time to model all life on earth. *Nature*, 493, 295–297.

93.

Raftery, A.E. (2014). Use and communication of probabilistic forecasts. *arXiv*, 1408.4812.

94.

Regan, H., Colyvan, M. & Burgman, M. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol Appl*, 12, 618–628.

95.

Ripa, J., Storlind, L., Lundberg, P. & Brown, J. (2009). Niche coevolution in consumer–resource communities. *Evol Ecol Res*, 11, 305–323.

96.

Salguero-Gómez, R., Jones, O.R., Archer, C.R., Buckley, Y.M., Che-Castaldo, J., Caswell, H., *et al.* (2014). The compadre Plant Matrix Database: an open online repository for plant demography. *J Ecol*, doi: 10.1111/1365–2745.12334.

97.

Salvino, L.W., Cawley, R., Grebogi, C. & Yorke, J.A. (1995). Predictability of time series. *Phys Lett A*, 209, 327–332.

98.

Schimel, D.S., Asner, G.P. & Moorcroft, P. (2013). Observing changing ecological diversity in the Anthropocene. *Front Ecol Environ*, 11, 129–137.

99.

Seferian, R., Bopp, L., Gehlen, M., Swingedouw, D., Mignot, J., Guilyardi, E., *et al.* (2014). Multiyear predictability of tropical marine productivity. *Proc Natl Acad Sci*, 111, 11646–11651.

100.

Shaman, J. & Karspeck, A. (2012). Forecasting seasonal outbreaks of influenza. *Proc Natl Acad Sci*, 109, 20425–20430.

101.

Shuman, F. (1978). Numerical weather prediction. *Bull Am Meteorol Soc*, 59, 5–17.

102.

Simmons, A. & Hollingsworth, A. (2002). Some aspects of the improvement in skill of numerical weather prediction. *Q J R Meteorol Soc*, 647–677.

103.

Spiegelhalter, D., Pearson, M. & Short, I. (2011). Visualizing uncertainty about the future. *Science*, 333, 1393–1400.

104.

Stevenson, T.C.E. (1864). New description of box for holding thermometers. *J Scottish Meteorol Soc*, 1, 122.

105.

Stigall, A.L. (2012). Using ecological niche modelling to evaluate niche stability in deep time. *J Biogeogr*, 39, 772–781.

106.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., *et al.* (2012). Detecting causality in complex ecosystems. *Science*, 338, 496–500.

107.

Sutherland, W.J. (2006). Predicting the ecological consequences of environmental change: A review of the methods. *J Appl Ecol*, 43, 599–616.

108.

Sutherland, W.J., Armstrong-brown, S., Paul, R., Brereton, T.O.M., Brickland, J., Colin, D., *et al.* (2006). The identification of 100 ecological questions of high policy relevance in the UK, 617–627.

109.

Takimoto, G. (2009). Early warning signals of demographic regime shifts in invading populations. *Popul Ecol*, 51, 419–426.

110.

Tallis, H.M. & Kareiva, P. (2006). Shaping global environmental decisions using socio-ecological models. *Trends Ecol Evol*, 21, 562–8.

111.

Turchin, P. (2003). *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. Princeton University Press.

112.

Vandermeer, J. (1969). The competitive structure of communities: An experimental approach with protozoa. *Ecology*, 50, 362–371.

113.

Vincenzi, S. (2014). Extinction risk and eco-evolutionary dynamics in a variable environment with increasing frequency of extreme events. *J R Soc Interface*, 11, 20140441.

114.

Ward, E.J., Holmes, E.E., Thorson, J.T. & Collen, B. (2014). Complexity is costly: A meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, 123, 652–661.

115.

Wollrab, S., Diehl, S. & De Roos, A.M. (2012). Simple rules describe bottom-up and top-down control in food webs with alternative energy pathways. *Ecol Lett*, 15, 935–946.

116.

Wootton, J.T. (2002). Indirect effects in complex ecosystems: Recent progress and future challenges. *J Sea Res*, 48, 157–172.

117.

Wootton, J.T. (2004). Markov chain models predict the consequences of experimental extinctions. *Ecol Lett*, 7, 653–660.

118.

Yachi, S. & Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proc Natl Acad Sci*, 96, 1463–1468.

119.

Yodzis, P. (1988). The indeterminacy of ecological interactions as perceived through perturbation experiments. *Ecology*, 69, 508–515.

120.

Yoshida, T., Jones, L.E., Ellner, S.P., Fussmann, G.F. & Hairston, N.G. (2003). Rapid evolution drives ecological dynamics in a predator-prey system. *Nature*, 424, 303–306.

121.

Ziehmann, C., Smith, L. a. & Kurths, J. (2000). Localized Lyapunov exponents and the prediction of predictability. *Phys Lett A*, 271, 237–251.

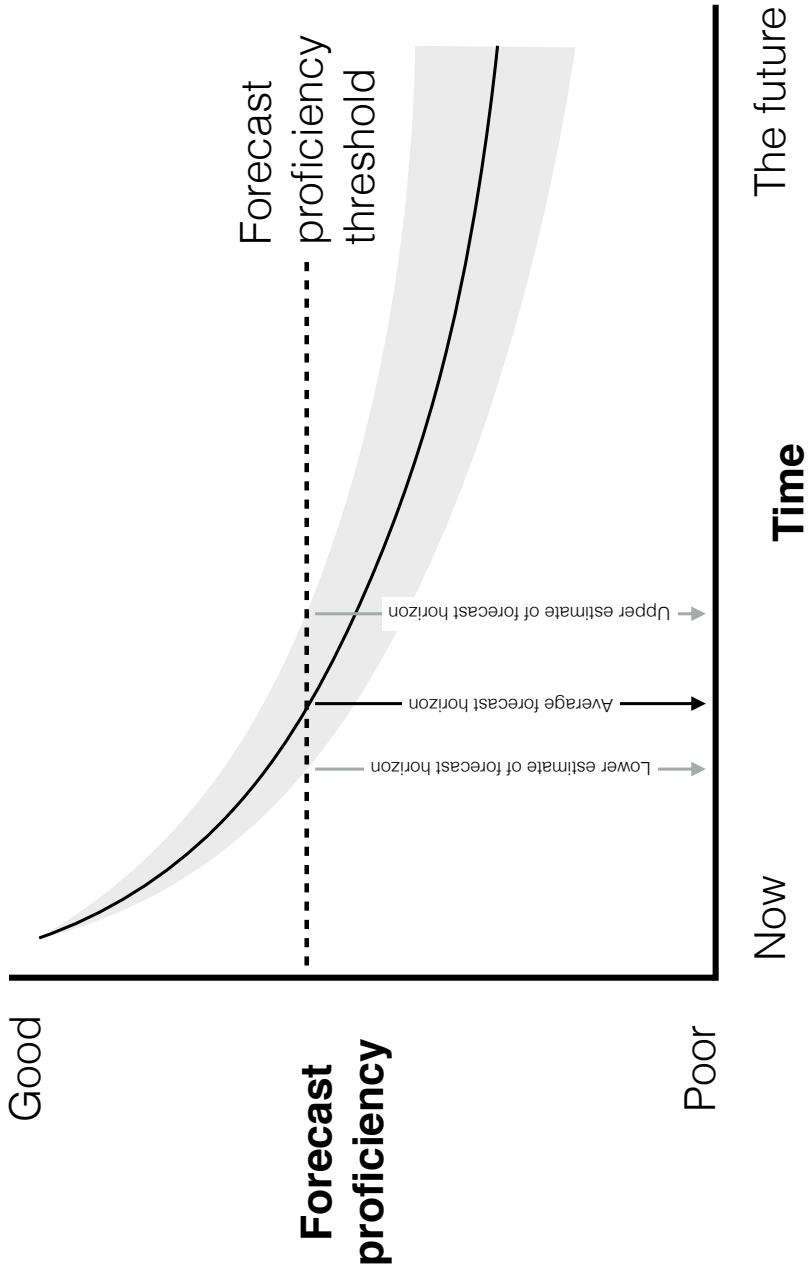


Figure 1. The forecast horizon is the time at which average forecast proficiency (black curved line) falls below the prediction proficiency threshold. Because forecast proficiency at any particular time will be a distribution (e.g., grey area) there will be a distribution of forecasts horizons that can be used as an estimate of uncertainty in forecast horizon (e.g., give a lower and upper estimate of the forecast horizon).

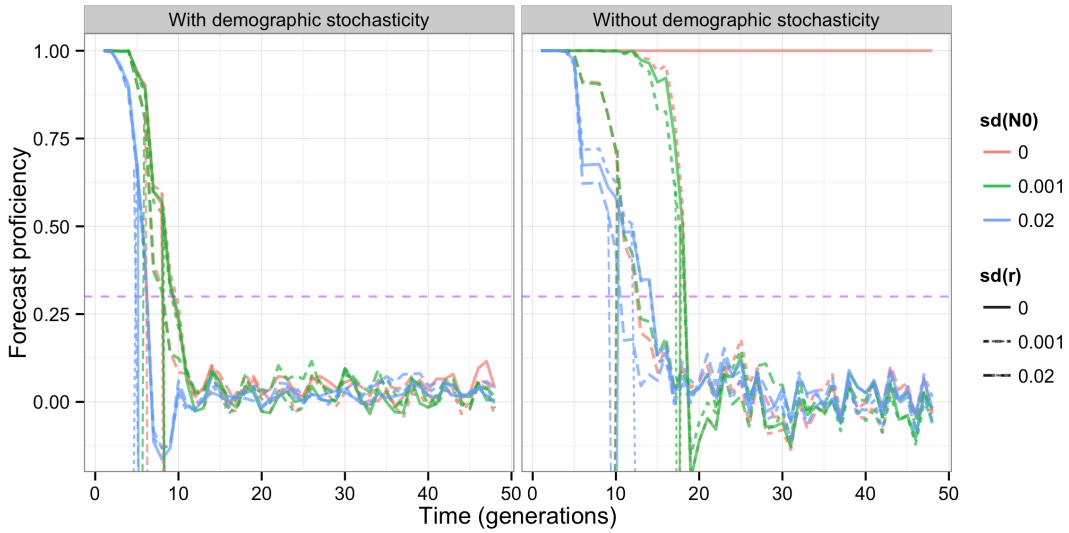


Figure 2. Forecast proficiency as a function of how far into the future forecasts are made, for different levels of uncertainty in the growth rate parameter [sd(r)] of the predictive model, and uncertainty in the initial population size [sd(N_0)] of the predictive model. Also shown is the effect of the presence or absence of demographic stochasticity in the true dynamics. The y-axis shows average forecast proficiencies across replicates. The horizontal purple dashed line is the forecast proficiency threshold (arbitrarily 0.3) and the vertical lines point show the furthest time into the future at which forecast proficiency is above the forecast proficiency threshold, i.e., vertical lines show the forecast horizon.

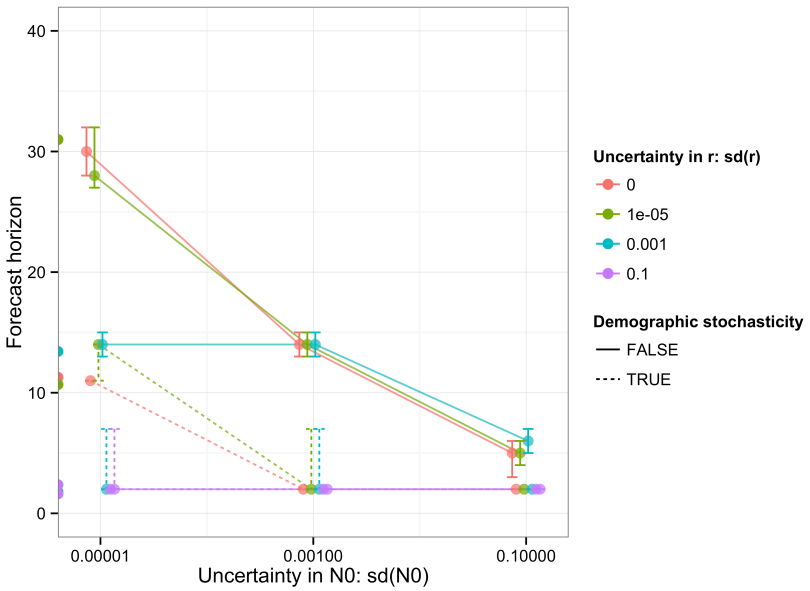


Figure 3. Median (+- 55th to 65th percentile) forecast horizon (number of generations) as a function of uncertainty in initial condition *N*0 and growth rate *r* for population dynamics with or without demographic stochasticity. The two apparently missing lines are under the purple line. Points on the y-axis are for when sd(N0) = 0 so that log10 of this is undefined; other points are moved slightly in the y direction to make them visible. The orange point for sd(N0) = 0 and without demographic stochasticity is not shown, since it has an unlimited forecast horizon. The 55th to 66th percentile was chosen to give reasonably small error bars, for clarity.

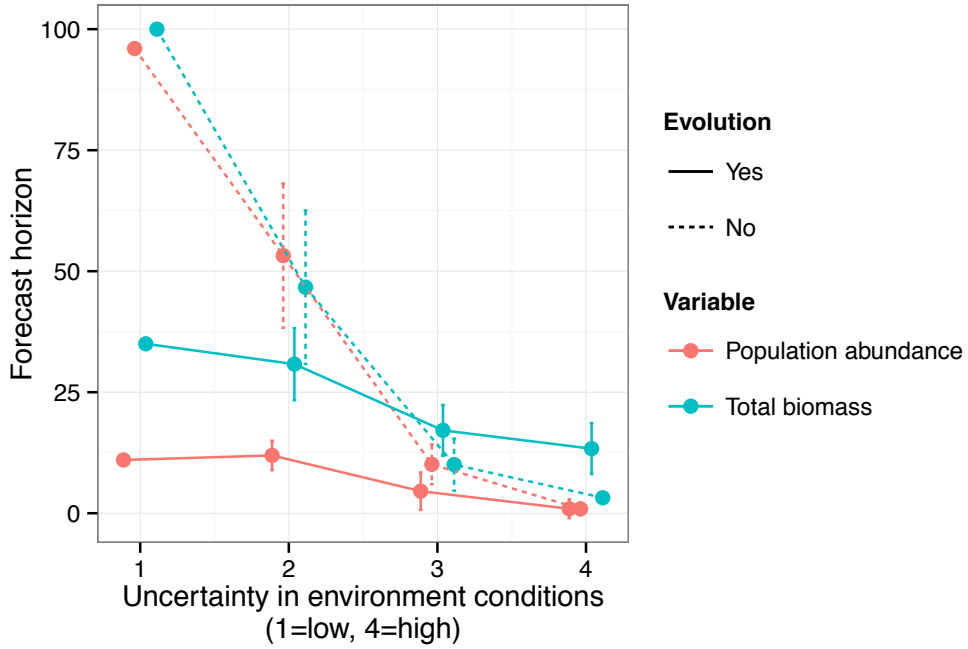


Figure 4. Effects of uncertainty about future environment (x-axis), of evolution, and of level of ecological organisation on forecast horizon (number of generations). Data come from a simulation study of a community of competitors. Error bars are one standard deviation.

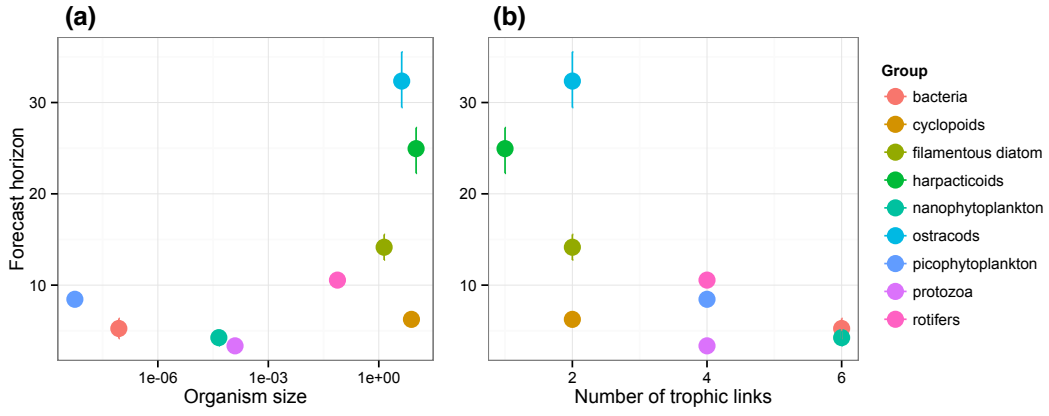


Figure 5 Forecast horizons (days) from Benincà *et al.* (2008) plotted against (a) approximate body size of the organisms in taxonomic groups (gathered from the literature) and (b) number of trophic links (taken from figure 1a of Benincà *et al.* (2008)). Y-error bars show the range of forecast horizons constructed from the 95% confidence intervals of curve fits to data in Figure 2 of Benincà *et al.* (2008).

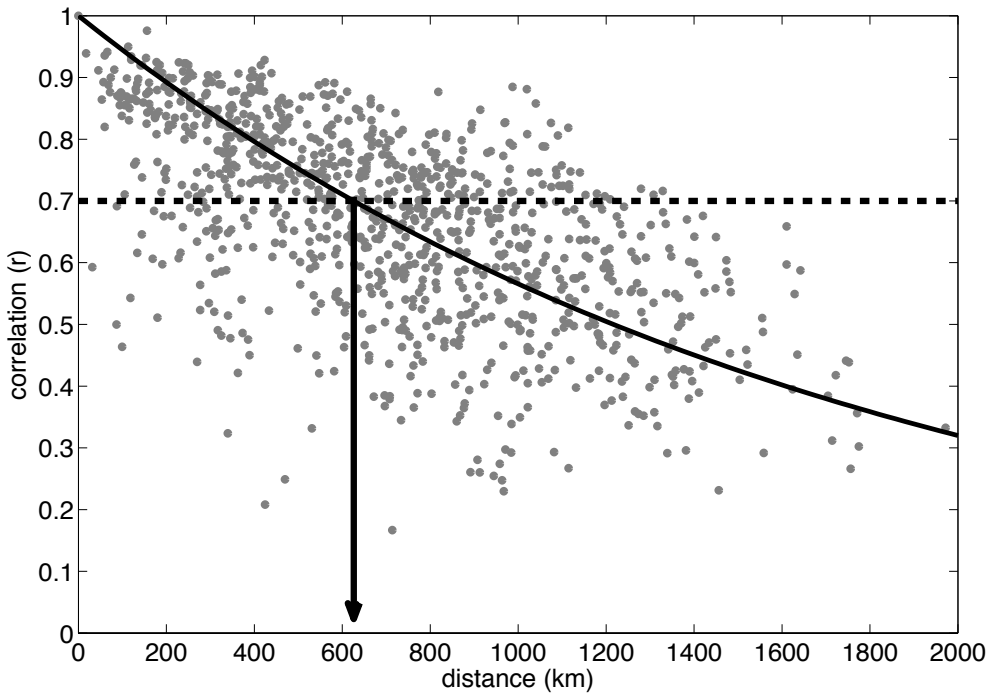
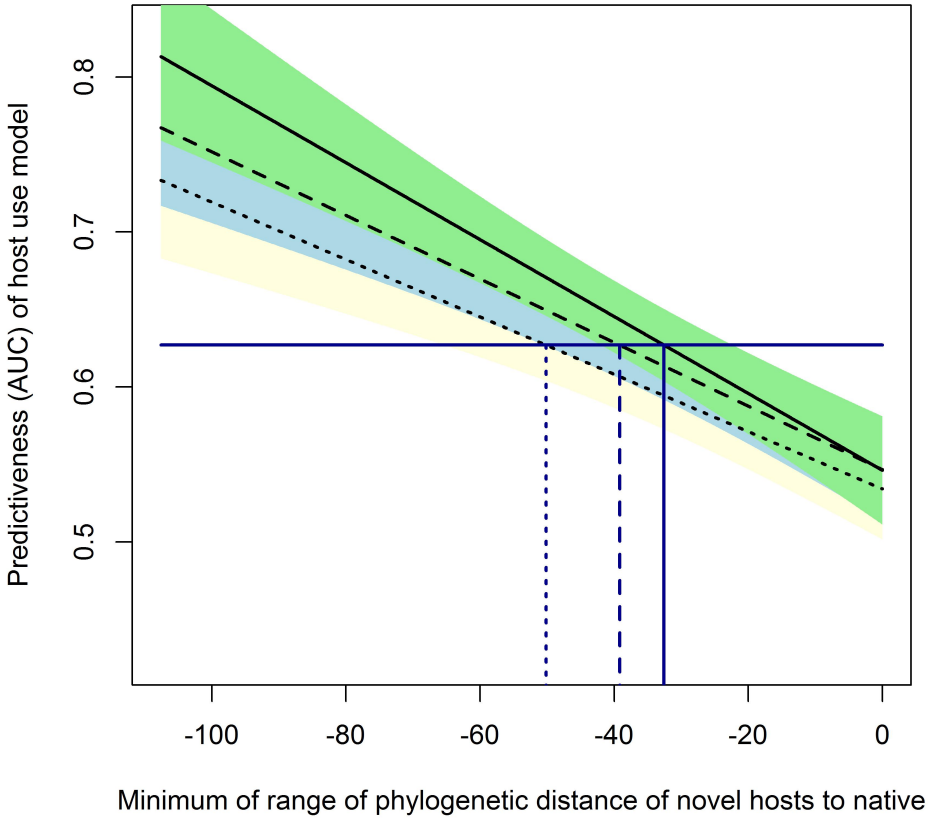


Figure 6 Distance-decay of similarity in community composition. With a forecast proficiency threshold of 0.7 correlation, there is a forecast horizon of just over 600km. I.e., the statistical model of the relationship will, on average, forecast with 0.7 correlation or greater up to 600km distance. This example uses Pearson correlation of square-root transformed abundances as a measure of similarity of relative abundance between pairs of routes from the North American Breeding Bird Survey.



1204

Figure 7 Fitted relationships between forecast proficiency (AUC) and phylogenetic distance
1206 (MYA) when all data were used to parameterise the forecasting model (solid line, green shading), when 2/3 of the data were used (dashed line, blue shading) and when 1/3 of the data
1208 were used (dotted line, yellow shading). The horizontal line is the median AUC for predictions from the full model. The prediction threshold for models built using reduced datasets occurred
1210 at a coarser phylogenetic distance, indicating that increased information allows finer predictions of host use over plant phylogeny. Fits are linear regressions and shaded areas the standard error
1212 of the regression.