Ecological Forecasting

Michael C. Dietze, **Colin Averill**, **John Foster**, **Kathryn Wheeler**

LAST MODIFIED: 25 SEPTEMBER 2018
DOI: 10.1093/OBO/9780199830060-0205

Introduction

Ecologists have long tried to predict how natural systems will behave in the future, or in response to manipulative treatments, for both basic and applied purposes. However, it is only recently that ecologists have begun to distinguish probabilistic forecasting from other forms of modeling. Clark, et al., "Ecological forecasts: An emerging imperative," *Science* (Vol. 293 [2001], pp. 657–660) first defined ecological forecasting as "the process of predicting the state of ecosystems, ecosystem services, and natural capital, with fully specified uncertainties, and is contingent on explicit scenarios of climate, land use, human population, technologies, and economic activity" (p. 657). A key thing that sets forecasting apart from just running a model (either statistical or process-based) into the future is a shift to seeing the world probabilistically—using probability distributions to capture our imperfect understanding of the natural world and to formally include and partition numerous sources of uncertainty and natural variability. This, combined with the emphasis on forecasting real-world systems over theory, means that ecological forecasting tends to be a data-intensive and statistically focused activity (even when employing process-based models). Much of the imperative for ecologists to focus on forecasting comes from the need to respond to the multitude of environmental problems facing society and the aspiration that environmental decisions be made with the best available science in hand. Because all decision-making is ultimately based on what will happen in the future, either under the status quo or different decision alternatives, environmental decision-making ultimately depends on forecasts. Ecological forecasters try to make those forecasts, and their uncertainties, explicit. At the same time, a number of authors have argued that making ecology more predictive is also key to advancing basic science and maturing as a discipline. Generating forecasts forces us to synthesize what we already understand about a system, embodying that understanding in a quantitative model of how we think things will be different in the future, at a new location, or under different conditions. By being quantitatively precise about both our predictions and how confident we are in them, forecasts are more open to direct refutation. Forecasting into yet-to-be measured times and places also emphasizes out-of-sample validation, which is fundamentally a stronger test of our theories and reduces the risk of overfitting and post hoc rationalization. This potential to simultaneously improve basic science and increase social relevance represents a promising win-win scenario.

General Overviews

The consideration of ecological forecasting as a distinct research area is relatively new in ecology, dating to the Clark, et al. 2001 call to arms. Within the United States, this kicked off ecological forecasting programs within both the National Oceanic and Atmospheric Administration and NASA followed, a few years later, by a National Science Foundation research coordination network that produced a number of important papers, such as the overview by Luo, et al. 2011. More recently Dietze 2017 is the first book on the subject, providing an overview of concepts, methods, and case studies from a wide range of ecological subdisciplines. Focused on long-term climate change responses, much of the initial research in ecological forecasting has focused on what MacCracken 2001 defines as projections—forecasts driven by scenarios. However, Dietze, et al. 2018 argues the importance of focusing on near-term predictions (forecasts based on what is known today), which can be iteratively confronted with new observations and updated, as key to accelerating the pace of learning and increasing social relevance. Houlahan, et al. 2017 similarly argues the central importance of making predictions that can be validated to advancing ecology. Finally, as ecological forecasting has matured it has begun to develop best practices, such as those outlined by Harris, et al. 2018 and discussed in more detail in the Methods.

Clark, J. S., S. R. Carpenter, M. Barber, et al. 2001. Ecological forecasts: An emerging imperative. *Science* 293:657–660.

Widely considered the seminal article in ecological forecasting; outlines the need to anticipate environmental change and develop policy based on such forecasts.

Dietze M. C. 2017. *Ecological forecasting*. Princeton, NJ: Princeton Univ. Press.

First and currently the only textbook on forecasting for ecologists, it synthesizes a wide range of forecasting concepts in one place and highlights how forecasting goes beyond traditional modeling. Emphasizes the Bayesian state space statistical framework, its iterative approximations (Kalman and particle filters), and uncertainty quantification. Also includes case study chapters from a wide range of ecological disciplines, decision support, and informatics.

Dietze, M. C., A. Fox, L. Beck-Johnson, et al. 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences* 115.7: 1424–1432.

Lays out the central arguments, in terms of both accelerating basic research and making ecology more relevant to decision making, for focusing on near-term forecasts that are iteratively updated as new data become available. Also identifies the major challenges and opportunities across five different focal areas (data, theory and methods, cyberinfrastructure, decision support, and training, culture and institutions) and provides a roadmap for tackling them.

Harris, D. J., S. D. Taylor, and E. P. White. 2018. Forecasting biodiversity in breeding birds using best practices. *PeerJ* 6:e4278.

This paper outlines a set of ten "best practices" for ecological forecasting that should be required reading for anyone working in this area. The paper also provides a nice case study in multimodel hindcasting and assessment, comparing a range of time-series and species distribution models for their ability to predict breeding bird survey data.

Houlahan J., S. McKinney, T. Anderson, and B. McGill. 2017. The priority of prediction in ecological understanding. *Oikos* 126:1–7.

Provides a strong argument about the value of forecasting for advancing basic science and the need to test quantitative, precise hypotheses. Links the lack of emphasis on prediction in ecology to the "crisis of reproducibility," slow progress, and lack of generality of results. Provides a provocative summary of "questions that are rarely asked in ecology," reasons ecologists don't predict more, and how we might change the way we do ecology.

Luo, Y., K. Ogle, C. Tucker, et al. 2011. Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications* 21.5: 1429–1442.

A product of the FORECAST Research Coordination Network, this paper provides a nice overview and summary of what ecological forecasting is, where it fits in the history of how we do science, and the role of models and data assimilation in ecological forecasting.

MacCracken, M. 2001. Prediction versus projection—forecast versus possibility. *WeatherZine* 26:3–4.

Formally defines the difference between a prediction (a forecast based on what is known today; e.g., weather forecast) and a projection (a forecast conditional on a scenario; e.g., Intergovernmental Panel on Climate Change climate projections).

Theory and Concepts

At certain spatial and temporal scales, most ecological processes are predictable given knowledge about the current state of the system (e.g., they will be mostly the same three seconds from now). Much of the theoretical and conceptual work in ecological forecasting is focused on understanding what limits the predictability of an ecological system as we extend out in space and time. Dietze 2017 argues that doing so requires understanding how five different sources of uncertainty limit predictability: initial condition uncertainty, driver uncertainty, parameter uncertainty, parameter variability, and process uncertainty. At the shortest scale, the predictability of most systems is limited by initial condition uncertainty. If a system is chaotic then this uncertainty grows with time and tends to dominate the forecast (e.g., weather, Shuman 1989). If we define the forecast horizon as the scale in time (or space) over which a forecast remains useful (Petchey, et al. 2015), the key to increasing the forecast horizon of a chaotic system is to reduce the initial condition uncertainty. However, if an ecological system has stabilizing feedbacks, then the relative contribution of initial condition uncertainty will decrease through time and the overall forecast uncertainty will be increasingly dominated by uncertainty in forcing variables, parameters and differences in competing model structures. Looking at earth system models, Bonan and Doney 2018 (cited under Case Studies: Terrestrial Ecosystems) find this trade-off, with the relative impact of initial condition uncertainty declining over time. Weng and Luo 2011 refines this further, finding the relative impact of parameter uncertainty declined, and structural (process) uncertainty increased, with time for a terrestrial carbon pool forecast. Luo, et al. 2014 focuses on the role of driver uncertainty and how this interacts with structural features of ecosystems to limit predictability. Finally, Shuman 1989 reminds us that the forecast horizon can grow with experience and that learning by doing is often more efficient than waiting until models are "good enough."

Dietze, M. C. 2017. Prediction in ecology: A first-principles framework. *Ecological Applications* 27:2048–2060.

Argues that determining which sources of uncertainty (initial conditions, drivers, parameter uncertainty, parameter variability/random effects, process error) control the predictability of different ecological processes is of central importance not just to ecological forecasting but also to seeking general understanding in ecology more broadly. Provides an overarching framework for comparing predictability across systems, illustrates its application in a case study and derives a number of more general hypotheses and predictions.

Luo, Y., T. F. Keenan, and M. Smith. 2014. Predictability of the terrestrial carbon cycle. *Global Change Biology* 21.5: 1737–1751.

While focused on the terrestrial carbon cycle, many of this paper's underlying points about the intrinsic predictability of ecological systems apply more broadly. Focuses on how five fundamental internal properties (compartmentalization, partitioning, first-order decay, etc.) structure the emergent constraints on how ecosystems respond to five different classes of external forcings (cyclic, pulse, gradual, etc.) and rank these in terms of their potential predictability.

Petchey, O. L., M. Pontarp, T. M. Massie, et al. 2015. The ecological forecast horizon, and examples of its uses and determinants. *Ecological Letters* 18:597–611.

Defines the concept of the "ecological forecast horizon" as "the distance in time, space or environmental parameters at which forecast proficiency drops below [a] threshold" (p. 601) of usefulness. Using a series of case studies, they highlight the need to search for generalities and study predictability as a concept. Lays out key ideas, such as the need to catalogue ecological predictability, explore how predictability scales, and invest in shared tools and infrastructure.

Shuman, F. 1989. History of numerical weather prediction at the National Meteorological Center. *Weather Forecast* 4:286–296.

A key insight from this paper surrounds the decision by numerical weather forecasters to start making forecasts immediately under a "learn as you go" approach, rather than to wait until their models were "good enough." Argues that the latter path would have not have led to success.

Weng, E., and Y. Luo. 2011. Relative information contributions of model vs. data to short- and long-term forecasts of forest carbon dynamics. *Ecological Applications* 21:1490–1505.

Interesting analysis partitioning the relative contribution of model structure versus data constraints on the uncertainty in a stand-level forecast of terrestrial carbon pools over different time scales. While some pools were always dominated by model or data, the overall pattern was one of data dominating short-term projections but decaying over decades such that long-term projections were driven more by model structure.

Methods

Ecological forecasts should provide probabilistic predictions with fully specified uncertainties. Because of this focus on the probability of the future conditional on data, most ecological forecasting papers make use of Bayesian approaches, though both traditional frequentist and new machine learning approaches are also employed. Clark 2005 introduces the basics of Bayesian analysis and hierarchical Bayes approaches to partitioning different sources of uncertainty and variability. While there are many Bayesian textbooks available that explain the ins and outs of fitting statistical models, van Oijen 2017 extends this thinking to the calibration of process-based ecological models. Once uncertainties have been quantified and decomposed, scientists are often left with not one, but many different models that describe the data in hand, that they need to validate and choose among. This process begins with assessing model performance and ensuring that models conform to statistical assumptions. Conn, et al. 2018 summarizes the approaches involved in doing this assessment for probabilistic predictions that go beyond traditional model fit statistics. Next, Hooten and Hobbs 2015 summarizes the techniques for selecting among multiple alternative probabilistics models. Medlyn, et al. 2015 approaches this same problem from a more conceptual perspective, focusing on a hypothesis-centered assessment of process-based models to address the question of what models need to be able to reproduce for us to have confidence that they are representing processes correctly rather than just reproducing patterns. Dormann, et al. 2018 summarizes the techniques for producing a multimodel forecast that averages across different competing hypotheses, an approach that often (but not always) improves forecasts by averaging out the different biases in different models. Beyond the impacts that different sources of uncertainty have on predictability (see also Dietze 2017, cited under Theory and Concepts), partitioning uncertainty informs future scientific investment that will maximize forecast improvement. Will our forecasts improve the most by collecting more data (reducing parameter uncertainty), collecting data more accurately (reducing observation uncertainty), or considering alternative model structures (reducing process uncertainty)? Decomposing uncertainty requires statistical tools, such as the uncertainty and sensitivity analyses described by Cariboni, et al. 2007 and Saltelli, et al. 2008. Forecasts can also open the door to more adaptive approaches to measurement and monitoring, described conceptually by Krause, et al. 2015 and statistically by Hooten, et al. 2009.

Cariboni, J., D. Gatelli, R. Liska, and A. Saltelli. 2007. The role of sensitivity analysis in ecological modelling. *Ecological Modelling*. 203.1–2: 167–182.

Sensitivity and uncertainty analyses are commonly performed on ecological models to understand (respectively) how model outputs change as a function of inputs, and how model predictive uncertainties are controlled by model sensitivity and input uncertainty. Provides a nice summary of the available methods for these analyses from an ecological perspective.

Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters* 8:2–14.

Beyond this paper’s historical role in introducing many ecologists to hierarchical Bayes, it remains an accessible introduction to these approaches and the critical need to distinguish between uncertainty and variability in ecological inference and prediction. While uncertainty declines asymptotically with more data, sources of variability in ecological systems do not and thus need to be identified and quantified in ecological predictions even if we cannot (yet) explain them.

Conn, P. B., D. S. Johnson, P. J. Williams, S. R. Melin, and M. B. Hooten. 2018. A guide to Bayesian model checking for ecologists. *Ecological Monographs*.

Validation and testing assumptions is critical to any modeling, but most of the training students receive focus on techniques that ignore the uncertainty in both models and data. This paper provides a nice summary (with case studies) of how to assess probabilistic models, and hierarchical models in particular, that has direct implications for how we assess probabilistic forecasts.

Dormann, C. F., J. M. Calabrese, G. Guillera-Arroita, et al. 2018. Model averaging in ecology: a review of Bayesian, information-theoretic and tactical approaches for predictive inference. *Ecological Monographs*.

Forecasting with multiple models is an important way of capturing model structural uncertainty (selecting a single “best” model always underestimates uncertainty) and often reduces biases. This paper reviews the mathematical foundations that underlie model averaging and the range of approaches available.

Hooten, M., and N. T. Hobbs. 2015. A guide to Bayesian model selection for ecologists. *Ecological Monographs* 85:3–28.

Using regularization as an umbrella concept, this paper provides a practical summary of available Bayesian methods for model selection and multimodel inference (model averaging). Notes that ecologists primarily favor in-sample approaches over stronger, predictive (out-of-sample) approaches to validation and model selection. For out-of-sample predictions (forecasts), they highlight the importance of using local and proper scores, such as log predictive density.

Hooten M., C. Wikle, S. Sheriff, and J. Rushin. 2009. Optimal spatio-temporal hybrid sampling designs for ecological monitoring. *Journal of Vegetation Science* 20:639–649.

Monitoring data are critical inputs to most ecological forecasts, but forecasts themselves have the potential to improve the efficiency of monitoring by focusing effort on the times and places with the greatest return on investment for reducing uncertainties. This paper explores the statistical methodology for dynamic, adaptive monitoring and provides a real-world case study of monitoring vegetation through a combination of fixed and “roving” monitoring points.

Krause, S., J. Lewandowski, C. Dahm, and K. Tockner. 2015. Frontiers in real-time ecohydrology—a paradigm shift in understanding complex environmental systems. *Ecohydrology* 8:529–537.

Complements Hooten, et al. 2009 by exploring adaptive monitoring from a broader perspective, focusing on conceptual and technological issues, rather than statistics.

Medlyn, B. E., S. Zaehle, M. G. De Kauwe, et al. 2015. Using ecosystem experiments to improve vegetation models. *Nature Climate Change* 5.6: 528–534.

From the perspective of summarizing the results of the Free-Air CO_2 Enrichment experiment Model-Data Synthesis activity, Medlyn raises deeper questions about how we can use the assessment of ensembles of process-based models within a hypothesis-testing framework. By focusing on “identifying and evaluating the main assumptions causing differences among models,” it reorients model assessment to focus on capturing processes rather than patterns in data.

Saltelli, A., M. Ratto, T. Andres, et al. 2008. *Global sensitivity analysis. The primer*. New York: Wiley.

Expands on the concepts in Cariboni, et al. 2007, going into much deeper detail on methods, their trade-offs, and how to implement them.

van Oijen, M. 2017. Bayesian methods for quantifying and reducing uncertainty and error in forest models. *Current Forestry Reports* 3.4: 269–280.

Written from a forest model perspective, the concepts here are relevant to how we fit process-based ecological models more generally. Some of the key suggestions specific to process models include moving beyond spin-up to Bayesian estimation of initial condition uncertainties, the need to account for model structural errors in calibration, and the need for more faster and more flexible tools for calibrating process models.

Iterative Forecasting and the Forecast-Analysis Cycle

Near-term forecasting often emphasizes the process of learning from, and updating, forecasts as new information becomes available. For simple models this can be done by refitting the model to data and rerunning the model, but for more complex models this is computationally challenging. Kalman 1960 opened the door to doing this updating iteratively through a forecast-analysis cycle. First, forecasts are made, projecting current states into the future with uncertainties. This forecast serves as our best prior understanding of the future state of the system before it is observed. Second, once new observations are made, a Bayesian analysis is performed to update this prior understanding of the state of the system given the new data. After this the cycle is repeated with new forecasts generated based on the updated states. A variety of approaches exist for performing this cycle that can largely be classified by the methods they use to propagate uncertainties: analytical propagation with linear models (Kalman filter), analytical propagation based on a linear approximation to a nonlinear model (extended Kalman filter), and ensemble-based approaches to numerically propagate uncertainties (ensemble Kalman filter, particle filter). Methods also differ in whether they assume the forecast prior follows a named probability distribution (Kalman variants, most typically a Gaussian assumption) or whether they use a Monte Carlo prior (particle filter). Wikle and Berliner 2007 provides a good first introduction to these approaches, while Kalnay 2002 and Lewis, et al. 2006 are useful textbooks from a physical science perspective (also see the textbook Dietze 2017, cited under General Overviews, for an ecological perspective). Ensemble approaches are particularly common in ecological forecasting, with Evensen (Evensen 2009a, Evensen 2009b) being the standard reference for the ensemble Kalman filter and Doucet and Johansen 2011 providing a nice summary of particle filters. For particularly computational intensive models, dimension reduction approaches such as the Unscented Kalman Filter (Uhlmann 1995) may provide a useful alternative.

Evensen, G. 2009a. *Data assimilation: The ensemble Kalman filter*. 2d ed. New York: Springer.

This textbook is the standard reference on the ensemble Kalman filter, a data assimilation method that is widely popular for ecological models because of its relative simplicity (doesn't require creating adjoint or linear tangent models), generality, and robustness.

Evensen, G. 2009b. The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine* 29.3: 83–104.

This IEEE paper provides a high-level summary of the ensemble Kalman filter that many will find more accessible than the Evensen 2009a textbook

Doucet, A., and A. M. Johansen. 2011. A tutorial on particle filtering and smoothing: Fifteen years later. In *The Oxford handbook of nonlinear filtering*. Vol. 12. Edited by D. Crisan and B. Rozovskii, 656–704. Oxford: Oxford Univ. Press.

This paper provides a concise review of particle filters, featuring accessible outlines of how many of these approaches are implemented. While their use in ecology is still limited, particle filters are an important ensemble-based statistical approach for iteratively updating forecasts that relaxes the strong distributional assumptions in the Kalman filter variants and replaces them with Monte Carlo posterior estimates, similar to more familiar Bayesian Markov chain Monte Carlo methods.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82:35.

Kalman's classic 1960 paper is the origin point for data assimilation and the forecast–analysis cycle that is at the heart of iterative forecasting. To achieve analytical tractability, the eponymous Kalman filter makes strong assumptions of normality and linearity, the latter of which was quickly relaxed, using a linear tangent approach (Taylor series), to the extended Kalman filter, which saw its first real-world application in the Apollo navigation system.

Kalnay, E. 2002. *Atmospheric modeling, data assimilation and predictability*. Cambridge, UK: Cambridge Univ. Press.

An atmospheric perspective on how ensemble forecasting (Chapter 6) and iterative data assimilation (Chapter 5) methods work, including extended Kalman filter, ensemble Kalman filter, four-dimensional variational data assimilation, and earlier approaches. Also has a chapter on the history of numerical weather prediction (Chapter 1) and provides a good perspective on the nature of the atmospheric forecasting problem (initial conditions) and the limits to predictability.

Lewis J. M., S. Lakshmivarahan, and S. Dhall. 2006. *Dynamic data assimilation: A least squares approach*. Cambridge, UK: Cambridge Univ. Press.

A broad, technical, but relatively accessible overview of data assimilation methods, with examples from a range of different physical systems. Presents a lot of foundational theory and techniques and then explores a wide range of methods, broadly organized by whether the model is static (one time) or dynamic (time-varying) and deterministic (adjoint methods) or stochastic (Kalman filter variants).

Uhlmann, J. K. 1995. Dynamic map building and localization: New theoretical foundations. PhD diss., University of Oxford, Department of Engineering Science.

Among the Kalman Filter variants, Uhlmann's unscented Kalman filter provides an important example of a larger class of filters that attempt to address a critical problem: How to reduce the dimensionality of forecast uncertainty propagation while still retaining the flexibility and nonlinearity of ensemble-based approaches? Rather than sampling the initial conditions randomly, the unscented Kalman filter transforms it deterministically, requiring only 2k+1 ensemble members for k dimensions.

Wikle, C., and L. Berliner. 2007. A Bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena* 230.1–2: 1–16.

A technical but accessible summary of data assimilation methods, derived from a Bayesian perspective, including Kalman filtering and smoothing, ensemble Kalman filter, particle filtering, hierarchical state-space models, and Markov chain Monte Carlo. Starts from the most basic case (univariate normal data and normal prior), generalizes this to the multivariate case, and makes explicit the connections between data assimilation and spatial statistics (e.g., Kriging), which are more familiar to many ecologists.

Decision Support

Much of traditional environmental management is based on the assumption of stationarity—that the future will be like the past—which allows decisions to be guided by fairly direct data analysis (yield curves, one hundred-year floodplains, etc.). However, as Milly, et al. 2008 notes, the assumption of stationarity is becoming less viable, necessitating that decisions be supported by more explicit projections. As our ability to forecast over multiple time scales, from the immediate to the long term, improves, Bradford, et al. 2018 argues the use of forecasts in environmental decision-making is becoming more and more important. Indeed, the goal of many ecological forecasts is to inform decision-making. But to be effective, it is important to understand how forecasts inform the decision-making process. Gregory, et al. 2012 provides an accessible overview of environmental decision-making, which emphasizes that while having the best available information is key to good decision-making, the process is also driven by human values and risk tolerance and can be strongly influenced by human misperceptions of probability and uncertainty. One important take-home of this is the acknowledgement that having a regulatory predefined threshold of when a forecast is “good enough” to be useful is the exception, rather than the norm. This is counter to the perception by many ecologists that our forecasts are not good enough to be useful and reinforces the need to start forecasting now and learn by doing (see also Shuman 1989, cited under Theory and Concepts). Coming back to the idea of the role of human perception of probability, Morgan 2014 summarizes the role of expert elicitation, which can play an important role at multiple stages in an ecological forecast. Kahneman 2013 provides critical insight into how the mind relies on biases and heuristics to process new information, which has large impacts on how we present probabilistic forecasts to both decision makers and the public. Taleb 2007 argues that are our ability to forecast is hindered by our ability to anticipate low-probability/high-impact events. On the more technical side, Williams and Hooten 2016 provides an overview of the statistical approaches involved in translating ecological forecasts into decision support. Finally, Ketz, et al. 2016 and Hobday, et al. 2016 provide real-world case studies on the incorporation of ecological forecasts into decision-making, looking at managed terrestrial and marine animal populations, respectively.

Bradford, J. B., J. L. Betancourt, B. J. Butterfield, S. M. Munson, and T. E. Wood. 2018. Anticipatory natural resource science and management for a dynamic future. *Frontiers in Ecology and the Environment*.

Argues that near-term ecological forecasts allow managers to improve seasonal and annual planning and decision making as most interventions are more effective in the early stages of a change. Simultaneously, long-term projections enable managers to enhance long-term management investments. They emphasize that “predictions need not be perfect to be useful; they only need to lead to decisions that are better than those based on historical conditions” (p. 297).

Gregory, R., L. Failing, M. Harstone, et al. 2012. *Structured decision making: A practical guide to environmental management choices*. New York: Wiley.

Accessible introduction to the science and practice of environmental decision-making. Emphasizes that while science provides facts and knowledge, most decisions cannot simply be optimized because of trade-offs among decision alternatives that are ultimately resolved based human values. Explains how uncertainties fit into the decision framework (as risk), how to communicate them, and when they matter (individuals vary in their risk tolerance).

Hobday, A., M. Claire, J. Eveson, and J. Hartog. 2016. Seasonal forecasting for decision support in marine fisheries and aquaculture. *Fisheries Oceanography* 25:45–56.

Discusses the development of iterative seasonal forecasts in marine fisheries and under what conditions such forecasts provide useful management information. Uses forecasts of two wild (tuna in southern and eastern Australia) and two aquaculture (prawn, salmon) population as illustration.

Kahneman, D. 2013. *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Any ecological forecaster hoping their forecasts will benefit society needs to understand how forecasts will be misinterpreted by the human mind (including their own). Written for a general audience, Kahneman's book gives critical insight into how the brain often misperceives statistics, uncertainties, and probabilities.

Ketz, A. C., T. L. Johnson, R. J. Monello, and N. T. Hobbs. 2016. Informing management with monitoring data: The value of Bayesian forecasting. *Ecosphere* 7.11:e01587.

This elk population forecast for Rocky Mountain National Park provides a real-world example of near-term iterative forecasting under different scenarios (cull levels) used to inform management (probability of population staying in target range). Introduction provides a nice summary of the connection between forecasting and adaptive management.

Milly, P. C. D., J. Betancourt, M. Falkenmark, et al. 2008. Stationarity is dead: Whither water management? *Science* 319:573–574.

A short policy commentary that raises the issue that much of our natural resource laws and management approaches are based, implicitly or explicitly, on the assumption that the environment is stationary—that natural systems fluctuate, but they do so within "an unchanging envelope of variability" that we can understand through historical observations. The path forward needs to be nonstationary probabilistic forecasts.

Morgan, M. G. 2014. Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Science of the United States of America*. 111:7176–7184.

Expert elicitation plays an important role in many stages of ecological forecasting, from eliciting Bayesian priors, to informing model structure, to interpreting forecasts, to constructing utility functions. This paper discusses when and how expert elicitation is used in decision analysis and many of the key issues involved, such as expert overconfidence and cognitive biases, elicitation protocols, and pooling judgement.

Taleb, N. N. 2007. *The black swan: The impact of the highly improbable*. 2d ed. New York: Random House.

By focusing on low probability events that had a disproportionate impact on society, Taleb emphasizes the challenges of the "unknown unknowns" and failures of imagination when trying to make forecasts, as well as the limitations of learning from confirmatory experience. Named for the assertion by 17th century Europeans that "all swans are white," which was proven false by the discovery of black swans in Australia.

Williams, P. J., and M. B. Hooten. 2016. Combining statistical inference and decisions in ecology. *Ecological Applications* 26:1930–1942.

Many ecological forecasts are done with the goal of informing decision. This paper provides a nice summary of the statistics involved in translating a prediction and its uncertainty into an "optimal" decision. Provides an overview of loss (or utility) functions in terms of what they are, how they transform predictions, and their construction, including common functional choices.

Case Studies: Aquatic Ecosystems

Aquatic ecologists have often been at the forefront of ecological forecasting, as they work on systems that have a clear societal relevance, and where the practical observational challenges necessitated early adoption of more sophisticated statistical techniques. This capacity, as well as the connection between ecological forecasts and management, are perhaps most obvious when it comes to fisheries. Kuikka, et al. 2014 summarizes fifteen years' experience using Bayesian approaches for fusing numerous data sources, each providing partial information about different life history stages, to assess and forecast Baltic salmon stocks. Similarly, Scales, et al. 2017 demonstrates the use of remote sensing and a physical ocean data assimilation system to develop an automated ecoinformatic workflow providing daily maps of swordfish presence and abundance. Lindegren, et al. 2010 develops a forecast of the cod, sprat, and herring food web to make projections of cod stocks under climate, harvest, and salinity scenarios. Hobday, et al. 2016 (cited under Decision Support) discusses the use of iterative seasonal forecasts for four Australian fisheries. However, aquatic ecological forecasts are not limited to fisheries. Smith, et al. 2015 discusses a nutrient–phytoplankton–zooplankton model of North Atlantic algal blooms using remote sensing data. Udevitz, et al. 2017 generates walrus population forecasts based on a model that captures behavioral and bioenergetic responses to sea ice. Finally, Helmuth, et al. 2006 reviews a range of efforts to forecast population dynamics and range shifts in rocky intertidal systems. Overall, while the subdisciplinary silos between aquatic and terrestrial ecologists can be entrenched, the ecological forecasting research in these areas suggests that there is much to be gained by greater dialog across disciplinary boundaries.

Helmuth, B., N. Mieszkowska, P. Moore, and S. J. Hawkins. 2006. Living on the edge of two changing worlds: Forecasting the responses of rocky intertidal ecosystems to climate change. *Annual Review of Ecology, Evolution, and Systematics* 37:373–404.

Intertidal communities are subject to both marine and terrestrial environments and thus are often used as early warning systems for the impacts of climate change. This paper reviews the effects of climate change on rocky intertidal habitats and describes a framework for using ecological forecasting to quantify future changes. While the paper focuses on one specific ecosystem type, the presented key questions are applicable for a variety of ecological forecasts.

Kuikka, S., J. Vanhatalo, and H. Pulkkinen. 2014. Experiences in Bayesian inference in Baltic salmon management. *Statistical Science* 29:42–49.

Summarizes fifteen years' experience using Bayesian approaches to assess and forecast Baltic salmon stocks. Provides a great example of using submodels for different processes and information (field data, literature synthesis, expert elicitation) to generate informative priors that feed into an overall stock model used to make predictions. Also discusses the practical challenges faced in implementing an ecological forecast and communicating it to decision makers.

Lindegren, M., C. Mollmann, A. Nielsen, et al. 2010. Ecological forecasting under climate change: The case of Baltic cod. *Proceedings of the Royal Society B* 277:2121–2130.

Applies a linear multivariate autoregressive state-space (MARSS) approach to model the cod, sprat, herring food web in the Baltic Sea. Projections under different scenarios of warming, salinity, and harvest found that declines in salinity increase the risk of population collapse (due to both direct impacts and food web interactions) and these declines can be partly offset by harvest reductions.

Scales, K. L., E. L. Hazen, S. M. Maxwell, et al. 2017. Fit to predict? Eco-informatics for predicting the catchability of a pelagic fish in near real time. *Ecological Applications* 27.8: 2313–2329.

Developed a near real-time daily spatial forecast for swordfish presence and abundance off the coast of California using boosted regression trees and a combination of satellite observations and outputs from an ocean data assimilation system.

Smith, M. J., D. P. Tittensor, V. Lyutsarev, and E. Murphy. 2015. Inferred support for disturbance-recovery hypothesis of North Atlantic phytoplankton blooms. *Journal of Geophysical Research: Oceans* 120:1–22.

Used a daily nutrient–phytoplankton–zooplankton model to predict North Atlantic phytoplankton blooms and test alternative hypotheses about the drivers of spring bloom growth. Uses Bayesian calibration to fit the model against SeaWiFS and then hindcasts five years of held out data as validation.

Udevitz, M. S., C. V. Jay, R. L. Taylor, et al. 2017. Forecasting consequences of changing sea ice availability for Pacific walruses. *Ecosphere* 8.11: e02014.

Fit a series of coupled, autoregressive Bayesian generalized linear mixed models to model walrus behavior (movement among regions, time in water, time foraging) in the Chukchi Sea as a function of sea ice and walrus bioenergetics (feeding, metabolism, reproduction) as functions of behavior. Uses this to forecast walrus responses to future sea ice projections.

Case Studies: Population Ecology

Population forecasts can take many forms and be implemented on almost any organism—livestock, amphibians, plants, and human and wildlife pathogens are just a few examples outlined here. Forecasts of threatened, invasive, and infectious populations play an important role in allowing conservation biologists and public health officials to make more informed management decisions. Bayesian hierarchical state-space models are used extensively in this section because of their ability to assimilate data and partition observation and process uncertainties. As habitat degradation coupled with global climate change threaten the world's biodiversity, it is increasingly necessary to evaluate population dynamics of threatened and endangered species in response to these exogenous forces. Evans, et al. 2010 combines multiple data sources within a Bayesian matrix model to assess the viability of an endangered plant, the Florida scrub mint, to fire frequency. Forecasts can also be used to determine management strategies for endangered species. Indeed, Chandler, et al. 2015 investigates the metapopulation dynamics of the threatened Chiricahua leopard frog using hierarchical Bayesian modeling. The authors forecasted the overall population to be viable for the next fifty years even though individual subpopulations appear to be in decline, showing that increasing movement between subpopulations is the best future management strategy. Forecasting the spread of disease can be particularly difficult as individuals can be infectious and asymptomatic, a subset of infected individuals will never be reported, and there is heterogeneity in pathogen transmission. Hierarchical state-space models handle these problems well with data assimilation and can be applied to a variety disease systems, including vector-borne (LaDeau, et al. 2011). The state-space forecasting framework has also been used for near real-time (updated daily) forecasting of the H1N1 outbreak in Singapore in 2009 (Ong, et al. 2010), and the foot-and-mouth disease outbreak affecting livestock in Great Britain in 2001 (Ferguson, et al. 2001). Ong was able to make forecasts available to the public via a website, which has untold benefits for public health, while Ferguson's forecasts aided public health officials in managing the epidemic. Like Ferguson, the forecast by Hobbs, et al. 2015 of brucellosis transmission in the Yellowstone bison herd led to an adaptive management strategy. Concerns about spread are not limited to disease—Ibanez, et al. 2014 lays out a hierarchical Bayes framework for forecasting invasive species based on assessing dispersal, colonization, and proliferation. Identifying which processes limit which species can be helpful for targeting management actions.

Chandler, R. B., E. Muths, B. H. Sigafus, et al. 2015. Spatial occupancy models for predicting metapopulation dynamics and viability following reintroduction. *Journal of Applied Ecology* 52.5: 1325–1333.

Fit a metapopulation model to spatiotemporal data on the threatened Chiricahua leopard frog and used this to make probabilistic fifty-year projections of population viability. Used a hierarchical Bayes framework to account for incomplete sampling and imperfect detection. Predicts a less than 3 percent chance of metapopulation extinction, despite over 80 percent of sites having high extinction probability (84 percent), because of dispersal and recolonization.

Evans, M., K. Holsinger, and E. Menges. 2010. Fire, vital rates, and population viability: A hierarchical Bayesian analysis of the endangered Florida scrub mint. *Ecological Monogram* 80:627–649.

Endangered species assessments represent an important class of ecological forecasts. This paper illustrates a Bayesian approach to population viability assessment, transforming a traditional stage-structured matrix model into matrix of fully coupled generalized linear mixed models. This example highlights the ability to combine multiple sources of information simultaneously and to formally quantify and partition unexplained variability, in this case including a shared year effect across different life history stages.

Ferguson, N. M., C. A. Donnelly, and R. M. Anderson. 2001. The foot-and-mouth epidemic in Great Britain: Pattern of spread and impact of interventions. *Science* 292:1155–1160.

The first in a series of papers on the 2001 foot-and-mouth epidemic in Great Britain, this body of research was one of the first examples of near real-time ecological forecasting that had a direct policy impact, in this case leading to the decision to aggressively "ring" cull animals suspected of infection.

Hobbs, N. T., C. Germia, J. Treanor, et al. 2015. State-space modeling to support management of brucellosis in the Yellowstone bison population. *Ecological Monograms* 85:525–556.

A state-space model was constructed that assimilated data from aerial observations, ground observations, capture history, and seroprevalence of individuals. The model includes state transitions (e.g., yearling to juvenile) and both horizontal and vertical transmission of brucellosis. Forecasts were evaluated with respect to managing the bison population.

Ibanez, I., J. M. Diez, L. P. Miller, et al. 2014. Integrated assessment of biological invasions. *Ecological Applications* 24.1: 25–37.

This paper provides a framework for forecasting biological invasions in terms of three processes (dispersal, colonization, and proliferation). Uses case studies to illustrate how these processes can be calibrated and predicted in a hierarchical Bayes frameworks using multiple sources of information simultaneously. This framework can be used to identify what processes may be limiting invasions and windows of opportunity that could be targeted by management action.

LaDeau, S. L., G. E. Glass, N. T. Hobbs, A. Latimer, and R. S. Ostfeld. 2011. Data-model fusion to better understand emerging pathogens and improve infectious disease forecasting. *Ecological Applications* 21:1443–1460.

Authors provide a review into the challenges of forecasting infectious diseases of varying type, including a vector-borne system (malaria), a zoonotic system (severe acute respiratory syndrome), and a vector-borne zoonotic system (Lyme disease and West Nile virus). State-space models that include the heterogeneity in pathogen transmission are advocated.

Ong, J. B. S., M. I. -C. Chen, A. R. Cook, et al. 2010. Real-time epidemic monitoring and forecasting of H1N1–2009 using influenza-like illness from general practice and family doctor clinics in Singapore. *PLoS One* 5:e10036.

In an excellent example of a real-world application of real-time iterative forecasting, this paper summarizes efforts to predict the 2009 H1N1 outbreak in Singapore. Forecast system was rapidly developed and deployed to a public website. The model itself was a fully Bayesian specification of a traditional SEIR model that was updated daily via a particle filter to allow both model states and parameter estimates to evolve.

Case Studies: Terrestrial Ecosystems

Forecasts of terrestrial ecosystem processes are often focused on issues surrounding vegetation productivity (both natural and agricultural), disturbance, the carbon cycle, and the coupling between the land surface and the atmosphere. A distinctive feature of terrestrial ecosystem forecasts is that they are frequently made using process-based models. At the largest scales, this includes the growing representation of ecological processes in earth system models, where model uncertainties continue to dominate long-term projections (Moorcroft 2006, Bonan and Doney 2018). At regional scales, Thomas, et al. 2017 uses hierarchical Bayesian methods to integrate data from both natural environmental gradients and experimental studies to generate forecasts of pine forest productivity in the southeastern United States. Similarly, Hufkens, et al. 2016 uses a data-informed vegetation-hydrological model to forecast potential climate impacts on North American grasslands, projecting that phenological shifts toward greater spring and fall productivity will offset summer declines. Using a more statistical model, Wilson, et al. 2015 assesses rates of postfire vegetation recovery in the Cape Floristic Region of South Africa and uses this to project changes in recovery rates under climate change. Over shorter time scales, Chen, et al. 2011 develops seasonal forecast of Amazon fire activity with a three- to five-month lead time based on remote sensing and climate oscillation indices. Cook, et al. 2005 looks at vegetation phenology in Europe and similarly found impacts of climate oscillations. As an alternative to equilibrium-based species distribution models, Tredennick, et al. 2016 develops a landscape-scale dynamic population model of sagebrush canopy cover to forecast the population distribution over both interannual and multi-decadal scales. Finally, Clark, et al. 2003 also reminds us that some things in ecology may not be easily forecast—looking at the impacts of long-distance seed dispersal of tree population spread they find that the inherent stochasticity in the process dominates over uncertainties in model structure or parameters.

Bonan, G. B., and S. C. Doney. 2018. Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. *Science* 359.6375: eaam8328.

Summarizes how ecological processes are included in the land and ocean components of global Earth system models. While focused mostly on processes, the authors also discuss the challenges of prediction, noting that “only recently have the necessary large, multimodel and multimember ensembles become available to distinguish [climate responses] from natural variability and model uncertainty” (p. 5). Land uncertainties are shown to be driven by differences among models (structure, parameters, etc.).

Chen, Y., J. T. Randerson, D. C. Morton, et al. 2011. Forecasting fire season severity in South America using sea surface temperature anomalies. *Science* 334:787–791.

The authors develop an empirical regression model that allows prediction of South American forest fire season severity with a lead time of three to five months. This is an example of a near-term ecological forecast (making an annual or subannual predictions). By making predictions in the near term, scientists gain feedback on model predictive performance in time. Scientists can use this feedback to further refine and improve model accuracy.

Clark, J. S., M. Lewis, J. S. McLachlan, and J. HilleRisLambers. 2003. Estimating population spread: What can we forecast and how well. *Ecology* 84.8: 1979–1988.

This paper investigates how useful increasing our knowledge of long-distance dispersal (LDD) is in making informative forecasts of spread velocity. They found that even a complete understanding of LDD parameters might not increase how informative estimates of spread velocity are because LDD is unpredictable. Inherent uncertainty, not parameter sensitivity, limits how informative forecasts of spread velocity are.

Cook, B. I., T. M. Smith, and M. E. Mann. 2005. The North Atlantic Oscillation and regional phenology prediction over Europe. *Global Change Biology* 11:919–926.

Because temperature is a primary cue in vegetation phenology, global warming is causing forecasting phenology to become more important. Numerous studies have analyzed how phenology is changing, but this one is particularly interesting because it not only links phenology with temperature but also investigates how much of the variance in the annual phenological cycle is correlated with the North Atlantic Oscillation, thus providing a mechanism to forecast multiyear phenology.

Hufkens, K., T. F. Keenan, L. B. Flanagan, et al. 2016. Productivity of North American grasslands is increased under future climate scenarios despite rising aridity. *Nature Climate Change* 6.7: 710–714.

Calibrated a simple model of daily grassland productivity across a network of North American phenocam sites and used this to project climate change responses, accounting for driver and parameter uncertainty. Forecasts a widespread increase in productivity, despite increases in aridity, because spring and fall increases are projected to more than compensate for declines due to summer droughts.

Moorcroft, P. R. 2006. How close are we to a predictive science of the biosphere? *Trends in Ecology & Evolution* 21.7: 400–407.

Framed primarily from the perspective of the representation of terrestrial ecosystems within earth system models, Moorcroft argues, “Our current understanding of biosphere–atmosphere feedbacks is a collection of interesting, but largely untested, hypotheses for the future state of terrestrial ecosystems” (p. 401). Summarizes some of the key uncertainties about climate feedbacks and then focuses on model calibration and validation and the need to improve model representations of subgrid heterogeneity and biodiversity.

Thomas, R. Q., E. B. Brooks, A. L. Jersild, et al. 2017. Leveraging 35 years of Pinus taeda research in the southeastern US to constrain forest carbon cycle predictions: Regional data assimilation using ecosystem experiments. *Biogeosciences* 14:3525–3547.

Case study forecasting the impact of climate change on southern pine forestry. Illustrates the fusion of information from a wide range of observations and experiments over a regional scale (294 stands) and more than thirty years of data within a process model. Takes a hierarchical Bayes approach that separates observation and process errors. Propagates uncertainties into regional-scale spatiotemporal projections under different CO_2, climate, and nutrient scenarios.

Tredennick, A. T., M. B. Hooten, C. L. Aldridge, et al. 2016. Forecasting climate change impacts on plant populations over large spatial extents. *Ecosphere* 7.10: e01525.

A spatially explicit landscape-scale sagebrush model was calibrated (with spatial and temporal random effects) against twenty-eight years of cover data and used to produce short- and long-term forecasts of population dynamics under alternate climate change scenarios. The use of dynamic models provides an important alternative to equilibrium-based distribution models on management-relevant spatial and temporal scales.

Wilson, A. M., A. M. Latimer, and J. A. Silander. 2015. Climatic controls on ecosystem resilience: Postfire regeneration in the Cape Floristic Region of South Africa. *Proceedings of the National Academy of Sciences* 112.29: 9058–9063.

Modeled postfire Normalized Difference Vegetation Index recovery in the Cape Floristic Region, where the rate of recovery, asymptote, and seasonal amplitude were each functions of climate, soil, and topographic variables, with a 25 percent hold-out validation. Analysis identifies strong gradients in post-fire recovery while forecasts under climate change projections indicate that warmer winters might accelerate recovery but also increase fire frequency.

back to top