

A Predictive Framework to Understand Forest Responses to Global Change

Sean M. McMahon,^a Michael C. Dietze,^b Michelle H. Hersh,^c
Emily V. Moran,^d and James S. Clark^{a,d,e}

^aNicholas School of the Environment, Duke University, Durham, North Carolina, USA

^bUniversity of Illinois at Urbana-Champaign, Urbana, Illinois, USA

^cUniversity Program in Ecology, Duke University, Durham, North Carolina, USA

^dDepartment of Biology, Duke University, Durham, North Carolina, USA

^eDepartment of Statistical Science, Duke University, Durham, North Carolina, USA

Forests are one of Earth's critical biomes. They have been shown to respond strongly to many of the drivers that are predicted to change natural systems over this century, including climate, introduced species, and other anthropogenic influences. Predicting how different tree species might respond to this complex of forces remains a daunting challenge for forest ecologists. Yet shifts in species composition and abundance can radically influence hydrological and atmospheric systems, plant and animal ranges, and human populations, making this challenge an important one to address. Forest ecologists have gathered a great deal of data over the past decades and are now using novel quantitative and computational tools to translate those data into predictions about the fate of forests. Here, after a brief review of the threats to forests over the next century, one of the more promising approaches to making ecological predictions is described: using hierarchical Bayesian methods to model forest demography and simulating future forests from those models. This approach captures complex processes, such as seed dispersal and mortality, and incorporates uncertainty due to unknown mechanisms, data problems, and parameter uncertainty. After describing the approach, an example by simulating drought for a southeastern forest is offered. Finally, there is a discussion of how this approach and others need to be cast within a framework of prediction that strives to answer the important questions posed to environmental scientists, but does so with a respect for the challenges inherent in predicting the future of a complex biological system.

Key words: Bayesian methods; climate; forecasts; forests; global change; models; prediction

Introduction

Anticipating forest diversity of the coming century is an important component of global-change research. Forests cover more than a third of the Earth's landmass. They regulate atmospheric gases, influence hydrological cy-

cles, and provide important ecosystem services to human populations, such as food, fiber, and fuel. Rising atmospheric CO₂, climate change, introduced species, and land cover modification by humans are dramatically altering the distribution, biodiversity, and biogeochemistry of forests. Our understanding of these changes remains rudimentary, despite their relevance for human society. Progress is limited because it has been difficult to assimilate the increasing information now available from field studies and

Address for correspondence: Sean M. McMahon, LRSC 210A, Nicholas School of the Environment, Duke University, Durham, NC 27708. seanmcm@duke.edu

models into a coherent framework for prediction. Here we consider the challenges of assimilation and use a specific example to illustrate how emerging tools can help.

The last decade has brought new approaches to modeling and computation, and it is timely to ask how greater exploitation of these advances might help scientists predict future environments. Can increased capacity to assimilate large amounts of information derived from diverse sources help us anticipate landscape change? Although direct calibration of biological responses to the full range of environmental variables is not expected anytime soon, substantial information already exists on how variation in climate, land cover, and CO₂ affect individual trees and how these effects can differ across species and communities. Much of this fine-scale information on tree physiology and ecology is hard to merge with the coarse-scale variation in climate, atmospheric CO₂, and land cover. The availability of new modeling frameworks and computational tools motivate a fresh look at the prediction problem, with a focus on how fine-scale processes such as the growth and mortality of individual trees are influenced by global-change variables and how those responses extrapolate to regional biodiversity.

In this chapter we discuss how prediction is used not so much as a way of forecasting the future, but rather as a means for synthesizing what is known. We briefly review some of the biological processes critical to how forests respond to global change, how these processes affect individuals through demographic rates (growth, survival, and reproduction), and how to place this understanding in a context that permits application of some new modeling and computational tools. We then illustrate a prediction approach with an example from the southeastern United States. We do not advocate this as a means for precise forecasting, but rather as a means for application of current understanding to explore the types of changes that are important for scientific progress and stewardship.

Global-Change Drivers

To address how forests might change over the next century ecologists need to focus on processes that are known to be affected by global change and that have been established as important to current forest structure. We review here some of the processes that have been well-studied and might be integrated into predictive models.

Climate Change

Climate change is among the most important drivers of forest change, yet ecologists find a daunting challenge to quantify and anticipate the exact ways in which changing climate will affect forests. Correlations between climate variables and species distributions are only a beginning (Iverson and Prasad 1998). Forest responses to climate change are complicated by biology, because tree populations interact (Canadell *et al.* 2007; Roe and Baker 2007). Competition, natural enemies, and availability of mutualists influence species distributions, and these influences will change as climate and tree species abundances change.

Forest composition will in large part be determined by species' ability to migrate. Changes in geographic distribution of some species are already evident (McCarty 2001; Parmesan and Yohe 2003), including northward range shifts and local extinctions at low latitudes and altitudes for small mammal, insect, and bird populations (Thomas *et al.* 2006; Hitch and Leberg 2007; Moritz *et al.* 2008). Because of the long generation time of trees, one of the few sources of information on potential tree range shifts comes from fossil-pollen data. These data show that glacial cycles have caused tree species to undergo repeated range shifts in the past (Williams *et al.* 2004). However, paleoecological data provide limited insight for contemporary climate change, because change is now more rapid than in the past (Huntley 1991), will result in higher global temperatures than were experienced over the last several million years (Crowley 1990), and is occurring in a different

biological setting. Postglacial migrations occurred in a poorly understood competitive setting, certainly different from that of today. Seeds from potential immigrant species must establish in competition with resident species of intact forests. Under modern conditions, migration capacity of many species will be limited by low fecundity, restricted dispersal, and reduced establishment opportunities (Clark *et al.* 2001, 2003; McLachlan *et al.* 2005).

Predicting changes in composition within current species ranges is also difficult. An important insight from paleoecology is the apparently idiosyncratic response of many species to past natural warming (Davis 1969; Van Devender and Spaulding 1979; Webb 1986; Prentice *et al.* 1991; Tinner and Lotter 2001). Novel combinations of climate and edaphic factors, together with dispersal limitation, lead us to expect the emergence of new forest communities supporting reorganized food webs (Saxon *et al.* 2005; Williams *et al.* 2007b). The rate and magnitude of change will be species-specific, and it will depend on life history (Meunier *et al.* 2007) as well as on interactions involving mutualists (e.g., pollinators, mycorrhizae), competitors, and natural enemies (Ibanez *et al.* 2006). Further climate change will interact with disturbance cycles and human land-use, as shifts in climate put pressure on how societies use land that is currently covered by forests.

Rising CO₂

The direct effects of climate change will be mediated by changes in photosynthetic rates and water use efficiencies under elevated atmospheric CO₂ (Long *et al.* 2004). These processes are observable at leaf scales, but extrapolation to entire forest stands is difficult. Increases in photosynthetic rate and productivity are documented for woody plants, the magnitude of increase depending on resource availability (Ogren 1984; Curtis and Wang 1998; Ainsworth and Long 2005). Both models and experiments suggest that species differences may be important for overall response

(Bolker *et al.* 1995; Reich *et al.* 2001; Mohan *et al.* 2007). Forest net primary productivity is expected to increase under elevated CO₂ (Norby *et al.* 2005), although free air CO₂ experiments (FACE) reveal that tree species' responses differ. For example, loblolly pines (*Pinus taeda*) treated with high CO₂ continued to add more wood biomass than control pines over 8 years, while sweetgum (*Liquidambar styraciflua*) at another site showed only a transient effect, likely due to differences in allocation of resources derived from CO₂ (Moore *et al.* 2006).

Rising CO₂ may affect forest composition as well as productivity, shifting community structure through differential effects on maturation and fecundity (LaDeau and Clark 2001) and seedling recruitment (Hattenschwiler and Korner 2000; Mohan *et al.* 2007; Sefcik *et al.* 2007). Pines exposed to high CO₂ mature at smaller diameters and produce, on average, three times as many cones as control trees (LaDeau and Clark 2001). Any population responding to changes in atmospheric gases to this degree would be expected to increase in abundance at the expense of those that do not.

Introduced and Invasive Species

Species invasions, especially large-scale outbreaks of exotic pests and pathogens, have occurred with increasing frequency during the past century (Orwig and Foster 1998) and can radically alter forest community structure. Chestnut blight (*Cryphonectria parasitica*), Dutch elm disease (*Ophiostoma* spp.), and gypsy moths (*Lymantria dispar*) are examples of introduced pathogens that have led to severe mortality and near extinction of their host-tree species. Current emerging pests and diseases such as sudden oak death (*Phytophthora ramorum*), balsam and hemlock woolly adelgids (*Adelges piceae* and *A. tsugae*), white pine blister rust (*Cronartium ribicola*), butternut canker (*Sirococcus clavignenti-juglandacearum*), dogwood anthracnose (*Discula* spp.), and beech-bark disease (*Cryptococcus fagisuga*) are causing heavy mortality of their

host species (Burdon *et al.* 2006; Lovett *et al.* 2006).

Invasive plant species often affect native species through competition for resources (Gorchov and Trisel 2003; Martin and Marks 2006), but some also alter soil properties (Henehan *et al.* 2006) or exude chemicals toxic to benign competitors or tree mutualists (Stinson *et al.* 2006). While many invasive species favor disturbed habitats, intact forests may be most vulnerable to invasion by late-successional, shade-tolerant exotics (Martin and Marks 2006). The full impact of invasive species, both plant and animal, can be difficult to assess because interactions between exotics and the native community change can lead to “invasion meltdowns” (Simberloff and Von Holle 1999), where the effects of one invader facilitate others.

Natural Enemies

Along with the critical threat of introduced pathogens, trees may interact with existing native or invading pests and pathogens in new ways under novel climate conditions (Harvell *et al.* 2002; Logan *et al.* 2003; Burdon *et al.* 2006). Some of these changes are already occurring. Recent outbreaks of pests in Europe and North America have been occurring at higher elevations and latitudes than in the past (Logan *et al.* 2003; Battisti *et al.* 2005; Woods *et al.* 2005). Pathogens previously limited by environmental conditions unfavorable to overwintering, reproduction, infection, or other aspects of their life cycles may become more prevalent (Harvell *et al.* 2002; Bergot *et al.* 2004; Garrett *et al.* 2006). Drought and storm damage can increase the impact of insect (Powers *et al.* 1999) and fungal pathogen outbreaks (Clinton *et al.* 1993; Desprez-Loustau *et al.* 2006). Because of their short life spans and physiological sensitivity to temperature, insects and pathogens may adapt more quickly to new climate regimes than their host plants (Ayers and Lombardero 2000; Garrett *et al.* 2006). However, plant defense responses to disease and herbivory will also not remain static

(Garrett *et al.* 2006; Stiling and Cornelissen 2007), and the rate at which tree populations can respond to outbreaks will be crucial to determining future patterns of pathogen and pest damage. Although many studies agree that forests generally face increased risks of pest and pathogen damage under climate change, responses will depend on the specific combination of host and pest/pathogen (Cannon 1998; Roy *et al.* 2004).

Hierarchical Modeling: A Pathway to a Predictive Framework

Models cannot capture the full complexity of interactions that result from the direct effects of climate change filtered through biological interactions. Instead, the only hope for learning about the potential responses of forests to climate change must come from models that can account for these potentially complex interactions without explicitly describing them. This is what is meant by “modeling uncertainty.” Precise specification of the complexity is not feasible, whereas ignoring the uncertainty and complexity will be misleading. Substantial information already exists to aid in model building, but there are also important gaps in our knowledge about forest processes. To exploit what is known, we must therefore develop model depictions of known processes that can be combined with appropriate structures for what is unknown. Hierarchical modeling provides one way of approaching the problem, and serves as a basis for Bayesian prediction. There are many other approaches to predicting forests. Although we focus here on Bayesian demographic modeling, we review other methods in the final section.

Anticipating how forests will respond to environmental changes requires models that relate environmental conditions, such as precipitation and temperature, to forest dynamics, such as population change due to mortality and growth. These models must also be able to project those relationships forward

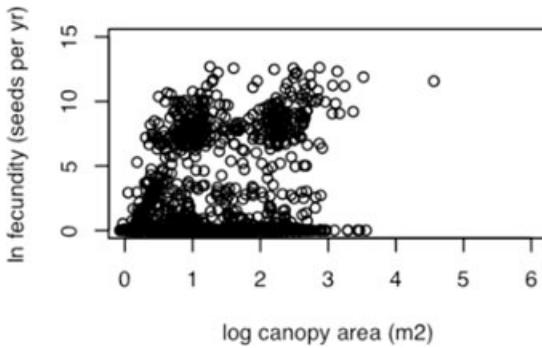


Figure 1. Predicted seed production across a light gradient for *Acer rubrum*. Light is measured as the log of canopy area exposed to direct sunlight estimated from aerial photographs and ground observations.

under scenarios of future conditions. Further, to make these predictions meaningful, ecologists must allow for the fact that the processes driving forest dynamics are only partially known. There is uncertainty about current conditions used to parameterize the models, unknown aspects of the processes themselves, and uncertainty about future conditions. All can influence predictive distributions for numbers, abundances, and geographical ranges of species. A predictive distribution combines the deterministic relationships that are sufficiently well known to be expressed as functional forms with the uncertainty that comes with observations, specification of the model, and estimation of parameters.

Prediction can apply not only to actual future states, but to model and data implications for any time or place where data have not been collected. For example, using a model that was fitted with light as a covariate (Fig. 1), we can construct predictive distributions of a demographic response to light intensities not included in the analysis. The data used for model fitting here include a range of light values but do not necessarily include the precise values where predictions will be made. The predictive distribution therefore needs to integrate over uncertainty in the estimates of the parameters, the observations, and in the model specification, and can thus predict the demographic re-

sponse over a more continuous range of light values. In this section, we summarize the basic conceptual framework for prediction. We then discuss how simulation models are used in prediction and how different sources of uncertainty enter into this analysis. This section concludes with remarks on how to evaluate the predictions produced by these models.

Although there are many important ways to quantify the processes important to forests, from biogeochemical changes to standing biomass, the prediction process we describe here focuses on demographic rates. Changes in demography have a long history in assessments of environmental impacts on population dynamics and health and can indirectly describe chemical storage and biomass. If we want to understand the past and future of our own species, we turn first to demographic rates. The per capita population growth rates used to summarize the population dynamics of nations are routinely broken down into birth, death, and migration rates. Epidemiological data are applied to problems of nutrition and disease, with demographic rates being the response variables.

Demographic rates of trees provide an important scale for analysis and prediction. Seed bank persistence, seed germination, tree growth, survival, fecundity, and dispersal are the basic demographic rates that can be used to understand the impact of global change on individual trees, and by extrapolation to populations and the forest community. By quantifying how these rates relate to important environmental variables as well as to one another (the number of seeds influences the number of future seedlings, and so on) we can begin to capture the behaviors of forests as dynamic systems, which is essential to substantive prediction. Forest ecologists have long understood the value of modeling at this scale, and a number of important forest simulators use demography as the key dynamic engine for their population forecasts (e.g., FORET, Sortie, ED). We extend this tradition to fully embrace the Bayesian structure for both inference and predictive simulation.

Coupling Bayesian Inference and Prediction

We focus on the Bayesian framework because it offers a conceptually coherent framework for prediction (Clark *et al.* 2007a). It is not restricted to a particular model but rather accommodates many types of data and processes. Traditionally, model forecasts have focused on forward simulation, starting with parameter values and scenarios of change as inputs and producing simulated data as output. Parameterization these models, however, was rarely conducted in a way that properly anticipates the role of uncertainty used for prediction (forward simulation from the model). For example, estimates of model parameters often assume a different model, sometimes even a different scale, from the model used to make predictions (e.g., temperature or CO₂ dependence at the level of leaves is often used as a basis for simulating canopy response to atmospheric CO₂).

Extrapolation is easily accomplished under a Bayesian framework where all model parameters are considered random variables. By describing the distributions of those random variables (such as a model for leaf response to CO₂), fine-scale relationships can be integrated to produce predictions for larger scales. The result is a set of predictions that are coherent with model fitting—we know where the uncertainty comes from, because there is a quantitative link between data, process, and prediction. The Bayesian approach encourages a view of uncertainty with specific sources, including process, parameter estimates, and observation errors. Process error is especially important because process models are typically crude, representing only a small subset of all known and important interactions. Process error uses stochasticity to account for the variation in processes due to interactions that have not been formally modeled.

This hierarchical framework involves specifying the relationships between different parts of the model in the form of conditional proba-

bilities. That is, we understand one component of the model *given* (or “conditioned on”) the status of other components. The data are conditional on an underlying process. Likewise, a process at one scale might be conditioned on a different process at another scale. For example, tree growth depends on light available to the tree. Light availability might, in turn, depend on interception by canopies of other trees. One of the key strengths of the Bayesian framework is the flexibility with which models can be constructed to represent complex problems by combining these relatively simple conditional probabilities. This approach has the flexibility to incorporate numerous sources of data, model output, indirectly observed variables, observation error, and prior information about all parts of the problem that have been gathered from sources such as literature or expert opinion. The key components of such analyses are here termed *process models*, *data models*, and *parameter models* (Clark *et al.* 2003).

In our case, process models represent our understanding of how environmental variables influence tree demography. A demographic rate that cannot be measured directly can be represented by a process model informed by relevant data. For example, estimates of fecundity in trees may be informed by seed-trap data, a model for how seeds disperse, by observations of flowering, a model of tree maturation, and by data on tree sizes (Clark *et al.* 2004). Geometric models of canopy light transmission can be related to remote sensing imagery and canopy-status observations to estimate the canopy light environment, and to hemispherical photography to estimate understory light conditions. Process models can be as simple as a linear relationship or as complex as a system of nonlinear differential equations, such as those used for the atmosphere or hydrological cycles. The process model itself is uncertain, and we would like that uncertainty to be reflected in predictive distributions. Model error is particularly hard to quantify, because the concept of a “correct” model rarely has clear meaning in environmental science. Processes are influenced

by many factors, only a subset of which will be specified.

Data models connect observations with an underlying process. Multiple data sets might inform the same process. For example, a process model for tree growth may be informed by both inventory data and tree increment cores, with each data set having its own data model accounting for the different biases and error distributions (Clark *et al.* 2007b).

Parameter models, the core of model estimation, describe distributions for all model parameters. The posterior distribution of parameters used for inference about parameter significance combines the likelihood distributions with *priors*, distributions based on prior knowledge (or uniform distributions when no prior knowledge is assumed). Further, parameter models can account for hierarchical relationships in the models that may have important influence on inference. For example, we expect the parameters linking growth to light to differ among individuals, which could warrant a random-effects prior that quantifies how likely two individuals are to differ in this growth–light relationship. Inference entails estimates of these different contributions to the overall model fit. Prediction combines these estimates with scenarios that involve the model predictors, or “explanatory variables.”

Using Models for Predicting Future States

Moving from inference to predicting future states adds yet another source of uncertainty to the predictive process—that associated with a hypothesized scenario. If climate variables enter as explanatory variables for tree growth, and we want to predict growth in a different climate setting, we must recognize the uncertainty associated with climate change. If the model is dynamic, then even at the outset of a predictive simulation, the initial conditions are uncertain. Initial conditions may be set to observations (see later in the chapter), but many variables may be unobserved. Boundary conditions are

often based on scenarios about the future state of the system. Scenarios may include the best estimates of experts or the output of other models, having uncertainties of their own. Bayesian analysis allows for uncertainty in initial states and boundary conditions. Ideally, predictive distributions integrate uncertainty from all of these sources.

BOX 1: A taxonomy of uncertainty could include sources of error summarized here:

- *Process error.* Uncertainty in the process model. Variation affects dynamics and can compound over time.
- *Measurement error.* Discrepancy between observations and true process state. Measurement error can, but does not have to be propagated to predictions.
- *Random individual and temporal effects (RITES).* Individual effects refer to uncertainty that is associated with individuals (such as differences created by genetic variation) and over time. Temporal uncertainty can apply to all individuals at once, but changes over time (e.g., seed masting in trees).
- *Initial condition uncertainty.* Error in the specification of the initial state of a predictive simulation.
- *Scenario uncertainty.* Uncertainty in the variables that describe predicted scenarios (e.g., temperature means over the next 100 years).

It is important to interpret predictive distributions appropriately. Some ecological processes are highly predictable, for reasons that are well understood. Leaf-area index can often be predicted based on local hydrology (Grier and Running 1977). Other processes, such as long-distance dispersal, are inherently unpredictable despite confident parameter estimates for dispersal kernels (Clark *et al.* 2003). Still other processes depend on uncertain scenarios, such as CO₂ acclimation.

Because many climate-change phenomena are complex, with many potential sources of error, predictive distributions may often be less useful for anticipating the future than they are for analyzing how different sources of

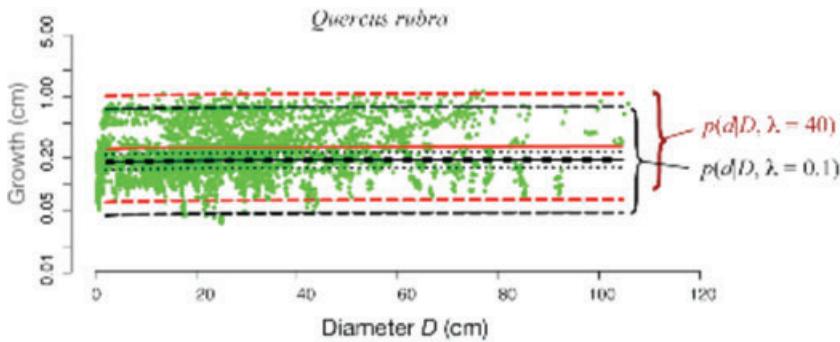


Figure 2. Posterior mean estimates of latent states (green dots) and predictive intervals for low (black) and high (red) canopy exposure (λ is an estimate of the canopy area exposed to full sunlight). Black predictive intervals for low light, from narrow to wide, are parameter uncertainty (dashed), random individual effects (dotted), year-to-year variation (dashed), and process error (not visible). For high light (red) only predictive mean and year-to-year variation are shown.

uncertainty contribute to our understanding of a process. This understanding can then be used to prioritize research needs and direct future field research and experiments designed to reduce forecast uncertainty.

The Inference Stage

Predicting forest change requires a series of steps that link observations on individual trees to computer simulations of entire forest stands under anticipated climate scenarios. To illustrate these concepts we discuss an application where hierarchical models were fitted to data, including estimation of measurement error, process error, and parameter error, all of which are propagated to predictive distributions based on scenarios for soil moisture.

Prediction begins with inference, the estimation of parameter values for the model. Figure 2 shows how different sources of uncertainty contribute to a predictive interval around the relationship between the diameter of a tree and its growth rate. In this figure growth rates of adult trees are predicted for a range of diameters at two light levels for *Quercus rubra*. For this example of diameter growth, the size of the predictive intervals illustrates how year-to-year variation dominates over individual differences, parameter uncertainty, and model uncertainty. In other

words, annual variation is large and not well described based on the covariates in the model, which include light availability, tree diameter, and growth the previous year. However, individual differences and parameter uncertainty are relatively small. This combination of process and uncertainty will be further discussed in an example that follows.

An analysis of uncertainty can help us understand population dynamics even when processes are poorly specified. Figure 3 shows how predictive intervals can be used to assess goodness of fit. The predictive distributions shown in Figure 3 can be used to simulate growth rates of trees, averaging over all sources of variation (dashed lines in Fig. 3). For comparison, the green lines in Figure 3 show growth trajectories of actual trees based on tree rings measured from increment core samples. The bounds of the predicted growth trajectories are not based on models of how age relates to diameter increment. Instead, these trajectories were estimated purely from the process model of growth over time, given uncertainty. Every year growth is estimated from covariates and the diameter change is calculated.

To incorporate species interactions in the model, we use a simulator where trees intercept light, thereby shading their neighbors. Scenarios involving soil moisture fluctuation allow us

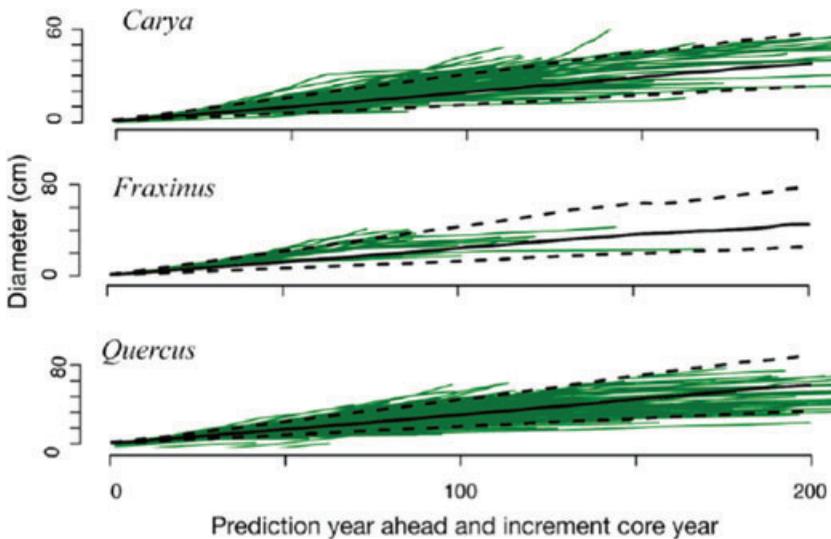


Figure 3. Comparison of increment data from tree rings (green) and predictive distributions of tree diameter.

to explore the consequences of climate change in the context of competition among trees for light. We use the Scalable Landscape, Inference and Prediction (SLIP) simulator (Govindarajan *et al.* 2007) which combines the demographic-rate models to project individuals forward. SLIP differs from other forest simulators (e.g., SORTIE, JABOWA, FORET) in that it incorporates uncertainty estimated at each stage of the model described earlier. In the SLIP predictions, interacting individuals grow, survive, and reproduce using the same model fitted to data over a landscape grid of 512-by-512 meters. We incorporate not only responses to covariates but also the estimated variation across years and among individuals (RITES, see Box 1) and uncertainty in the processes model. Trees compete for light, and the largest trees create the light environment experienced by smaller trees.

To complete the path from data to prediction under changing climate, we can compare model runs with current climate variables (temperature and soil moisture) with those run under scenarios that reflect anticipated shifts in southeastern climate patterns. This final section describes how SLIP simulated forest dynamics under a scenario of periodic extreme drought.

A Drought Example

Droughts are predicted to occur in the southeastern United States with greater frequency and intensity in the coming century (Mearns *et al.* 2003). The structure of forests in the Southeast may depend on whether or not and which deep-rooted forest trees are buffered from increasingly arid climate. If adults remain connected to a deep water supply, while shallow-rooted seedlings fail to survive, we could experience a gradual shift to Mediterranean-like conditions.

For a simple demonstration, we vary soil moisture stochastically, with an underlying periodic signal (a simple sine function with random noise). As a control, we assumed a mean and variance for the observed soil-moisture data that are uncorrelated across years. In these simulations we include soil moisture effects only at the stages of germination and seedling growth and mortality. These simulations address the extent to which large trees might be buffered against increased aridity, whereas shallow-rooted seedlings are not.

Figure 4 shows modeled relationships between the seedlings of three different species (corresponding to the three columns), *Acer*

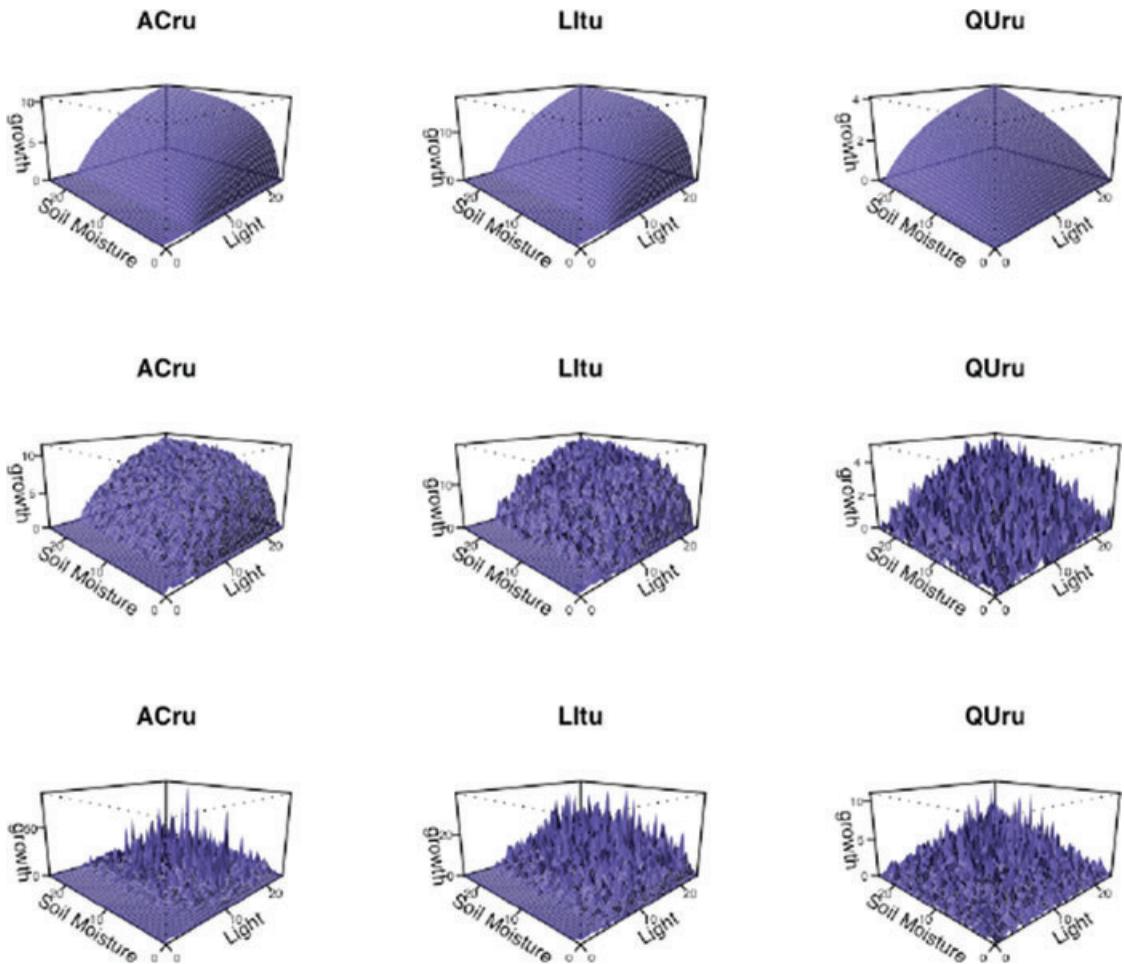


Figure 4. Example simulations of the response of seedling growth to soil moisture and light for three species—A: rubrum, L: tulipifera, and Q: rubra. The three rows show **(A)** the deterministic relationship between covariates and growth, **(B)** an instance where process error has been drawn for each individual from the process-error distribution estimated from data, and **(C)** where instances of individual effects as well as process error are drawn from their respective distributions. Error is represented by drawing random deviations from the error distributions, and therefore are examples of the uncertainty and not the actual estimated error terms in order to visualize the variance due to error.

rubrum, *Liriodendron tulipifera*, and *Q. rubra*, and light and soil moisture. These relationships were estimated from data in field experiments (Ibanez *et al.* 2007) using a logistic regression. The z -axis marks predicted annual seedling growth in centimeters of height, and the x - and y -axes mark percent of total solar radiation reaching the understory and percent moisture content in the soil, respectively. The first row of Figure 4 shows the process model (the deterministic model without uncertainty). Light has a strong positive effect on growth,

while soil moisture has a more subtle effect. These relationships differ for different species. In row 2, every point on the graph was changed by adding a random variable drawn from the distribution of process error, so that for every combination of light and soil moisture, possible changes in expected height due to process error are illustrated. Under poor conditions, even with positive values for process error, there is no growth (the flat plateaus under low light or low soil moisture values). Note, too, that different species have different levels

of process uncertainty. *Q. rubra* shows greater differences from the deterministic process we modeled. That is, for *Q. rubra* light and soil moisture fail to capture many of the determinants of seedling growth, whereas for *L. tulipifera*, soil moisture and light explain a good deal of seedling growth. In the third row random draws from both process error and individual effects have been added to the deterministic value of growth predicted with the environmental variables. This illustrates potential deviations from the deterministic kernel due to unmodeled processes.

Given the large number of seedlings and the highly dynamic understory-light environment of intact forests, inferring large-scale population dynamics under periodic droughts from the estimated growth of seedlings alone would be unrealistic. Seedlings grow to become adult trees, which shade seedlings, and individuals attain adulthood (or not) due to many processes, all of which differ between species. To understand how the entire forest system will respond to drought, therefore, we need to integrate these effects over the behavior of all forest trees, not just seedlings.

We ran five replicates of the control and the drought treatment over 200 simulated years, with a starting forest based on the observed abundances and sizes of 30 tree species from the Duke Forest in central North Carolina. We found a clear response to drought in the forest, but also resilience in key community-level patterns (Fig. 5). Seedling densities (Fig. 5a), as predicted by our inference based on process models, oscillated with soil moisture levels. Although the roughly 50% drop in seedling densities did translate into fewer saplings (which includes stems between 0.5-m and 2-m tall) (Fig. 5b), the duration of the drop in sapling number and the speed of recovery at the end of drought showed a diminished response as compared to that of seedling densities. This modeled tolerance of periodic drought is even more pronounced as the size class of trees increases. Adult densities and growth showed little response to drought cycles (Fig. 5c). Although stem numbers did

fluctuate, they did so very close to the range of the control runs. Thus, adult stems were not reduced to savannah-like densities, even though seedling mortality during droughts was high.

In order to draw inference for this experiment, we return to how uncertainty entered the model and the simulations. In this example, we ran five replicates of the two treatments. These repeated runs, although not extensive, show how sensitive our inference is to the model's stochastic components. Process error and RITES were included in these runs, so that at the scale of individuals and years, uncertainty entered in every vital rate, and did so differently for different species. Despite these sources of uncertainty, the abundances and densities of tree life-history stages were remarkably consistent over the runs.

To generate more informed predictive distributions, we would need to run the simulations more times, but the trend is obvious in this example. The patterns we see in densities over time also seem robust to the high degree of stochasticity modeled here—drought treatments behave similarly, and control runs behave similarly across replicates. Importantly, however, we included no uncertainty in our initial conditions or predictive scenarios. We assumed a specific starting forest composition and constructed a soil-moisture regime as input that implied no uncertainty in either soil moisture over the next 200 years or in the forest's initial composition. We know that changes in sequence, frequency, and intensity of the drought scenarios modeled would strongly influence how the forest behaves (which will be discussed in more detail in the concluding section).

A single experiment does not have to accommodate every potential source of uncertainty. Any inference drawn from model results, however, must be restricted by how uncertainty enters the model. The drought simulation just given tests a hypothesis focused on a specific mechanism: Could adult trees that are not affected by rare drought events buffer the overall forest from droughts that damage

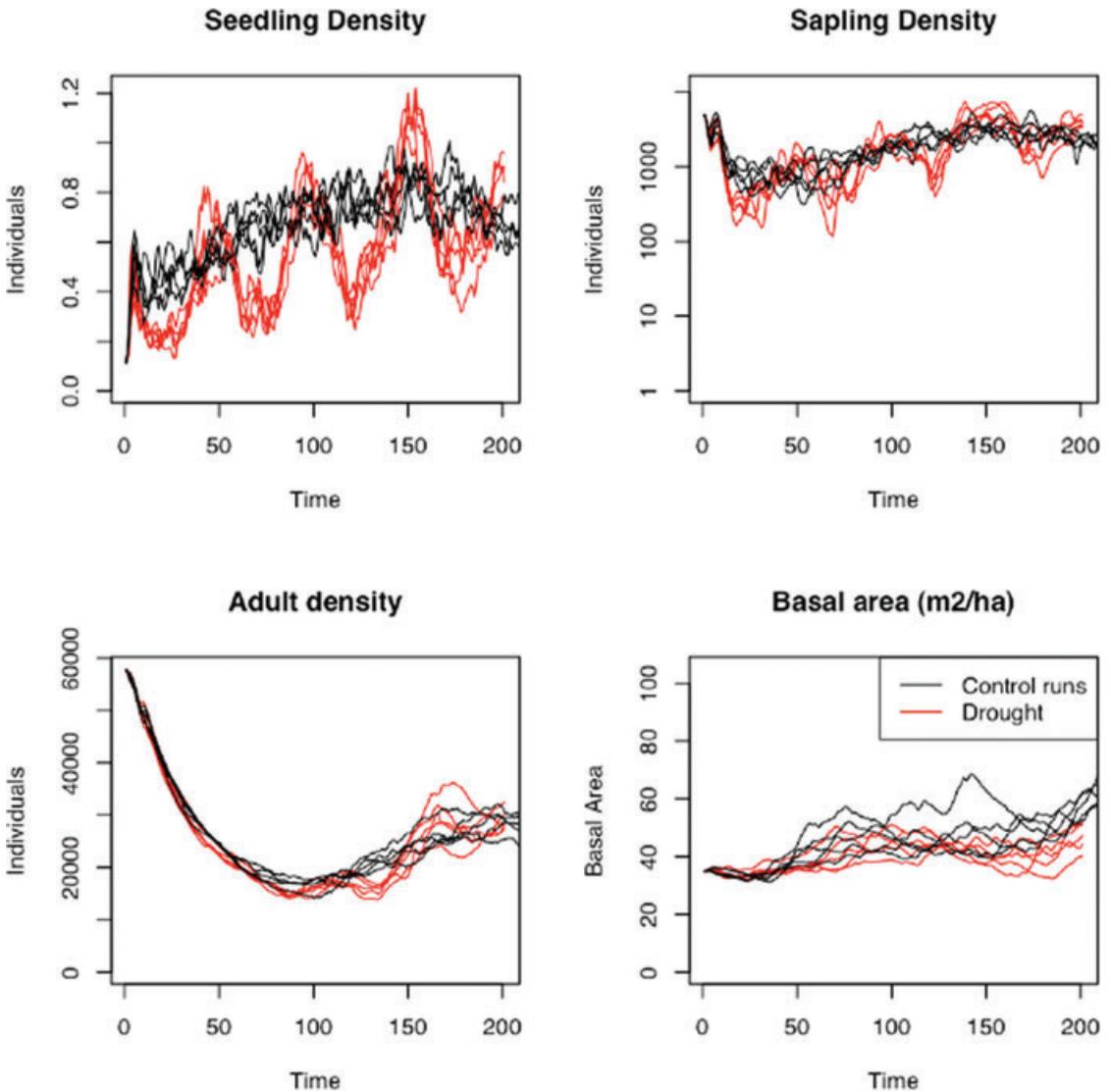


Figure 5. Time series for five separate 200-year simulation runs with drought treatments (*red lines*) and without (*black lines*). Note log y -axes for saplings.

seedlings and saplings? We can deduce from these experiments that more frequent drought or longer drought periods might eventually influence adult population and forest biomass. The important inference for this simulation is that the forest as a whole may be buffered from extreme drought events that only affect early life-history stages. Results such as this one do not contribute much to an overall prediction of forest change, but can contribute to a toolbox of information essential for developing forecasts.

The final section discusses how forecasting can combine data, models of ecological processes, and climate scenarios to create a protocol for prediction.

A New Forecasting Protocol Based on Process

The root of model-based prediction lies in model-based inference. The goal of the

inference stage of analysis is to fit a deterministic core of a model (e.g., “how trees respond to the environment”) and to partition unexplained variance in a way that can help the model realize the variety of possible futures based on how little we know now. The importance of the deterministic portion of a model for prediction is well understood by ecologists (it’s our job), but the impacts of the error and how it is partitioned are less appreciated. Most ecological models intended for forecasting drop the residual error entirely, or treat the error as additive Gaussian white noise. There is a growing recognition of the need to include autocorrelation in space and time, but modern statistical methods also allow for the inclusion of nonadditive effects, such as the variation of model parameters in space and time as well as hierarchical random effects at all levels, down to every tree having its own parameters, as shown in the preceding example. Also, as demonstrated in our example, data models allow the separation of various forms of measurement error from process error, which is critical in ecology, as variability is often dominated by true differences among individuals, sites, or samples rather than measurement error. The treatment of variability may not only affect the predictive variance, but also the mean and even our understanding of the dynamics of the process, such as the mechanisms for species coexistence (Clark *et al.* 2007a).

Moving from inference to prediction introduces a number of additional sources of uncertainty. Forward simulation requires the specification of both *initial conditions*, which describe the initial state of the model, and *boundary conditions*, which describe the time-evolution of model constraints. The specification of initial conditions is often based on data, but the initial values of all state variables are rarely observed completely, and there are often multiple data sets that can inform the initial state. Determining the initial conditions can therefore be a challenging problem of statistical inference at the start. Although this type of statistical modeling is not currently common in the ecological literature, it is a backbone of inference

in meteorology where it is referred to as *data assimilation*. The classic tools for this inference in data assimilation have been variants of the Kalman filter (Kalman 1960), which estimates the dynamics of a system based on sparse measurements of its characteristics. Newer hierarchical Bayes approaches are both more flexible and less complicated (Wikle and Berliner 2006). Boundary conditions, on the other hand, are often developed based on scenarios about the future state of the system, and thus all forecasts are made conditional on these scenarios. Scenarios may include the best estimates of experts, maximum and minimum range bounding scenarios, or the output of other models with their own uncertainty. For example, the Intergovernmental Panel on Climate Change (IPCC) uses expert opinion and socioeconomic models to develop a standard set of climate-change scenarios that are boundary conditions for gas emissions, land-use change, and development within general circulation models (GCMs) (IPCC 2000, 2007). The outputs of GCMs routinely become the input scenarios for forest biogeochemistry and biodiversity models (e.g., VEMAP). Obviously, scenarios and the forecasts made from them cannot be expected to be true, since the information provided by forecasts can alter the behavior of the agents involved, but they are useful precisely because they provide us with information that allows us to alter our behavior.

Another critical component of forecast uncertainty comes from model choice. Model choice error is particularly hard to quantify, since all probability statements are conditional on the model implemented or upon a finite set of alternate models, rather than the infinite set of all possible models. Information theory and model selection can help us choose between alternate model forms, or to weight models for model averaging (Burnham and Anderson 2002); however, such statistical measures are not a substitute for an experts understanding of the system nor do they guard against errors in our conceptual understanding of the system. It is often a useful exercise to compare forecasts

generated by different models produced by different research teams; however, such models are never truly independent samples since they often encapsulate the same understanding and even share submodels as a whole [e.g., almost all ecosystem models estimate photosynthesis using the Farquhar *et al.* (1980) model].

Finally, a critical component of forecasting is to understand the predictability of the process and the factors limiting better prediction. Some ecological processes are very predictable. For example, leaf-area index tends to vary little for closed-canopy forests around the world. Other processes, such as long-distance dispersal, are inherently unpredictable: once the parameters involved are reasonably constrained, collecting reams of additional data will not drastically improve prediction. Ecological processes also vary widely in the temporal horizon for accurate prediction. Some processes are clearly sensitive to initial conditions, while others, such as forest succession, tend to have strong feedbacks that cause dynamics to converge over time. Processes also differ in the extent to which we can consider their parameters fixed versus varying in time and space, especially in the face of novel environmental change. The analysis of models with the proper propagation of error will itself be a critical component of understanding the main sources of uncertainty. This understanding can then be used to prioritize research needs and direct future field research and experiments designed to reduce forecast uncertainty.

It is easy to become overwhelmed by the urgency and complexity of current changes in ecological systems. By working together with climate scientists, forest ecologists and managers, statisticians, and computer scientists, and paying careful attention to details of ecological systems and how they extrapolate to produce patterns in communities and regions, we can develop forecasts, learn from them, improve them, and redevelop them in the hope that along the way, our informed answers will have a positive impact on how our society responds to these changes.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- Ainsworth, E.A. & S.P. Long. 2005. What have we learned from 15 years of free-air CO₂ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy. *New Phytol.* **165**: 351–371.
- Ayers, M.P. & M.J. Lombardero. 2000. Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *Sci. Total Environ.* **262**: 263–286.
- Battisti, A., M. Stastny, S. Netherer, *et al.* 2005. Expansion of geographic range in the pine processionary moth caused by increased winter temperatures. *Ecol. Appl.* **15**: 2084–2096.
- Bergot, M., E. Cloppet, V. Perarnaud, *et al.* 2004. Simulation of potential range expansion of oak disease caused by *Phytophthora cinnamomi* under climate change. *Global Change Biol.* **10**: 1539–1552.
- Bolker, B.M., S.W. Pacala, F.A. Bazazz, *et al.* 1995. Species diversity and ecosystem response to carbon dioxide fertilization: conclusions from a temperate forest model. *Global Change Biol.* **1**: 373–381.
- Burdon, J.J., P.H. Thrall & L. Ericson. 2006. The current and future dynamics of disease in plant communities. *Annu. Rev. Phytopathol.* **44**: 19–39.
- Buizza, R. & T.N. Palmer. 1995. The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.* **52**: 1434–1456.
- Burnham, K.P. & D.R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer. New York.
- Canadell, J.G., M.U.F. Kirschbaum, W.A. Kurz, *et al.* 2007. Factoring out natural and indirect human effects on terrestrial carbon sources and sinks. *Environ. Sci. Policy* **10**: 370–384.
- Cannon, R.J.C. 1998. The implications of predicted climate change for insect pests in the UK, with emphasis on non-indigenous species. *Global Change Biol.* **4**: 785–796.
- Clark, J.S., J. Mohan, M. Dietze & I. Ibanez. 2003. Coexistence: how to identify trophic tradeoffs. *Ecology* **84**: 17–31.
- Clark, J.S., S.R. Carpenter, M. Barber, *et al.* 2001. Ecological forecasts: an emerging imperative. *Science* **293**: 657–660.
- Clark, J.S., M. Dietze, P. Agarwal, *et al.* 2007a. Resolving the biodiversity debate. *Ecol. Lett.* **10**: 647–662.
- Clark, J.S., M. Wolosin, M. Dietze, *et al.* 2007b. Tree growth inference and prediction from diameter censuses and ring widths. *Ecol. Appl.* **17**: 1942–1953.

- Clark, J.S., S. LaDeau & I. Ibanez. 2004. Fecundity of trees and the colonization-competition hypothesis. *Ecol. Monogr.* **74**: 415–442.
- Clinton, B.D., L.R. Boring & W.T. Swank. 1993. Canopy gas characteristics and drought influences in oak forests of the Coweeta basin. *Ecology* **74**: 1551–1558.
- Crowley, T.J. 1990. Are there any satisfactory geologic analogs for a future greenhouse warming? *J. Clim.* **3**: 1282–1292.
- Curtis, P.S. & X.Z. Wang. 1998. A meta-analysis of elevated CO₂ effects on woody plant mass, form, and physiology. *Oecologia* **113**: 299–313.
- Davis, M.B. 1969. Climatic changes in southern Connecticut recorded by pollen deposition at Rogers lake. *Ecology* **50**: 409–422.
- Desprez-Loustau, M.L., B. Marcais, L.M. Mageleisen *et al.* 2006. Interactive effects of drought and pathogens in forest trees. *Ann. Forest Sci.* **63**: 597–612.
- Farquhar, G.D., S. von Caemmerer & J.A. Berry. 1980. A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* **149**: 78–90.
- Garrett, K.A., S.P. Dendy, E.E. Frank, *et al.* 2006. Climate change effects on plant disease: genomes to ecosystems. *Annu. Rev. Phytopathol.* **44**: 489–509.
- Govindarajan, S., M.C. Dietze, P.K. Agarwal & J.S. Clark. 2007. A scalable algorithm for dispersing population. *J. Intel. Inform. Syst.* **29**: 39–61.
- Gorchov, D.L. & D.E. Trisel. 2003. Competitive effects of the invasive shrub, *Lonicera maaackii* (Rupr.) Herder (Caprifoliaceae), on the growth and survival of native tree seedlings. *Plant Ecol.* **166**: 13–24.
- Grier, C.G. & S.W. Running. 1977. Leaf area of mature northwestern coniferous forests: relation to site water balance. *Ecology* **58**(4): 893–899.
- Harvell, C.D., C.E. Mitchell, J.R. Ward, *et al.* 2002. Ecology—Climate warming and disease risks for terrestrial and marine biota. *Science* **296**: 2158–2162.
- Hattenschwiler, S. & C. Korner. 2000. Tree seedling responses to in situ CO₂-enrichment differ among species and depend on understorey light availability. *Global Change Biol.* **6**: 213–226.
- Heneghan, L., F. Fatemi, L. Umek, *et al.* 2006. The invasive shrub European buckthorn (*Rhamnus cathartica*, L.) alters soil properties in Midwestern US woodlands. *Appl. Soil Ecol.* **32**: 142–148.
- Hitch, A.T. & P.L. Leberg. 2007. Breeding distributions of North American bird species moving north as a result of climate change. *Cons. Biol.* **21**: 534–539.
- Huntley, B. 1991. How plants respond to climate change: migration rates, individualism and the consequences for plant communities. *Ann. Bot.* **67**: 15–22.
- Ibanez, I., J.S. Clark, M. Dietze, *et al.* 2006. Predicting biodiversity change: outside the climate envelope, beyond the species-area curve. *Ecology* **87**: 1896–1906.
- Ibanez, I., J.S. Clark, S. LaDeau & J.H.R. Lambers. 2007. Exploiting temporal variability to understand tree recruitment response to climate change. *Ecol. Monogr.* **77**: 163–177.
- Intergovernmental Panel on Climate Change. 2000. *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. N. Nakicenovic & R. Swark, Eds. Cambridge University Press, Cambridge, UK.
- Intergovernmental Panel on Climate Change. 2007. *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, Cambridge, UK.
- Iverson, L.R. & A. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecol. Monogr.* **68**: 465–485.
- Kalman, R. 1960. A new approach to linear filtering and prediction problems. *J. Basic Engineering* **82**: 35–45.
- LaDeau, S. & J.S. Clark. 2001. Rising CO₂ levels and the fecundity of forest trees. *Science* **292**: 95–98.
- Logan, J.A., J. Regniere & J.A. Powell. 2003. Assessing the impacts of global warming on forest pest dynamics. *Front. Ecol. Environ.* **1**: 130–137.
- Long, S.P., E.A. Ainsworth, A. Rogers, & D.R. Ort. 2004. Rising atmospheric carbon dioxide: plants face the future. *Annu. Rev. Plant Biol.* **55**: 591–628.
- Lovett, G.M., C.D. Canham, M.A. Arthur, *et al.* 2006. Forest ecosystem responses to exotic pests and pathogens in eastern North America. *Bioscience* **56**: 395–405.
- Martin, P.H. & P.L. Marks. 2006. Intact forests provide only weak resistance to a shade-tolerant invasive Norway maple (*Acer platanoides* L.). *J. Ecol.* **94**: 1070–1079.
- McCarty, J.P. 2001. Ecological consequences of recent climate change. *Conserv. Biol.* **15**: 320–331.
- McLachlan, J.S., J.S. Clark & P.S. Manos. 2005. Molecular indicators of tree migration capacity under rapid climate change. *Ecology* **86**: 2088–2098.
- Mearns, L.O., F. Giorgi, L. McDaniel & C. Shields. 2003. Climate scenarios for the southeastern U.S. based on GCM and Regional model simulations. *Clim. Change* **60**: 7–35.
- Meunier, C., L. Sirois & Y. Begin. 2007. Climate and *Picea mariana* seed maturation relationships: a multi-scale perspective. *Ecol. Monogr.* **77**: 361–376.
- Mohan, J.E., J.S. Clark & W.H. Schlesinger. 2007. Long-term CO₂ enrichment of a forest ecosystem: Implications for forest regeneration and succession. *Ecol. Appl.* **17**: 1198–1212.
- Moore, D.J.P., S. Aref, R.M. Ho, *et al.* 2006. Annual basal area increment and growth duration of *Pinus taeda* in response to eight years of free-air carbon dioxide enrichment. *Global Change Biol.* **12**: 1367–1377.

- Moritz, C. *et al.* 2008. Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science* **322**: 261–264.
- Norby, R.J., E.H. DeLucia, B. Gielen, *et al.* 2005. Forest response to elevated CO₂ is conserved across a broad range of productivity. *Proc. Nat. Acad. Sci.* **102**: 18052–18056.
- Ogren, W.L. 1984. Photorespiration: pathways, regulation, and modification. *Annu. Rev. Plant Physiol.* **35**: 415–452.
- Orwig, D.A. & D.R. Foster. 1998. Forest response to the introduced hemlock woolly adelgid in southern New England, USA. *J. Torrey Bot. Soc.* **125**: 60–73.
- Parmesan, C. & G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**: 37–42.
- Powers, J.S. *et al.* 1999. Plant-pest interactions in time and space: a Douglas-fir bark beetle outbreak as a case study. *Lands. Eco.* **14**: 105–120.
- Prentice, I.C., P.J. Bartlein & T. Webb III. 1991. Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology* **72**: 2038–2056.
- Reich, P.B., J. Knops, D. Tilman, *et al.* 2001. Plant diversity enhances ecosystem responses to elevated CO₂ and nitrogen deposition. *Nature* **410**: 809–812.
- Roe, G.H. & M.B. Baker. 2007. Why is climate sensitivity so unpredictable? *Science* **318**: 629–632.
- Roy, B.A., S. Gusewell & J. Harte. 2004. Response of plant pathogens and herbivores to a warming experiment. *Ecology* **85**: 2570–2581.
- Saxon, E., B. Baker, W. Hargrove, *et al.* 2005. Mapping environments at risk under different global climate change scenarios. *Ecol. Lett.* **8**: 53–60.
- Sefcik, L.T., D.R. Zak & D.S. Ellsworth. 2007. Seedling survival in a northern temperate forest understory is increased by elevated atmospheric carbon dioxide and atmospheric nitrogen deposition. *Global Change Biol.* **13**: 132–146.
- Simberloff, D. & B. Von Holle. 1999. Positive interactions of nonindigenous species: invasional meltdown? *Biol. Invasions* **1**: 21–32.
- Stiling, P. & T. Cornelissen. 2007. How does elevated carbon dioxide (CO₂) affect plant-herbivore interactions? A field experiment and meta-analysis of CO₂-mediated changes on plant chemistry and herbivore performance. *Global Change Biol.* **13**: 1823–1842.
- Stinson, K.A., S.A. Campbell, J.R. Powell, *et al.* 2006. Invasive plant suppresses the growth of native tree seedlings by disrupting belowground mutualisms. *PLoS Biol.* **4**: e140.
- Thomas, C.D., A.M.A. Franco & J.K. Hill. 2006. Range retractions and extinction in the face of climate warming. *TREE* **21**: 415–416.
- Tinner, W. & A.F. Lotter. 2001. Central European vegetation response to abrupt climate change at 8.2 ka. *Geology* **29**: 551–554.
- Van Devender, T.R. & W.G. Spaulding. 1979. Development of vegetation and climate in the southwestern United States. *Science* **204**(4394): 701–710.
- Webb, T., III. 1986. Is vegetation in equilibrium with climate? How to interpret late-quaternary pollen data. *Plant Ecol.* **67**: 75–91.
- Wikle, C.K. & L.M. Berliner. 2006. A Bayesian tutorial for data assimilation. *Physica D.* **230**(1): 1–16.
- Williams, A.L., K.E. Wills, J.K. Janes, *et al.* 2007a. Warming and free-air CO₂ enrichment alter demographics in four co-occurring grassland species. *New Phytol.* **176**: 365–374.
- Williams, J.W., B.N. Schuman, T. Webb, *et al.* 2004. Late-quaternary vegetation dynamics in North America: scaling from taxa to biomes. *Ecol. Monogr.* **74**: 309–334.
- Williams, J.W., S.T. Jackson & J.E. Kutzbach. 2007b. Projected distributions of novel and disappearing climates by 2100 AD. *Proc. Natl. Acad. Sci. USA* **104**: 5738–5742.
- Woods, A., K.D. Coates & A. Hamann. 2005. Is an unprecedented *dothistroma* needle blight epidemic related to climate change? *Bioscience* **55**: 761–769.