Synthesizing multiple data sources to understand the population and community ecology of California trees

by

Melissa Viola Eitzel Solera

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Science, Policy, and Management

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Perry de Valpine, Chair
Professor John J. Battles
Professor Mary E. Power

Fall 2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Synthesizing multiple data sources to understand the population and community ecology of California trees

Abstract

Synthesizing multiple data sources to understand the population and community ecology of California trees

by

Melissa Viola Eitzel Solera

Doctor of Philosophy in Environmental Science, Policy, and Management

University of California, Berkeley

Professor Perry de Valpine, Chair

In this work, I answer timely questions regarding tree growth, tree survival, and community change in California tree species, using a variety of sophisticated statistical and remote sensing tools.

In Chapter 1, I address tree growth for a single tree species with a thorough explanation of hierarchical state-space models for forest inventory data. Understanding tree growth as a function of tree size is important for a multitude of ecological and management applications. Determining what limits growth is of central interest, and forest inventory permanent plots are an abundant source of long-term information but are highly complex. Observation error and multiple sources of shared variation (spatial plot effects, temporal repeated measures, and a mosaic of sampling intervals) make these data challenging to use for growth estimation. I account for these complexities and incorporate potential limiting factors (tree size, competition, and resource supply) into a hierarchical state-space model. I estimate the diameter growth of white fir (*Abies concolor*) in the Sierra Nevada of California from forest inventory data, showing that estimating such a model is feasible in a Bayesian framework using readily available modeling tools. In this forest, white fir growth depends strongly on tree size, total plot basal area, and unexplained variation between individual trees. Plot-level resource supply variables (representing light, water, and nutrient availability) do not have a strong impact on inventory-size trees. This approach can be applied to other networks of permanent forest plots, leading to greater ecological insights on tree growth.

In Chapter 2, I expand my state-space modeling to examine survival in seven tree species, as well as investigating the results of modeling them in aggregate (at the community level) and comparing with the individual species models. Declining tree survival is a complex, well-recognized problem, but studies have been largely limited to relatively rare old-growth forests or low-diversity systems, and to models which are species-aggregated or cannot easily accommodate yearly climate variables. I estimate survival models for a relatively diverse second-growth forest in the Sierra Nevada of California using a hierarchical state-space

framework. I account for a mosaic of measurement intervals and random plot variation, and I directly include yearly stand development variables alongside climate variables and topographic proxies for nutrient limitation. My model captures the expected dependence of survival on tree size. At the community level, stand development variables account for decreasing survival trends, but species-specific models reveal a diversity of factors influencing survival. Species time trends in survival do not always conform to existing theories of Sierran forest dynamics, and size relationships with survival differ for each species. Within species, low survival is concentrated in susceptible subsets of the population and single estimates of annual survival rates do not reflect this heterogeneity in survival. Ultimately only full population dynamics integrating these results with models of recruitment can address the potential for community shifts over time.

In Chapter 3, I combine statistical modeling with remote sensing techniques to investigate whether topographic variables influence changes in woody cover. In the North Coast of California, changes in fire management have resulted in conversion of oak woodland into coniferous forest, but the controls on this slow transition are unknown. Historical aerial imagery, in combination with Object-Based Image Analysis (OBIA), allows us to classify land cover types from the 1940s and compare these maps with recent cover. Few studies have used these maps to model drivers of cover change, partly due to two statistical challenges: 1) appropriately accounting for spatial autocorrelation (ideally without throwing away data) and 2) appropriately modeling percent cover which is bounded between 0 and 100 and not normally distributed. I study the change in woody cover in California's North Coast using historical (1948) and recent (2009) high-spatial-resolution imagery. I classify the imagery using eCognition Developer and aggregate the resulting maps to the scale of a Digital Elevation Model (DEM) in order to understand topographic drivers of woody cover change. I use Generalized Additive Models (GAMs) with a quasi-binomial probability distribution to account for spatial autocorrelation and the boundedness of the percent woody cover variable. I explore the relative roles of elevation, topographic slope, aspect (Northness/Eastness), topographic wetness index, profile curvature, historical percent woody cover, and geographical coordinates in influencing current percent woody cover. I estimate these models for scales of 20, 30, 40, 50, 60, 70, 80, 90, and 100 m, reflecting both tree neighborhood scales and stand scales. I find that historical woody cover has a consistent positive effect on current woody cover, and that the spatial term in the model is significant even after controlling for historical cover. Specific topographic variables emerge as important for different sites at different scales, but no overall pattern emerges across sites or scales for any of the topographic variables I tested. This GAM framework for modeling historical data is flexible and could be used with more variables, more flexible relationships with predictor variables, and larger scales. Modeling drivers of woody cover change from historical ecology data sources can be a valuable way to plan restoration and enhance ecological insight into landscape change.

I conclude that these techniques are promising but a framework is needed for sensitivity analysis, as modeling results can depend strongly on variable selection and model structure.

For those that sustain me: my ancestors, my elders, and my family – chosen and by blood, human and more-than-human.

Contents

List of Figures

List of Tables

Acknowledgments

There are so many wonderful, brilliant, fun people who have been part of my time at Berkeley. I am so grateful to my committee, Mary Power, John Battles, and Perry de Valpine. Perry has supported me even from far away on sabbatical and found time to read drafts and investigate computational problems while managing a busy home life. He is a role model of finding time for both family and career, and has given me advice in all those areas as well. Mary has always been willing to brave the wilds of the statistics in my papers to give me excellent feedback on drawing out the ecology, and calling me on jargon which would be difficult for those less statistically inclined. John has supported me in not just my research efforts but also my efforts to create a program for graduate students to do cooperative extension and outreach. He found time and money for me when I wasn't even his student, and I have always enjoyed sitting and discussing my research projects with him. I also thank Maggi Kelly for support and advice on everything geospatial and cooperative extension-oriented, and Louise Fortmann for interdisciplinary explorations, collaborations, networking, and tea. There are so many other faculty members who have been supportive of my projects, including Katie Suding, Justin Brashares, Rick Standiford, and the instructors of the many excellent courses I've taken. I am also grateful for the interdisciplinary joint teaching experience I shared with Tina Mendez and Kurt Spreyer.

My fellow graduate students have also made my experience at Berkeley some of the most inspiring, engaging, fun, and productive years of my life. I want particularly to thank Sibyl Diver, Jesse Williamson, and Karen Andrade for many delicious interdisciplinary conversations and projects. My cohort, the incoming class of 2008, were wonderful compatriots over the years, and in particular Justin Kitzes has been a great collaborator on modeling projects. I want to thank the new Graduate Student Extension students, Kevi Mace, Stella Cousins, Luke Macaulay, Adam Calo, Kate Wilkin, and Hillary Sardiñas, for carrying on and making real my hopes for the GSE program. And there are so many more people in our extended community in ESPM who have been such a wonderful part of my graduate experience. I also want to recognized Dr. Paul Vojta of the Math Department, who is currently maintaining the LaTeXfiles for the ucbthesis class, which made it easy for me to typeset this dissertation. When I emailed him out of nowhere asking a few questions, he responded promptly though I am not in his department!

Finally, I am deeply grateful for the support of my friends and family. My grandfather, who passed away before I had even completed a semester at Berkeley, would be so proud to see me finish my dissertation. He and my grandmother, as well as her parents, and my father, are all proud Berkeley graduates and it has been an honor to continue in their footsteps. My parents, grandparents, and partner, Jon, have been supportive in every way as I have worked my way through courses, oral exams, and research projects. Jon is a tireless copy-editor as well as making sure I eat. Thank you to him and to my parents for everything. *Y

Specific acknowledgments for each chapter follow.

For chapter 1 of my dissertation, I thank my collaborators and coauthors, John Battles, Jonas Knape, Robert York, and Perry de Valpine for their support and feedback on the research and writing. This chapter is modified from a previously published paper, Eitzel et al., in *Ecological Applications*, 23(6), pp. 1288-1296, ©2013 by the Ecological Society of America. I also thank the technicians who have conducted inventories at Blodgett for the last 40 years. Comments from Mary Power, James Clark, Jon Snitow, Andrew Finley, and two anonymous reviewers improved the manuscript. This project was partially funded by the Sierra Nevada Adaptive Management Project (SNAMP). SNAMP is funded by USDA Forest Service Region 5, USDA Forest Service Pacific Southwest Research Station, US Fish and Wildlife Service, California Department of Water Resources, California Department of Fish and Game, California Department of Forestry and Fire Protection, and the Sierra Nevada Conservancy. The published version of this chapter (see references) is contribution number 20 from SNAMP.

For chapter 2 of my dissertation, I thank my collaborators and coauthors, John Battles, Robert York, and Perry de Valpine for their support and feedback on the research and writing. I also thank the crews at Blodgett for decades of inventory work. Comments from Mary Power, Jon Solera, Philip van Mantgem, and Nathan Stephenson improved the manuscript. Advice from Jonas Knape was helpful in early modeling efforts.

For chapter 3 of my dissertation, I thank my collaborators and coauthors, Maggi Kelly and Perry de Valpine for their support and feedback on the research and writing. Thanks also go to Micha Salomon and Chuck Striplen at the San Francisco Estuary Institute for assistance with orthorectifying and georegistering historical imagery, to Madelinn Schriver, Rosemary Sherriff, Lenya Quinn-Davidson, and Yana Valachovic for field insight on the study areas, to the Geospatial Innovation Facility at UC Berkeley for access to and support on using eCognition and Leica Photogrammetry Suite, to Jon Solera for Python scripting assistance, and to Iryna Dronova for advice on multi-threshold segmentation and modeling spatial autocorrelation. Comments from Mary Power, John Battles, and Jon Solera improved this manuscript. University of California Agricultural and Natural Resources funded this work and the California Geological Survey scanned and provided the historical imagery.

Overture

In organizing my dissertation, I have chosen a musical metaphor. This is an apt metaphor because, like many classical musical works, dissertations are long, complex, and have recurring themes. In this case, there are three movements (the three chapters), with intermissions in between, the themes are introduced in an overture, and the piece ends with a new idea in a coda. Some of the themes that emerge throughout this work include:

1. The importance of individual variation and population heterogeneity in forest trees in demographic relationships (Ellner and Rees 2006, Clark et al. 2010, Knape et al. 2011).

2. The uniqueness of drivers of community change and tree population demography for different species or sites.

3. The potential for using existing data sources (e.g. forest inventories, historical aerial imagery) to answer questions about long-term changes in ecosystems.

4. The methodological complexities of combining analytical tools to make best use of these complex and idiosyncratic long-term datasets.

5. The increasing importance of global change studies (Menzel et al. 2014) and the corresponding need for these kinds of ecoinformatic tools to synthesize existing and emerging datasets (Michener and Jones 2012) to inform policy and management.

In my dissertation, I focus on how these themes play out in California woodland and forest ecosystems. I begin with sophisticated statistical models which can appropriately account for individual-based measurements (as in forest inventories), asking in Chapter 1 how such models can address growth limitation in a common species in fire-excluded systems in the Sierra Nevada. In Chapter 2, I then broaden this modeling strategy to address urgent concerns about trends in tree survival for a second-growth forest, and ask whether results differ between individual-species models and aggregated-species models. Finally, in Chapter 3, I investigate the fusion of high-spatial-resolution imagery remote sensing technologies with historical aerial imagery and statistical modeling to address questions concerning the topographic drivers of ecosystem conversion from oak woodland to coniferous forest in the North Coast.

Chapter 1

Estimating Tree Growth Models from Complex Forest Monitoring Data

Understanding the limitations on tree growth is important in many ecological and management applications. Not only is tree growth a basic demographic process that profoundly influences tree population dynamics (Harcombe 1987, Metcalf et al. 2009*a*), but it is also one of the primary means of evaluating forest management goals (Chojnacky 2001). Tree growth rates partly determine tree mortality (Das et al. 2007), and individual-based forest simulators require growth data to parameterize their models (e.g. SORTIE, Pacala et al. 1996). Cambial growth underlies estimates of carbon sequestration (Mohan et al. 2007, Berner et al. 2011) and tree growth as a function of size is an important element in dendrochronological analysis (Bunn 2008). A tree's growth is metabolically limited by its size (Macfarlane and Kobe 2006, Coomes et al. 2011) and competition is a fundamentally limiting factor in closed-canopy conditions (Lines et al. 2010, Kunstler et al. 2011). In addition, the supply of energy, water, and nutrients (typically measured by proxies such as insolation, water deficit, and soil type) can strongly limit tree growth and can mask density dependence (He and Duncan 2000). Permanent forest plot data are an ideal way to learn about growth limitation. Typically these data include tagged trees whose diameter at breast height (DBH, breast height = 1.37 m) is measured at regular intervals. Changes in DBH are often used to measure tree growth (e.g. US Forest Service Forest Inventory and Analysis network[1] and Smithsonian Center for Tropical Forest Science network[2]).

Unfortunately, long-term monitoring data are typically highly complex, including observation error, missing data and uneven time intervals, spatial nesting and autocorrelation, and repeated measures on the same individuals. Most typical statistical models cannot account for all of these issues. More sophisticated hierarchical models can incorporate these shared sources of variation as well as the error inherent in the observations (e.g. Royle and Dorazio 2008, Cressie et al. 2009, Ponciano et al. 2009, Clark et al. 2010). Hierarchical models

[1] http://www.fia.fs.fed.us/

[2] http://www.ctfs.si.edu/

can accommodate nested random effects to account for correlations between measurements from the same site, the same plot within site, or the same individual, as well as incorporating crossed random effects for different years. In addition, one can explicitly model measurement error by treating the unmeasured, true sizes as 'latent states' that are statistically related to measured sizes. Freely available software (e.g. lme4 in R, Bates and Maechler 2010; OpenBUGS, Lunn et al. 2009) and multiple textbooks on hierarchical modeling (Clark 2007, Royle and Dorazio 2008, Kery 2010, Kery and Schaub 2012) place this set of tools in the hands of managers and ecologists, allowing them to make better use of complex long-term datasets.

Previous work using hierarchical models with forest inventories has addressed many spatial and temporal particulars of these data, as well as inevitable observation error. Spatial autocorrelation is a chief issue in permanent plot data, and has been accounted for using correlated spatial random effects in Banerjee and Finley (2007) and Finley (2011). One temporal issue in inventory data is the difficulty of inferring annual growth rates from tree diameter censuses taken at longer and sometimes uneven time intervals. Gregoire et al. (1995) used a continuous-time temporal autocorrelation between measurements to account for these uneven intervals, but comment that their measure of elapsed time as a distance measure may not be meaningful. This temporal mosaic of sampling intervals is common in permanent plot networks as it can be difficult to maintain regular measurement intervals. Even if regular intervals are maintained in any particular network, mismatched time intervals are inevitable when analyses include data from different plot networks. Diameter censuses are easy to execute and typically can include many trees, but are prone to errors due to improper diameter tape placement and bark loss. Therefore, unrealistic negative growth increments are common and many ad hoc methods are used to account for this issue (e.g. by adding an arbitrary amount to all growth increments, which biases estimates towards larger annual growth rates). A way of realistically accounting for observation error is needed (Clark et al. 2007). Previous studies have balanced the strengths and weaknesses of diameter censuses by estimating hierarchical models combining diameter measurements on all trees in some years with tree ring data on some trees in all years in order to infer annual growth (Clark et al. 2007, 2010). Some of these studies include tree size and limiting factors such as canopy cover and climate variables (Clark et al. 2010, 2011, 2012). Other studies have incorporated random effects for spatial nesting with measures of competition and methods of accounting for uneven time intervals (Weiskittel et al. 2007). But none of these existing models have demonstrated how to infer annual growth from a mosaic of sampling intervals by explicitly estimating the unmeasured sizes as well as modeling the repeated measures on individual trees as a random effect.

In this study I developed a hierarchical model to infer annual growth rates from a mosaic of sampling intervals while incorporating multiple sources of unexplained variation. This state-space model (de Valpine 2003) includes: growth as a function of tree size, resource supply, and competition; random effects to account for year-to-year variation, repeated mea-

sures on trees, and spatial nesting; and explicit modeling of unmeasured tree sizes and the associated observation error. The model incorporates the conceptual sophistication of many previous models in such a way that these effects can be estimated simultaneously and compared. While previous studies do include limiting factors, I include interactions between tree size and both fixed and random variables to more completely incorporate size dependence in every aspect of growth, explaining which of these has the biggest impact. In addition, my model handles time by including random year effects as well as latent states. This approach results in crossed random effects between time and spatial nesting factors, which are technically challenging to estimate. Due to the combination of these crossed and nested random effects, unequal time intervals, and observation error, standard statistical analysis software packages will not suffice and I use Markov chain Monte Carlo (MCMC) to estimate parameters in OpenBUGS. I then compare the relative magnitudes of each estimated source of variation and of the effects of different explanatory variables. This case study illustrates the utility of these tools for forest monitoring data, and suggests how these methods can be extended to other such datasets.

1.1 Methods

Study Site

Blodgett Forest Research Station (BFRS) is located in the central western slopes of the Sierra Nevada (38 °52′ N; 120 °40′ W). The research station's 1,780 hectares (ha) are divided into compartments (8-80 ha each) containing plots (0.04 ha each) measured periodically to monitor forest composition and structure. (See Appendix A for number of trees, plots, and compartments in this study as well as sampling intervals.) The data I use in this study comes from the reserve compartments, which have seen no management other than fire suppression in the last 100 years. As such, the reserve compartments at Blodgett are representative of much of the western slopes of the Sierra Nevada, which have seen a similar management history and experience similar abiotic conditions.

BFRS currently consists of mature second-growth mixed conifer forest. Trees have not yet reached their maximum size, and diameter growth appears to be an approximately linear function of diameter. Due to fire exclusion in the reserve compartments, later-successional shade-tolerant coniferous species are most common. White fir (*Abies concolor*) is one of these dominant species and is becoming more dominant over time throughout the Sierra Nevada (Ansley and Battles 1998, Collins et al. 2011). Due to the fact that the forest is still recovering from a clearcut in the early 1900s (Battles et al. 2008), tree diameter is a good representation of tree canopy position. The explanatory variables I include in my model are the following: tree size (DBH) in cm, plot basal area in m^2/ha, insolation in Wh/m^2, topographic slope in %, elevation in m, annual water deficit in mm (see Table 1.1), and categorical soil type (as a proxy for soil nutrients). I detail both the measurement methods and the auditing of these data in Appendix B.

Covariate Name	Mean	Std.Dev	Min	Max	Units	Level
Insolation	628688.10	36106.87	499531.88	692878.06	Wh/m^2	Plot
Slope	15.19	10.09	1.00	43.00	%	Plot
Elevation	1315.73	35.48	1272.54	1450.85	m	Plot
Basal Area	62.02	26.17	0.90	135.48	m^2/ha	Plot/Year
Annual Water Deficit	-174.87	66.63	-311.93	-66.67	mm	Year
Tree Size (DBH)	30.75	20.56	0.25	131.32	cm	Tree/Year

Table 1.1: Explanatory variables; soil is categorical and is not shown here, but most soil types originate from granites, and Cohasset is more developed than the Holland soil family – see Appendix B for details on soils and other explanatory variables.

Statistical Model and Estimation

Given the importance of annual growth in assessing tree performance, and that the data have a mosaic of sampling intervals for different trees, I choose an annual time step for this model. The statistical model for tree growth in each year is hierarchical at several levels. In all the following formulae, subscript i is for compartment, j is for plot, and k is for tree. Superscript m indexes the explanatory variables in Table 1.1. First, I represent the observation process by modeling observations of tree diameter in cm, $y_{ijk}(t)$, as a function of latent (unknown) tree size $x_{ijk}(t)$ in cm with normally distributed observation error with variance σ^2_{DBH}: $y_{ijk}(t) \sim \mathcal{N}(x_{ijk}(t), \sigma^2_{\text{DBH}})$. Next, the process model representing annual tree growth is:

$$\begin{aligned} x_{ijk}(t+1) \sim \mathcal{N}(&\alpha_{ijk}(t) + \beta_{ijk}(t) x_{ijk}(t) \\ &+ \sum_m \gamma^m z^m_{ij}(t) + \sum_m \kappa^m z^m_{ij}(t) x_{ijk}(t), \sigma^2_\varepsilon) \end{aligned} \tag{1.1}$$

I assume that size in the next year is a linear function of several other explanatory variables, which are denoted $z^m_{ij}(t)$ and have parameters for slope γ^m and interaction with size κ^m. The modeling framework does allow for more complex functional forms (e.g. Weiskittel et al. 2007) as needed, but the data do not warrant this complexity. The z^m are centered and scaled based on these variables as measured in the BFRS inventory. The scale of the latent states x is established using the measured inventory sizes y (see Appendix C for details on standardization). The z^m are measured at plot and/or year level: insolation, topographic slope, elevation, and soil category are all measured at plot level, i.e. z^{insol}_{ij}, z^{tslope}_{ij}, z^{elev}_{ij}, and a group of five indicator variables representing a tree's soil type (z^C_{ij}, z^H_{ij}, z^{HB}_{ij}, z^{HM}_{ij}, and z^J_{ij}); basal area is measured at plot and year level, $z^{ba}_{ij}(t)$; and annual water deficit is measured at year level, $z^{def}(t)$ (Table 1.1). I assume that size in the next year $x_{ijk}(t+1)$ is a linear function of size in the previous year, $x_{ijk}(t)$, with soil type-dependent slope and intercept (e.g., for Jocal soil type: slope $a_J = \beta_{ijk}(t) + \kappa^J$ and intercept $b_J = \alpha_{ijk}(t) + \gamma^J$). In Results, I report the average growth increment $\bar{b}$ and average effect of size on growth increment $\bar{a} - 1$,

which are weighted averages over soil types and for average values of explanatory variables. Residual error, with variance σ^2_ε, accounts for additional unexplained variation in growth.

At the next hierarchical level, I model the collective random effects on intercept $\alpha_{ijk}(t)$ and slope $\beta_{ijk}(t)$ with respect to size x as a combination of random effects for tree (q^α_{ijk} and q^β_{ijk}), plot (p^α_{ij} and p^β_{ij}), compartment (c^α_i and c^β_i), and year ($w^\alpha(t)$ and $w^\beta(t)$). The intercept effects reflect differences in overall growth increment while the slope effects reflect differences in growth as a function of size. The random effect intercept for a specific tree is determined by the random tree, compartment, plot, and year effects as follows: $\alpha_{ijk}(t) = q^\alpha_{ijk} + c^\alpha_i + p^\alpha_{ij} + w^\alpha(t)$; and the slope is similar: $\beta_{ijk}(t) = q^\beta_{ijk} + c^\beta_i + p^\beta_{ij} + w^\beta(t)$. The random effects for tree, compartment, plot, and year follow normal distributions, and in Results, I display the standard deviations for each of these random effects (e.g. $\sigma_{\alpha,c}$ for the standard deviation of compartment intercept effects). At each level of nesting, random effects are assumed to be independent (see Appendix C for more details on model specification).

I estimated the parameters, random effects, and latent states in a Bayesian framework using MCMC sampling techniques in OpenBUGS (Lunn et al. 2009). While the MCMC needed some adjustments to produce useful results, ultimately I was able to estimate the full, complex model with all explanatory variables and sources of variation. I used R (R Development Core Team 2009) to format the data and generate initial values for random effects. I could not use completely arbitrary initial values generated by BUGS because I encountered difficulties with slow mixing due to the complexity of the model. Instead I used spline-interpolated sizes (Wood 2006) in a linear mixed effects model (lmer: Bates and Maechler 2010) to generate plausible starting values for random effects and latent sizes. Initial values for intercept and slope parameters for explanatory variables were set to zero.

I used uninformative priors, with the exception of observation error standard deviation. When I included very small values in my prior distributions for this parameter, MCMC chains mixed poorly. As I expect at least a small amount of observation error, I chose an inverse gamma or a uniform prior with a nonzero minimum, based on the minimum rounding error inherent in the diameter tape. In Results, I report estimates from the uniform prior model for all parameters other than observation error (for which I report results for both priors). I assessed convergence both visually and using Gelman-Rubin diagnostics in the coda package (Plummer et al. 2010). See Appendix D for details on priors, initial values, trace plots, and convergence tests. Finally, while I do not demonstrate a formal model selection procedure here, I do estimate several simpler models to check the robustness of my estimates to removal of other model components and to different choices for observation error priors (Appendix G).

1.2 Results

The mean for growth increment $\bar{b}$ was 0.463 (with a 95% credible interval of 0.339, 0.600) cm per year, and for change in growth increment with size $\bar{a} - 1$ was 0.013 (0.006, 0.020) per year. I thus confirm that size has a significant effect on growth and should be included in the model. For all explanatory variable parameters on the same standardized, unitless scale, I list means and 95% credible intervals in Appendix E. The credible intervals of parameters for insolation, elevation, and annual water deficit overlap zero. On the other hand, basal area's effect on growth increment γ^{ba} and its interaction with size κ^{ba} do not significantly overlap zero. The effect of topographic slope on growth increment γ^{tslope} is borderline significant as well, though its interaction with tree size κ^{tslope} is not (Figure 1.1). Note that some of the parameters such as the effect of annual water deficit (γ^{def} and κ^{def}) are poorly determined (have a flatter posterior and consequently a broader credible interval). Though several of the soil type effects also overlap each other, some are distinct from each other and from their average, and all of their intercepts are significantly different from zero. Of the soil types, Cohasset has the highest slope κ^{C} and the second highest intercept γ^{C}, consistent with the typically high productivity of these soils. (See Appendix E for parameter estimates.) The estimates for these significant variables (basal area, topographic slope, and soil effects) are robust to the removal of the others (insolation and elevation), and vice versa (e.g. removal of topographic slope does not render insolation significant). In contrast, when annual deficit is removed, the magnitude of random year effects increases (see Appendix G for comparisons between the full model and several simplified models).

The way these explanatory variables interact with size is more apparent when growth increment as a function of size is shown for low and high values of the explanatory variables (Figure 1.2). These results on the original size scale allow biological interpretation (see Appendix C for algebra underlying the rescaling of covariates after estimation). The slope of all lines is significantly greater than zero, indicating that size is significant for all values of the other explanatory variables. Several of the effects of the high and low values of specific covariates make sense: for higher water deficits, growth is lower, and growth is higher overall in Cohasset soils. However, variables other than basal area show high overlap between high and low values and thus do not substantially affect tree growth. (Soils overlap considerably as well.) Observation error standard deviation σ_{DBH} is estimated at 0.149 cm (0.082, 0.218) with a uniform prior and 0.111 cm (0.073, 0.175) with an inverse gamma prior. Residual standard deviation σ_{ε} is estimated at 0.387 cm (0.358, 0.414). Because standard deviations can never be less than zero, determination of significance is not simple. However for my results a practical choice was to calculate a ratio of the lower credible bound to the width of the credible interval to represent the posterior's separation from zero. The posteriors which were visually distinct from zero had ratios ranging from 0.22 to 3.76, while those which were not separated from zero had ratios on the order of 10^4 or 10^3. I consider those in the first group to be significant and those in the second group not to be significant (see Table E.2 in Appendix E for these ratios). Based on this reasoning, among the intercept random

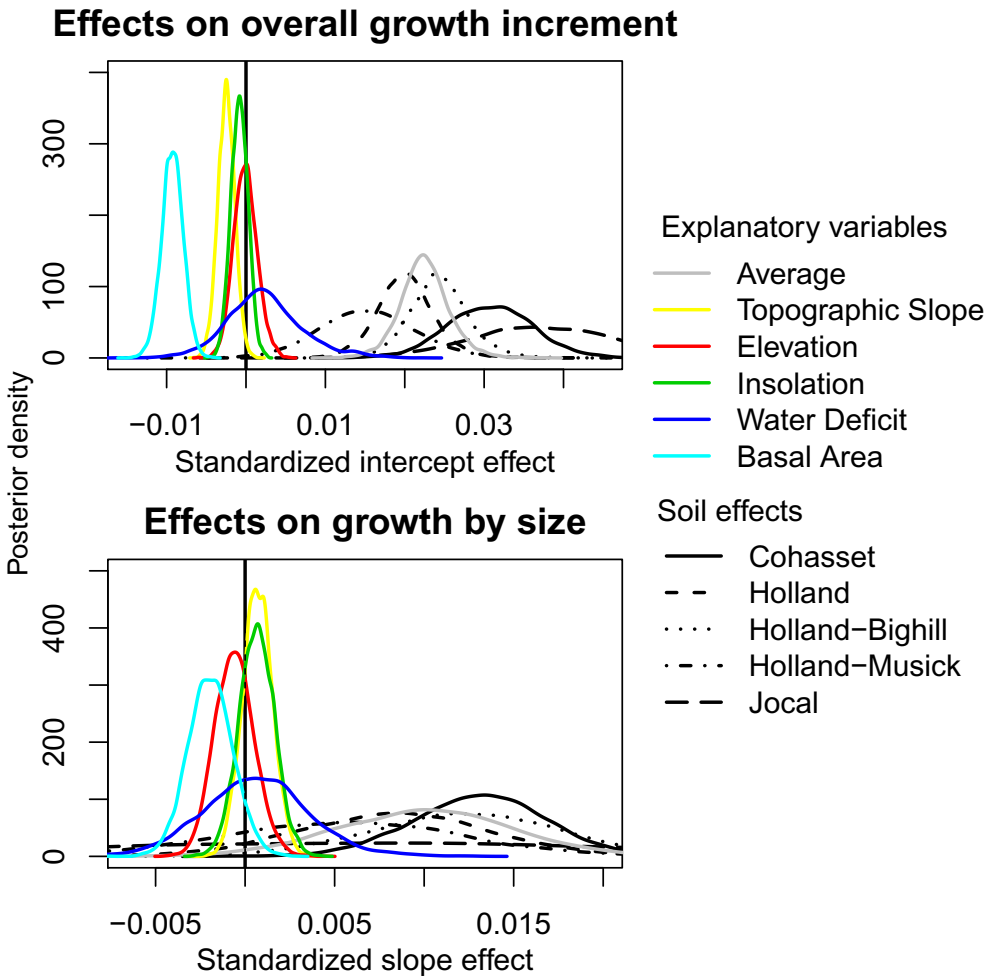

Figure 1.1: Parameter posterior densities for relationships between growth increment with size and other explanatory variables (topographic slope γ^{tslope} and κ^{tslope}, elevation γ^{elev} and κ^{elev}, insolation γ^{insol} and κ^{insol}, annual water deficit γ^{def} and κ^{def}, basal area γ^{ba} and κ^{ba}), all of which are for the standardized versions of these variables. In addition to these continuous explanatory variables, the effects due to different categorical soil types (γ^{C}, γ^{H}, γ^{HB}, γ^{HM}, γ^{J} and κ^{C}, κ^{H}, κ^{HB}, κ^{HM}, κ^{J}) are shown in black. The gray line indicates the average growth increment (upper figure, $\bar{b}$) and effect of tree size on growth increment (lower figure, $\bar{a} - 1$). (Upper) Effects of continuous explanatory variables and soil types on the overall growth increment ($\bar{b}$ and γs). (Lower) Effects of explanatory variables on the slope of future size with respect to current size ($\bar{a} - 1$ and κs).

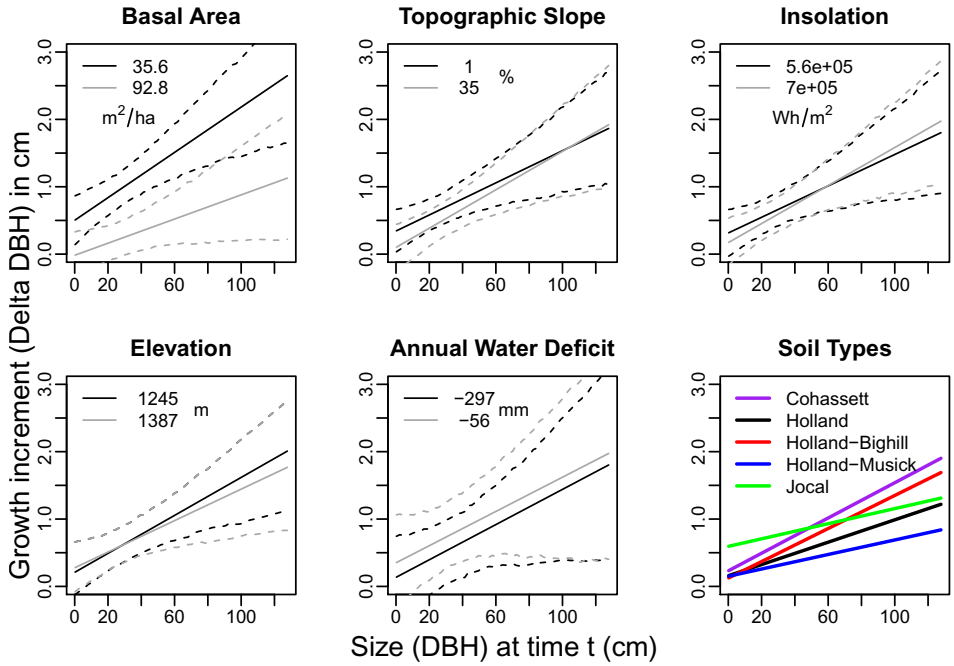

Figure 1.2: Explanatory variable effects on growth increment as a function of size, rescaled to cm. All plots show that growth increases with tree size in the previous year. Solid lines are the means from posteriors of parameter estimates; dashed lines are 95 % credible intervals. For all explanatory variables other than soil type, black shows growth increment for a low value of the explanatory variable (-2 standard deviations) and dark gray for a high value (+2 standard deviations). Credible intervals for soil type overlap a great deal and are not shown for clarity.

effect standard deviations, the year, tree and plot effects are significant ($\sigma_{\alpha,w}$, $\sigma_{\alpha,q}$, and $\sigma_{\alpha,p}$, Figure 1.3). Of the slope random effect standard deviations, only plot ($\sigma_{\beta,p}$) is significant (see Appendix E and Figure 1.3). Though several of the sources of unexplained variation have a significant effect on the overall growth increment, the variation they introduce is substantially less than the variation introduced by overall residual variation. Several of the random effect standard deviations are not well determined: year and compartment intercept effects have particularly broad, flat posteriors. Finally, I also examined explicitly-estimated latent tree sizes $x_{ijk}(t)$ for nine trees in the inventory. Some trees' unmeasured sizes are better constrained by the data than others (small trees in particular are poorly estimated, with wide credible bands); but for several larger trees the annual growth between censuses has narrow credible bands and reasonable values (Figure F.1, Appendix F).

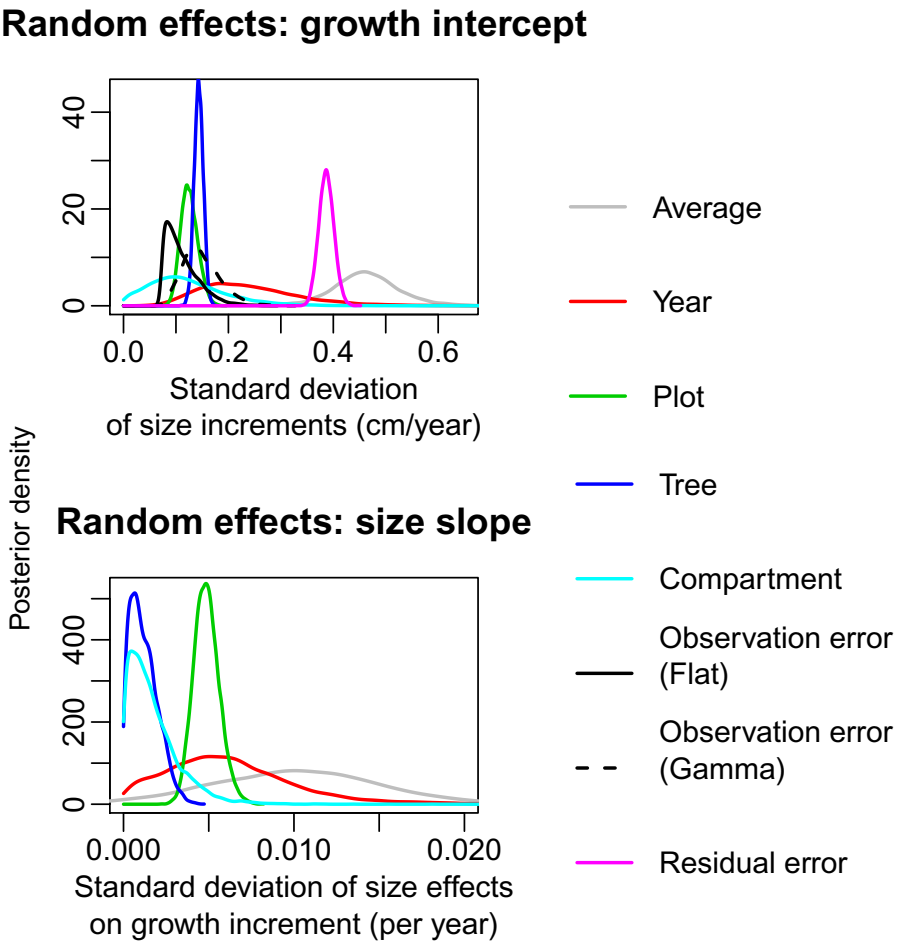

Figure 1.3: Random effect standard deviation posteriors. (Upper) The intercept standard deviations (compartment $\sigma_{\alpha,c}$, year $\sigma_{\alpha,w}$, tree $\sigma_{\alpha,q}$, and plot $\sigma_{\alpha,p}$) on the cm scale, with the average growth increment $\bar{b}$ from Figure 1.1 shown for scale in grey. Observation error standard deviation σ_{DBH} and residual error standard deviation σ_ε are shown on the same plot for comparison. Only for observation error, I show posteriors for both the inverse gamma (dashed) and uniform (solid) prior models. (Lower) The slope standard deviations ($\sigma_{\beta,c}$, $\sigma_{\beta,w}$, $\sigma_{\beta,q}$, and $\sigma_{\beta,p}$); again, the growth increment slope $\bar{a} - 1$ from Figure 1.1 is shown for scale in grey.

1.3 Discussion

This study demonstrates the feasibility of estimating a hierarchical model from forest inventory data with the full richness of both categorical and continuous explanatory variables, many sources of variation, and observation error. I have successfully estimated annual growth

from a mosaic of sampling intervals. Backward model selection strategies require one to start with the most complex model available, and I have shown that this most complex model can be estimated. Forest inventory datasets with this kind of sampling structure and with individually marked trees are becoming more common, so this modeling approach could be applied to many other forest dynamics problems.

Tree Size, Basal Area, and Resource Supply

These models confirm that dependency of tree growth on size and competition intensity cannot be ignored when modeling growth, e.g. in other applications such as population dynamics or dendrochronology. The estimated annual average growth increment of 0.463 cm per year at average conditions on these plots is high but reasonable, as BFRS is a productive site and its second-growth forests are still increasing in biomass. I chose a linear model for diameter growth for this study, but the modeling framework easily accommodates other functional forms for dependency on size and competition. (See Cao (2000), Nord-Larsen (2006), Weiskittel et al. (2007), Cao and Strub (2008), who have incorporated complex functional forms with uneven inventory time intervals.) In particular, one could explore nonlinear functions that allow tree growth to slow as trees reach very large sizes (as in Clark et al. 2007, 2010, 2011, 2012), although forests at BFRS are young enough not to show this behavior.

Basal area's dominance among the remaining explanatory variables has two possible explanations. First, competition is a likely limiting factor in forests like BFRS. As in many Sierra Nevada forests, fire suppression has allowed shade-tolerant species to dominate younger cohorts and regeneration. Tree density is high and recruitment to the canopy is only possible in distinct canopy gaps. Second, although elevation and insolation appear not to be important at plot level, these variables occur in a narrow range at BFRS relative to the species' fundamental niche (Lutz et al. 2010). Variation in these variables at BFRS is small relative to the US Forest Service's permanent plot data in a similar latitude range (see Appendix H for details of comparison to Forest Inventory and Analysis data), and other studies show larger responses when larger ranges of covariates are available (Clark et al. 2012). The fact that annual water deficit did not have a strong effect on growth is surprising given the documented relationship between growth and climate for white fir trees in northern California (Yeh and Wensel 2000). One explanation for this result is that water deficits at BFRS tend to be lower than comparable forests in the region (see Appendix H). Also recall that when water deficit is removed from the model, the intercept year standard deviation increases (see Appendix G). So although the effect of annual water deficit cannot be precisely estimated, its magnitude may be biologically important. Finally, this dataset does not include many trees smaller than 11.4 cm DBH, and these smallest individuals may be more strongly influenced by resource supply than larger, more established individuals, and also at a spatial scale below plot level.

Sources of Variation

The behavior of the model implies that observation error standard deviation may not be independently estimable in this kind of model. Priors allowing the standard deviation to approach zero, including uniform and half-t priors (Gelman 2006, Knape et al. 2008), result in bad mixing as the system struggles to move away from a scenario in which observations perfectly match latent states. This behavior could be due to the MCMC samplers in OpenBUGS; different software may not have this problem. Some observation error in this system is expected, however, and including it should improve other estimates. I chose a minimal lower bound on the observation error standard deviation, based on the rounding error in the measurement tape (0.073 cm; see Appendix D for details of this calculation). My estimates are larger than this minimum, but still much smaller than the 1 cm (0.8 %) reported by Clark et al. in a study combining tree cores with DBH measurements (2007) and the 2.7 % reported by Gonzalez et al. (2010) for repeated DBH measurements made in similar forests (though this latter study included outliers and obvious errors, which I have removed from the inventory; see Appendix B). Models like these for similar datasets should check to confirm that observation error can be estimated and may likewise consider an informative prior to ensure a minimum amount of observation error. Analysts may also try different MCMC samplers, which may be less sensitive to this behavior.

Even after accounting for basal area, the significant plot random effect standard deviations essentially give each plot its own size relationship with growth. Plots at BFRS are approximately 2-3 neighborhoods in size (Canham et al. 2006, Das et al. 2011), so neighborhood effects could be acting below the plot level. Density dependence in closed canopy forests can be complex, so an effect due to plot in addition to an overall basal area measurement at the plot level is reasonable. The importance of slope random effects in addition to intercept random effects implies that models may need to include variation in more than just the overall growth intercept. The broad, flat posteriors of some random effect standard deviations may be the result of a small sample size to estimate them; in particular, the broad posterior for the year effect variance is not surprising as some years do not have data and the annual increment cannot always be precisely estimated.

The nontrivial tree intercept standard deviation, even when size is included in the model, implies that some trees have a growth advantage over their entire lifetime in the inventory. This slightly different average growth of different trees is ecologically important: for example, in population models, individual quality (Ellner and Rees 2006) can have important impacts on population growth. Unfortunately it is impossible with these data to determine whether this individual quality reflects genetic superiority, a favorable microsite when the tree first established, or some other neighborhood factor. The significance of this individual variation highlights the importance of studying forests at the individual tree level (Clark et al. 2012).

1.4 Future Work and Implications for Practice

I have demonstrated the use of a complex hierarchical model including explanatory variables and multiple random effects on long-term forest inventory data. Though this model accounts for uneven time intervals effectively, it does not include the sophistication of explicitly including underlying spatial autocorrelation (Banerjee and Finley 2007, Finley 2011). In addition, though I have included a variety of explanatory variables, there is a constellation of possible variables that could have been included. Ultimately, one would want to use both knowledge of the system and a model selection procedure to determine what functional form to use and which explanatory variables and random effects to keep in the model, which I do not demonstrate here.

The results of these models can be directly used to parameterize population models (e.g. integral projection population models, Metcalf et al. 2009*a*, Ghosh et al. 2012) or forest simulators (e.g. SORTIE). Comparative studies using rich long-term datasets can illuminate patterns in ecological processes over large geographical areas. These studies may test basic ecological theories such as resource limitation and niche theory, or applied questions about management practices (e.g. how the effectiveness of fuel hazard reduction treatments could be affected by climate change). Data on individually tagged trees, rather than plot-level data, are becoming the norm, and long-term monitoring data are maturing. Forest permanent plots are common in Long Term Ecological Research sites (over a third of the International Long Term Ecological Research sites are listed as "forest").[3] Since 1999 when the US Forest Service Forest Inventory and Analysis program moved to annualized inventory, their program includes remeasurement of tagged trees, and numerous other long-term datasets are available from a variety of sources (e.g. the Smithsonian Center for Tropical Forest Studies).

As policymakers, managers, and ecologists alike call for adaptive management strategies,[4] including long-term monitoring as a key component in assessing ecosystem interventions, sophisticated modeling is needed to appropriately analyze these data so that inference can feed back into management planning and therefore complete the adaptive management cycle. Monitoring data are often rich in covariates and highly complex in sources of variation, and uneven time intervals are a common problem. This study, including the appendices which detail the modeling process and the supplements which include the code and data to run the model,[5] should help more ecologists and managers to try these types of sophisticated models and open up new ways of using inventory data from across the world.

[3] http://www.ilternet.edu/

[4] http://www.fs.usda.gov/detail/planningrule/home?cid=stelprdb5359471

[5] Available at Ecological Archives: http://www.esapubs.org/archive/appl/A023/067/

Intermission 1

Tree growth and survival are tightly linked (Das et al. 2007, Metcalf et al. 2009*b*). I now move from a thorough exploration of tree growth using a sophisticated statistical method for complex data to the application of the method to an important ecological question: trends in and drivers of forest tree mortality. I make use of the methods in Chapter 1 for fitting hierarchical state-space models using Bayesian MCMC in OpenBUGS, with the knowledge of where these methods can go wrong (e.g. ensuring that a prior is non-informative when the variable is transformed through a nonlinear function), and also investigate a method for model selection where more traditional methods are questionable (Millar 2009). I also broaden the question from a single species as a proof-of-concept to a study of survival and mortality in multiple species at the same site with different life histories, seeking to understand potential drivers of tree survival and mortality. I investigate the potential for time trends both at the individual-species level and at the species-aggregated (community) level.

Chapter 2

Can't see the trees for the forest: complex factors influence survival in a temperate second-growth forest

Concern about forest decline is high given extensive forest insect outbreaks, rising levels of pollution, and recent increases in global temperature and greater variability in precipitation (Likens and Franklin 2009, FAO 2010, Zeppel et al. 2013). This concern is motivated by the important suite of ecosystem services provided by forest trees as foundation species (Ellison et al. 2005). A number of studies and reviews have highlighted regional (van Mantgem et al. 2009, Peng et al. 2011, Luo and Chen 2013) and global (Allen et al. 2010, Wang et al. 2012) decreases in tree survival, particularly in large trees (Dolanc et al. 2013), and some of these studies implicate climate change. In the arid West of the United States, studies of declining tree survival focus on old-growth forests, where endogenous changes in forest structure and species composition are minimal and assumed not to influence demographic trends (van Mantgem and Stephenson 2007, van Mantgem et al. 2009; but see Lutz et al. 2014). While old-growth forests are particularly important because they hold great ecological and social significance, they represent only 36% of the global forest (FAO 2010) and are particularly rare in temperate regions. For example, only 6% of forests in the United States are classified as primary (FAO 2010). Thus most forests are simultaneously developing following a significant disturbance as well as potentially responding to perturbations in the global environment. Under these circumstances, it is essential to disentangle the impact of non-stationary environmental changes (e.g. climate change) from demographic trends due to stand development (Luo and Chen 2013). To date, the few studies that have incorporated stand dynamics into mortality assessments for second-growth forests have focused on boreal forests following wildfire (Thorpe and Daniels 2012, Luo and Chen 2013). To better understand temperate tree survival in the face of increasing threats, I investigate a productive, second-growth mixed conifer forest in California.

To further understand the complex factors influencing tree survival in more diverse sys-

tems, I also investigate both species-aggregated models and species-specific models. Studies of forest decline typically aggregate species together (van Mantgem and Stephenson 2007, van Mantgem et al. 2009) or fit models for systems with one (Hurst et al. 2011, Uzoh and Mori 2012) or only a few individual species (Thorpe and Daniels 2012, Luo and Chen 2013). Decline of forest trees at the community level (e.g. species-aggregated) is important for evaluating the risk to fundamental forest ecosystem services, and many carbon sequestration estimates are indiscriminate of species. However, species-specific models of tree survival reveal ecological principles at play, highlight dangers to particular species of cultural and commercial interest, and improve parameterization of individual-based forest simulators and gap models (Bigler and Bugmann 2004, Larocque et al. 2011). Trends in survival at a coarser, community-level scale could mask trends at a finer, species-specific scale – but trends at both scales could be ecologically, culturally, and economically important.

Finally, intraspecific heterogeneity in survival may be important. Survival processes are spatially complex (Das et al. 2011), and small trees (Igor et al. 2009) or trees in crowded stands (Das et al. 2007) may have a reduced chance of survival relative to their larger counterparts in less crowded stands. The concept of a single annual survival probability becomes less meaningful for a heterogeneous population: a high 'mean' annual survival may miss susceptible subsets of the population with substantially lower survival. The importance of susceptible individuals can be important in applications of survival models, e.g. population models where individual quality can have a strong impact on population-level results (Ellner and Rees 2006). As Clark et al. (2012) point out, survival happens at the scale of individuals, but our concerns address the species and landscape scale. In this study I pay special attention to the drivers of mortality for susceptible subsets of the population.

Specifically, I ask the following questions regarding tree survival in a temperate second-growth forest: 1) Is there evidence for a secular trend in survival even when accounting for the confounding trend in stand development, and does climate change account for such a trend? 2) How strongly do species-specific results differ from species-aggregated results, and how much do species differ from each other in their responses to explanatory variables? 3) Within a species, how heterogeneous is survival and what are the drivers of that heterogeneity?

Forest inventory data can be used to address these questions. These data can introduce their own complexities: for example, census intervals are often greater than one year and are also variable, causing uncertainty in identifying when a tree died during the interval (Nord-Larsen 2006). This leads to a fundamental difficulty in modeling survival due to annual variables (Thorpe and Daniels 2012, Luo and Chen 2013), including climate or stand development. Hierarchical state-space models address these challenges by explicitly modeling the latent unmeasured survival status of the tree in each year (Clark et al. 2012). If the uneven census intervals provide some information in most years, this model structure may be able to borrow across data-poor years in order to investigate the possibility of time trends in survival while incorporating yearly changes in stand characteristics.

This chapter examines the annual survival probabilities for seven tree species in a second-growth temperate forest in the Sierra Nevada of California. Recovering from a clearcut in the early 1900s, this forest in the arid west of the United States is relatively healthy and has yet to be seriously impacted by obvious causes of mortality (e.g. extensive bark beetle outbreaks, Stark et al. 2013). Under a century of consistent fire suppression, canopy species and recruitment are dominated by shade-tolerant species, with shade-intolerant species remaining as a minor element in the stands (Ansley and Battles 1998, Collins et al. 2011). A state-space model can incorporate stand structure, tree size, climate, and resource availability proxies, as well as spatial heterogeneity among plots. This model is able to appropriately represent the complexity of forest tree survival and address my proposed questions.

2.1 Methods

Study Site, Species, and Variables

Data for this study come from Blodgett Forest Research Station (BFRS), located on the western slope of the Sierra Nevada at 38 °52′ N; 120 °40′ W. BFRS is a second-growth, mixed conifer, closed canopy forest in a Mediterranean climate. Data for this study are from "reserve" (control) compartments ranging in size from 11 to 45 ha. Since widespread logging 100 years ago, no management other than fire exclusion has been applied. This is a common management history for much of Sierran forests. All compartments contain 0.04 ha plots, and each plot is surveyed periodically but not always in the same year, creating a mosaic of measurement intervals ranging from 3-14 years. Altogether, inventories account for the time period from 1976 – 2010. Crews measure the diameter at breast height (DBH, breast height = 1.37 m) of living trees greater than 11.4 cm DBH and record information on standing dead trees. Smaller trees (greater than 0.254 cm DBH) are tracked in subplots.

This study includes all of the native canopy tree species. Black oak (*Quercus kelloggii*) and tanoak (*Notholithocarpus densiflorus*) are both fire-adapted hardwoods capable of vigorous vegetative resprouting, though black oak is extremely shade-intolerant while tanoak is considered shade tolerant (Niinemets and Valladares 2006). Ponderosa pine (*Pinus ponderosa*) and sugar pine (*Pinus lambertiana*) also rely on disturbance: ponderosa pine is capable of surviving low-intensity fires due to thicker bark, and sugar pine recruits well in forest gaps (Burns and Honkala 1990). As the fire regime has shifted from frequent, low-intensity fires set by native peoples to fire suppression, these species have become less abundant at BFRS. White fir (*Abies concolor*) and incense-cedar (*Calocedrus decurrens*), on the other hand, are shade-tolerant and have become the main canopy constituents at BFRS (Niinemets and Valladares 2006). Douglas-fir (*Pseudotsuga menziesii*) has intermediate shade-tolerance and also benefits from more frequent fire (Burns and Honkala 1990), but still remains an important component at BFRS. By stem count, the most abundant species are incense-cedar, Douglas-fir, white fir, and ponderosa pine. These species together

comprise 88.9% of the stems in the control plots. Less abundant are sugar pine, black oak, and tanoak; these species account for another 9.5% of stems (Appendix I).

In order to assess the relative impact of different biotic and abiotic factors on tree survival, these models include variables representing competition, tree size, climate, and topographic proxies for nutrient, light, and water availability.

The models include two biotic variables: tree size (DBH in cm) and plot basal area (m^2/ha). Plot basal area is calculated using the cross-sectional area of trees in the plot at breast height, divided by the area of the plot. The second-growth forest at BFRS has undergone heavy density-dependent mortality (Battles et al. 2008), with plot basal area increasing over time through the inventory period (Appendix I). When a tree was identified as dead in a census, no size was recorded in the database. There are several ways to estimate a missing size, but model results proved not to be sensitive to the different methods; ultimately a simple linear model based on tree diameter and census interval was adquate to estimate a final size (see Appendix J). Each tree's size and plot's basal area were interpolated over time using smoothing splines (Wood 2006). See Olson and Helms (1996) for more information on inventory procedures.

The abiotic variables represent climate and topography. The model includes three topographic variables (measured at the plot level) as proxies for resource availability (i.e. light, nutrients, and water): insolation, topographic slope, and elevation. Annual insolation (Wh/m^2) is calculated from a 1/3 arc second digital elevation model from the USGS Seamless Map Database (USGS 2011), using the solar radiation calculator in ArcGIS 9.3 (ESRI 2011). Topographic slope (in %) is measured in the field using a clinometer, and elevation (m) is measured from a topographic map to the nearest 3.05 m. The model also includes annual climatic water deficit, as well as year (to account for a time trend in survival). Annual water deficit is calculated in mm from weather records at BFRS, using AET 1.0 (Gavin 2007) and a modified Thornthwaite method; note that deficit is a negative number and more negative numbers indicate more stressful conditions. Deficit incorporates precipitation and temperature variables and has been correlated with tree survival in Sierran forests (van Mantgem and Stephenson 2007). Because trees are measured in summer, the climatic water deficit from the previous year reflects the drought stress from the previous summer and is most likely to affect tree survival in the current year (McDowell et al. 2008). Water deficit at BFRS has a slight but non-significant increase (more negative: -1.06 $mm/year$, p=0.29).

There are many possible climate variables, as well as many possible stand development variables and individual tree variables (Monserud and Sterba 1999). To make modeling tractable, I have chosen representative variables to account for the range of factors which have been shown to relate to tree survival: tree size reflects canopy position at BFRS (Eitzel et al. 2013), plot basal area is a good representation of crowding (Lines et al. 2010), and many studies in Mediterranean forests use annual climatic water deficit to represent climate effects (van Mantgem and Stephenson 2007, Thorne et al. 2012). See Table 1.1 and Appendix I for

summary statistics on and plots of these explanatory variables and Appendix B for details and more information on data auditing.

State-Space Model

The state-space formulation explicitly models the unknown yearly status of individual trees; because tree status is a discrete state, this model falls in the category of hidden Markov models. This model structure represents ignorance of when the tree died between inventories. In this sense, this approach is similar to a mark-recapture model as seen in wildlife applications (Gimenez et al. 2007), but with perfect detection. As the plots are not mapped, spatial heterogeneity in the forest is modeled using a random plot effect.

The alive/dead status $z_{ij}(t+1)$ for tree i in plot j is represented by a Bernoulli random variable (1 is 'alive'), conditional on status at time t, in the following way: $z_{ij}(t+1) \sim Bern(\phi_{ij}(t)z_{ij}(t))$ where $\phi_{ij}(t)$ is the probability of survival from year t to year $t+1$. Multiplying by the status at the previous time ensures that dead trees stay dead. The full model for the biotic and abiotic factors influencing survival probability is

$$\begin{aligned} logit(\phi_{ij}(t)) &= \beta_{0_j} + \sum_k \beta^k x_{ij}^k(t) \\ \beta_{0_j} &= b + p_j \end{aligned} \tag{2.1}$$

where b is an overall mean survival for an average tree, and p_j is a random effect for plot j with $p_j \sim N(0, \sigma_p^2)$. k indexes one of the explanatory variables: insolation (x_j^{insol}), topographic slope (x_j^{slope}), and elevation (x_j^{elev}) are all measured at plot level; plot basal area ($x_j^{ba}(t)$) is measured at plot and year level; tree size ($x_{ij}^{DBH}(t)$) is measured at the tree and year level; and annual water deficit is measured at year level ($x^{def}(t-1)$). The β^k are coefficients for each of these explanatory variables. In order to test for an unexplained linear time trend (often referred to as a 'secular trend'), β^t, the model also includes the measurement year $x^t = t$. To allow a more flexible survival relationship with size, both a linear (β^{DBH}) and quadratic (β^{DBH^2}) term for tree diameter are included. See Appendix K for full model formulae and a diagram representing the relationship of model components to each other. I validated a simple version of this model (with only an intercept, b) against mortality ($m = 1 - survival$) calculations (Sheil et al. 1995, Appendix L).

I have standardized (centered and scaled) the explanatory variables; centering improves mixing and clarifies interpretation, while scaling allows the relative impact of variables within a species to be assessed. When comparing between species, I have returned the parameter estimates to their original scale but left them centered. See Appendix K for algebra regarding standardizing and rescaling these parameter estimates.

To fit the state-space models in a Bayesian framework, I used Markov chain Monte Carlo (MCMC) estimation techniques in OpenBUGS (Lunn et al. 2009), run through R2WinBUGS

(Sturtz et al. 2005) in R (R Development Core Team 2009), and "`coda`" (Plummer et al. 2010) to calculate credible intervals. I chose uninformative priors: the random effect standard deviation σ_p was uniform from zero to 100, and the priors for the β^ks were normal with precision 10^{-6}. For survival of an average tree $expit(b)$, care in parameterization of priors was required ('expit' is the inverse logit). Due to the nonlinearity of the logit, a prior which is flat on the scale of the linear predictors (e.g. b) is U-shaped on the probability scale ($expit(b)$). Because tree survival is high, the estimated mean survival can be quite close to one, where the nonlinearity has the strongest effect (Van Dongen 2006). Therefore, as in Buoro et al. (2012), I used a uniform distribution from zero to one for $expit(b)$ rather than a normal distribution for b.

Model Selection

Due to the lengthy run times for each model and the number of candidate variables, it was not feasible to fit all possible models, requiring a stepwise approach. Backward stepwise selection from full models was also not feasible due to poor mixing and computational limitations. Therefore, I took a forward stepwise selection approach to navigating possible models. Although forward selection has limitations in some situations, it has been shown to have similar predictive accuracy to other model selection strategies (Murtaugh 2009).

Choosing a criterion to rank candidate variables to add to the model in my selection process is not straightforward in a Bayesian framework, and the common choice of the Deviance Information Criterion (DIC) is controversial for models like these with many latent states and random effects (Celeux et al. 2006, Millar 2009, Kery and Schaub 2012, Appendix M). Therefore I referred directly to posterior densities for a criterion to use in selecting the variable x^k: at each step, I add the variable whose β^k posterior is most separated from zero. I represent 'separation from zero' using a tail probability analogous to a two-tailed p-value. This metric is justified because it will be similar to a p-value, which, when comparing single variable additions, will be ranked in the same order as AIC differences (Murtaugh 2014, Appendix M). Tail probability ties are broken by choosing the variable with the larger mean effect size.

I proceeded using this tail probability criterion by fitting separate models with each individual variable by itself, choosing the variable with the smallest tail probability, and then fitting another round of models with that variable and each of the others. To determine when to stop adding variables, I used a traditional 5% threshold, but typically also went one step beyond that and checked one additional variable (and the result usually agreed with the 16% threshold more typical of AIC reasoning in these cases, de Valpine 2014). I emphasize that I am not engaged in significance testing, especially conditioned on model selection, but rather that I am approximating the AIC ordering from Bayesian posteriors for single-variable additions. I also compared the results with DIC for the models where DIC should be least problematic, and found that the final model in that case did have the lowest DIC among

all the models I calculated as part of the forward selection (Appendix M). To select the plot random effect standard deviation, I used slightly different criteria and model selection strategies (Appendix N).

I fitted models for incense-cedar, Douglas-fir, white fir, ponderosa pine, sugar pine, black oak, and tanoak, as well as for all species aggregated together. I referred to intermediate models in the model selection procedure to investigate confounding between secular trends and stand development (Question 1). I used the results of the model selection procedure to compare between species-aggregated results and individual-species results, and to compare among individual-species results (Question 2). In order to address heterogeneity within each species (Question 3), I transformed models back to the probability scale to examine the range of survival probabilities for trees under different conditions. I also compared a simple model with only the survival probability for an average tree $expit(b)$ to the more complex models created during model selection (which include random plot effects and other variables). (See Appendix L.) However, differences between $expit(b)$ in a simple state-space model and the same parameter in the complex models may be partly attributable to the nonlinearity of the logit and Jensen's inequality. I therefore weight comparisons from within the complex models more heavily in my discussion than comparisons between simple models and complex models.

2.2 Results

Question 1: Secular Time Trends and Stand Development

For the species-aggregated result, including plot basal area in the model renders the time trend non-significant: the posterior heavily overlaps zero (Figure 2.1a), and the model selection procedure does not select this variable once plot basal area is included. Ponderosa pine does not initially show a time trend, but when tree size and topographic slope are included in the model, the time trend emerges and is selected for the best model for this species (Figure 2.1b). Adding basal area to the best model for ponderosa pine renders the trend not significant, possibly due to collinearity. In the case of incense-cedar, the time trend becomes less significant with the addition of size and then barely not significant when basal area is added (i.e. using the AIC-like threshold, the time trend would be retained, Figure 2.1c). For sugar pine (Figure 2.1d), the time trend is significant when initially added during the model selection procedure, and remains separated from zero when tree size and plot basal area are included in later models.

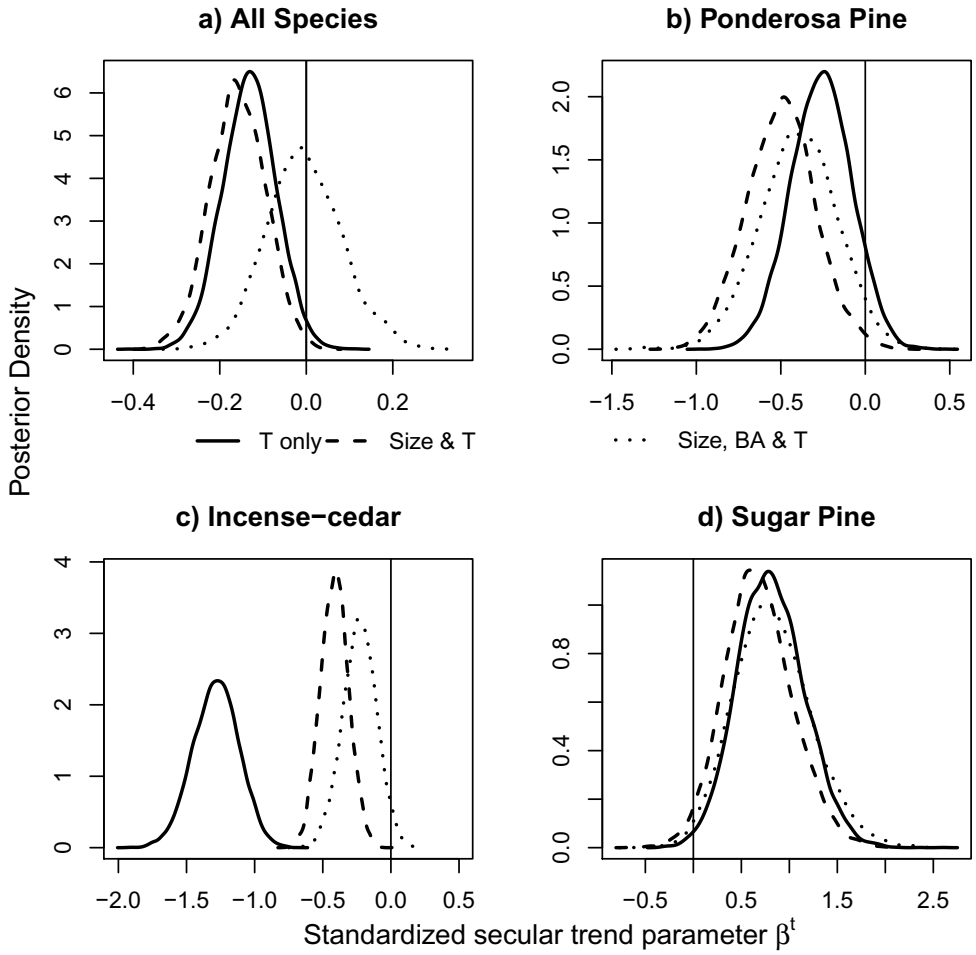

Figure 2.1: Parameter posteriors for the time trend in different models for all species aggregated together (a), ponderosa pine (b), incense-cedar (c), and sugar pine (d). Solid lines indicate models with only a time trend, while dashed lines represent models with tree size (including the quadratic term if selected for that species) along with the time trend, and dotted lines indicate models with size, basal area, and the time trend. For ponderosa pine, the topographic slope is also included in the latter two models as it was consistently important in the model selection.

Question 2: Species-Aggregated Results Versus Species-Specific Results, and Differences Between Species-Specific Results

For the most part, species-aggregated results do not mirror species-specific results (Table 2.1). The only variable which is consistent between species-aggregated models and individual-species models is the increasing linear size effect. Species-aggregated models do pick up on the importance of the quadratic size effect and the effect of plot basal area, but these effects are not consistent across all individual-species models. The species-aggregated models entirely miss the importance of other variables to individual species (e.g. elevation for tanoak). In some cases, species-aggregated parameter estimates appear to be an average of species-

	β^{DBH}	β^{DBH^2}	β^{ba}	β^{insol}	β^{elev}	β^{slope}	β^{t}	β^{def}
All Species	⇑	⇓	⇓					
Incense-cedar	⇑		⇓					
White Fir	⇑	⇑		⇑				
Douglas-fir	⇑							
Ponderosa Pine	⇑	⇓				⇑	⇓	
Sugar Pine	⇑						⇑	
Black Oak	⇑	⇓						
Tanoak	⇑				⇓	⇓		

Table 2.1: Summary of model selection results for each parameter and species. An arrow indicates the sign of the coefficient β^k for variable k (⇑ - positive coefficient; ⇓ - negative coefficient).

specific results (Figure 2.2a and 2.2c), but even here the species-aggregated result misses important details, e.g. incense-cedar's plot basal area effect is very different from the species-aggregated results (Figure 2.2b). All species individually show increased survival for larger trees, and none of them shows an effect of annual water deficit (Table 2.1, Figure 2.2). For all other variables, however, each species has a different response (Table 2.1, Figure 2.3, Appendix O). For several variables, only one species responds (i.e. insolation, elevation, plot basal area), and for other variables, species have opposite responses (i.e. quadratic size, topographic slope, secular time trend).

Question 3: Heterogeneity Within a Species

A single annual survival probability does not represent the heterogeneous effects of tree size and other variables on survival (Figure 2.3). For example, in ponderosa pine, the single estimate of survival probability $expit(b)$ (98.4% from a model without plot and explanatory variables; 99.3% for the final model; see Table L.2 in Appendix L) does not reflect the range of survival shown in Figure 2.3: from 65.4% for small trees (1.52 cm DBH) on shallow slopes

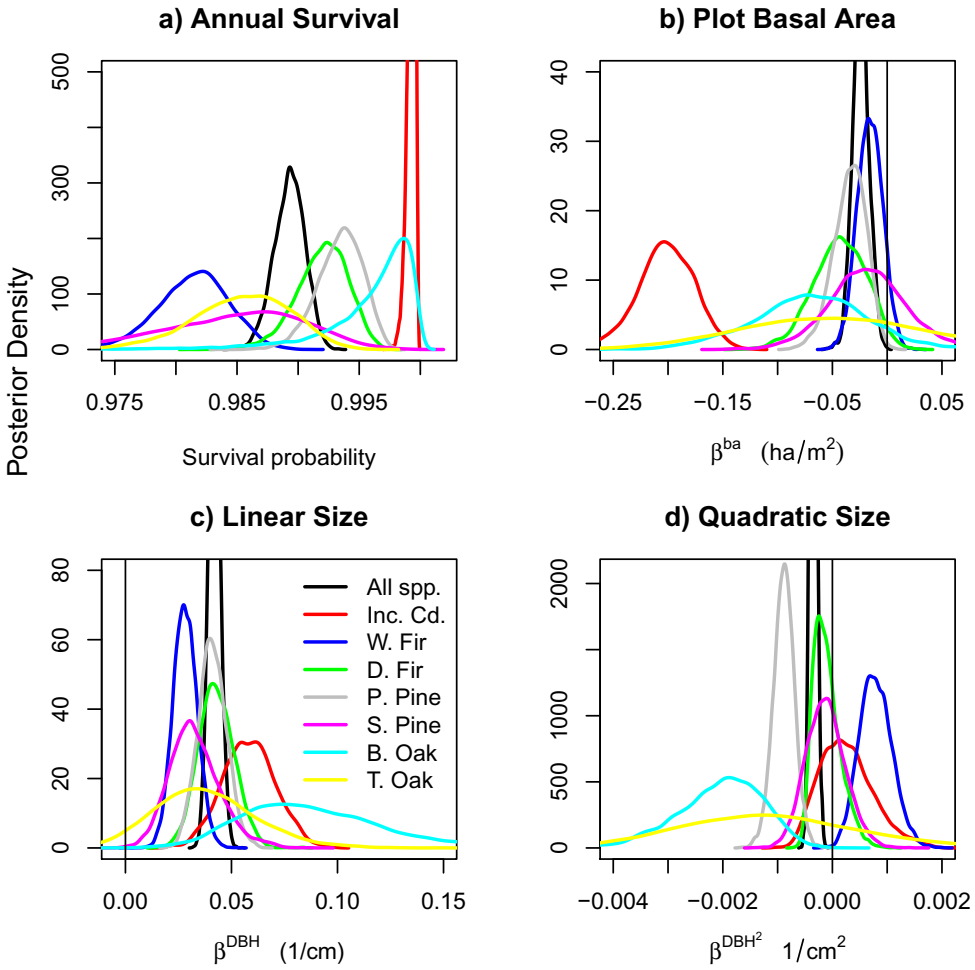

Figure 2.2: Rescaled parameter posteriors from similar models for all species. These models include a random plot effect, size (linear and quadratic), and basal area. For tanoak, basal area and size posteriors are from separate models (tanoak's $expit(b)$ in Figure 2.2a is from the basal area model). Posteriors have been rescaled in order to compare estimates between species. (a) Survival of an average tree in an average plot ($expit(b)$) for each species. (b)-(d) Estimates of the effects of basal area (b), linear size term (c), and quadratic size term (d).

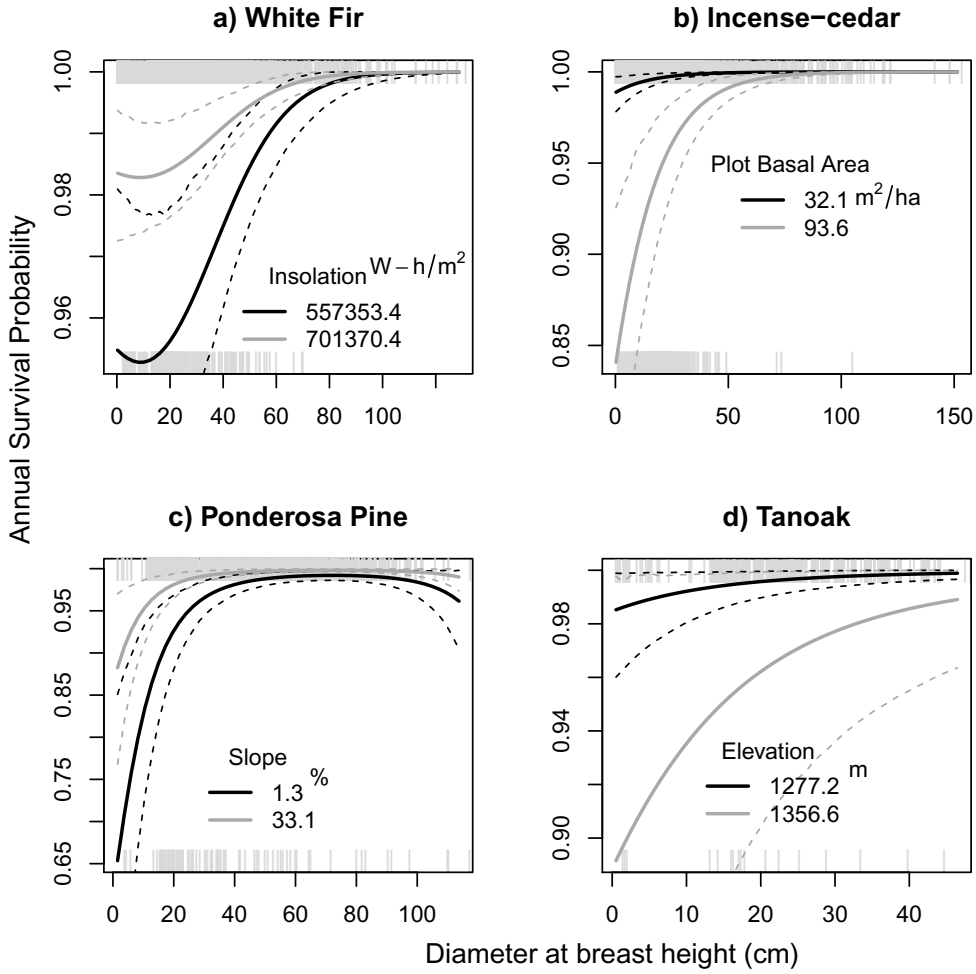

Figure 2.3: Size relationships with survival probability for four species: white fir (a), incense-cedar (b), ponderosa pine (c), tanoak (d). Black and dark gray lines show the effects of additional variables for each species. Black lines are two standard deviations below the mean value of the covariate, and dark gray are two standard deviations above. Dashed lines indicate 95% credible intervals due to uncertainty in other parameters. Data are shown as light gray tick-marks with live trees at the top of the plot and dead trees at the bottom. The parameter estimates in these figures are from 'full models' including all the forward-selected variables, as indicated in Table 2.1 (see Appendix O for other species and variables).

to 99.8% for medium-sized trees (71.09 cm DBH) on steep slopes. Other species show similar heterogeneity due to size and other variables (Figure 2.3, Appendix O).

2.3 Discussion

The answers to my questions point to a complex story, demonstrating the need for a rich set of species-specific models which incorporate stand development variables and tree size as well as climate variables, time trends, and topographic variables. Species differences emerge from these complex models, as well as intraspecific heterogeneity and temporal changes in survival.

Species Differences in Survival

Even in a second-growth forest, there is a diversity of responses to different factors controlling survival (Figures 2.2 and 2.3). Some of these patterns are surprising while others are not.

Of all my explanatory variables, tree size has the most consistent impact on survival, as in Hurst et al. (2011), but each species has a different relationship with size (Lines et al. 2010). Unsurprisingly, all species, both individually and aggregated, show higher survival for larger trees, but only some species show a decline in the largest individuals (black oak and ponderosa pine, Table 2.1 and Figure 2.2), as seen in other studies in the region (Dolanc et al. 2013). Tanoak's negative response to elevation is consistent with its current observed range limits: at BFRS, tanoak is at the high-elevation limit of its range (Burns and Honkala 1990).

Species' responses to plot basal area, on the other hand, show some surprising relationships. Plot basal area is increasing at BFRS (Appendix I), and reflects increasing competition for light, which in theory should affect all of the canopy species. Taking shade intolerance as an indication of competitive sensitivity, based on Niinemets and Valladares (2006), I expect the following ranking from least shade-tolerant (competitively weakest) to most shade-tolerant (competitively strongest): black oak, ponderosa pine, sugar pine, Douglas-fir, incense-cedar, tanoak, and white fir. Directly comparing basal area effects for all species (in a model with size and basal area, regardless of the best model, Figure 2.2b), black oak does show a relatively larger negative effect, Douglas-fir an intermediate effect, and white fir the smallest negative effect, as in Niinemets and Valladares (2006). Surprisingly, however, incense-cedar shows the largest negative effect, while tanoak is unusually intolerant of greater basal area and ponderosa pine and sugar pine are unusually tolerant relative to the other species. Note that the strong impact of basal area on incense-cedar would not have been apparent if I had limited myself to species-aggregated models, highlighting the importance of the species-specific models. Looking only at the best models, only incense-cedar selects the basal area effect. I have no explanation for this, though I note that the basal area effect for Douglas-fir would have been selected using AIC-like reasoning (de Valpine 2014).

From the two species' best models, incense-cedar is more negatively affected than Douglas-fir (Appendix M), reversed from what Niinemets and Valladares (2006) would predict.

Heterogeneity and Susceptible Subsets

At BFRS, mortality rates for an average tree (the complement of annual survival probability, or $1 - \phi = 1 - (expit(b))$) from simple models (1.21% – 2.3%; see Tables L.1 and L.2 in Appendix L) are similar to reports of background mortality rates in many forests (e.g. Canadian boreal forests: 1.4% per yr from Bond-Lamberty et al. 2014, 1-5% per yr from Luo and Chen 2013, 1.9-4.9% per yr from Peng et al. 2011; Russian deciduous forests: 1-3% per yr from Igor et al. 2009; New Zealand Nothofagus forests: 1.8-2.2% per yr from Hurst et al. 2011).

However, the answers to my questions demonstrate the need for a rich set of species-specific models which incorporate a diversity of factors. For each species, trees under various conditions are more likely to die than others, e.g. those in low-survival plots, of small size (Igor et al. 2009), in high basal area environments (Das et al. 2007), or at the edge of their range (e.g. tanoak at high elevation). The concept of a single annual survival probability becomes less meaningful for a heterogeneous population: a 'mean' annual survival probability may not represent these subsets of the population, whose differences in survival are exposed in a more complex model. The complex model including explanatory variables and plot effects is also valuable not only in providing information about the direction and magnitude of those components, but also in providing for better prediction and improving estimates of secular trends.

Even accounting for susceptibility via size or plot basal area does not accommodate all the heterogeneity in some species. Even after including these variables, incense-cedar, Douglas-fir, and ponderosa pine still have plot random effects which are important (Appendix N). In particular, for incense-cedar, which has the strongest plot basal area effect, the species' plot random effect has the largest standard deviation, potentially indicating spatial dynamics below the plot level, perhaps at the level of a neighborhood (Das et al. 2008).

I have modeled the population-level factors affecting tree survival through individual-based measurements. Individual heterogeneity strongly affects survival processes as lower-quality individuals are removed early on (Vaupel and Yashin 1985), so simple cohort analysis may be misleading (Appendix L). Though I do not model an individual tree quality random effect (Knape et al. 2011), I do include a tree-level variable (size) in my model (Clark et al. 2010). In addition, I have incorporated new individuals throughout the inventory period, so if the quality of the new recruits is well distributed, I may avoid this problem to some degree. Regardless of the source of individual variation, the importance of susceptible individuals may need to be accounted for in further applications of these results, e.g. structured population models, where individual quality can have a strong impact on population-level results (Ellner and Rees 2006).

Time Trends and Climate Effects

The secular time trend for the species-aggregated models is elusive, while for some individual species the trend is robust (Figure 2.1). My species-aggregated models are most similar to the existing models of survival for this ecosystem (van Mantgem and Stephenson 2007). This failure to ultimately find a time trend at the community level implies that incorporating stand development in a second-growth forest accounts for decreasing survival at the community level (Thorpe and Daniels 2012).

Where a secular time trend was selected for individual species, however, it does not always agree with existing thinking about the roles of these species in forest community dynamics. In these forests, a century of fire exclusion has resulted in increasing abundances of shade-tolerant fir and cedar at the expense of shade-intolerant pines, which require large canopy gaps such as those created by flare-ups during mixed-severity fires in order to regenerate successfully (York et al. 2013). Stands are increasingly characterized by larger numbers of small diameter trees (primarily white fir, Douglas-fir, and incense-cedar), increased canopy cover, and decreased gap size and gap abundance (Minnich et al. 1995). These structural qualities provide a positive feedback for increasing fir survival and decreasing pine survival. The models do indicate that ponderosa pine's survival is decreasing, but sugar pine shows an increase in survival. This unexpected result is particularly intriguing given that white pine blister rust (*Cronartium ribicola*) has been found in the Sierra Nevada.

Though there is no sign of an annual water deficit effect at this site (unlike van Mantgem and Stephenson 2007), in some ways this is unsurprising, as BFRS's weather records do not yet show a strong climate trend (Appendix I), and drought stress can take long time periods to play out (Bond-Lamberty et al. 2014).

2.4 Conclusions

I have shown that survival processes in second-growth temperate forests of the arid West are complex. Heterogeneity appears at every level, as species-specific stories differ from each other and from species-aggregated stories, and susceptible subsets emerge within species. Decreases in survival do occur for some species, but are not as widespread as other studies in old-growth Sierran forests or Canadian boreal forests indicate, and there is no evidence in these models of annual water deficit as a factor in tree survival at BFRS. The best way to assess the true biological significance of species differences in survival would be a full population model for each species (Zuidema et al. 2010). These models do not estimate recruitment rates (Suarez and Kitzberger 2008), and without modeling recruitment and growth as well, I cannot conclude anything at the population level based on these estimates of survival. Ultimately full population dynamics will determine how changes in survival will affect long-term community dynamics. Finally, this study is observational rather than experimental. As with many ecoinformatics studies, my results are clues pointing to possible relationships which could be more systematically examined.

Intermission 2

I now broaden my view from individual-based measurements to mapping tree communities using remote sensing. Rather than aggregating individual measurements of trees up to the species and then the community level, I start from a top-down approach of using high-spatial-resolution remotely sensed imagery to investigate topographic drivers of cover change. Mapping and statistical analysis are complementary methods, and I demonstrate their synthesis in Chapter 3. Here again, the theme of long-term data with challenging characteristics emerges, now with historical aerial imagery. In this study of woodland-to-forest transition in the North Coast, the story shifts from the heterogeneity between species to heterogeneity between sites. Finally, I explicitly focus on spatial scale in an effort to understand whether woodland-to-forest transitions are driven by topographic proxies for moisture more at the tree neighborhood level or the stand level. Though the ecosystem is different, the questions regarding the use of long-term data to guide current management are similar.

Chapter 3

Examining Topographic Drivers of Woody Cover Change in a Woodland-Forest Transition Zone Using Historical Aerial Imagery and Generalized Additive Models

In California, the shift from management by native peoples using frequent, low-intensity fire to a policy of fire exclusion has led to increases in coniferous forests at the expense of oak woodlands as shade-tolerant conifers recruit under oak canopies and eventually replace them (Cocking et al. 2012). Oak woodlands are widely valued for their many amenities (Standiford and Scott 2001), cultural values (Starrs 2001), support of rural livelihoods and lifestyles (Huntsinger et al. 2010), and ecological functions (Roche et al. 2012, Huntsinger and Oveido 2014), and already face many threats (e.g. the spread of introduced pathogen *Phytophthera ramorum*, Meentemeyer et al. 2011; climate change and reduced habitat suitability, Kueppers et al. 2005, mortality from introduced insects such as gold-spotted oak borer, Coleman et al. 2014; conversion to agriculture or residential land use, Huntsinger et al. 2010). Therefore, understanding the extent of encroachment by conifers into woodlands is an urgent issue. Unlike many of these other threats, conversion from oak woodland to coniferous forest is a slow process: conifer replacement of oaks can take 60-100 years (Cocking et al. 2014). Long-term information, in particular for sites with transitional woodland-forest areas, is essential in evaluating the extent of this problem.

One potential source of long-term data on forest cover is historical aerial imagery. This kind of imagery, typically dating back to the 1930s and 1940s, is available throughout North America (Morgan and Gergel 2013). Though the imagery can be difficult to find and often requires orthorectification and georegistration, classification of high-spatial-resolution black-and-white images has become more common with the commercial availability of Object-

Based Image Analysis (OBIA) software. OBIA allows analysts to use textural and contextual information in classifying single-band images, and the use of OBIA with historical aerial photos has expanded in the last 10 years (Laliberte et al. 2004, Marignani et al. 2008, Pringle et al. 2009, Allard et al. 2012, Martha et al. 2012). Generally speaking, most of the OBIA change detection literature is focused on innovations in mapping techniques and their application to many different systems (Desclée et al. 2006, Stow et al. 2008, Conchedda et al. 2008, de Chant and Kelly 2009, Dronova et al. 2011).

The goal, however, is not just to map the change but to understand the drivers of the mapped change, an interdisciplinary project involving both modeling and historical ecology highlighted by Gimmi and Bugmann (2013). Among studies using OBIA to classify historical aerial imagery, only a few model the drivers of change (Platt and Schoennagel 2009, Cserhalmi et al. 2011, Levick and Rogers 2011, Garbarino et al. 2013, Newman et al. 2014*a*,*b*). These studies encounter two fundamental issues with the data: 1) spatial autocorrelation and 2) percent of any given cover type is bounded between 0 and 100 – not normally distributed. The relative advantages and disadvantages inherent in different methods of accounting for spatial autocorrelation in statistical models have been much debated recently (Miller et al. 2007, Dormann et al. 2007, Betts et al. 2009, Dormann 2009, Beale et al. 2010, Hawkins 2012, Kuhn and Dormann 2012). In addition, the issue of transforming or otherwise appropriately working with proportion data has been highlighted for a long time, with recent developments favoring binomial or beta distributions with logit link functions (Warton and Hui 2011, Schmid et al. 2013).

Studies modeling historical change have used various simplifications and other strategies to account for these problems. Levick and Rogers (2011) broke up the response variable into categories and used a Canonical Correspondence Analysis in order to model the bounded variable (percent of a certain class) without dealing with its non-normal distribution. Platt and Schoennagel (2009) broke up predictor variables into categories to capture flexible relationships using ANOVAs, but this strategy ignores the inherently ranked nature of categories created from continuous variables. And while Cserhalmi et al. (2011) mention regressions without addressing autocorrelation, Newman et al. (2014*a*,*b*) resample their data to avoid it, using only a fraction of the area they have classified. These statistical challenges (autocorrelation and non-normality of proportion data) can be better addressed using appropriate statistical models without simplification or data reduction in order to investigate drivers of cover change.

In Humboldt County, California, Douglas-fir (*Pseudotsuga menziesii*) is encroaching into oak woodlands composed primarily of California black oak (*Quercus kellogii*) and Oregon white oak (*Quercus garryana*). Both oak species, and Oregon white oak in particular, are drought-tolerant and able to out-compete Douglas-fir in hotter, drier locations (Burns and Honkala 1990, Niinemets and Valladares 2006); therefore one expects moisture availability to strongly control the progress of encroachment. Historical soil moisture measurements are

not available, however, so topographic proxies can represent moisture controls on woodland-to-forest conversion. Topographic variables (including slope, aspect, elevation, curvature, radiation, and topographic wetness index) have long been used to predict vegetation characteristics (Franklin 1995, Deng et al. 2007, Jenkins and Coops 2011), and are unlikely to have changed in the last 60 years.

Though Douglas-fir forest is replacing oak woodland (Cocking et al. 2012), and is likely to do so more effectively in moist areas, identifying species from historical imagery proves difficult and mapping is restricted to a more general assessment of changes in woody versus herbaceous cover (Eitzel et al. 2014). Change in woody cover in these transitional sites, however, is largely due to conversion from woodland to forest, with some systematic increase in woody cover for all species and forest types (Schriver and Sherriff 2014). If encroachment and woodland-to-forest conversion has occurred since the 1940s, woody cover is expected to increase in wetter parts of the site. Models demonstrating little effect of moisture-related topographic variables would imply that woody cover increase is consistent across the site and that encroachment is not the most important process in woody cover change.

One final important complicating factor remains: the spatial scale of the process may not match the spatial scale of the data. Different ecological processes may act at different spatial scales and different hierarchical organization levels (e.g. individual tree, neighborhood, stand, site, landscape), and these hierarchical levels may not match spatial scales; additionally, thresholds may appear where emergent properties arise (Bissonette 1997). The data one has available are often at an arbitrary resolution that is more constrained by data acquisition than the process of interest (Deng et al. 2007). Recent calls for forest management to address multiple scales use simulation as a way to study changing forest responses at different scales (Seidl et al. 2013). Empirical studies on scaling relationships for vegetation patterns so far have only correlated topographic variables with vegetation indices at a range of spatial scales rather than testing multiple variables at once while incorporating spatial autocorrelation (Deng et al. 2007).

For the process of interest, conifer encroachment, it would be ideal to explore the potential for detecting increased woody cover characteristic of woodland-to-forest transition at a variety of scales, from the tree neighborhood to the stand level. Das et al. (2011) calculated a tree's neighborhood, "an area big enough to allow at least two of the largest trees to interact," (p. 1204) based on allometric equations for tree species in the nearby Sierra Nevada mixed-conifer forest. They found that a 9-m radius was a reasonable area to "capture local processes affecting trees both large and small" (p. 1204). As Douglas-fir is a major component in the densest parts of these woodland-to-forest transitional ecosystems, as well as a component in Sierran forests (Burns and Honkala 1990), I use 18 m as a minimum linear distance to include neighborhood dynamics. Oak species are likely to have larger and more variable crown diameters as well as wider spacing (Burns and Honkala 1990), so scales up to 40 or 50 m may be more appropriate as neighborhoods for the unencroached areas of these

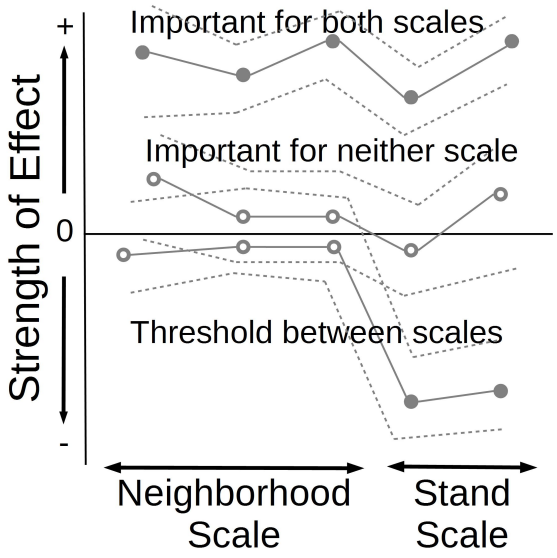

Figure 3.1: Diagram of scaling effects. Dashed lines show confidence intervals, solid dots are significant, while hollow dots are not significant. Some variables may be important for both scales, while others are important for neither; and still others might indicate a threshold where the variable is important at the stand scale but not the neighborhood scale.

systems. A single stand in these systems is likely to have a scale 50-100 m across (Lenya Quinn-Davidson, personal communication).

The process of woodland-to-forest conversion is happening at the neighborhood scale where individual Douglas-fir trees are out-competing individual oak trees, while overall tree response to moisture in terms of growth following the removal of fire is happening at the stand scale. Are there differences in the effects of moisture between neighborhood scales (20-50 m) and stand scales (50-100 m)? Are there stronger responses at the finer scales (where neighborhood dynamics actually play out) or at stand scale? Is there a threshold above or below which topographic variables representing moisture become more important in predicting woody cover (Figure 3.1), indicating an emergent property at neighborhood or stand scale? Do stand-scale and larger topographic variables miss enough variability not to show any effect at all?

In this study, I use high-spatial-resolution images to map 1948 (historical) and 2009 (re-

cent) woody cover at four woodland-forest transitional sites in Humboldt County. I model recent cover as a function of topographic variables and historical cover, addressing the statistical challenges outlined above by 1) choosing an appropriate statistical distribution (quasi-binomial) and link function (logit) to model non-normal data, 2) making use of a Generalized Additive Model (GAM) with a smoothing term on the spatial coordinates and an adequate number of spline knots to account for spatial autocorrelation, and 3) including the historical woody cover in an appropriately transformed way. I then answer the following questions: 1) Does woody cover increase more at wetter sites? 2) Do topographic variables representing greater moisture availability demonstrate significant and positive relationships with recent woody cover, after controlling for historical woody cover? 3) Are these relationships stronger at the neighborhood (20-40 m) scale or at the stand (50-100m) scale and is a threshold effect apparent between the two scales?

3.1 Methods

Study Areas

The North Coast of California contains a variety of ecosystem types, including redwood forest, coastal sage scrub, oak woodland, and prairie. The climate is Mediterranean, with cool, wet winters and hot, dry summers. Oak woodland at these sites in Humboldt County is characterized by California black oak and Oregon white oak with an understory predominantly composed of grasses and forbs. The transition between woodland (defined as more than 30 percent cover with 150-300 trees/ha, Agee 1993) and closed canopy forest (greater than 300 trees/ha, Agee 1993) occurs when evergreen species, typically Douglas-fir, encroach into woodland over time, forming a dense, shaded forest with little to no herbaceous understory.

These sites were selected to reflect different latitudes throughout northern Humboldt County, as well as several coast-to-interior locations for oak woodlands (Figure 3.2. They have considerable Douglas-fir encroachment, but also contain portions of intact oak woodland with minimal or no conifer in the canopy. Though parts of these sites have been harvested, areas were carefully selected within them that have no history of tree harvest since the 1940s. When drawing polygons for analysis, I avoided locations with skid trails or evidence of clearcuts. The areas defined by these polygons also have no known fires since the '40s (Schriver and Sherriff 2014). Therefore loss of woody cover can be assumed to result from disease or insect damage, and gain in woody cover from growth of established individuals and recruitment of some species into gaps. I further adjusted the polygons to remove artifacts such as marks on the historical images. Of the four sites, Iaqua Buttes (IB), Bald Hills (BH), Willow Creek (WC), and Blake Mountain (BM), Iaqua Buttes and Bald Hills are privately owned, while Blake Mountain and Willow Creek are found within Six Rivers National Forest (managed by the US Forest Service). Because I have selected areas where no trees have been

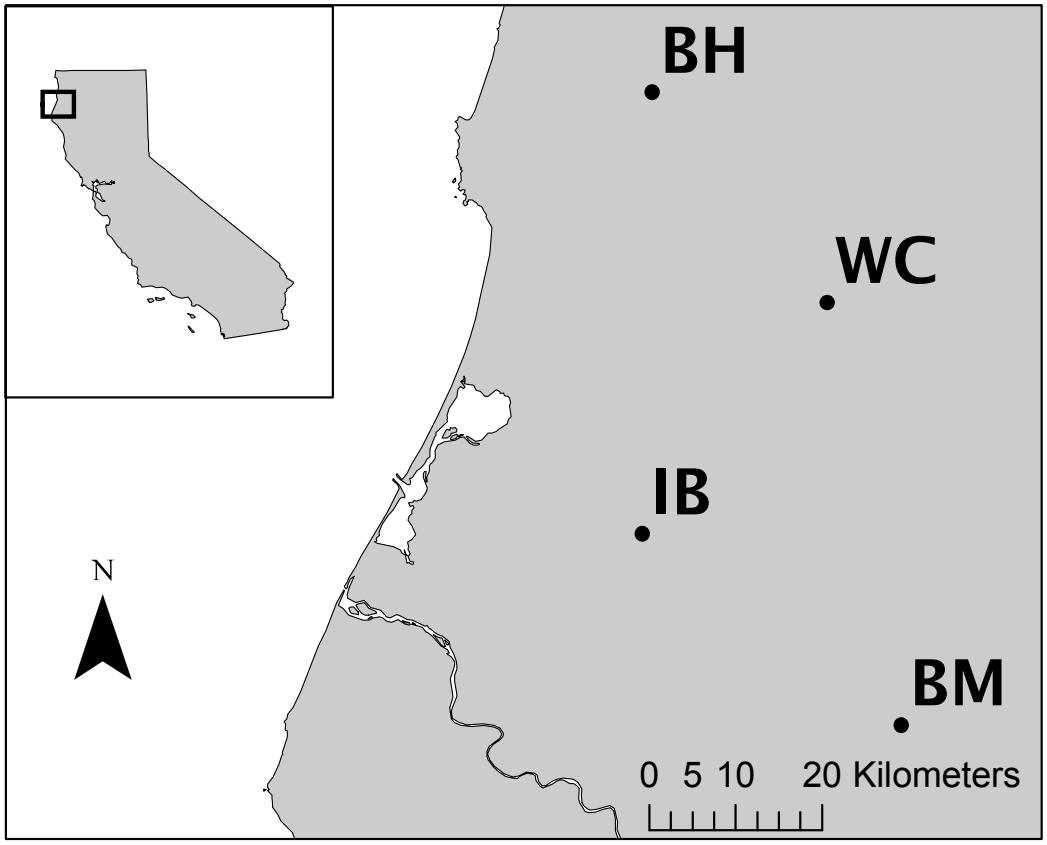

Figure 3.2: Geographical context. The four sites (Iaqua Buttes – IB; Bald Hills – BH; Blake Mountain – BM; Willow Creek – WC) are located in upper Humboldt County in California, USA.

cut and no known fires have occurred, land ownership is not expected to affect encroachment and increase in woody cover.

The majority of these sites are south-facing and locally relatively dry, because they were selected to include oak woodlands and oaks are drought tolerant (Niinemets and Valladares 2006). For Iaqua Buttes, Bald Hills, and Blake Mountain, conifers are present at lower elevations in moist valleys and tend to encroach upwards (Lenya Quinn-Davidson, personal communication). At Willow Creek, the conifers are encroaching from above the oaks, coming over a ridgeline. Iaqua Buttes and Bald Hills are closer to the coast and experience frequent foggy conditions, while Willow Creek and Blake Mountain are further inland with hotter, drier conditions overall. Given that Douglas-fir encroachment is driving the majority of woody cover increases in this system, cover is expected to increase more at Iaqua Buttes and Bald Hills, and topographic variables representing greater moisture are expected to have a

positive impact. Elevation is also expected to be important as it represents the location of the coniferous forest at each site: low elevations for Iaqua Buttes, Bald Hills, and Blake Mountain, and high elevations for Willow Creek.

Imagery and Data Sources

Historical images were flown for the USDA Forest Service in 1948. The black and white (panchromatic) image frames were scanned by Humboldt County's California Geological Survey at 800 dots per inch, which after orthorectification resulted in one meter by one meter square pixels. For the analysis, I use frames CDF2-17-007 (Iaqua Buttes), CDF2-15-093 (Blake Mountain), CDF2-15-153 (Willow Creek), and CDF2-19-196 (Bald Hills). For image pre-processing, I also used neighboring frames in each flightline, though these additional images were not included in this analysis. Metadata was available for these flights at the UC Santa Barbara Map Library.[1]

Recent images are from the 2009 survey of the US Department of Agriculture's National Agricultural Imagery Program (NAIP), as downloaded and tiled by Cal-Atlas.[2] NAIP is high-spatial-resolution (1-m x 1-m pixels) and has four bands: red ($\sim$ 635 nm), green ($\sim$ 560 nm), blue ($\sim$ 460 nm), and near infra-red ($\sim$ 860 nm). NAIP imagery is already orthorectified and projected to UTM Zone 10N.

The digital elevation model (DEM) is from USGS' National Elevation Dataset, as downloaded and tiled by Cal-Atlas.[3]. I projected the DEM to UTM zone 10N coordinates, and used it both for orthorectification and for calculating topographic predictor variables. The topographic variables in the statistical models are: 1) elevation from the DEM in meters, 2) topographic slope in degrees (maximum change in elevation between a cell and any of its neighboring cells), 3) profile curvature (curvature in the direction of steepest descent/ascent), 4) 'northness,' the cosine of aspect (measured from north, with '1' indicating north and '-1' indicating south, 5) 'eastness,' the sine of aspect ('1' indicating east and '-1' indicating west) as used in Levick and Rogers (2011) and 6) topographic wetness index (sometimes referred to as topographic moisture index or compound topographic index). Topographic wetness index is defined as $ln(\alpha/tan(\beta))$. α is the catchment area collecting to that pixel (offset by one in order to avoid taking the log of zero), calculated from a watershed delineation tool and divided by the cell width; β is the slope in degrees. Topographic wetness index has been shown to explain variation in vegetation metrics (Jenkins and Coops 2011, Wang et al. 2013). Raster operations and calculations were conducted in ArcGIS 10.2 (ESRI 2013). Topographic predictor variables were standardized (centered and scaled by their respective standard deviations) within each site-scale combination in order to compare their relative impact on woody cover percentage. Based on the hypothesis that greater moisture enables

[1] http://mil.library.ucsb.edu/apcatalog/report/report.php?filed_by=CDF2

[2] http://www.atlas.ca.gov/download.html#/casil/imageryBaseMapsLandCover/imagery/naip

[3] http://www.atlas.ca.gov/download.html#/casil/elevation

conifer encroachment and woodland-to-forest transition, I predict that northness will have a positive effect on woody cover, eastness will have a small but positive effect (as afternoon sun on west-facing slopes is more drying than morning sun on eastern slopes, Deng et al. 2009), topographic moisture index will have a positive effect, steeper slopes will have a negative effect, and positive curvature (defined in ArcGIS as convex upward) will have a negative effect on current cover. I predict that historical cover will have a positive effect, and elevation will have a positive effect for Iaqua Buttes, Blake Mountain, and Bald Hills, and a negative effect for Willow Creek (see "site description" section). I predict that these relationships should be stronger at the neighborhood scales (20-40 m), as the topographic predictor variables lose their representation of the range and variability of moisture conditions as they are aggregated to larger cell sizes. Finally, I expect these relationships to be more apparent at the wetter sites (Iaqua Buttes and Bald Hills), as it has been shown that in drier years topography matters less for vegetation (Dorman et al. 2013); I extend this logic to drier/wetter locations.

Pre-processing of Imagery

I used Leica Photogrammetry Suite (Intergraph 2012) to orthorectify and georegister the imagery using the 2009 NAIP imagery as a horizontal reference and the 10-m DEM as a vertical reference. I collected 50-150 ground control points for each site and used cubic convolution resampling for the orthorectification. I used ArcGIS to mask and mosaic the historical images to each site's polygon. I used package "`glcm`" (Zvoleff 2014) in R (R Development Core Team 2009) to calculate six different per-pixel gray-level co-occurrence matrix (GLCM) textures with a 7x7 moving window (Haralick et al. 1973). I calculated mean, variance, contrast, dissimilarity, entropy, and second moment. I chose a 7x7 window because it produced a visually smoother result, which was better for later segmentation and classification. For the 2009 NAIP imagery, I calculated a per-pixel Normalized Difference Vegetation Index (NDVI) layer using ArcGIS's Raster Calculator to use in segmentation.

Segmentation and Classification

I then used eCognition Developer 8 (Trimble 2013) to segment and classify each image for each site and year (total of eight images). I used a simple classification scheme at two scales. Because I am interested in the transition from open woodland to closed-canopy forest, I first classified land cover as 'forest' (characterized by dominance of woody species) or 'prairie' (characterized by open grassland area with very little tree canopy cover), and masked out 'prairie' according to the 2009 forest-prairie edge. There is little recruitment of woody species in the middle of the prairie, and any advancement of woody cover into the prairie at the forest edge between 1948 and 2009 is captured by using the 2009 image as the mask. Following this classification, within the 'forest' type, I classify areas as 'woody' (oak, Douglas-fir, shrubs, or other trees) and 'herbaceous' (open clearings). I proceed with analysis of 'woody' versus 'herbaceous' only within the 2009 'forest' area.

For the 2009 NAIP imagery, I used all four bands (RGB and IR) and per-pixel NDVI, and for historical imagery I used the image itself and the six GLCM texture measures. I first used multiresolution segmentation with a large scale parameter to mask forest and prairie from each other, and then multi-threshold segmentation on various bands to classify within those areas (Gärtner et al. 2014). Using ArcGIS' 'classify' tool for raster symbology, I examined the histogram of values for the band in question for breakpoints in order to choose thresholds. Different bands were helpful for segmenting and classifying different images. For 2009 NAIP images, NDVI as well as mean values of different bands (e.g. green, infra-red), or overall image brightness were often helpful (and for one site, Bald Hills, per-pixel texture variables calculated for the red band in addition to the image bands and NDVI were needed in order to discriminate between classes); for historical imagery the image brightness value worked well, in addition to contrast, homogeneity, dissimilarity and second moment.

Accuracy Assessment

Ultimately I conduct further analysis of woody cover change on only the forest region, due to the difficulty of classifying woody vs. herbaceous cover in the prairie area for the historical imagery. I also restrict the analysis to areas with no more than 5 m of offset in registration between the recent and historical images. This resulted in areas of 52 hectares (Iaqua Buttes), 63 hectares (Willow Creek), 71 hectares (Bald Hills), and 77 hectares (Blake Mountain). Within these regions, I assessed classification accuracy for each image by selecting 100 random points using ArcGIS's random point generating utility and visually assessing the class at that point. Because herbaceous cover was much less common in the forest, I used a proportionally stratified sampling strategy but ensured that at least 10 points of classified herbaceous cover were assessed for accuracy. Like most historical image classification studies, I lacked alternative imagery to validate the historical imagery, nor did I have alternative imagery to validate the recent high-resolution imagery, and thus made a determination by eye of actual class with which to compare the assigned class (Platt and Schoennagel 2009, Cserhalmi et al. 2011, Levick and Rogers 2011). In a few cases random points fell into pixels that were obviously mixed between herbaceous and woody cover; these were thrown out due to the impossibility of assigning a class to the point (6 cases out of 800 points). Field data to validate the historical imagery was impossible to obtain, and obtaining field data to validate the recent imagery was not feasible as part of this project. I considered accuracy assessment based on objects, but because multi-threshold segmentation results in few very large objects (unlike multi-resolution segmentation which results in fairly similarly-sized objects due to the constraint of the scale parameter), there were not appropriate objects to sample from and I determined that a point-based accuracy assessment was best (Müllerová et al. 2013). I report standard accuracies and kappa values to enable comparison with other studies (Jensen 2005, Müllerová et al. 2013).

Statistical Models

Spatial autocorrelation is incorporated in the models using a smoothing function of geographical coordinates (UTM Northing and Easting) in a Generalized Additive Model (GAM) in the R package "`mgcv`" (Mixed GAM Computation Vehicle; Wood 2006). For each site, I aggregated the data at the raster level to cell sizes of 20, 30, 40, 50, 60, 70, 80, 90 and 100 m and fit models for each cell size (total of 40 models). Note that I aggregate the topographic variables at those scales from the 10-m DEM using the 'average' method (rather than median) because it has been shown to have the most predictable statistical properties (Gotway et al. 2002). The GAM is ideal for these data for two reasons: 1) a Markovian random field (MRF) basis is available for the smoothing function, which is designed for use with polygons with potentially irregular sizes (which some of the cells have after applying the forest-only analysis mask); and 2) `mgcv` has extensions for non-normal response variables. The quasi-binomial family is a good choice, as it allows for overdispersion, which is appropriate here. The response variable is then framed as a number of 1-m by 1-m cells within the 10-m by 10-m cell of the DEM which are classified as 'woody cover' as binomial successes, and those classified as 'herbaceous cover' as binomial failures.

Though Dormann et al. (2007) reject the GAM as an adequate way to represent spatial autocorrelation, they use the default knot basis dimension (k) of 10, which is likely not to account for spatial autocorrelation in a dataset with small distances between points. Wood provides a function in `mgcv` called "`gam.check`" which gives a p-value for a test of whether k is large enough, and when tested with the dataset, gam.check indicated that a much higher k was needed. Therefore, k was set at $n/10$ (where n is the number of records), the GAM was fit with a smooth term for Northing and Easting, and the Moran test (package "`spdep`," Bivand 2013) was used to test for global autocorrelation in the model residuals. For those sites and scales which indicated significant autocorrelation, the model was run again with $2n/10$. I also tried a beta distribution as suggested by Schmid et al. (2013), but even at higher k values the Moran test indicated significant autocorrelation for most cell sizes. A thin plate regression spline basis, which is more typical of geographic modeling with GAMs as they are rotationally invariant, produced no appreciable changes to parameter estimates as compared with the MRF basis (see Figure P.1, Appendix P).

I fit models of recent woody cover as a function of topographic variables (elevation, slope, profile curvature, topographic wetness index, and aspect), historical woody cover, and smooth functions of Northing and Easting. Note that I logit-transformed historical woody cover before standardizing it to ensure that its scale would correspond to the scale of the response variable after applying the link function. I use a quasi-binomial "distribution" for the proportion of image pixels y_i within a DEM pixel i which are classified as woody cover in the 2009 NAIP image, with a link function $logit(E[y_i]) = \eta_i$ and $Var[y_i] = \phi E[y_i]$, where ϕ is the dispersion parameter. The covariates x are linear predictors with regression coefficients

β:

$$\eta_i = \beta_0 + \beta_{slope}x_i^{slope} + \beta_{elev}x_i^{elev} + \beta_{TWI}x_i^{TWI} + \beta_{curv}x_i^{curv} + \beta_{northn}x_i^{northn} + \beta_{eastn}x^{eastn} + \beta_{histcov}logit(x_i^{histcov}) + f(Northing_i, Easting_i) \quad (3.1)$$

$f()$ is a smooth function using the MRF basis. I fit models at all scales for all four sites independently, and corrected for multiple comparisons using the 'false discovery rate' method in the R function "`p.adjust`" (which uses the method of Benjamini and Yekutieli 2001). Figure 3.3 graphically illustrates the workflow associated with the processing and classification of the imagery, gridding at the scale of the DEM, scaling up, and fitting GAMs.

3.2 Results

Returning to my questions, I find that 1) woody cover does increase more for wetter sites (Iaqua Buttes and Bald Hills); 2) across sites and scales, historical woody cover is consistently significant and has a positive effect on recent woody cover, but results for each topographic variable haphazardly agree or disagree with my hypotheses; and 3) there is no consistent pattern between sites or scales for each topographic variable.

Preprocessing and Classification

Root mean square error (RMSE) for orthorectification in pixels was 8.24 for Iaqua Buttes, 4.16 for Bald Hills, 7.54 for Willow Creek, and 7.93 for Blake Mountain; recall that a pixel in this case is 1 m square. Though Jensen (2005) suggests that RMSE should be less than half a pixel, these RMSE values are within the limits of best practices according to the San Francisco Estuary Institute's historical ecology group which frequently orthorectifies historical imagery where topography makes exact registration difficult (Micha Solomon, personal communication). I show Iaqua Buttes' imagery and classifications as an example in Figure 3.4; see Appendix Q Figures Q.1 through Q.3 for the other three sites' imagery and classifications. Note my avoidance of the skid trails at the southern end of the image when drawing the polygon for Iaqua Buttes. Accuracy for Iaqua Buttes in 1948 was 92% (with a kappa statistic of 0.78); for the 2009 image, accuracy was 98% (kappa = 0.89). For Bald Hills, the 1948 image had a classification accuracy of 96% (kappa = 0.86); for the 2009 image, classification accuracy was 98% (kappa = 0.89). For Blake Mountain, the 1948 image had a classification accuracy of 94% (kappa = 0.74); for the 2009 image, classification accuracy was 99% (kappa = 0.95). For Willow Creek, the 1948 image had a classification accuracy of 99% (kappa = 0.94); for the 2009 image, classification accuracy was 99% (kappa = 0.94). See Appendix S Tables S.1 and S.2 for complete error matrices for each image. These accuracies and kappa values are similar to other OBIA-based classifications of historical and recent imagery (Marignani et al. 2008, van Lier et al. 2009, Cserhalmi et al. 2011, Allard et al. 2012, Martha et al. 2012).

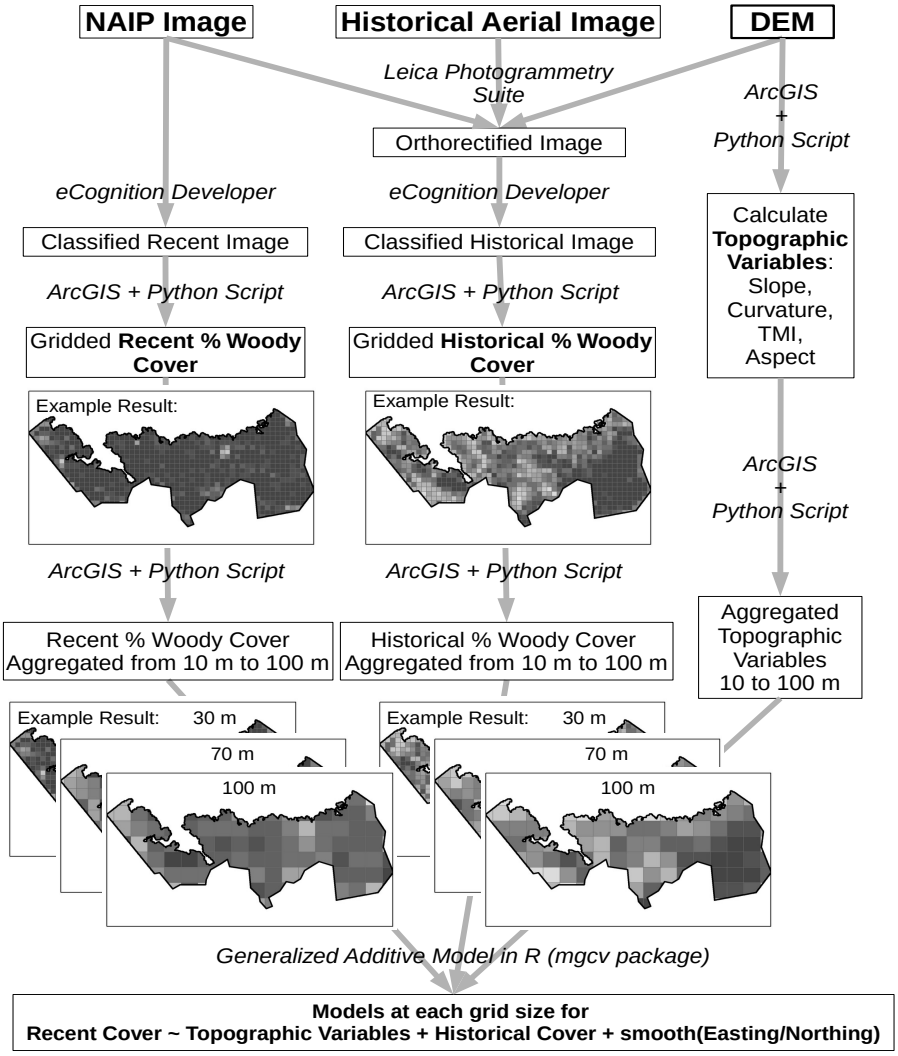

Figure 3.3: Workflow. The digital elevation model (DEM) and the 2009 National Agricultural Imagery Program (NAIP) image are used to orthorectify the historical image. The DEM is also used to create topographic variables: slope, profile curvature, aspect (northness/eastness), and topographic wetness index (TWI). The historical image and recent NAIP image are then each individually classified using OBIA techniques in eCognition (with the help of additional per-pixel layers, e.g. Normalized Difference Vegetation Index and texture variables; see text for details). The accuracy is assessed, and then each classified image is gridded based on the DEM's cell size of 10 m; then aggregated up to 100 m grid cells. At each scale and for each site, Generalized Additive Models are fit to explore the relationship between recent cover and previous cover, geographical coordinates, and topographic variables.

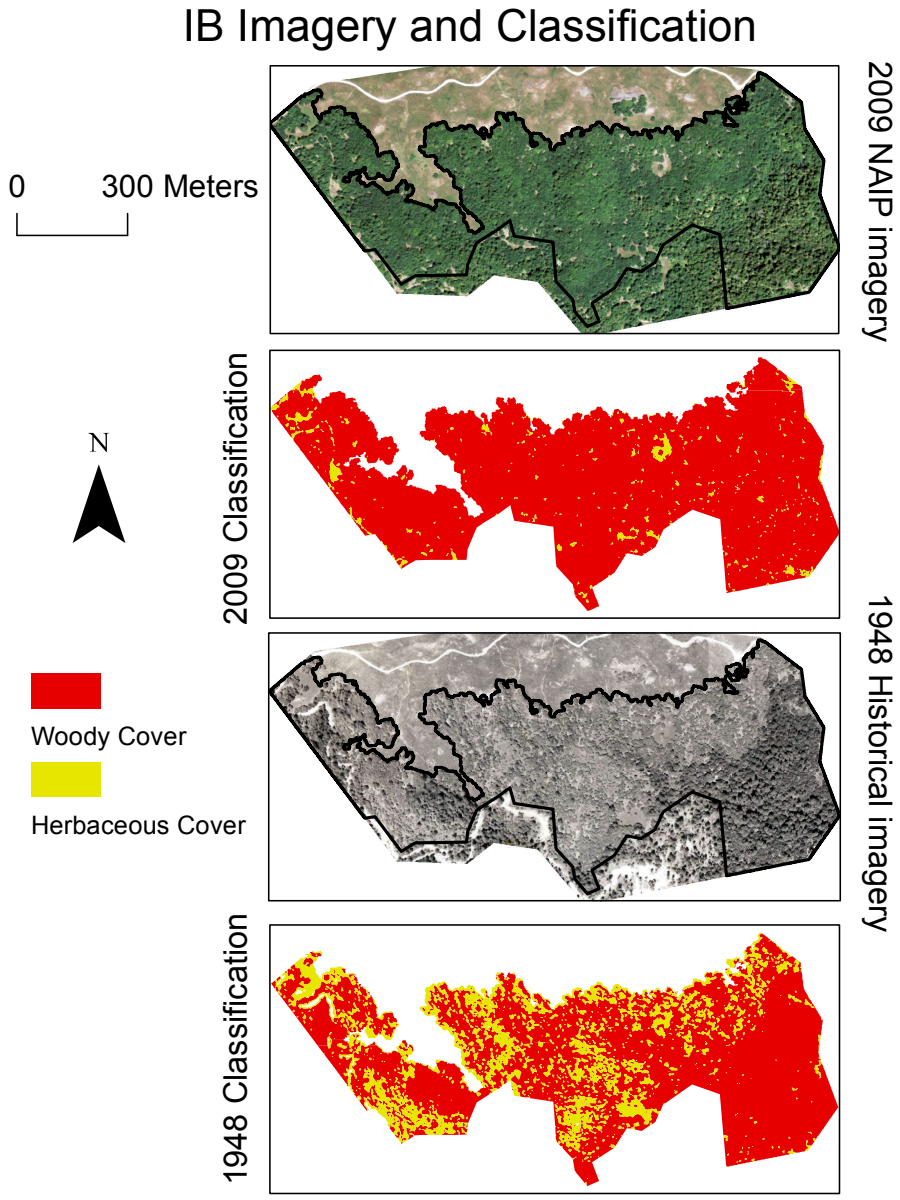

Figure 3.4: Example imagery and classification for the Iaqua Buttes site.

Basic Site Differences

In keeping with moisture limitation on Douglas-fir encroachment and woodland-forest conversion, for the two wetter sites, woody cover increased the most: Iaqua Buttes increased 21.3 percentage points and Bald Hills it increased by 11.3 percentage points. Meanwhile, woody cover at Willow Creek increased by 1.88 percentage points, and at Blake Mountain it decreased by 1.71 percentage points. Willow Creek is nearly all woody cover even in 1948, and therefore can only increase slightly. Blake Mountain has a large amount of woody cover in 1948 and is a very dry site, with canyon live oak (*Quercus chrysolepis*) dominating rocky slopes (Lenya Quinn-Davidson, personal communication), so some losses in woody cover over time are possible as Douglas-fir encroachment and overall vegetation growth may be limited by moisture.

For all models from 20 m to 100 m, the Moran test indicated no additional spatial autocorrelation in the residuals, after increasing the number of spline knots when needed. Dispersion as represented in the quasi-binomial's scale parameter was always estimated much greater than one, indicating that the quasi-binomial was a good choice for the data's distribution and a binomial without dispersion would not have been adequate. The scale parameter typically increased as the cell size went down, implying that the overdispersion becomes more extreme for smaller cells. The spatial term (the smooth function of Northing/Easting) was significant for three quarters of the site-scale combinations. See Appendix T Table T.1 for model diagnostics and results.

Sites and Topographic Variables

Figure 3.5 summarizes the changes in parameter estimates over different scales for each covariate. Across sites and scales, historical woody cover is consistently significant and has a positive effect on recent woody cover. For the topographic variables, no consistent pattern between sites or scales emerges. See below for results of significant variables for each site. Note that variables were standardized within each scale-site combination, and therefore the parameter estimates are not strictly speaking comparable across scales or sites; however, the relationships in Figure 3.5 do not change when using the non-standardized variables (see Appendix R Figure R.1 for a version of Figure 3.5 with parameter estimates not scaled by their standard deviations).

At Iaqua Buttes, the effect of profile curvature becomes positive for stand scales (80-100 m), opposite of what was expected. Topographic wetness index is positive for a neighborhood scale (40 m), and eastness is positive for stand scales (70 and 80 m), as predicted. Elevation at neighborhood scale is opposite of what was expected at neighborhood scale (negative effect), but positive at one of the stand scales (90 m); similarly, northness is positive for one stand scale (70 m) and negative for one neighborhood scale (40 m) and two stand scales (90 and 100 m).

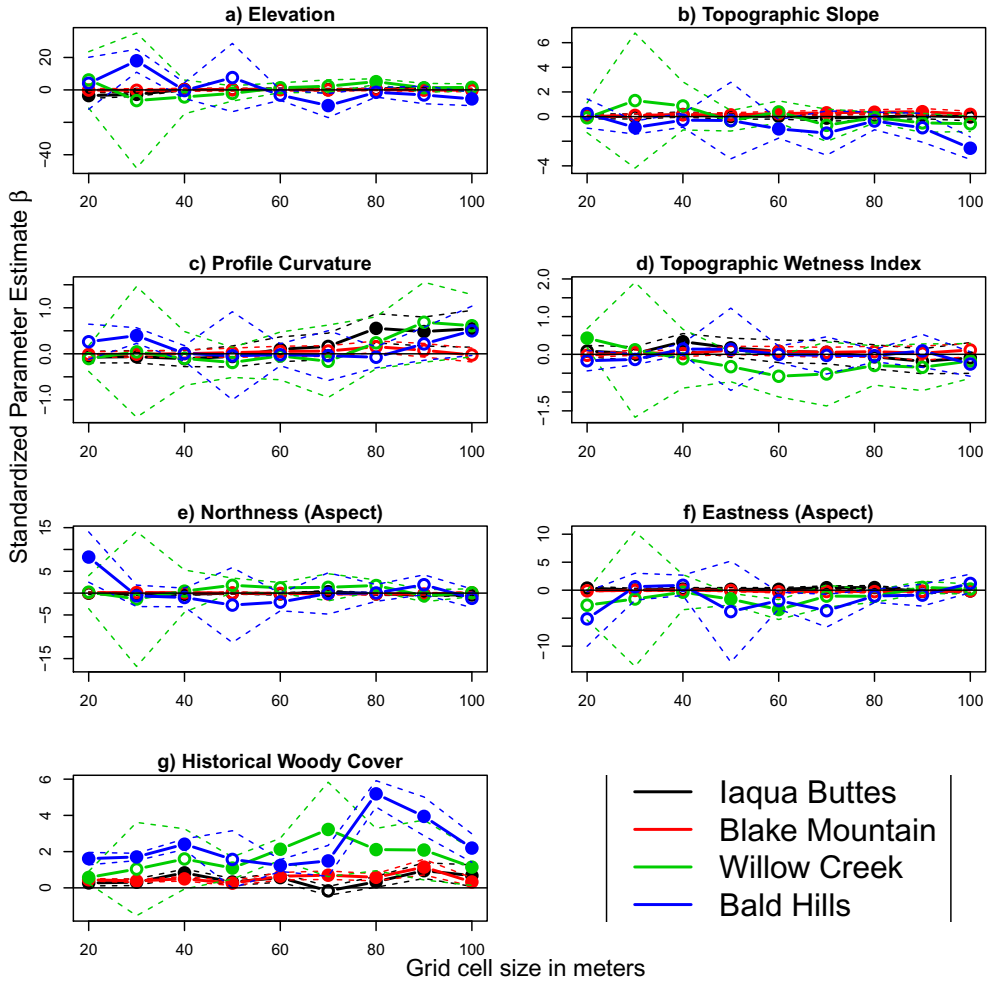

Figure 3.5: Results of scaling from 20 m grid cells to 100 m grid cells. Each color represents one site, with solid lines associated with parameter estimates and dashed lines showing 95% confidence intervals; parameter estimates which are significant at p <0.05 have solid circles and those that are not have open circles. Parameters are for standardized variables so the magnitude of the effects can be directly compared.

For Bald Hills, topographic slope is negative at both neighborhood and stand scales (30, 60, and 100 m), as expected. Northness is positive at the neighborhood scale (20 m), as expected. Profile curvature is positive at a stand scale (30 m), opposite from my predictions. Elevation is positive at neighborhood scale (30 m) and negative for two stand scales (70 and 100 m).

At Blake Mountain, topographic slope was positive at stand scale (60 through 90 m), inconsistent with my hypothesis. No other variables had significant effects but historical woody cover, which was positive at all scales as predicted.

At Willow Creek, elevation is positive at the stand scale (80 m), contrary to my hypothesis, while topographic wetness index is positive at the neighborhood scale (20 m), as predicted. Eastness is negative at stand scale (50 and 60 m), opposite of my hypothesis.

3.3 Discussion

I have tackled a number of technical problems in order to address the question of what topographic drivers influence woody cover and at what scales these influences are important. Generally, despite a messy dataset with many potential sources of error (including misregistration, classification error, and error associated with statistical models), as expected, the importance of controlling for historical woody cover is clear (Figure 3.5g). Otherwise, site- and scale-specific stories dominate the parameter estimates of the topographic variables, though there is limited evidence for the positive effect of topographic wetness index at the neighborhood scale for two sites. This result implies either that woody cover change is not driven by encroachment in moist areas, or that other factors control the transition from a more open woodland to a closed-canopy forest, e.g. other topographic variables, soil characteristics, or land-use history such as grazing.

Flexibility and Constraints of the Modeling Approach

The modeling approach that I demonstrate is flexible in many ways. If the extent of the available imagery is large enough and sufficient computer power is available, the GAM can be used at larger scales (e.g. Levick and Rogers 2011). One reason no clear relationship between woody cover and topographic variables emerged over scales or sites could be the assumption of the linear model for those variables; a GAM could also allow smooth functions of the other variables as well as the spatial coordinates. Smooth functions could capture nonlinear relationships in a more informative way than breaking the explanatory variables up into categories as Platt and Schoennagel (2009) did.

Another possible reason that no consistent pattern across sites emerged is that at some sites (Willow Creek, Blake Mountain) there was little total change in woody cover, i.e. not enough variation in the response variable for the predictors to show a relationship (indeed,

Blake Mountain shows very few significant variables at any scale). In addition, eCognition segmentation and classification strategies are still rapidly evolving: methods and rule sets for selecting segmentation parameters are under development (Dragut et al. 2010, Martha et al. 2012, Dragut et al. 2014) but there is not yet consensus on best practices. Better classifications (which then better capture the process of interest) may be possible as these tools develop further. It is also difficult to scale up this analysis to more sites or larger areas. Like other studies e.g. Platt and Schoennagel (2009), I found that different strategies for classification were needed for each image and generalization was difficult: texture variables were important for the historical images and for Bald Hills, while NDVI was most helpful for the recent images, and different combinations of multi-threshold segmentation and multi-resolution segmentation were most productive for different images (see Methods).

Topography, Disturbance, and Future Work

No systematic relationship between topographic variables and scale emerged between sites. This result contrasts with Platt and Schoennagel (2009) who demonstrate some evidence for relationships between woody cover change and topographic variables across their study area; however, they do not include geographical coordinates or sites in their model. Similarly, Levick and Rogers (2011) show some evidence that increasing topographic slope corresponds to decreased woody cover across their study area at their 10 ha and 100 ha scales, though these scales are far larger than the scales of my analysis. Newman et al. (2014*b*) show some impacts of slope across their study area, but the change in woody cover in their system is due to deforestation so slope has a strong impact on human access to forests.

I note that my two snapshots of the pattern of woody cover do not allow me to consider disturbances which have happened and then resolved during the time range 1948 through 2009. I am unable to discern whether slower woody cover change relates to poor growing conditions (e.g. thin, dry soils on steeper slopes, Dietrich et al. 1995) or small disturbances which occur and are recovered from quickly (e.g. shallow landslides, Dietrich et al. 2001). More frequent photos comprising a timeseries would help to discriminate between these two options, but would require flight metadata and additional analyst time for processing and classification. Even with more frequent photos, the disturbances could be small enough not to be visible in the images (Dietrich et al. 2001).

Conifer encroachment and subsequent woodland-to-forest transition is widespread throughout the North Coast of California (Cocking et al. 2012) and is maintained by continuing fire exclusion and forest practice rules that prevent the removal of encroaching conifers (Valachovic et al. 2014). I was unable to discern tree species in these images (Eitzel et al. 2014) and therefore the result that woody cover is increasing at three of the sites are only suggestive of the encroachment of Douglas-fir into oak woodland. In addition, some areas of these sites show an increase in woody cover where there are no known Douglas-fir (Schriver and Sherriff 2014), so some of the woody cover increase is due to growth of adult oak trees

as well as recruitment of Douglas-fir seedlings and saplings. Increasing woody cover is both a consequence of fire exclusion and a cause of extreme fire severity when fires do occur, causing the potential conversion of forested landscapes into shrublands (Collins and Roller 2013). If the forest practice rules are changed, landowners may wish to know where stands had been more open, and historical aerial imagery may be useful for planning restoration through conifer removal (Whipple et al. 2011, Grossinger 2012). Oaks respond well to conifer removal (Devine and Harrington 2006, 2013). Though this study using historical and recent high-spatial-resolution imagery produced inconclusive results regarding woody cover change and underlying encroachment processes, this simply highlights the importance of field work to document the problem and complement the historical imagery. In future remote sensing studies on this problem, higher spectral resolution may allow analysts to map species, but the process inherently involves a mixed conifer-oak class and field work will still be necessary to validate any such models.

Note that there was no apparent difference between privately (Iaqua Buttes and Bald Hills) and publicly (Blake Mountain and Willow Creek) owned sites but unknown site-specific factors may have more to do with this lack of pattern than more recent land ownership. For example, I have no knowledge of the grazing history at any of these sites. Private lands as well as public lands in this area of California have historically been subjected to a variety of land uses, including grazing as well as timber harvest, and grazing (or its removal) can have as large an impact on a system as fire (Yana Valachovic, personal communication). Other studies have shown that both topography and grazing history affect woody cover development following alteration of an existing disturbance regime (Garbarino et al. 2013), so future work should involve information on historical land uses at each site. Generally speaking, the uniqueness of the sites' responses to these variables implies that drivers of woody cover change are very site specific, and it is likely that additional variables are needed to explain spatial variation at each scale. The topographic variables I explored do not consistently explain the changes in woody cover. In future work, I could explore other variables representing grazing history, soil characteristics, moisture, and solar radiation (for example, annual insolation or a different transformation of aspect). I could also consider a single multiscale model which explicitly includes underlying processes and observation error from misclassification and misregistration (e.g. a state-space model, Eitzel et al. 2013).

Coda[1]

In this work, I have applied sophisticated statistical modeling and remote sensing techniques to ubiquitous and complex pre-existing data. I was able to distill from them new insights on drivers of tree growth and survival, sources of individual and intra-specific heterogeneity, and decadal changes in tree communities in both the North Coast and the Sierra Nevada of California. The application of sophisticated analysis techniques to learn about management-relevant ecology is only the first step, however. There are many calls for ecology to connect with policy (Schlesinger 2010), with practitioners (Palmer 2008), and with the public (Shirk et al. 2012), and ecologists increasingly feel the urgency of environmental degradation (Vitousek 1997, Robbins and Moore 2013). Many pressing ecological concerns involve many different geographical areas, organizational units, and communities. There is a need for organizations which can span boundaries to address urgent concerns (e.g. fire management Kocher et al. 2012, Moritz et al. 2014).

In order for ecological research to be truly useful and available to these stakeholders, the work becomes necessarily interdisciplinary. In Chapter 3 the interplay between community ecology and policy is clear: due to fire exclusion policies, Douglas-fir encroaches into woodland due to shade-tolerance and superior competition in closed canopy configurations, but the process is also maintained by the California Board of Forestry's forest practice rules restriction on cutting of conifers, which prevents restoration to woodland via conifer removal. To address environmental management challenges, ecological studies and restoration plans must be infused with social science (Eitzel et al. 2011), as social goals and constraints can be just as important as ecological goals and constraints (Hallett et al. 2013). The insights in the previous chapters have the potential to be useful for managers, regulators, policy-makers, planners, and land-owners, but further interdisciplinary work remains to be done to guide management. Fortunately interdisciplinary training is also increasing, and the next generation of scientists and practitioners have many opportunities to learn how to do interdisciplinary research (Oberg 2011, Andrade et al. 2014).

[4]Plummer, M., N. Best, K. Cowles, and K. Vines, 2010. coda: Output analysis and diagnostics for MCMC. URL `http://cran.r-project.org/package=coda`.

References

Agee, J. K. 1993. Fire ecology of Pacific Northwest forests. Island Press, Washington, DC.

Allard, M., R. a. Fournier, M. Grenier, J. Lefebvre, and J.-F. Giroux. 2012. Forty Years of Change in the Bulrush Marshes of the St. Lawrence Estuary and The Impact of the Greater Snow Goose. Wetlands **32**:1175–1188.

Allen, C. D., A. K. Macalady, H. Chenchouni, D. Bachelet, N. McDowell, M. Vennetier, T. Kitzberger, A. Rigling, D. D. Breshears, E. T. Hogg, P. Gonzalez, R. Fensham, Z. Zhang, J. Castro, N. Demidova, J.-H. Lim, G. Allard, S. W. Running, A. Semerci, and N. Cobb. 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. Forest Ecology and Management **259**:660–684.

Andrade, K., C. Corbin, S. Diver, M. V. Eitzel, J. Williamson, J. Brashares, and L. Fortmann. 2014. Finding your way in the interdisciplinary forest: notes on educating future conservation practitioners. Biodiversity and Conservation URL `http://link.springer.com/article/10.1007/s10531-014-0818-z`.

Ansley, J., and J. J. Battles. 1998. Forest composition, structure, and change in an old-growth mixed conifer forest in the northern Sierra Nevada. Journal of the Torrey Botanical Society **125**:297–308.

Banerjee, S., and A. Finley. 2007. Bayesian multi-resolution modeling for spatially replicated data sets with application to forest biomass data. Journal of Statistical Planning and Inference **137**:3193–3205.

Bates, D., and M. Maechler, 2010. lme4: Linear mixed-effects models using S4 classes. URL `http://cran.r-project.org/package=lme4`.

Battles, J. J., T. Robards, A. Das, K. Waring, J. K. Gilless, G. Biging, and F. Schurr. 2008. Climate change impacts on forest growth and tree mortality: a data-driven modeling study in the mixed-conifer forest of the Sierra Nevada, California. Climatic Change **87**:S193–S213.

Beale, C. M., J. J. Lennon, J. M. Yearsley, M. J. Brewer, and D. a. Elston. 2010. Regression analysis of spatial data. Ecology letters **13**:246–64.

Benjamini, Y., and D. Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. Annals of Statistics **29**:1165–1188.

Berner, L. T., P. S. A. Beck, A. G. Bunn, A. H. Lloyd, and S. J. Goetz. 2011. High-latitude tree growth and satellite vegetation indices: Correlations and trends in Russia and Canada (1982 - 2008). Journal of Geophysical Research **116**:1–13.

Betts, M. G., L. M. Ganio, M. M. P. Huso, N. a. Som, F. Huettmann, J. Bowman, and B. a. Wintle. 2009. Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review". Ecography **32**:374–378.

Bigler, C., and H. Bugmann. 2004. Assessing the performance of theoretical and empirical tree mortality models using tree-ring series of Norway spruce. Ecological Modelling **174**:225–239.

Bissonette, A., 1997. Scale-Sensitive Ecological Properties: Historical Context, Current Meaning. Chapter 1, pages 3–31 *in* Wildlife and landscape ecology : effects of pattern and scale.

Bivand, R. S. 2013. Applied Spatial Data Analysis with R. Springer.

Bolker, B., 2012. emdbook: Ecological Models and Data in R, R package version 1.3.2.

Bond-Lamberty, B., A. V. Rocha, K. Calvin, B. Holmes, C. Wang, and M. L. Goulden. 2014. Disturbance legacies and climate jointly drive tree growth and mortality in an intensively studied boreal forest. Global change biology **20**:216–27.

Bunn, A. 2008. A dendrochronology program library in R (dplR). Dendrochronologia **26**:115–124.

Buoro, M., E. Prévost, and O. Gimenez. 2012. Digging through model complexity: using hierarchical models to uncover evolutionary processes in the wild. Journal of evolutionary biology **25**:2077–90.

Burns, R. M., and B. H. Honkala. 1990. Silvics of North America: 1. Conifers; 2. Hardwoods. U.S. Department of Agriculture, Forest Service, Washington, DC. URL `http://www.na.fs.fed.us/spfo/pubs/silvics_manual/table_of_contents.htm`.

Canham, C. D., M. J. Papaik, M. Uriarte, W. H. McWilliams, J. C. Jenkins, and M. J. Twery. 2006. Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. Ecological Applications **16**:540–54.

Cao, Q., and M. Strub. 2008. Evaluation of Four Methods to Estimate Parameters of an Annual Tree Survival and Diameter Growth Model. Forest Science **54**:617–624.

Cao, Q. V. 2000. Prediction of Annual Diameter Growth and Survival for Individual Trees from Periodic Measurements. Forest Science **46**:127–131.

Celeux, G., F. Forbes, C. P. Robert, and D. M. Titterington. 2006. Deviance information criteria for missing data models. Bayesian Analysis **1**:651–673.

Chojnacky, D. C., 2001. On FIA Variables For Ecological Use. Pages 102–105 *in* G. A. Reams, R. E. McRoberts, and P. C. Van Deusen, editors. Proceedings of the second annual Forest Inventory and Analysis symposium, 2000 October 17-18; Salt Lake City, UT. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC.

Clark, J. S. 2007. Models for ecological data: An introduction. Princeton University Press, Princeton, N.J.

Clark, J. S., D. Bell, C. Chu, B. Courbaud, M. Dietze, M. Hersh, J. HilleRisLambers, I. Ibáñez, S. LaDeau, S. McMahon, J. Metcalf, J. Mohan, E. Moran, L. Pangle, S. Pearson, C. Salk, Z. Shen, D. Valle, and P. Wyckoff. 2010. High-dimensional coexistence based on individual variation: a synthesis of evidence. Ecological Monographs **80**:569–608.

Clark, J. S., D. M. Bell, M. H. Hersh, and L. Nichols. 2011. Climate change vulnerability of forest biodiversity: Climate and competition tracking of demographic rates. Global Change Biology **17**:1834–1849.

Clark, J. S., D. M. Bell, M. Kwit, A. Stine, B. Vierra, and K. Zhu. 2012. Individual-scale inference to anticipate climate-change vulnerability of biodiversity. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **367**:236–46.

Clark, J. S., M. Wolosin, M. Dietze, I. Ibáñez, S. LaDeau, M. Welsh, and B. Kloeppel. 2007. Tree growth inference and prediction from diameter censuses and ring widths. Ecological Applications **17**:1942–53.

Cocking, M. I., J. M. Varner, and R. L. Sherriff. 2012. California black oak responses to fire severity and native conifer encroachment in the Klamath Mountains. Forest Ecology and Management **270**:25–34.

Cocking, M. J., J. M. Varner, and E. A. Engber, 2014. Conifer Culprits: An Analysis of Conifer Encroachment in California Oak Woodlands. Page (in press) *in* Proceedings of the 7th Oak Symposium: Managing Oak Woodlands in a Dynamic World. General Technical Report - Pacific Southwest Research Station, USDA Forest Service, Visalia, CA.

Coleman, T. W., Y. Chen, A. D. Graves, S. M. Hishinuma, N. E. Grulke, M. L. Flint, S. J. Seybold, T. O. M. W. Coleman, Y. Chen, A. D. Graves, S. M. Hishinuma, N. E. Grulke, M. L. Flint, and S. J. Seybold. 2014. Developing Monitoring Techniques for the Invasive Goldspotted Oak Borer (Coleoptera: Buprestidae) in California **43**:729–743.

Collins, B., R. Everett, and S. Stephens. 2011. Impacts of fire exclusion and recent managed fire on forest structure in old growth Sierra Nevada mixed-conifer forests. Ecosphere **2**:Article: 51.

Collins, B. M., and G. B. Roller. 2013. Early forest dynamics in stand-replacing fire patches in the northern Sierra Nevada, California, USA. Landscape Ecology **28**:1801–1813.

Conchedda, G., L. Durieux, and P. Mayaux. 2008. An object-based method for mapping and change analysis in mangrove ecosystems. ISPRS Journal of Photogrammetry and Remote Sensing **63**:578–589.

Coomes, D. A., E. R. Lines, and R. B. Allen. 2011. Moving on from Metabolic Scaling Theory: Hierarchical models of tree growth and asymmetric competition for light. Journal of Ecology **99**:748–756.

Cressie, N., C. A. Calder, J. S. Clark, J. M. Ver Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. Ecological Applications **19**:553–570.

Cserhalmi, D., J. Nagy, D. Kristóf, and D. Neidert. 2011. Changes in a Wetland Ecosystem: A Vegetation Reconstruction Study Based on Historical Panchromatic Aerial Photographs and Succession Patterns. Folia Geobotanica **46**:351–371.

Csilléry, K., M. Seignobosc, V. Lafond, G. Kunstler, and B. Courbaud. 2013. Estimating long-term tree mortality rate time series by combining data from periodic inventories and harvest reports in a Bayesian state-space model. Forest Ecology and Management **292**:64–74.

Das, A., J. Battles, N. L. Stephenson, and P. J. van Mantgem. 2011. The contribution of competition to tree mortality in old-growth coniferous forests. Forest Ecology and Management **261**:1203–1213.

Das, A., J. Battles, P. J. van Mantgem, and N. L. Stephenson. 2008. Spatial elements of mortality risk in old-growth forests. Ecology **89**:1744–56.

Das, A. J., J. J. Battles, N. L. Stephenson, and P. J. van Mantgem. 2007. The relationship between tree growth patterns and likelihood of mortality: A study of two tree species in the Sierra Nevada. Canadian Journal of Forest Research **37**:580–597.

de Chant, T., and M. Kelly. 2009. Individual Object Change Detection for Monitoring the Impact of a Forest Pathogen on a Hardwood Forest. Photogrammetric Engineering & Remote Sensing **75**:1005–1013.

de Valpine, P. 2003. Better inferences from population-dynamics experiments using Monte Carlo state-space likelihood methods. Ecology **84**:3064–3077.

de Valpine, P. 2009. Shared challenges and common ground for Bayesian and classical analysis of hierarchical models. Ecological Applications **19**:584–588.

de Valpine, P. 2014. The common sense of P values. Ecology **95**:617–21.

Deng, Y., X. Chen, E. Chuvieco, T. Warner, and J. P. Wilson. 2007. Multi-scale linkages between topographic attributes and vegetation indices in a mountainous landscape. Remote Sensing of Environment **111**:122–134.

Deng, Y., M. F. Goodchild, and X. Chen. 2009. Using NDVI to define thermal south in several mountainous landscapes of California. Computers & Geosciences **35**:327–336.

Desclée, B., P. Bogaert, and P. Defourny. 2006. Forest change detection by statistical object-based method. Remote Sensing of Environment **102**:1–11.

Devine, W. D., and C. A. Harrington. 2006. Changes in Oregon white oak (Quercus garryana Dougl. ex Hook.) following release from overtopping conifers. Trees **20**:747–756.

Devine, W. D., and C. A. Harrington. 2013. Restoration release of overtopped Oregon white oak increases 10-year growth and acorn production. Forest Ecology and Management **291**:87–95.

Dietrich, W. E., D. Bellugi, and R. Real de Asua, 2001. Validation of the Shallow Landslide Model , SHALSTAB , for Forest Management. Pages 195–227 *in* Land Use and Watersheds: Human Influence on Hydrology and Geomorphology in Urban and Forest Areas Water Science and Application Volume 2, volume 2. American Geophysical Union.

Dietrich, W. E., R. Reiss, M.-l. Hsu, and D. R. Montgomery. 1995. A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. Hydrological Processes **9**:383–400.

Dolanc, C. R., J. H. Thorne, and H. D. Safford. 2013. Widespread shifts in the demographic structure of subalpine forests in the Sierra Nevada, California, 1934 to 2007. Global Ecology and Biogeography **22**:264–276.

Dorman, M., T. Svoray, and A. Perevolotsky. 2013. Homogenization in forest performance across an environmental gradient — The interplay between rainfall and topographic aspect. Forest Ecology and Management **310**:256–266.

Dormann, C. F. 2009. Response to Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review". Ecography **32**:379–381.

Dormann, C. F., J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. D. Kissling, I. Kühn, R. Ohlemüller, P. R. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography **30**:609–628.

Dragut, L., O. Csillik, C. Eisank, and D. Tiede. 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. ISPRS Journal of Photogrammetry and Remote Sensing **88**:119–127.

Dragut, L., D. Tiede, and S. R. Levick. 2010. ESP : a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. International Journal of Geographical Information Science **24**:859–871.

Dronova, I., P. Gong, and L. Wang. 2011. Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. Remote Sensing of Environment **115**:3220–3236.

Eitzel, M., M. Kelly, and L. Quinn-Davidson, 2014. Lessons Learned in Historical Mapping of Conifer and Oak in the North Coast. Page (in press) *in* Proceedings of the 7th Oak Symposium: Managing Oak Woodlands in a Dynamic World. General Technical Report - Pacific Southwest Research Station, USDA Forest Service, Visalia, CA.

Eitzel, M. V., J. Battles, R. York, J. Knape, and P. de Valpine. 2013. Estimating tree growth from complex forest monitoring data. Ecological Applications **23**:1288–1296.

Eitzel, M. V., S. Diver, H. Sardiñas, L. M. Hallett, J. J. Olson, A. Romero, G. D. L. T. Oliveira, A. T. Schuknecht, R. Tidmore, and K. N. Suding. 2011. Insights from a Cross-Disciplinary Seminar: 10 Pivotal Papers for Ecological Restoration. Restoration Ecology **20**:147–152.

Ellison, A. M., M. S. Bank, B. D. Clinton, E. A. Colburn, K. Elliott, C. R. Ford, D. R. Foster, B. D. Kloeppel, J. D. Knoepp, G. M. Lovett, J. Mohan, D. A. Orwig, N. L. Rodenhouse, W. V. Sobczak, K. A. Stinson, J. K. Stone, C. M. Swan, J. Thompson, B. V. Holle, and J. R. Webster. 2005. Loss of foundation species: Consequences for the structure and dynamics of forested ecosystems. Frontiers in Ecology and the Environment **3**:479–486.

Ellner, S. P., and M. Rees. 2006. Integral projection models for species with complex demography. American Naturalist **167**:410–428.

ESRI, 2011. ArcGIS. URL `http://www.esri.com/`.

ESRI, 2013. ArcGIS. URL `http://www.esri.com/`.

FAO, 2010. Global Forest Resources Assessment 2010. Technical report, FAO, Rome.

Finley, A. O. 2011. A hierarchical model for quantifying forest variables over large heterogeneous landscapes with uncertain forest areas. Journal of the American Statistical Association **106**:31–48.

Finley, A. O., S. Banerjee, and D. W. MacFarlane. 2011. A Hierarchical Model for Quantifying Forest Variables Over Large Heterogeneous Landscapes With Uncertain Forest Areas. Journal of the American Statistical Association **106**:31–48.

Franklin, J. 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. Progress in Physical Geography **19**:474–499.

Frazer, G., C. Canham, and K. Lertzman, 1999. Gap Light Analyzer (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs, user's manual and program documentation.

Garbarino, M., E. Lingua, P. J. Weisberg, A. Bottero, F. Meloni, and R. Motta. 2013. Land-use history and topographic gradients as driving factors of subalpine Larix decidua forests. Landscape Ecology **28**:805–817.

Gärtner, P., M. Förster, A. Kurban, and B. Kleinschmit. 2014. Object based change detection of Central Asian Tugai vegetation with very high spatial resolution satellite imagery. International Journal of Applied Earth Observation and Geoinformation **31**:110–121.

Gavin, D., 2007. AET. URL `http://geography.uoregon.edu/gavin/software.html`.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). Bayesian Analysis **1**:515–534.

Ghosh, S., A. E. Gelfand, and J. S. Clark. 2012. Inference for Size Demography From Point Pattern Data Using Integral Projection Models. Journal of Agricultural, Biological, and Environmental Statistics **17**:641–677.

Gimenez, O., V. Rossi, R. Choquet, C. Dehais, B. Doris, H. Varella, J.-P. Vila, and R. Pradel. 2007. State-space modelling of data on marked individuals. Ecological Modelling **206**:431–438.

Gimmi, U., and H. Bugmann. 2013. Preface: integrating historical ecology and ecological modeling. Landscape Ecology **28**:785–787.

Gonzalez, P., G. Asner, J. Battles, M. Lefsky, K. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from Lidar, QuickBird, and field inventories in California. Remote Sensing of the Environment **114**:1561–1575.

Gotway, C. A., L. J. Young, L. J. Young, and C. A. Gotway. 2002. Combining Incompatible Spatial Data. Journal of the American Statistical Association **97**:632–648.

Gregoire, T. G., O. Schabenberger, and J. P. Barrettt. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. Canadian Journal of Forest Research **25**:137–156.

Grossinger, R. 2012. Napa Valley Historical Ecology Atlas: Exploring a Hidden Landscape of Transformation and Resilience. Univ of California Press.

Hallett, L. M., S. Diver, M. V. Eitzel, J. J. Olson, B. S. Ramage, H. Sardinas, Z. Statman-Weil, and K. N. Suding. 2013. Do We Practice What We Preach? Goal Setting for Ecological Restoration. Restoration Ecology **21**:312–319.

Haralick, R., K. Shanmugam, and I. Dinstein. 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics. **3**:610–620.

Harcombe, P. 1987. Tree life tables. Bioscience **37**:557–568.

Hawkins, B. A. 2012. Eight (and a half) deadly sins of spatial analysis. Journal of Biogeography **39**:1–9.

He, F. L., and R. P. Duncan. 2000. Density-dependent effects on tree survival in an old-growth Douglas fir forest. Journal of Ecology **88**:676–688.

Huntsinger, L., M. Johnson, M. Stafford, and J. Fried. 2010. Hardwood Rangeland Landowners in California from 1985 to 2004: Production, Ecosystem Services, and Permanence. Rangeland Ecology & Management **63**:324–334.

Huntsinger, L., and J. L. Oveido. 2014. Ecosystem Services are Social – ecological Services in a Traditional Pastoral System: the Case of California's Mediterranean Rangelands. Ecology And Society **19**:8.

Hurst, J. M., R. B. Allen, D. A. Coomes, and R. P. Duncan. 2011. Size-specific tree mortality varies with neighbourhood crowding and disturbance in a Montane Nothofagus forest. PloS one **6**:e26670.

Igor, D., D. Alexander, and N. Vasiliy. 2009. Tree mortality in a mixed deciduous forest in Northwestern Russia over 22 years. Annals of Forest Science **66(411)**:1–11.

Intergraph, 2012. ERDAS Imagine 2013.

Jenkins, R. B., and N. C. Coops. 2011. Landscape Controls on Structural Variation in Eucalypt Vegetation Communities: Woronora Plateau, Australia. Australian Geographer **42**:1–17.

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. 3rd edition. Pearson Prentice Hall, Upper Saddle River, NJ.

Kery, M. 2010. Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses. Elsevier, Amsterdam; Boston.

Kery, M., and M. Schaub. 2012. Bayesian Population Analysis Using WinBUGS: A Hierarchical Perspective. Academic Press, Waltham, MA.

Knape, J., N. Jonzén, M. Sköld, J. Kikkawa, and H. McCallum. 2011. Individual heterogeneity and senescence in silvereyes on Heron Island. Ecology **92**:813–20.

Knape, J., M. Sköld, N. Jonzén, M. Akesson, S. Bensch, B. Hansson, and D. Hasselquist. 2008. An analysis of hatching success in the great reed warbler Acrocephalus arundinaceus. Oikos **117**:430–438.

Kocher, S. D., E. Toman, S. F. Trainor, V. Wright, J. S. Briggs, C. P. Goebel, M. Montblanc, A. Oxarart, D. L. Pepin, T. A. Steelman, A. Thode, and T. A. Waldrop. 2012. How Can We Span the Boundaries between Wildland Fire Science and Management in the United States? Journal of Forestry **110**:421–428.

Kueppers, L. M., M. a. Snyder, L. C. Sloan, E. S. Zavaleta, and B. Fulfrost. 2005. Modeled regional climate change and California endemic oak ranges. Proceedings of the National Academy of Sciences of the United States of America **102**:16281–6.

Kuhn, I., and C. F. Dormann. 2012. Less than eight (and a half) misconceptions of spatial analysis. Journal of Biogeography **39**:995–1003.

Kunstler, G., C. H. Albert, B. Courbaud, S. Lavergne, W. Thuiller, G. Vieilledent, N. E. Zimmermann, and D. A. Coomes. 2011. Effects of competition on tree radial-growth vary in importance but not in intensity along climatic gradients. Journal of Ecology **99**:300–312.

Laliberte, A. S., A. Rango, K. M. Havstad, J. F. Paris, R. F. Beck, R. McNeely, and A. L. Gonzalez. 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. Remote Sensing of Environment **93**:198–210.

Larocque, G. R., L. Archambault, and C. Delisle. 2011. Development of the gap model ZELIG-CFS to predict the dynamics of North American mixed forest types with complex structures. Ecological Modelling **222**:2570–2583.

Levick, S. R., and K. H. Rogers. 2011. Context-dependent vegetation dynamics in an African savanna. Landscape Ecology **26**:515–528.

Likens, G. E., and J. F. Franklin. 2009. Ecosystem Thinking in the Northern Forest - and Beyond. BioScience **59**:511–513.

Lines, E. R., D. A. Coomes, and D. W. Purves. 2010. Influences of forest structure, climate and species composition on tree mortality across the eastern US. Plos One **5**:Article: e13212.

Lunn, D., D. Spiegelhalter, A. Thomas, and N. Best. 2009. The BUGS project: Evolution, critique and future directions. Statistics in Medicine **28**:3049–3067.

Luo, Y., and H. Y. H. Chen. 2013. Observations from old forests underestimate climate change effects on tree mortality. Nature Communications **4**:1655.

Lutz, J. A., A. J. Larson, T. J. Furniss, D. C. Donato, J. A. Freund, M. E. Swanson, K. J. Bible, J. Chen, and J. F. Franklin. 2014. Spatially nonrandom tree mortality and ingrowth maintain equilibrium pattern in an old-growth Pseudotsuga-Tsuga forest. Ecology **95**:2047–2054.

Lutz, J. A., J. W. van Wagtendonk, and J. F. Franklin. 2010. Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. Journal of Biogeography **37**:936–950.

Macfarlane, D. W., and R. K. Kobe. 2006. Selecting models for capturing tree-size effects on growth-resource relationships. Canadian Journal of Forest Research **36**:1695–1704.

Marignani, M., D. Rocchini, D. Torri, A. Chiarucci, and S. Maccherini. 2008. Planning restoration in a cultural landscape in Italy using an object-based approach and historical analysis. Landscape and Urban Planning **84**:28–37.

Martha, T. R., N. Kerle, C. J. van Westen, V. Jetten, and K. Vinod Kumar. 2012. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. ISPRS Journal of Photogrammetry and Remote Sensing **67**:105–119.

McDowell, N., W. T. Pockman, C. D. Allen, D. D. Breshears, N. Cobb, T. Kolb, J. Plaut, J. Sperry, A. West, D. G. Williams, and E. a. Yepez. 2008. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? The New phytologist **178**:719–39.

Meentemeyer, R. K., N. J. Cunniffe, A. R. Cook, J. a. N. Filipe, R. D. Hunter, D. M. Rizzo, and C. a. Gilligan. 2011. Epidemiological modeling of invasion in heterogeneous landscapes: spread of sudden oak death in California (1990–2030). Ecosphere **2**:art17.

Menzel, A., M. Matiu, and T. H. Sparks. 2014. Twenty years of successful papers in Global Change Biology. Global change biology **20**:3587–90.

Metcalf, C., C. Horvitz, S. Tuljapurkar, and D. Clark. 2009*a*. A time to grow and a time to die: a new way to analyze the dynamics of size, light, age, and death of tropical trees. Ecology **90**:2766–2778.

Metcalf, C. J. E., S. M. McMahon, and J. S. Clark. 2009*b*. Overcoming data sparseness and parametric constraints in modeling of tree mortality: a new nonparametric Bayesian model. Canadian Journal of Forest Research **39**:1677–1687.

Michener, W. K., and M. B. Jones. 2012. Ecoinformatics: supporting ecology as a data-intensive science. Trends in ecology & evolution **27**:85–93.

Millar, R. B. 2009. Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' factors. Biometrics **65**:962–9.

Miller, J., J. Franklin, and R. Aspinall. 2007. Incorporating spatial dependence in predictive vegetation models. Ecological Modelling **202**:225–242.

Minnich, R. A., M. G. Barbour, J. H. Burk, and R. F. Fernau. 1995. Sixty Years of Change in Californian Conifer Forests of the San Bernardino Mountains. Conservation Biology & Philosophy **9**:902–914.

Mohan, J. E., J. S. Clark, and W. H. Schlesinger. 2007. Long-term CO2 enrichment of a forest ecosystem: implications for forest regeneration and succession. Ecological Applications **17**:1198–212.

Monserud, R. a., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. Forest Ecology and Management **113**:109–123.

Morgan, J. L., and S. E. Gergel. 2013. Automated analysis of aerial photographs and potential for historic forest mapping. Canadian Journal of Forest Research **43**:699–710.

Moritz, M. a., E. Batllori, R. a. Bradstock, a. M. Gill, J. Handmer, P. F. Hessburg, J. Leonard, S. McCaffrey, D. C. Odion, T. Schoennagel, and A. D. Syphard. 2014. Learning to coexist with wildfire. Nature **515**:58–66.

Müllerová, J., J. Pergl, and P. Pyšek. 2013. Remote sensing as a tool for monitoring plant invasions: Testing the effects of data resolution and image classification approach on the detection of a model plant species Heracleum mantegazzianum (giant hogweed). International Journal of Applied Earth Observation and Geoinformation **25**:55–65.

Murtaugh, P. 2009. Performance of several variable-selection methods applied to real ecological data. Ecology letters **12**:1061–8.

Murtaugh, P. 2014. In defense of P values. Ecology **95**:611–7.

Newman, M. E., K. P. McLaren, and B. S. Wilson. 2014*a*. Assessing deforestation and fragmentation in a tropical moist forest over 68 years; the impact of roads and legal protection in the Cockpit Country, Jamaica. Forest Ecology and Management **315**:138–152.

Newman, M. E., K. P. McLaren, and B. S. Wilson. 2014*b*. Long-term socio-economic and spatial pattern drivers of land cover change in a Caribbean tropical moist forest, the Cockpit Country, Jamaica. Agriculture, Ecosystems & Environment **186**:185–200.

Niinemets, U., and F. Valladares. 2006. Tolerance to shade, drought, and waterlogging of temperate northern hemisphere trees and shrubs. Ecological Monographs **76**:521–547.

Nord-Larsen, T. 2006. Modeling Individual-Tree Growth from Data with Highly Irregular Measurement Intervals. Forest Science **52**:198–208.

Oberg, G. 2011. Interdisciplinary environmental studies: A primer. 1st edition. Wiley-Blackwell, West Sussex, UK.

Olson, C. M., and J. A. Helms, 1996. Forest Growth and Stand Structure at Blodgett Forest Research Station 1933-95, Sierra Nevada Ecosystem Project: Final report to Congress, vol. III, Assessments and scientific basis for management options. Technical report, University of California, Centers for Water and Wildland Resources, Davis, CA. URL `http://pubs.usgs.gov/dds/dds-43/VOL_III/VIII_C16.PDF`.

Pacala, S., C. Canham, J. Saponara, J. Silander, R. Kobe, and E. Ribbens. 1996. Forest models defined by field measurements: Estimation, error analysis and dynamics. Ecological Monographs **66**:1–43.

Palmer, M. a. 2008. Reforming Watershed Restoration: Science in Need of Application and Applications in Need of Science. Estuaries and Coasts **32**:1–17.

Peng, C., Z. Ma, X. Lei, Q. Zhu, H. Chen, W. Wang, S. Liu, W. Li, X. Fang, and X. Zhou. 2011. A drought-induced pervasive increase in tree mortality across Canada's boreal forests. Nature Climate Change **1**:467–471.

Platt, R., and T. Schoennagel. 2009. An object-oriented approach to assessing changes in tree cover in the Colorado Front Range 1938–1999. Forest Ecology and Management **258**:1342–1349.

Plummer, M. 2008. Penalized loss functions for Bayesian model comparison. Biostatistics (Oxford, England) **9**:523–39.

Plummer, M., N. Best, K. Cowles, and K. Vines, 2010. coda: Output analysis and diagnostics for MCMC. URL `http://cran.r-project.org/package=coda`.

Ponciano, J. M., M. L. Taper, B. Dennis, and S. R. Lele. 2009. Hierarchical models in ecology: Confidence intervals, hypothesis testing, and model selection using data cloning. Ecology **90**:356–362.

Pringle, R. M., M. Syfert, J. K. Webb, and R. Shine. 2009. Quantifying historical changes in habitat availability for endangered species: use of pixel- and object-based remote sensing. Journal of Applied Ecology **46**:544–553.

PRISM Climate Group. 2011. Climate Grids URL `http://www.prism.oregonstate.edu/products/`.

R Development Core Team, 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robbins, P., and S. a. Moore. 2013. Ecological anxiety disorder: diagnosing the politics of the Anthropocene. Cultural Geographies **20**:3–19.

Roche, L. M., K. J. Rice, and K. W. Tate. 2012. Oak conservation maintains native grass stands in an oak woodland-annual grassland system. Biodiversity and Conservation **21**:2555–2568.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities. 1st edition. Academic, Amsterdam; Boston.

Schlesinger, W. H. 2010. Translational Ecology. Science **329**:609.

Schmid, M., F. Wickler, K. O. Maloney, R. Mitchell, N. Fenske, and A. Mayr. 2013. Boosted beta regression. PloS one **8**:e61623.

Schriver, M., and R. Sherriff, 2014. Tree Establishment in Oregon White Oak and California Black Oak Woodlands in Northwestern California. Page (in press) *in* Proceedings of the 7th Oak Symposium: Managing Oak Woodlands in a Dynamic World. General Technical Report - Pacific Southwest Research Station, USDA Forest Service, Visalia, CA.

Seidl, R., C. S. Eastaugh, K. Kramer, M. Maroschek, C. Reyer, J. Socha, G. Vacchiano, T. Zlatanov, and H. Hasenauer. 2013. Scaling issues in forest ecosystem management and how to address them with models. European Journal of Forest Research **132**:653–666.

Sheil, D., D. F. Burslem, and D. Alder. 1995. The interpretation and misinterpretation of mortality rate measures. Journal of Ecology **83**:331–333.

Sheil, D., and R. M. May. 1996. Mortality and Recruitment Rate Evaluations in Heterogeneous Tropical Forests. The Journal of Ecology **84**:91–100.

Shirk, J. L., H. L. Ballard, C. C. Wilderman, T. Phillips, A. Wiggins, and R. Jordan. 2012. Public Participation in Scientific Research : a Framework for Deliberate Design. Ecology and Society **17**:art29.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **64**:583–639.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2014. The deviance information criterion: 12 years on. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **76**:485–493.

Standiford, R. B., and T. Scott. 2001. Value of oak woodlands and open space on private property values in southern california. Investigación agraria. Sistemas y recursos forestales **10**:137–152.

Stark, D. T., D. L. Wood, A. J. Storer, and S. L. Stephens. 2013. Prescribed fire and mechanical thinning effects on bark beetle caused tree mortality in a mid-elevation Sierran mixed-conifer forest. Forest Ecology and Management **306**:61–67.

Starrs, P. F., 2001. Perspectives on cultural values of California oaks. Pages 21–30 *in* K. L. Standiford, R. B.; McCreary, D.; Purcell, editor. Proceedings of the fifth symposium on oak woodlands: oaks in California's changing landscape. General Technical Report - Pacific Southwest Research Station, USDA Forest Service PSW-GTR-184, San Diego, CA.

Stow, D., Y. Hamada, L. Coulter, and Z. Anguelova. 2008. Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. Remote Sensing of Environment **112**:1051–1061.

Sturtz, S., U. Ligges, and A. Gelman. 2005. R2WinBUGS: A Package for Running WinBUGS from R. Journal of Statistical Software **12**:1–16.

Suarez, M. L., and T. Kitzberger. 2008. Recruitment patterns following a severe drought: long-term compositional shifts in Patagonian forests. Canadian Journal of Forest Research **38**:3002–3010.

Thorne, J., R. Boynton, L. Flint, A. Flint, and T.-N. Le, 2012. Development and Application of Downscaled Hydroclimatic Predictor Variables for Use in Climate Vulnerability and Assessment Studies. Technical report, University of California, Davis. URL `http://uc-ciee.org/downloads/Development and Application of Downscaled Hydroclimatic Predictor Variables.pdf`.

Thorpe, H. C., and L. D. Daniels. 2012. Long-term trends in tree mortality rates in the Alberta foothills are driven by stand development. Canadian Journal of Forest Research **42**:1687–1696.

Trimble, 2013. eCognition Developer.

USDA Forest Service. 2011. Forest Inventory and Analysis National Program URL `http://www.fia.fs.fed.us/tools-data/`.

USGS, 2011. Seamless Data Warehouse. URL `http://seamless.usgs.gov/`.

Uzoh, F. C., and S. R. Mori. 2012. Applying survival analysis to managed even-aged stands of ponderosa pine for assessment of tree mortality in the western United States. Forest Ecology and Management **285**:101–122.

Valachovic, Y., R. Standiford, and L. Quinn-Davidson, 2014. Can the California Forest Practice Rules Adapt to Address Conifer Encroachment? Page (in press) *in* Proceedings of the 7th Oak Symposium: Managing Oak Woodlands in a Dynamic World. General Technical Report - Pacific Southwest Research Station, USDA Forest Service, Visalia, CA.

Van Dongen, S. 2006. Prior specification in Bayesian statistics: three cautionary tales. Journal of theoretical biology **242**:90–100.

van Lier, O. R., R. a. Fournier, R. L. Bradley, and N. Thiffault. 2009. A multi-resolution satellite imagery approach for large area mapping of ericaceous shrubs in Northern Quebec, Canada. International Journal of Applied Earth Observation and Geoinformation **11**:334–343.

van Mantgem, P. J., and N. L. Stephenson. 2007. Apparent climatically induced increase of tree mortality rates in a temperate forest. Ecology Letters **10**:909–916.

van Mantgem, P. J., N. L. Stephenson, J. C. Byrne, L. D. Daniels, J. F. Franklin, P. Z. Fulé, M. E. Harmon, A. J. Larson, J. M. Smith, A. H. Taylor, and T. T. Veblen. 2009. Widespread increase of tree mortality rates in the western United States. Science **323**:521–4.

Vaupel, J. W., and A. I. Yashin. 1985. Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics. The American Statistician **39**:176–185.

Vitousek, P. M. 1997. Human Domination of Earth's Ecosystems. Science **277**:494–499.

Wang, W., C. Peng, D. D. Kneeshaw, and G. R. Larocque. 2012. Drought-induced tree mortality: ecological consequences, causes, and modeling. Environmental Reviews **20**:109–121.

Wang, Y., X. Hou, M. Wang, L. Wu, L. Ying, and Y. Feng. 2013. Topographic controls on vegetation index in a hilly landscape: a case study in the Jiaodong Peninsula, eastern China. Environmental Earth Sciences **70**:625–634.

Warton, D. I., and F. K. C. Hui. 2011. The arcsine is asinine: the analysis of proportions in ecology. Ecology **92**:3–10.

Weiskittel, A. R., S. M. Garber, G. P. Johnson, D. A. Maguire, and R. A. Monserud. 2007. Annualized diameter and height growth equations for Pacific Northwest plantation-grown Douglas-fir, western hemlock, and red alder. Forest Ecology and Management **250**:266–278.

Whipple, A. A., R. M. Grossinger, and F. W. Davis. 2011. Shifting Baselines in a California Oak Savanna: Nineteenth Century Data to Inform Restoration Scenarios. Restoration Ecology **19**:88–101.

Willmott, C., C. Rowe, and Y. Mintz. 1985. Climatology of the terrestrial seasonal water cycle. Journal of Climatology **5**:589–606.

Wood, S. N. 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton, FL.

Yang, Y., and S. Huang. 2013. A Generalized Mixed Logistic Model for Predicting Individual Tree Survival Probability with Unequal Measurement Intervals. Forest Science **59**:177–187.

Yeh, H.-Y., and L. C. Wensel. 2000. The relationship between tree diameter growth and climate for coniferous species in northern California. Canadian Journal of Forest Research **30**:1463–1471.

York, R., J. Battles, R. Wenk, and S. Saah. 2013. A gap-based approach for regenerating pine species and reducing surface fuels in multi-aged mixed conifer stands in the Sierra Nevada, California. Forestry **85**:203–213.

York, R. A., J. J. Battles, A. K. Eschtruth, and F. G. Schurr. 2011. Giant Sequoia (Sequoiadendron giganteum) Regeneration in Experimental Canopy Gaps. Restoration Ecology **19**:14–23.

York, R. A., J. J. Battles, and R. C. Heald. 2003. Edge effects in mixed conifer group selection openings: tree height response to resource gradients. Forest Ecology and Management **179**:107–121.

Zeppel, M. J., W. R. Anderegg, and H. D. Adams. 2013. Forest mortality due to drought: latest insights, evidence and unresolved questions on physiological pathways and consequences of tree death. New Phytologist **197**:372–374.

Zuidema, P. A., E. Jongejans, P. D. Chien, H. J. During, and F. Schieving. 2010. Integral Projection Models for trees: a new parameterization method and a validation of model output. Journal of Ecology **98**:345–355.

Zuur, A. F., E. N. Ieno, N. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Statistics for Biology and Health, Springer New York, New York, NY.

Zvoleff, A., 2014. glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R. URL `http://cran.r-project.org/package=glcm`.

Appendix A

Number of Trees and Plots in Each Compartment, and Number and Length of Time Intervals

Compartment	Plots	Trees
210	4	31
211	2	21
220	20	123
221	1	60
292	27	176
310	5	22
390	7	58
510	9	104
600	7	18
630	4	68
650	3	8
651	1	1
652	1	1
Total:	91	691

Table A.1: Number of plots and trees in each compartment

Interval	YearSpan	Count
1984-1987	3	48
1977-1981	4	2
1976-1981	5	9
1979-1984	5	60
1980-1985	5	52
1993-1998	5	60
1998-2003	5	54
2004-2009	5	19
1976-1982	6	67
1978-1985	7	36
1994-2001	7	65
1985-1993	8	56
1995-2004	9	122
1985-1995	10	21
1994-2004	10	262
1994-2005	11	8
1982-1994	12	74
1983-1995	12	1
1981-1994	13	75
1980-1994	14	56

Table A.2: Number and length of time intervals

Appendix B

Measurement Methods and Auditing of Explanatory Variables

This appendix details measurement methods for both tree diameter and the other covariates, as well as how these data were audited for outliers or checked against other measurements for accuracy.

B.1 Biotic Factors: Tree Size (Diameter) and Competition (Basal Area)

Much of BFRS was logged and then burned to reduce surface fuels between 1900 and 1913 before the land was given to the University of California, Berkeley. A network of inventory plots (0.04 ha) have been periodically sampled since 1976. All trees 11.4 cm DBH and larger in the plot are tagged and measured with a diameter tape to the nearest 0.1 inch (0.254 cm). Smaller trees (minimum size = 0.254 cm DBH) are tracked in subplots.

Different plots were measured in different years, giving rise to 20 different combinations of start and end years for intervals (see Appendix A). This has resulted in measurements for some trees having years in common and others not. There are three intervals (four censuses) available for most trees. Altogether, the records span 34 years from 1976 to 2010. See Olson and Helms (1996) for more detail on the sampling design.

I audited the data by identifying trees showing potential errors (negative growth, outliers, and inconsistent species identification or survival status). Records with obvious mistakes were corrected (e.g. erroneous species identification) and in some cases, discarded. Outliers above and below three standard deviations were checked against growth chronologies developed from annual tree rings from BFRS (Battles, unpublished data). I discarded records whose growth exceeded the maximum observed from the chronologies using a moving ten-year window (approximately 1 cm/year increase in DBH). Applying this upper bound

effectively removed all trees showing a presumably spurious increase in DBH of more than 10% during the interval. I therefore applied a -10% lower bound to match.

A total of 200 out of 5762 records (these numbers include all species) were corrected, and 76 records were removed as outliers for white fir. Other than removing these records, I kept trees showing negative growth in the dataset to avoid biased parameter estimates and because I explicitly modeled observation error.

I calculated total basal area directly from the inventory data for all species for each plot. I did not have individual tree locations and was therefore unable to use neighborhood techniques (Canham et al. 2006) or spatially correlated random effects (Banerjee and Finley 2007, Finley 2011), but this aggregation of basal area by plot and year is a reasonable way to represent the effect of competition on each plot (Lines et al. 2010).

B.2 Abiotic Factors: Insolation, Annual Water Deficit, Topographic Slope, Elevation, and Soil Type

I represented light availability using yearly insolation values calculated using ArcGIS 9.3 (solar radiation calculator, ESRI (2011)) from a 1/3 arc second digital elevation model from the USGS Seamless Map Database (USGS 2011). These insolation values were checked for consistency against values calculated from a local model using the Gap Light Analyzer (Frazer et al. 1999), parameterized using measurements from a LiCor 2000 Pyranometer (as in York et al. 2003, 2011). Both absolute and relative (solstice/equinox ratios) measurements were consistent.

Moisture availability was represented using annual water deficit, which incorporates both water supply (precipitation) and evaporative potential (temperature) (Willmott et al. 1985). I used data on monthly precipitation and temperature from the Blodgett weather station and day length (calculated from latitude) to estimate annual water deficit from 1963-2010. From these data, I used AET 1.0 (Gavin 2007) to calculate the annual climatic water deficit using the modified Thornthwaite method.

Elevation was measured from a topographic map to the nearest 10 feet (3.05 m). Slope was measured to the nearest % in the field using a clinometer, averaging uphill and downhill measurements. The only categorical covariate included was soil type, which reflects potential nutrient availability. Five soil types are found at Blodgett and all are reflected in the reserve compartments. (Cohasset: Andesite parent rock with high development; Jocal: metasedimentary parent rock; Holland & Holland-Musick: granodiorite parent with intermediate development, Holland-Bighill: granodiorite with little development.) Cohasset soils are typically more productive than the other types.

Soil type, insolation, elevation, and topographical slope were measured at the plot level; I assume that these do not change significantly from year to year. Annual water deficit was calculated for every year, and basal area was calculated for each plot in each year.

For the analysis, all continuous explanatory variables, including tree DBH (size), were centered (mean subtracted) and scaled (divided by standard deviation) before estimation. This standardization was important to improve mixing of the models. For tree sizes, the mean and standard deviation for standardizing were calculated from all measured sizes used in the analysis, and the additional unmeasured (latent) sizes in the model were then defined on this scale. Resulting parameter estimates related to size (DBH) were unscaled (multiplied by standard deviation of size measurements) to provide better interpretability (see Appendix C for algebra on standardizing and unscaling).

Appendix C

Additional Details on Statistical Model and Standardization of Explanatory Variables

This appendix includes a more detailed version of the model specification, including explicitly stating normal distributions for random effects, as well as the algebra for how the variables were centered and scaled, and then later unscaled (but left centered) in the results.

C.1 Full Model Description

Recall that subscript i is for compartment, j is for plot, and k is for tree; superscript m indicates one of several explanatory variables. The process model relating size at $(t+1)$ to the other variables is:

$$x_{ijk}(t+1) \sim \mathcal{N}(\alpha_{ijk}(t) + \beta_{ijk}(t)x_{ijk}(t) + \sum_m \gamma^m z_{ij}^m(t) + \sum_m \kappa^m z_{ij}^m(t)x_{ijk}(t), \sigma_\varepsilon^2) \qquad \text{(C.1)}$$

x_{ijk} is size, while $z_{ij}^m(t)$ are the other explanatory variables. These are measured at plot and/or year level: insolation, topographic slope, elevation, and soil category are all measured at plot level, i.e. z_{ij}^{insol}, z_{ij}^{tslope}, z_{ij}^{elev}, and a group of five indicator variables representing a tree's soil type (z_{ij}^C, z_{ij}^H, z_{ij}^{HB}, z_{ij}^{HM}, and z_{ij}^J) ; basal area is measured at plot and year level, $z_{ij}^{ba}(t)$; and annual water deficit is measured at year level, $z^{def}(t)$. The variables z^m, x, and y are centered and scaled as detailed in the next section.

I assume that size in the next year is a linear function of the $z_{ij}^m(t)$ and have parameters for slope γ^m and interaction with size κ^m. I also assume that size in the next year $x_{ijk}(t+1)$ is a linear function of size in the previous year, $x_{ijk}(t)$, with soil type-dependent slope and intercept (e.g., for Cohasset: slope $a_C = \beta_{ijk}(t) + \kappa^C$ and intercept $b_C = \alpha_{ijk}(t) + \gamma^C$). I report the average growth increment $\bar{b}$ and average effect of size on growth increment $\bar{a} - 1$,

which are weighted averages over soil types and for average values of explanatory variables. Residual error, with variance σ^2_ε, accounts for additional unexplained variation in growth.

At the next hierarchical level, I model the collective random effects on intercept $\alpha_{ijk}(t)$ and the random effects on size slope $\beta_{ijk}(t)$ as a combination of random effects for tree (q^α_{ijk} and q^β_{ijk}), plot (p^α_{ij} and p^β_{ij}), compartment (c^α_i and c^β_i), and year ($w^\alpha(t)$ and $w^\beta(t)$). The intercept effects reflect differences in overall growth increment while the slope effects reflect differences in growth as a function of size. The random effect intercept and slope for a specific tree is determined by the random tree, plot, compartment and year effects as follows:

$$\alpha_{ijk}(t) = q^\alpha_{ijk} + c^\alpha_i + p^\alpha_{ij} + w^\alpha(t) \quad \beta_{ijk}(t) = q^\beta_{ijk} + c^\beta_i + p^\beta_{ij} + w^\beta(t) \tag{C.2}$$

The random effects for years, compartments, plots, and individuals follow normal distributions,

$$\begin{array}{llll} w^\alpha(t) \sim \mathcal{N}(0, \sigma^2_{\alpha,w}), & c^\alpha_i \sim \mathcal{N}(0, \sigma^2_{\alpha,c}), & p^\alpha_{ij} \sim \mathcal{N}(0, \sigma^2_{\alpha,p}), & q^\alpha_{ijk} \sim \mathcal{N}(0, \sigma^2_{\alpha,q}), \\ w^\beta(t) \sim \mathcal{N}(0, \sigma^2_{\beta,w}), & c^\beta_i \sim \mathcal{N}(0, \sigma^2_{\beta,c}), & p^\beta_{ij} \sim \mathcal{N}(0, \sigma^2_{\beta,p}), & q^\beta_{ijk} \sim \mathcal{N}(0, \sigma^2_{\beta,q}) \end{array} \tag{C.3}$$

respectively. At each level of nesting, random effects are assumed to be independent.

$x_{ijk}(t)$ is not measured perfectly (a latent state), and so $y_{ijk}(t)$ is the size actually measured with some observation error:

$$y_{ijk}(t) \sim \mathcal{N}(x_{ijk}(t), \sigma^2_{\text{DBH}}) \tag{C.4}$$

C.2 Standardizing and Unscaling of Explanatory Variables

One can combine equations C.1 and C.4 in a more algebraic form:

$$y_{ijk}(t+1) = \alpha_{ijk} + \beta_{ijk} x_{ijk}(t) + \sum_m \gamma^m z^m_{ij}(t) + \sum_m \kappa^m z^m_{ij}(t) x_{ijk}(t) + \epsilon_{res} + \epsilon_{obs} \tag{C.5}$$

(recall that ϵ_{res} and ϵ_{obs} are normally distributed with mean zero and variance σ^2_ε and σ^2_{DBH}, respectively)

One can expand this using equation C.2:

$$\begin{aligned} y_{ijk}(t+1) = {} & q^\alpha_{ijk} + c^\alpha_i + p^\alpha_{ij} + w^\alpha(t) + \left(q^\beta_{ijk} + c^\beta_i + p^\beta_{ij} + w^\beta(t) \right) x_{ijk}(t) \\ & + \sum_m \gamma^m z^m_{ij}(t) + \sum_m \kappa^m z^m_{ij}(t) x_{ijk}(t) + \epsilon_{res} + \epsilon_{obs} \end{aligned} \tag{C.6}$$

In reality, I am estimating those slopes, intercepts, random effects, and random effect standard deviations for standardized covariates. For true latent size x and measured size y, I have standardized by subtracting the mean and dividing by the standard deviation for the measured sizes y (μ_y and σ_y). I have done the same for the covariates z^m (with σ_{z^m} and μ_{z^m}). When I unscale (but leave the variables centered) for interpretation, each of the continuous covariates will need to be unscaled by its respective standard deviation.

This is how I have scaled the original variables:

$$x'_{ijk}(t) = \frac{x_{ijk}(t)-\mu_y}{\sigma_y} \quad y'_{ijk}(t) = \frac{y_{ijk}(t)-\mu_y}{\sigma_y} \quad {z^m}'_{ij}(t) = \frac{z^m_{ij}(t)-\mu_{z^m}}{\sigma_{z^m}} \tag{C.7}$$

The equation I have estimated the parameters for is in the primed variables:

$$\begin{aligned} y'_{ijk}(t+1) = {q^\alpha_{ijk}}' + c^{\alpha\prime}_i + {p^\alpha_{ij}}' + w^{\alpha\prime}(t) + \left({q^\beta_{ijk}}' + c^{\beta\prime}_i + {p^\beta_{ij}}' + w^{\beta\prime}(t)\right) x'_{ijk}(t) \\ + \sum_m \gamma^{m\prime} z^{m\prime}_{ij}(t) + \sum_m \kappa^{m\prime} z^{m\prime}_{ij}(t) x_{ijk}(t)' + \epsilon'_{res} + \epsilon'_{obs} \end{aligned} \tag{C.8}$$

Therefore, I wish to know the relationship between these primed parameters and equations in unscaled variables. When I unscale the covariates, I only wish to multiply by the standard deviation to restore the units of each covariate. I leave the variables centered, such that interpretation of the parameter estimates refers to the covariates at their mean values. Thus, I want an equation in the following double-primed variables:

$$x''_{ijk}(t) = x_{ijk}(t) - \mu_y \quad y''_{ijk}(t) = y_{ijk}(t) - \mu_y \quad z^m_{ij}(t)'' = z^m_{ij}(t) - \mu_{z^m} \tag{C.9}$$

First I write the equation whose parameters I estimated (equation C.8) in terms of the unprimed original variables:

$$\begin{aligned} \frac{y_{ijk}(t+1)-\mu_y}{\sigma_y} = {q^\alpha_{ijk}}' + c^{\alpha\prime}_i + {p^\alpha_{ij}}' + w^{\alpha\prime}(t) + \left({q^\beta_{ijk}}' + c^{\beta\prime}_i + {p^\beta_{ij}}' + w^{\beta\prime}(t)\right) \frac{x_{ijk}(t)-\mu_y}{\sigma_y} \\ + \sum_m \gamma^{m\prime} \frac{z^m_{ij}(t)-\mu_{z^m}}{\sigma_{z^m}}(t) + \sum_m \kappa^{m\prime} \frac{x_{ijk}(t)-\mu_y}{\sigma_y} \frac{z^m_{ij}(t)-\mu_{z^m}}{\sigma_{z^m}} + \epsilon'_{res} + \epsilon'_{obs} \end{aligned} \tag{C.10}$$

Substituting x'', y'' and $z^{m\prime\prime}$ into the primed equation, I have:

$$\begin{aligned} \frac{y''_{ijk}(t+1)}{\sigma_y} = {q^\alpha_{ijk}}' + c^{\alpha\prime}_i + {p^\alpha_{ij}}' + w^{\alpha\prime}(t) + \left({q^\beta_{ijk}}' + c^{\beta\prime}_i + {p^\beta_{ij}}' + w^{\beta\prime}(t)\right) \frac{x''_{ijk}(t)}{\sigma_y} \\ + \sum_m \gamma^{m\prime} \frac{{z^m_{ijk}}''(t)}{\sigma_{z^m}} + \sum_m \kappa^{m\prime} \frac{x''_{ijk}(t)}{\sigma_y} \frac{{z^m_{ijk}}''(t)}{\sigma_{z^m}} + \epsilon'_{res} + \epsilon'_{obs} \end{aligned} \tag{C.11}$$

Now, multiply through by σ_y:

$$y''_{ijk}(t+1) = \sigma_y \left(q^{\alpha}_{ijk}{}' + c^{\alpha\prime}_i + p^{\alpha\prime}_{ij} + w^{\alpha\prime}(t)\right) + \sigma_y \left(q^{\beta}_{ijk}{}' + c^{\beta\prime}_i + p^{\beta\prime}_{ij} + w^{\beta\prime}(t)\right) \frac{x''_{ijk}(t)}{\sigma_y} + \sigma_y \sum_m \gamma^{m\prime} \frac{z^{m\prime\prime}_{ij}(t)}{\sigma_{z^m}} + \sigma_y \sum_m \kappa^{m\prime} \frac{x''_{ijk}(t)}{\sigma_y} \frac{z^{m\prime\prime}_{ij}(t)}{\sigma_{z^m}} + \sigma_y \epsilon'_{res} + \sigma_y \epsilon'_{obs} \tag{C.12}$$

After canceling σ_y, explanatory variable slope and intercept parameters scaled as follows:

$$\alpha_{ijk}{}''(t) = \sigma_y \alpha_{ijk}{}'(t) \quad \beta_{ijk}{}''(t) = \beta_{ijk}{}'(t) \quad \gamma^{m\prime\prime} = \tfrac{\sigma_y}{\sigma_{z^m}} \gamma^{m\prime} \quad \kappa^{m\prime\prime} = \tfrac{1}{\sigma_{z^m}} \kappa^{m\prime} \tag{C.13}$$

Observation error and residual error epsilons scale like intercepts:

$$\epsilon''_{res} = \sigma_y \epsilon'_{res} \quad \epsilon''_{obs} = \sigma_y \epsilon'_{obs} \tag{C.14}$$

Random effects scale as follows (slope effects re unscaled, intercept effects are multiplied by size standard deviation σ_y):

$$\begin{array}{cccc} w^{\alpha}(t)'' = \sigma_y w^{\alpha}(t)', & c^{\alpha\prime\prime}_i = \sigma_y c^{\alpha\prime}_i, & p^{\alpha}_{ij}{}'' = \sigma_y p^{\alpha\prime}_{ij}, & q^{\alpha}_{ijk}{}'' = \sigma_y q^{\alpha}_{ijk}{}', \\ w^{\beta}(t)'' = w^{\beta}(t)', & c^{\beta\prime\prime}_i = c^{\beta\prime}_i, & p^{\beta}_{ij}{}'' = p^{\beta}_{ij}{}', & q^{\beta}_{ijk}{}'' = q^{\beta}_{ijk}{}' \end{array} \tag{C.15}$$

I assume that standard deviations for random effects (i.e. $\sigma^2_{\alpha,w}$, $\sigma^2_{\beta,w}$, $\sigma^2_{\alpha,q}$, $\sigma^2_{\beta,q}$, $\sigma^2_{\alpha,c}$, $\sigma^2_{\beta,c}$, $\sigma^2_{\alpha,p}$, $\sigma^2_{\beta,p}$, equation C.3) scale in the same manner as the random effects themselves, which I confirmed through simulations (not shown).

Appendix D

Details on Model Estimation and the Evaluation of Model Results

This appendix gives details on model estimation and the evaluation of model results, including specification of priors and selection of initial values for chains, results of Gelman-Rubin convergence assessments; comparisons of posteriors and priors; and tradeoffs between observation and residual standard deviations.

D.1 Prior Specification

All but one of the priors used in estimation of the model were uninformative. Uninformative priors for means were a normal distribution with a high standard deviation (low precision: 1.0e-6) and for standard deviations, uniform priors over the range [0,100]. The first latent size of each tree was given a non-informative (normal, low precision) prior.

For the observation error, I observed very poor mixing for priors which allowed small values (including zero), as the system struggled to move away from the scenario in which the latent states exactly matched the observations. The resulting observation error posteriors were sensitive to the degree to which the priors excluded small values (Figure D.1).

Because some observation error is certainly present in my system, both due to the precision of the diameter tape and due to bark loss and human error, I chose two different reasonably flat, otherwise uninformative priors with a nonzero lower limit. I chose this lower limit based on the rounding error inherent in the diameter tape, which I considered to be a minimum. Reasoning that the 0.1 inch increment labeled on the tape leads to true observations that are 0.05 inches in either direction being uniformly attributed to the mark on the tape in the center of that interval, and assuming x is the truth, while y is the measurement, I have:

$$\begin{aligned} y &= x + u \\ u &\sim \text{uniform}(-a/2, a/2) \end{aligned} \tag{D.1}$$

(Where $a = 0.1$) I am essentially taking the case where the only error is due to rounding on the tape to mean that $\epsilon_{obs} \sim \text{uniform}(-a/2, a/2)$. To establish the correct standard deviation associated with this model for observation error (to then use as a lower bound in a prior for this node in the larger model), I turn to the definition of the uniform distribution from $-a/2$ to $a/2$:

$$f(u) = \begin{cases} 1/a, & \text{for } -a/2 \leq u \leq a/2 \\ 0, & \text{otherwise} \end{cases} \tag{D.2}$$

And the definition of variance as the second central moment:

$$\begin{aligned} V[u] &= \textstyle\int_{-\frac{a}{2}}^{\frac{a}{2}} u^2 f(u) du \\ V[u] &= \textstyle\int_{-\frac{a}{2}}^{\frac{a}{2}} u^2 \frac{1}{a} du \\ V[u] &= \left. \frac{u^3}{3} \frac{1}{a} \right|_{-\frac{a}{2}}^{\frac{a}{2}} \\ V[u] &= \frac{a^2}{24} - \frac{-a^2}{24} \\ V[u] &= \frac{a^2}{12} \end{aligned} \tag{D.3}$$

So, the standard deviation is $a/\sqrt{12}$. Plugging in $a = 0.1$ and converting from inches to my standardized scale, I have the lower limit for the standard deviation set at 0.0036. For my two candidate priors, I therefore use a uniform distribution from [0.0036, 100] on the observation error standard deviation and a gamma prior with shape and rate (.01,0.0001) on the observation error precision. These parameters for the gamma on precision gave rise to a threshold in the standard deviation at approximately the right location. The gamma prior on the precision results in an inverse gamma prior on the standard deviation.

I also tried a uniform prior on observation error standard deviation from [0,100] on a simplified model with several covariates left out (elevation, insolation, and annual deficit), which has the effect of allowing values as low as zero (and which mixes slightly better due to the simpler covariate structure). Finally, I tried a gamma with shape and rate (0.0001,0.0001), which has the effect of excluding values below 0.005 on the standardized scale (Figure D.1).

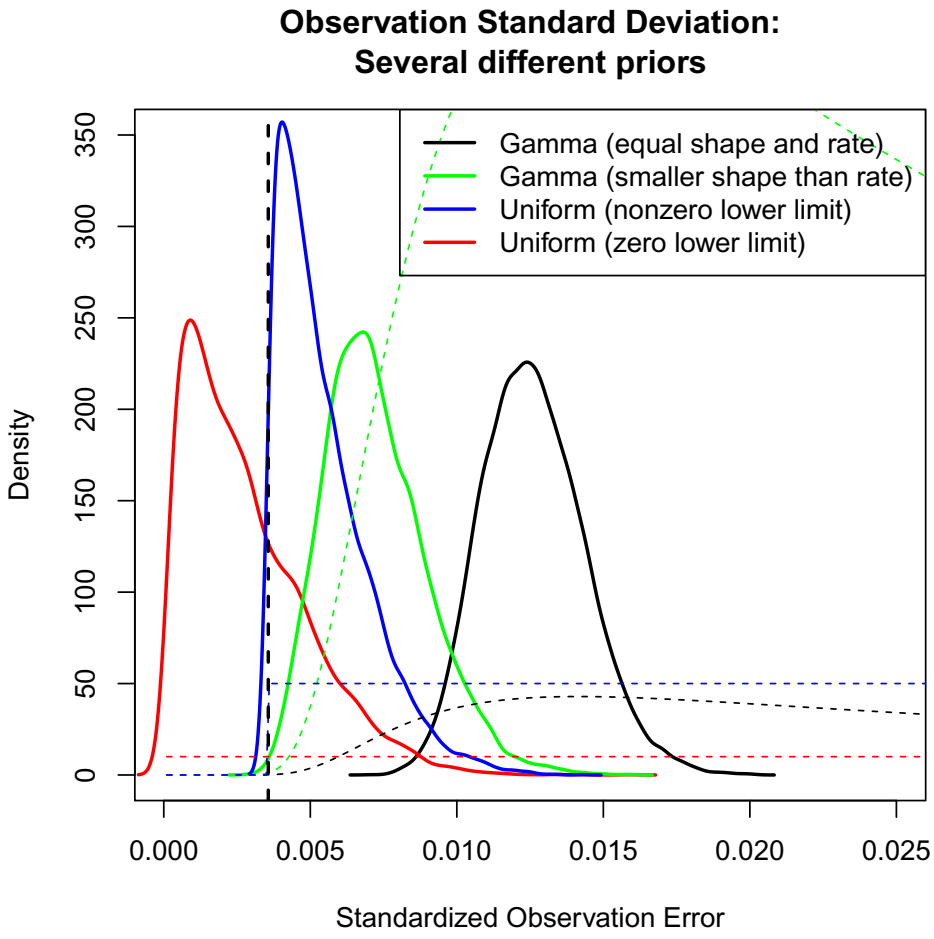

Figure D.1: Four different models with different observation error σ_{DBH} priors. The blue, black, and green lines are priors (dashed) and posteriors (solid) for the full model described in Appendix C; the red lines are for a reduced model which does not include annual water deficit, elevation, or insolation, but is otherwise similar to the other models. The green and black posteriors correspond to gamma priors, and the red and blue to uniform priors. The priors (dashed lines) have been multiplied by powers of ten in order to be visible on the same scale as the posteriors; they are not all multiplied by the same factor in order to be visible alongside each other as well. Note that the green gamma prior results in a much more quickly increasing allowance of small values, hence the position of the green posterior closer to zero than the black posterior. Similarly, note that the red posterior (corresponding to the prior which does not exclude zero) is very close to zero, while the blue posterior is very close to the diameter tape-based lower limit established above, which is shown in a dashed black vertical line. The model and estimation procedure biases the observation error as close to zero as it is allowed to be.

Though these appendices do not include chains for all the models, the general behavior was that mixing was better the higher the threshold of exclusion (the more small values excluded). The lower the threshold, the worse the mixing. So mixing was best with the inverse gamma with equal shape and rate, and worst for the uniform prior allowing zero values.

This can be interpreted to mean that the observation error cannot truly be estimated with my model, data, and MCMC sampler (OpenBUGS); but for the results in Chapter 1 I chose the uniform prior with the cutoff at 0.0036 because it is reasonable to assume at least some observation error is reasonable and will improve other estimates.

D.2 Estimation

OpenBUGS was run through R using R2WinBUGS (Sturtz et al. 2005) long enough to achieve sufficient mixing. The results in chapter 1 are from a five day run on a Dell Studio XPS with an Intel Core i7 CPU, 2.93 GHz (4 cores) and 12 GB of RAM, running 64-bit Windows Vista Ultimate. In order to use R2WinBUGS, I had to use the 32-bit version of R. See online supplement for model code, data, and initial values and Figure D.2 for trace plots of MCMC chains. With the included supplements,[1] it is possible to reproduce my parameter estimates, though the MCMC may need to be run for several weeks depending on computer speeds. 95% credible intervals were calculated using the "`coda`" package (Plummer et al. 2010).

D.3 Initial Values

Due to slow mixing, randomly generated initial values from BUGS tended to produce chains that would not converge. Thus, initial values were primarily generated from interpolating the sizes of trees using smoothing splines between inventories (Wood 2006) and fitting linear mixed effects models for random effects using R package "`lme4`" (Bates and Maechler 2010). For the two chains used to test convergence, initial values were generated by adding a small amount of randomly generated variation to the original initial values (normal in the case of the slopes and intercepts, uniform in the case of the variances).

D.4 Assessment of Convergence: Gelman-Rubin Statistics and Trace Plots

Assessing the mixing and convergence of chains is central to MCMC. If a chain is not mixing and has not converged, then the results are not interpretable. Gelman-Rubin diagnostics

[1] Available at Ecological Archives: http://www.esapubs.org/archive/appl/A023/067/

require running multiple chains and comparing them to assess convergence through a plot or single diagnostic value. These functions are built into OpenBUGS; I do not show the plots here but do display the convergence statistics, which should be close to 1 (Table D.1). I also give examples in Figure D.2 of some of my trace plots which show the actual values in the chains. In this case, the analyst is looking for a lack of pattern in the trace: no trends and a trace that looks like noise. Based on the Gelman-Rubin diagnostics and my trace plots, I conclude that mixing is adequate.

Name	PointEstimate	UpperCredibleInterval
res.sd	1.00129	1.00611
obs.sd	1.00028	1.00156
i.overall	1.0049	1.02105
s.size	1.0005	1.00072
i.year.sd	1.00633	1.03022
i.tree.sd	1.00004	1.00006
i.plot.sd	1.00015	1.0008
i.comp.sd	1.00396	1.00886
s.year.sd	1.00259	1.01073
s.tree.sd	1.0281	1.07996
s.plot.sd	1	1.00005
s.comp.sd	1.00029	1.00149
s.slope	0.999966	0.999998
s.ba	1.00008	1.00059
soil.mean	1.00418	1.00423
s.anndef	1.0018	1.00714
int.size.ba	1.00001	1.00012
int.insol.anndef	1.0036	1.00934
s.elev	0.99999	1.00006
s.insol	1.00046	1.00238
int.size.insol	1.00073	1.00326
int.size.slope	0.999984	1.00002
int.size.elev	0.999958	0.999959
int.size.anndef	1.02702	1.11724
int.size.soil	1.00132	1.0006
deviance	1.00029	1.00164

Table D.1: Gelman-Rubin statistics for tracked nodes. Names of variables are internal to the model. "s" refers to a main effect γ^m for a covariate z^m, or to a slope random effect standard deviation σ_β; "int.size" refers to interactions κ^m between size x and covariates z^m, and "i" refers to an intercept random effect σ_α or to the overall mean for Cohasset soils, "i.overall" (γ^C). See supplement for model code, where these names are used.

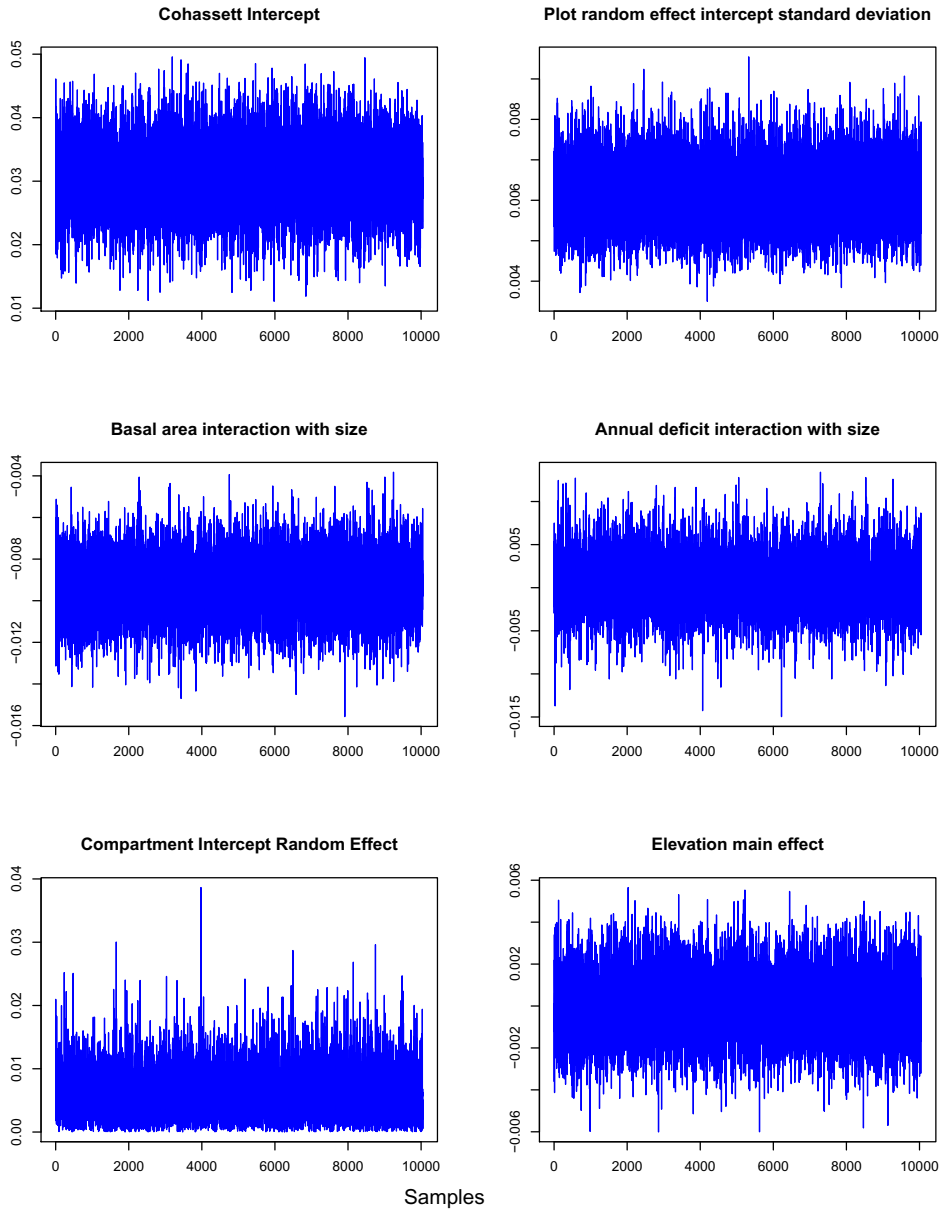

Figure D.2: Trace plots showing the "white noise" behavior indicative of good mixing for six parameters.

D.5 Comparison of Posteriors and Priors: Bayesian Learning Figures

One way to assess the information in the data and the relative influence of the priors is to compare posterior distributions for parameters with their prior distributions and initial values. All but one of my priors (observation error) are uninformative, and the posteriors and priors for that estimate are shown in Figure D.1. All the other posteriors are clearly distinguished from their priors. Initial values, especially for the standard deviations, are typically clearly outside the main support of the posterior (the range of values where the posterior is concentrated), indicating that sensitivity to initial values was minimal within this range. The posteriors for both chains tend to overlap. When the initial values are very different from the posteriors, the two chain posteriors are indistinguishable due to their overlap.

Figure D.3 shows residual error and observation error standard deviations, Figures D.4 and D.5 show random effect standard deviations, and Figures D.6 and D.7 show covariate main effects and interactions with size.

D.6 Tradeoff Between Observation and Residual Standard Deviation

Observation standard deviation and residual standard deviation, while mathematically identifiable, can be weakly identifiable in practice, depending on the prior used for the observation error. For more informative priors (with a higher threshold, excluding a wider range of small values), these two nodes in the model do trade off in the MCMC. Figure D.8 shows MCMC samples from the two variables exhibiting this tradeoff, that when one node becomes larger, the other becomes smaller (they are slightly inversely related), but this effect is most pronounced in the model with the gamma prior with equal shape and rate. This correlation between observation and residual error is not extreme, however: the tradeoff is never so large that one node goes to zero when the other node is large, not even in the model where this tradeoff is most pronounced.

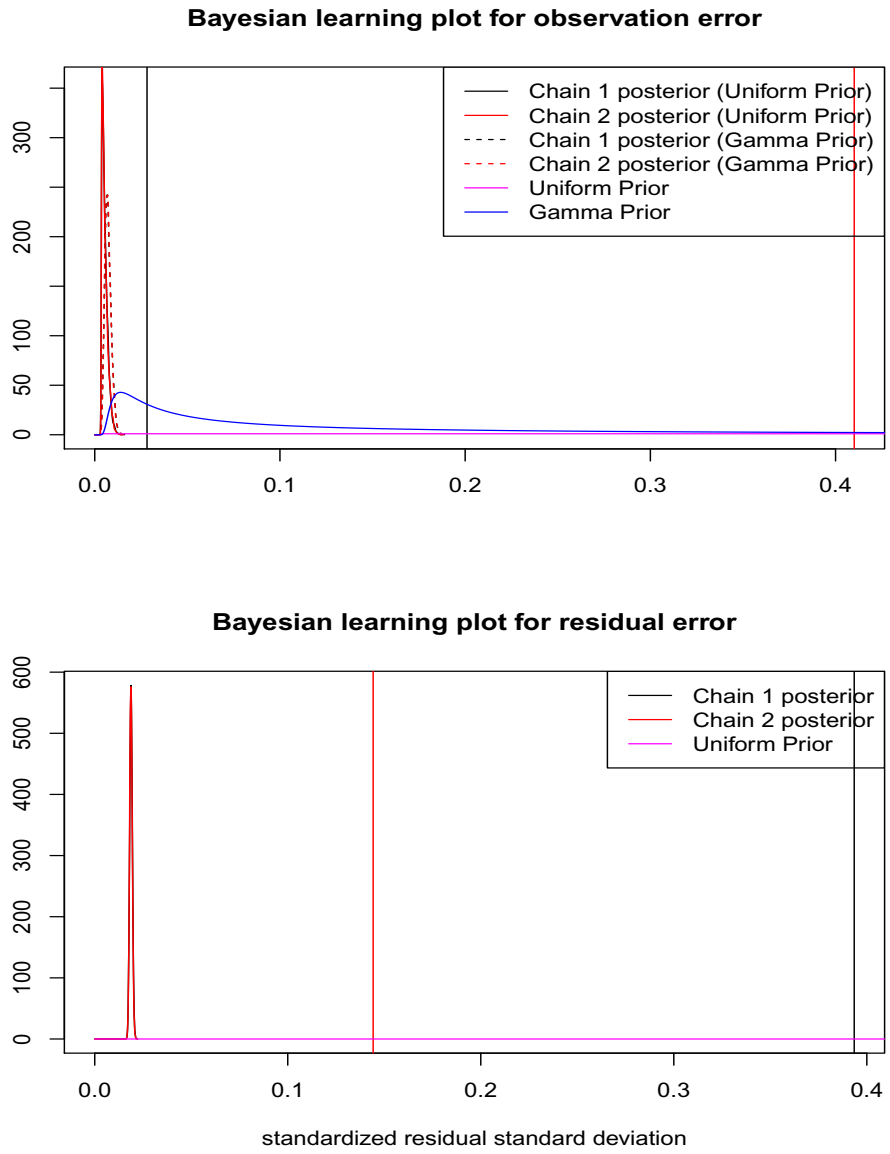

Figure D.3: Bayesian learning figures for observation error σ_{DBH} and residual error σ_{ε}. Priors in the observation error figure have been multiplied by 100 for visibility. The details of the priors for observation error are difficult to discern (See Figure D.1 for a closer look), but the initial values can clearly be quite large and the model will still converge to values close to zero. The red and black are two different chains; posteriors are shown with corresponding initial values as vertical lines in matching colors. Priors are shown in magenta (uniform), or blue (gamma). All posteriors are given on the standardized unitless scale. The two chain posteriors are indistinguishable due to their overlap on both figures, even with the different priors on observation error.

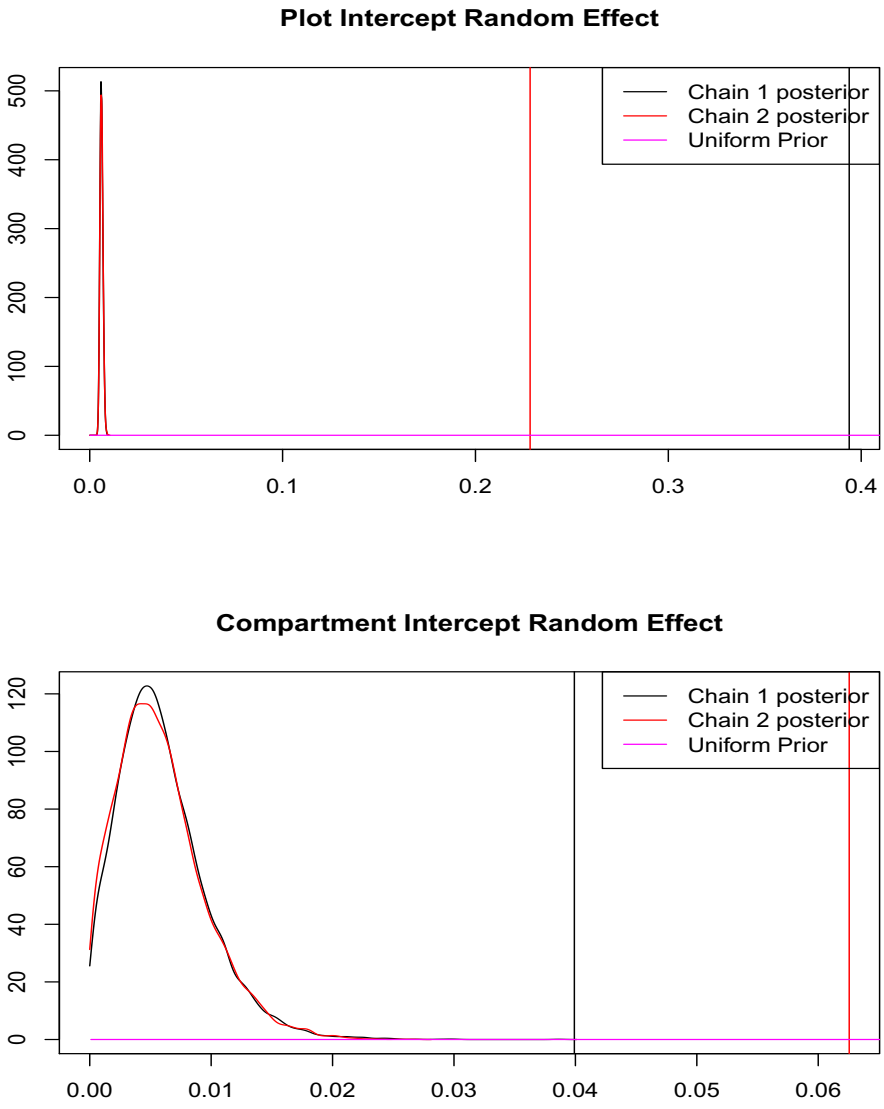

Figure D.4: Bayesian learning figures for plot and compartment intercept random effect standard deviations ($\sigma_{\alpha,p}$ and $\sigma_{\alpha,c}$). The red and black are two different chains; posteriors are shown with corresponding initial values as vertical lines in matching colors. Uniform prior is shown in magenta. All posteriors are given on the standardized unitless scale. The two chain posteriors are indistinguishable due to their overlap for the plot intercept standard deviation.

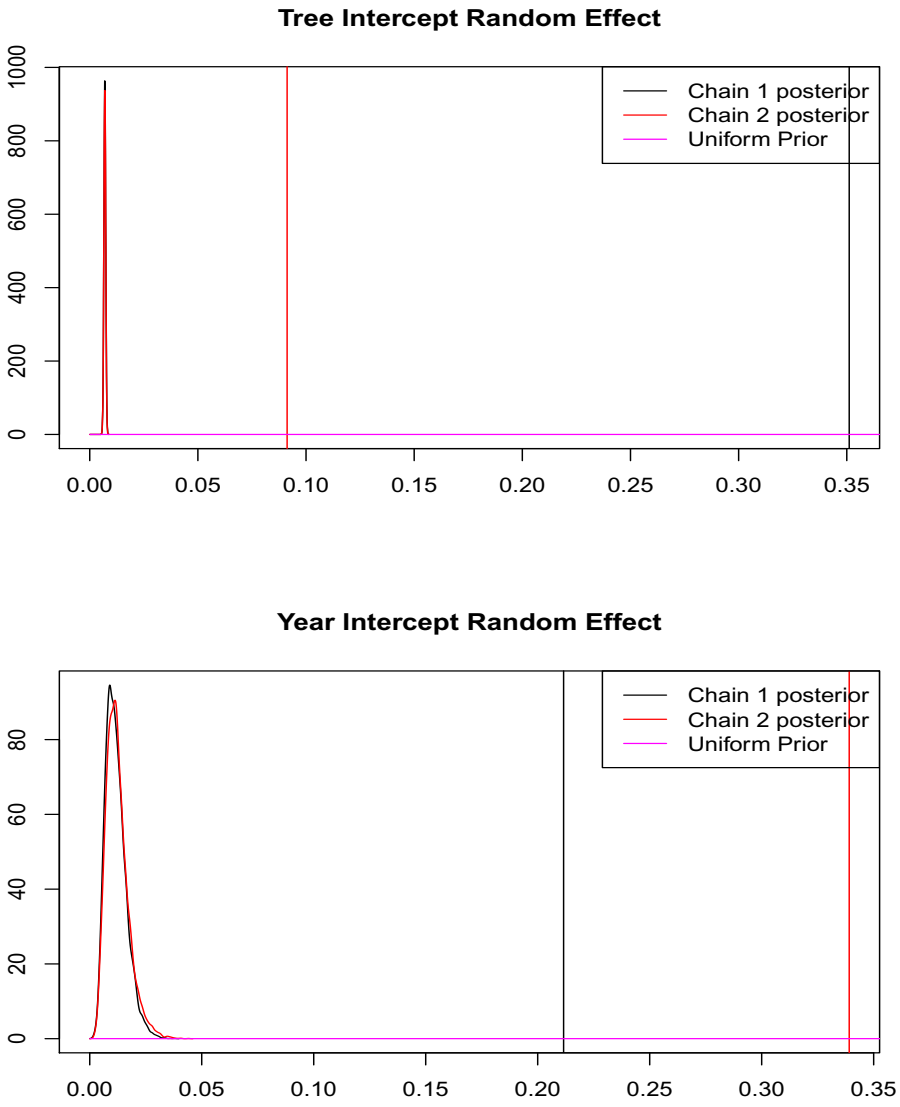

Figure D.5: Bayesian learning figures for tree and year intercept random effect standard deviations ($\sigma_{\alpha,q}$ and $\sigma_{\alpha,w}$). The red and black are two different chains; posteriors are shown with corresponding initial values as vertical lines in matching colors. Uniform prior is shown in magenta. All posteriors are given on the standardized unitless scale. The two chain posteriors are indistinguishable due to their overlap for the tree intercept standard deviation.

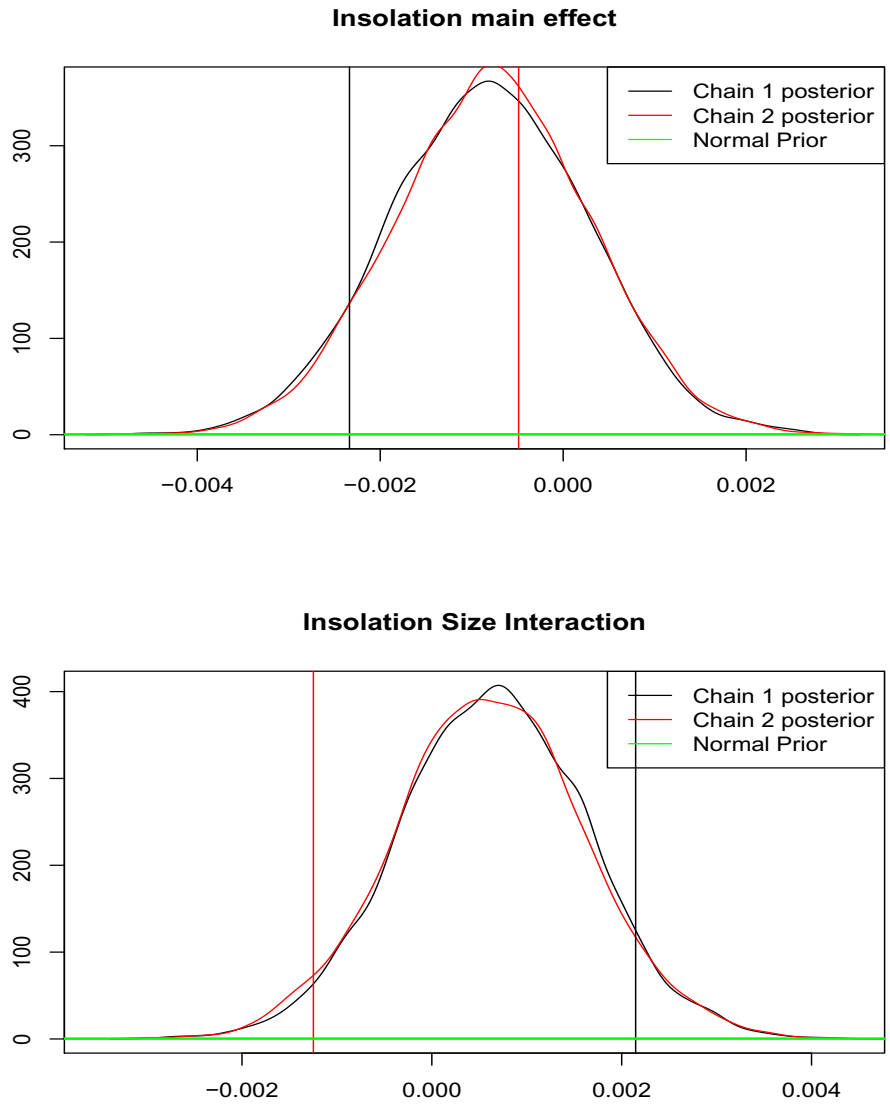

Figure D.6: Bayesian learning figures for insolation main effect γ^{insol} and size interaction κ^{insol}. The red and black are two different chains; posteriors are shown with corresponding initial values as vertical lines in matching colors. Normal prior is shown in green. All posteriors are given on the standardized unitless scale.

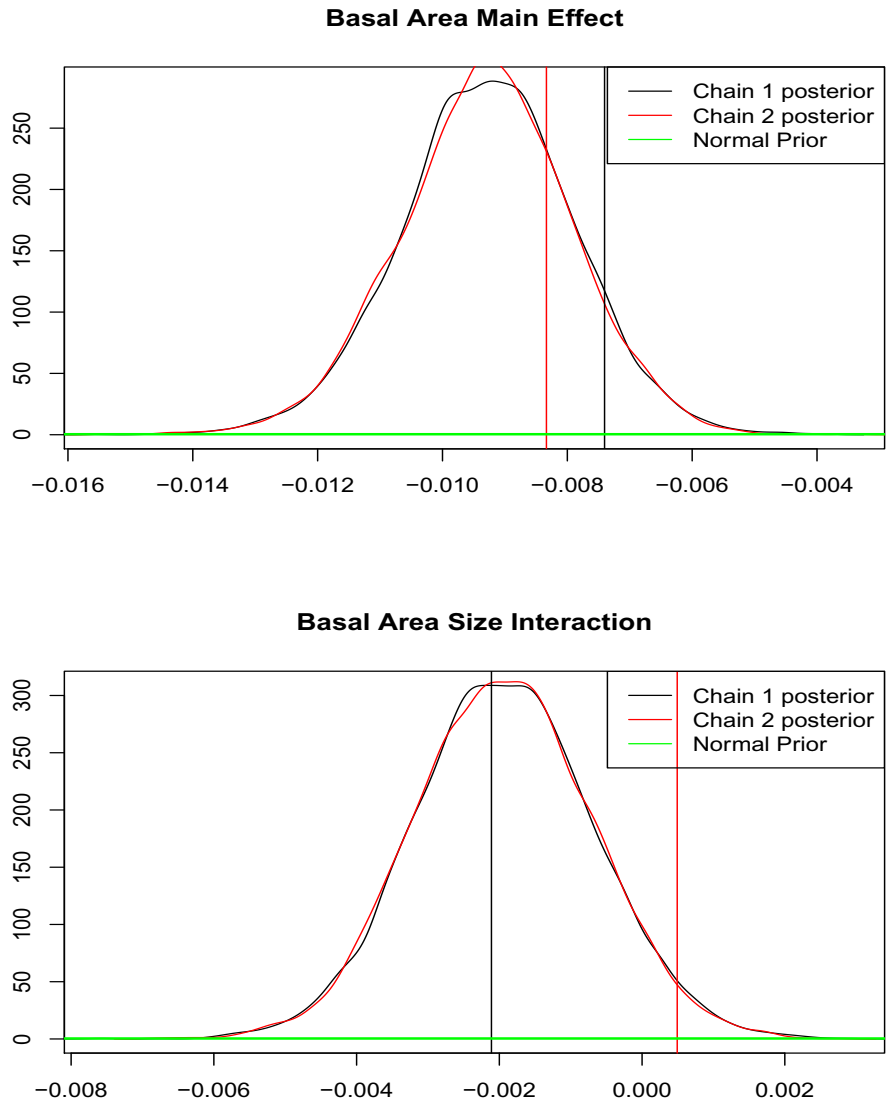

Figure D.7: Bayesian learning figures for basal area main effect γ^{ba} and interaction with size κ^{ba}. The red and black are two different chains; posteriors are shown with corresponding initial values as vertical lines in matching colors. Normal prior is shown in green. All posteriors are given on the standardized unitless scale.

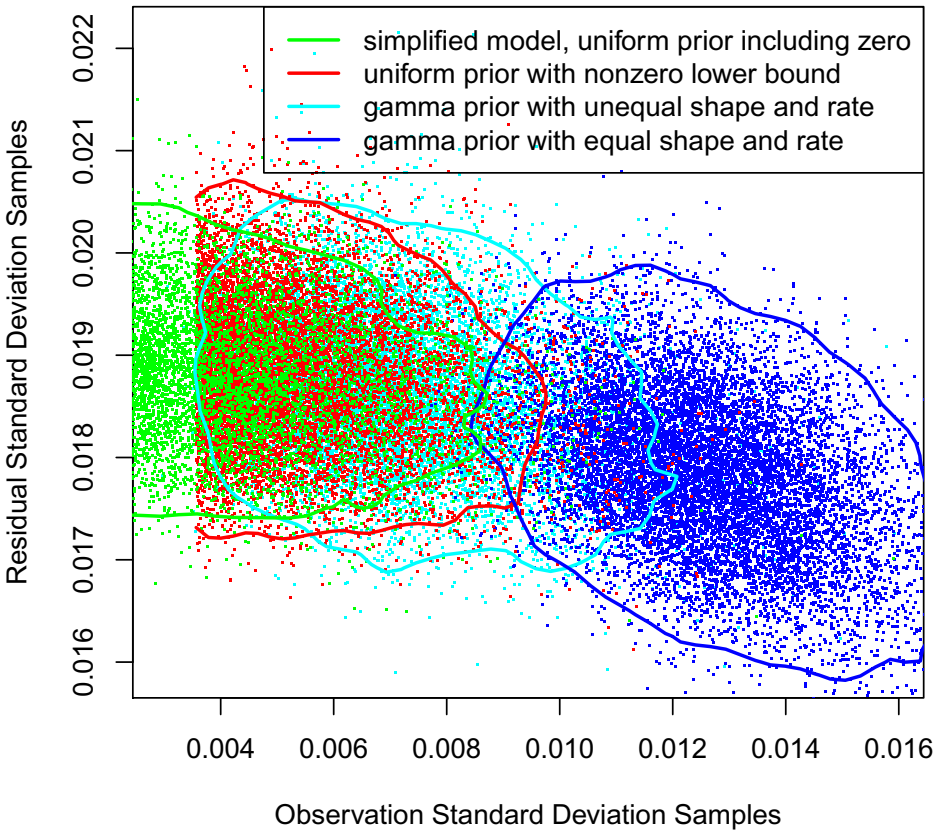

Figure D.8: Scatterplots of MCMC samples for residual standard deviation σ_ε and observation standard deviation σ_{DBH} for four different models with different observation error priors. Contour lines are 95% credible contours (Bolker 2012). Cyan is the model with a gamma prior with unequal shape and rate; red is the model with a uniform prior with a nonzero lower bound (clearly visible!); blue is the model with a gamma prior with equal shape and rate; and green is the simplified model (missing annual water deficit, insolation, and elevation) with a uniform prior including zero. Clearly the tradeoff between residual and observation error is affected by the prior chosen for observation error, and how close it is allowed to get to zero.

Analysts should watch for these kinds of tradeoffs in their models as well. In this case, the tradeoff relationship means that when observation error standard deviation is estimated with different priors, residual standard deviation is affected, though not extremely so: the biggest differences in mean estimates for residual standard deviation in the four models is 0.001 on the standardized scale, or about 0.02 cm (Figure D.9), and the significance of the parameter does not change (it is still well separated from zero). In fact, it is a mark of the robustness of the residual standard deviation estimate. Despite this tradeoff relationship with observation error, the residual error estimate does not change substantially when the observation error standard deviation does. None of the other parameters in the model are affected by observation error's prior.

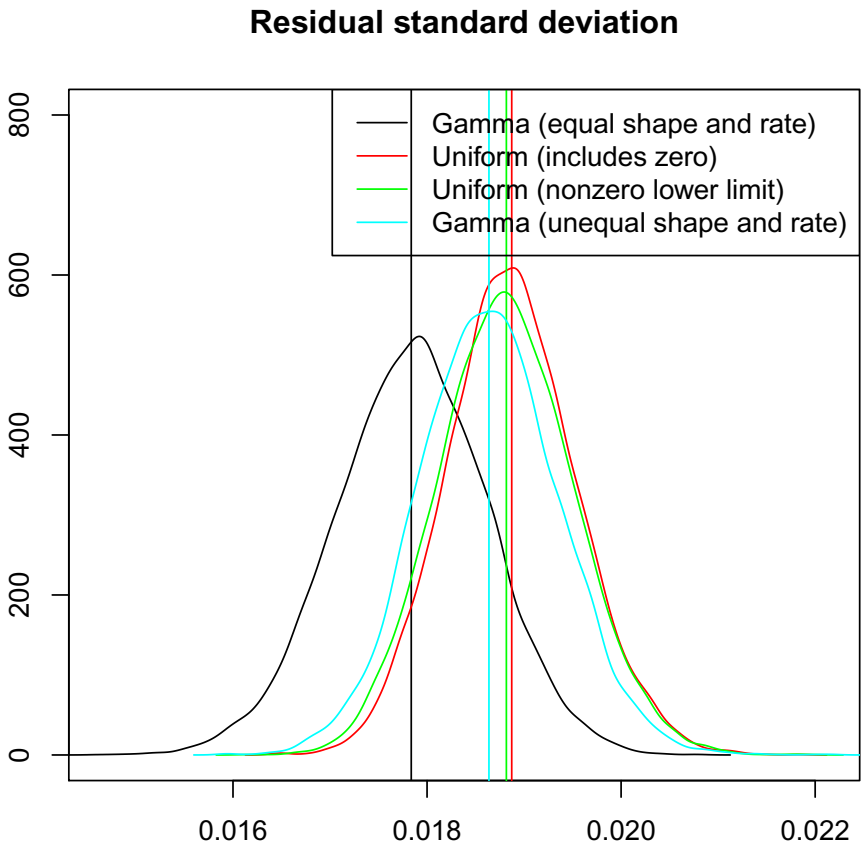

Figure D.9: Posterior estimates for residual standard deviation σ_ε for four different observation error posteriors. Cyan is the model with a gamma prior with unequal shape and rate; green is the model with a uniform prior with a nonzero lower bound; black is the model with a gamma prior with equal shape and rate; and red is the simplified model (missing annual water deficit, insolation, and elevation) with a uniform prior including zero. Vertical lines are means for each posterior.

Appendix E

Table of Parameter Estimates for Growth Model

Fixed continuous covariate effects				
Parameter	Intercept Mean		Slope Mean	
Insolation	-0.001	(-0.003, 0.001)	0.001	(-0.001, 0.003)
Elevation	0	(-0.003, 0.003)	-0.001	(-0.003, 0.002)
Annual water deficit	0.002	(-0.008, 0.012)	0.001	(-0.005, 0.007)
Topographic slope	-0.002*	(-0.005, 0)	0.001	(-0.001, 0.002)
Basal area	-0.009*	(-0.012, -0.007)	-0.002	(-0.004, 0.001)
Average size	0.023*	(0.016, 0.029)	0.01*	(0, 0.02)
Fixed soil effects				
Cohasset	0.031*	(0.02, 0.041)	0.013*	(0.006, 0.02)
Holland	0.02*	(0.013, 0.027)	0.008	(-0.003, 0.019)
Holland-Bighill	0.024*	(0.017, 0.032)	0.012*	(0.002, 0.022)
Holland-Musick	0.015*	(0.004, 0.027)	0.005	(-0.008, 0.018)
Jocal	0.037*	(0.02, 0.055)	0.006	(-0.028, 0.037)
Random effects				
Parameter	Intercept Mean		Slope Mean	
Year	0.012*	(0.004, 0.021)	0.006	(0, 0.013)
Compartment	0.006	(0, 0.013)	0.002	(0, 0.005)
Plot	0.006*	(0.005, 0.008)	0.005*	(0.003, 0.006)
Tree	0.007*	(0.006, 0.008)	0.001	(0, 0.003)

Table E.1: Parameter point estimates (means) and 95% credible intervals. Asterisks indicate parameters which are significantly separated from zero. For random effects, which can never overlap zero, I use ratios of the lower credible bound to the width of the credible interval to represent a posterior's separation from zero, shown in table E.2. For parameters which could overlap zero in theory (e.g. soil and other fixed effects), I consider parameters which either do not overlap zero at all or have a credible bound exactly at zero to be significant. All values given are on the standardized, unitless scale. Soil effects are marked significant if they individually do not overlap zero; some which are not marked significant may also be significantly different from each other; pairwise comparisons are not shown here. Note that here a mean is reported for each soil, which matches Figure 1.1, but in the code to estimate the model Cohasset was used as a reference and estimated differences from Cohasset for the other four soils.

	Intercept SD	Slope SD
Year	0.2275*	0.0002
Compartment	0.0008	0.0001
Plot	1.5175*	1.1700*
Tree	3.7637*	0.0036

Table E.2: Ratios of lower credible bound to credible band width for random effect standard deviation parameter estimates, used to determine significance of these parameters (significant parameters marked with asterisks). Because they cannot overlap zero, this metric can be used to capture a posterior's separation from zero. Several parameters have posteriors (and MCMC chains) that are clearly attracted to zero, while others are clearly stable away from zero. This ratio captures this behavior, with the significant parameters having both a high lower credible bound and a narrower credible interval.

Appendix F

Latent Unmeasured Sizes for Nine Trees

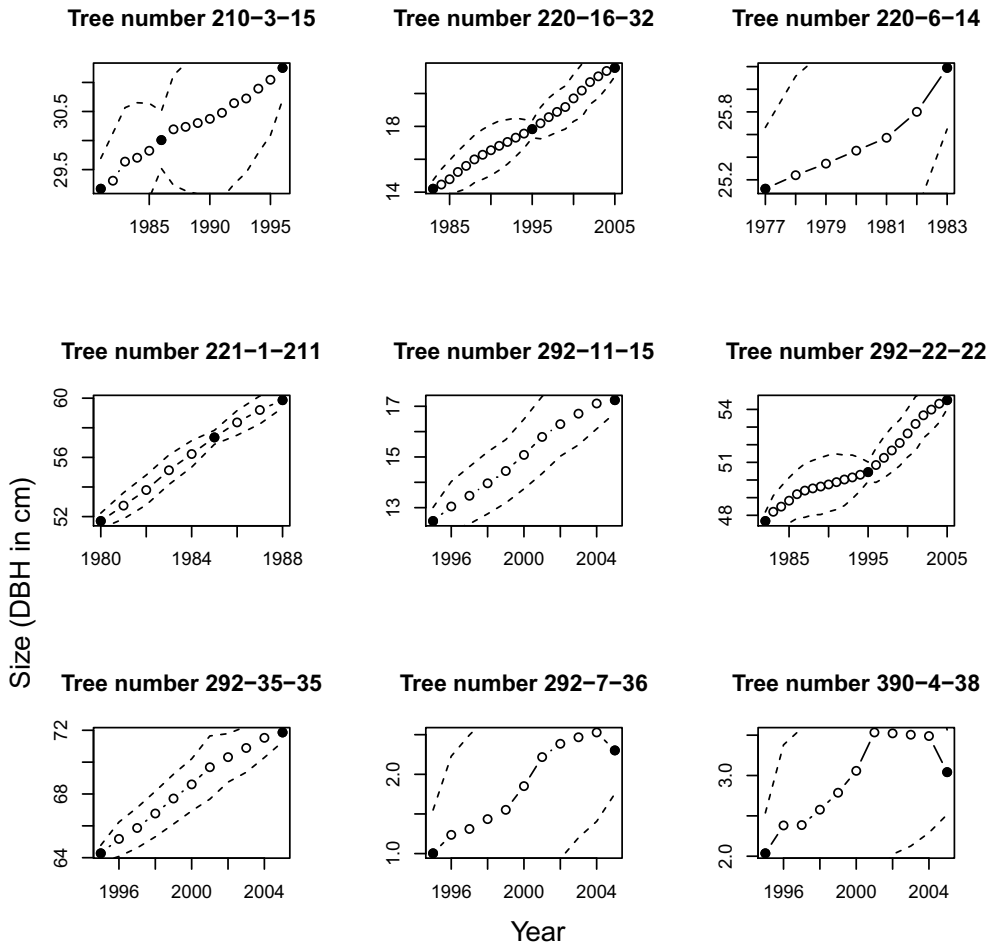

Figure F.1: The latent states x of nine trees sampled from the model are shown below with 95% credible intervals to show that the estimates give positive growth with no other constraints, and that the credible bands tighten up in the neighborhood of measured data points (as in Clark et al. 2007). Note that for smaller trees, the credible bands are wider and point estimates do show the potential for negative growth; the estimation procedure is less precise for slower growth. For white fir, there are 60 trees out of 691 that are 3 cm or smaller at the beginning of the inventory, so this is fewer than 10% of the trees in the model. The remaining trees show reasonable point estimates for annual growth increments. These results are from the model with a gamma prior on observation error with equal shape and rate, which gives the largest estimates for observation error.

Appendix G

Examples of Simpler Models

To test the robustness of the parameter estimates for basal area and topographic slope, I estimated simpler models omitting other covariates. I tried various combinations of omitting significant or not significant covariates, and checked the two observation error priors which give the most different observation error estimates to see if these different priors affected other parameter estimates. When a covariate was included, both the main effect and the interaction was included, and likewise when it was removed from the model. The models I used are shown in Table G.1, which indicates which variables were included or omitted. See Table G.2 for point estimates (means) of parameters in different models; see Figures G.1 through G.5 for posteriors for selected cases.

Model Number	TD	BA	TS	SE	WD	E	I	OEP
Model 1	X	X	X	X	X	X	X	UPNZ
Model 2	X	X	X	X				UPZ
Model 3	X	X	X	X				GP
Model 4	X	X		X			X	UPZ
Model 5	X	X	X	X	X			UPZ
Model 6*	X	X	X	X	X			UPZ

Table G.1: Which variables are omitted or included in simpler models. TD = tree diameter; BA = basal area; TS = topographic slope; SE = soil effects; WD = annual water deficit; E = elevation, I = Insolation; OEP = Observation Error Prior: UPNZ (uniform prior, no zero allowed - the prior used in Chapter 1), UPZ (uniform prior, zero allowed), GP (gamma prior with equal shape and rate). *Model 6 differs from Model 5 in that Model 6 has had its year random effects removed.

Note that all models contain tree diameter, basal area, and soil effects, and are omitted from figure legends below. In general, parameter estimates are quite robust. Comparing models 2 and 3 (same covariates, different prior for observation error) shows that the changes

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Plot Intercept Effect	0.006	0.006	0.006	0.006	0.006	0.006
Plot Slope Effect	0.005	0.005	0.005	0.005	0.005	0.005
Tree Intercept Effect	0.007	0.007	0.007	0.007	0.007	0.006
Tree Slope Effect	0.001	0.001	0.001	0.001	0.001	0.001
Comp. Intercept Effect	0.006	0.006	0.007	0.008	0.006	0.006
Comp. Slope Effect	0.002	0.002	0.002	0.002	0.002	0.002
Year Intercept Effect	0.012	0.108	0.104	0.1	0.012	X
Year Slope Effect	0.006	0.006	0.005	0.005	0.006	X
Cohasset Main Effect	0.031	0.029	0.034	0.029	0.031	0.033
Cohasset-Size Int.	1.013	1.013	1.013	1.013	1.013	1.013
Basal Area Main Effect	-0.009	-0.01	-0.01	-0.009	-0.009	-0.008
Basal Area-Size Int.	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
Insolation Main Effect	-0.001	X	X	0	X	X
Insolation-Size Int.	0.001	X	X	0	X	X
Top. Slope Main Effect	-0.002	-0.002	-0.002	X	-0.002	-0.002
Top. Slope-Size Int.	0.001	0	0	X	0	0
Water Def. Main Effect	0.002	X	X	X	0.003	0.001
Water Def.-Size Int.	0.001	X	X	X	-0.001	0.001
Elevation Main Effect	0	X	X	X	X	X
Elevation-Size Int.	-0.001	X	X	X	X	X
Obs. Error	0.005	0.003	0.013	0.002	0.003	0.004
Residual Error	0.019	0.019	0.018	0.019	0.019	0.024

Table G.2: Parameter point estimates (means) for model parameters from different models. "X" indicates that variable is not in the model. Most estimates are stable across models; see Figures G.1 through G.5 for posteriors for some nodes which shift considerably when other variables are removed from the model.

in priors for observation error do not affect other parameter estimates (Figures G.2, G.3, and G.4). While some parameters show some shifts when others are removed, the significance of the estimates is not changed, and the shifts are typically quite small (don't change the point estimate much). For example, when comparing model 2 (with topographic slope included and insolation excluded) and model 4 (insolation included and topographic slope excluded) with model 1 which includes them both (Figures G.4 and G.5), topographic slope remains significant when insolation is removed and insolation remains not significant when topographic slope is removed; it did not take up the variation unaccounted for by the removal of topographic slope.

The only exception was the tradeoff between annual water deficit and intercept year effects: when annual deficit was removed from the model (Models 2, 3, and 4 as compared

with 1 and 5), the year intercept value was consistently an order of magnitude higher. So even though the annual deficit does not appear significant (overlaps zero), it clearly has explanatory power (Figure G.2). Model 6 omitted the year random effects, and annual deficit still did not become significant though its estimate was shifted slightly and improved in precision (the posteriors were less broad; compare with Model 5 (Figure G.1). All other estimates are unaffected by the removal of other covariates in these simplified models.

All results in these figures are on the standardized unitless scale.

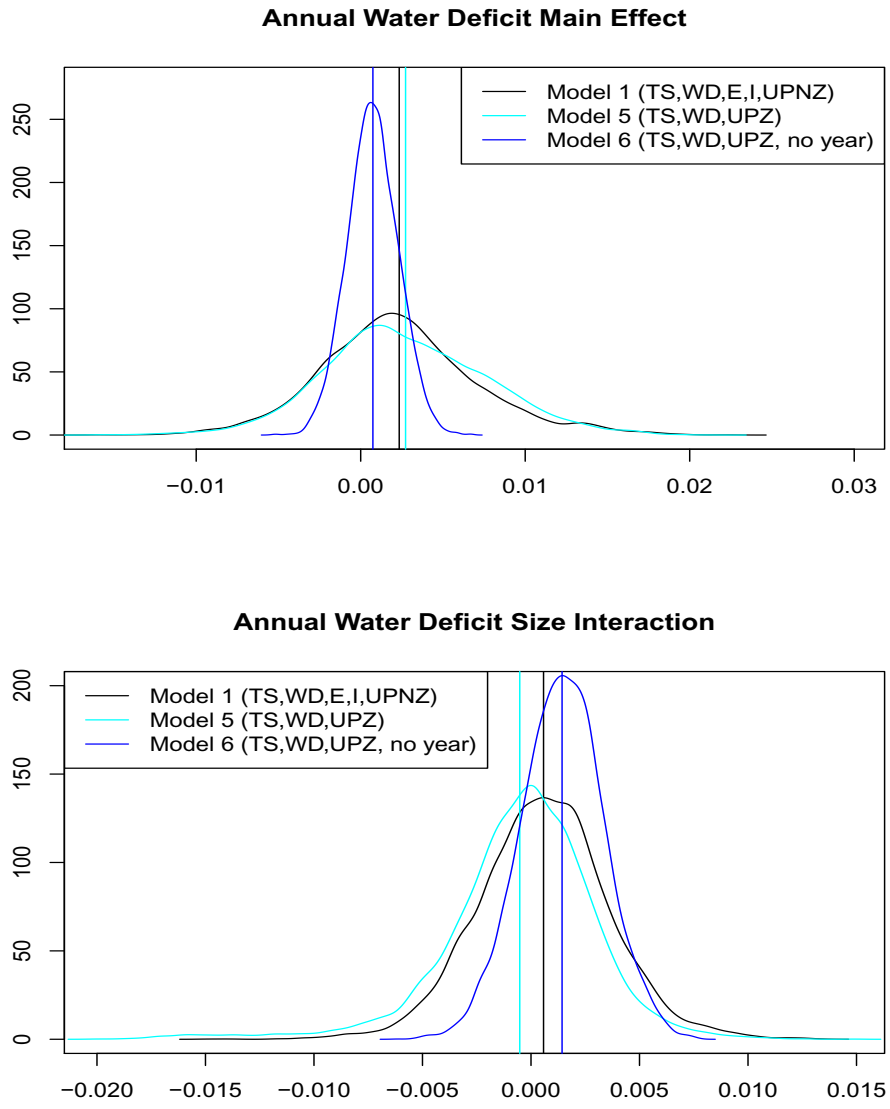

Figure G.1: Removing year random effects $\sigma_{\alpha,w}$ and $\sigma_{\beta,w}$ tightens up the estimates for annual water deficit and shifts them slightly; however, annual deficit does not take up significant explanatory power (still overlaps zero). (Upper) Main effect for annual water deficit γ^{def}. (Lower) Interaction between size and annual water deficit κ^{def}.
TS = topographic slope; WD = annual water deficit; E = elevation, I = Insolation; Observation Error Prior: UPNZ (uniform prior, no zero allowed — the prior used in Chapter 1), UPZ (uniform prior, zero allowed), GP (gamma prior with equal shape and rate).

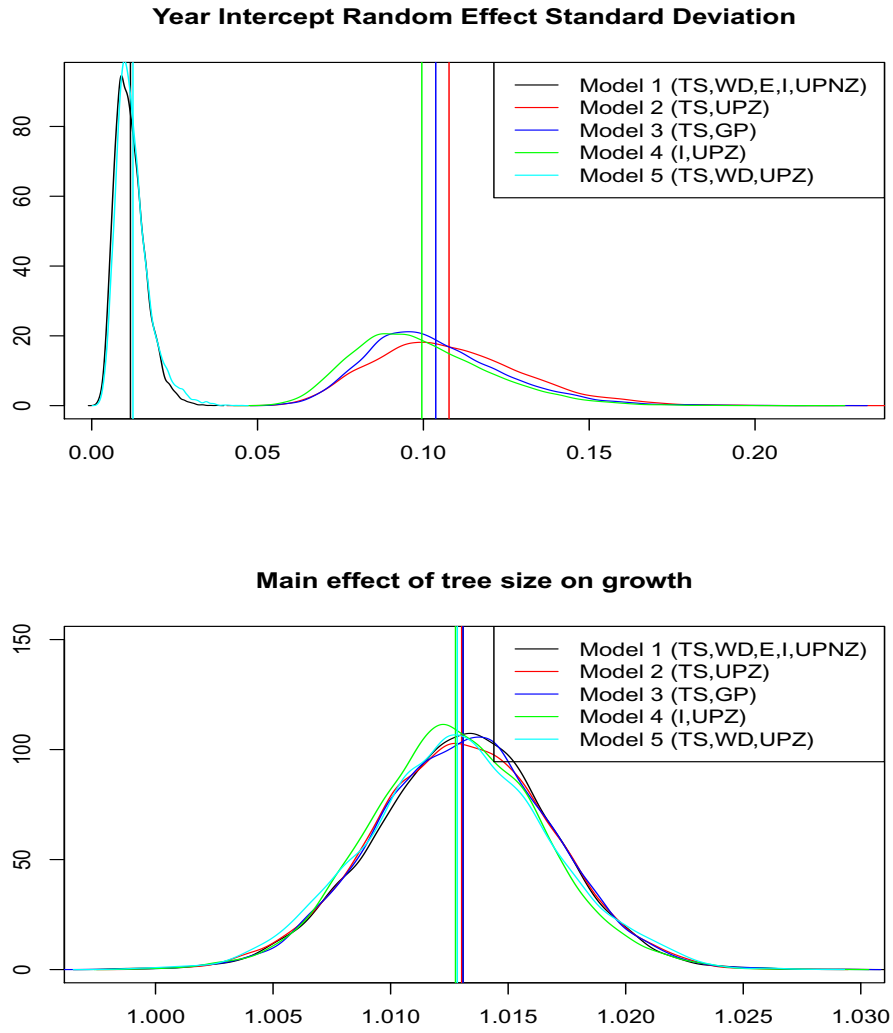

Figure G.2: (Upper) The models which lack annual water deficit all have a higher estimate for the year intercept random effect standard deviation $\sigma_{\alpha,w}$. Though annual water deficit's parameters overlapped zero significantly, at least the main effect γ^{def} does seem to have some explanatory power, because the random effect is taking up some variation from it. The year random effect is still significantly separated from zero when annual deficit is in the model, implying that there is other year-to-year variation that annual deficit does not account for. (Lower) An example of a parameter which is unchanged regardless of other covariates are included or not: the tree size main effect for Cohasset soil κ^C.
TS = topographic slope; WD = annual water deficit; E = elevation, I = Insolation; Observation Error Prior: UPNZ (uniform prior, no zero allowed — the prior used in Chapter 1), UPZ (uniform prior, zero allowed), GP (gamma prior with equal shape and rate).

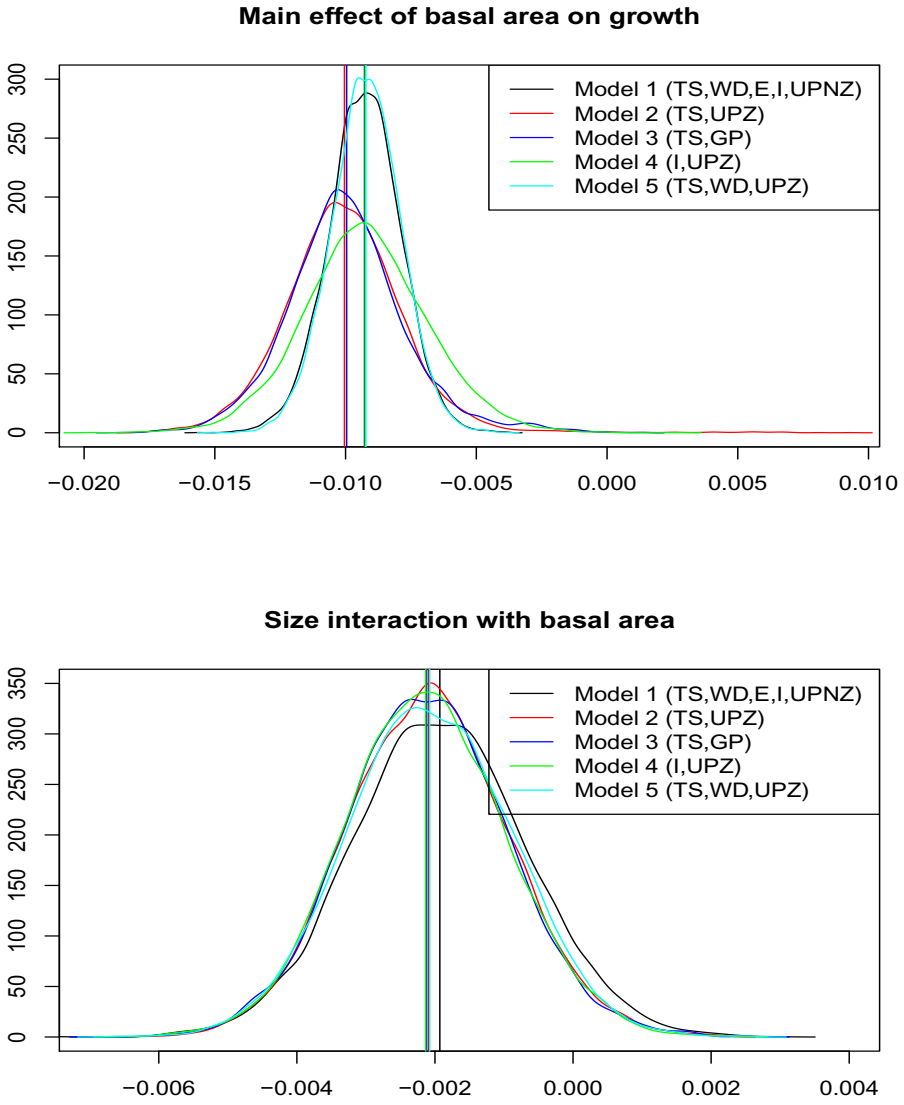

Figure G.3: For basal area, the removal of other covariates does not change the parameter estimates much. (Upper) The main effect of basal area on tree growth γ^{ba} is affected only slightly by the removal of various other covariates. Differences in means for different models are no larger than 0.0008 (on the standardized scale). (Lower) There is very little difference in basal area's interaction with size κ^{ba} parameter for models with different covariates included.
TS = topographic slope; WD = annual water deficit; E = elevation, I = Insolation; Observation Error Prior: UPNZ (uniform prior, no zero allowed — the prior used in Chapter 1), UPZ (uniform prior, zero allowed), GP (gamma prior with equal shape and rate).

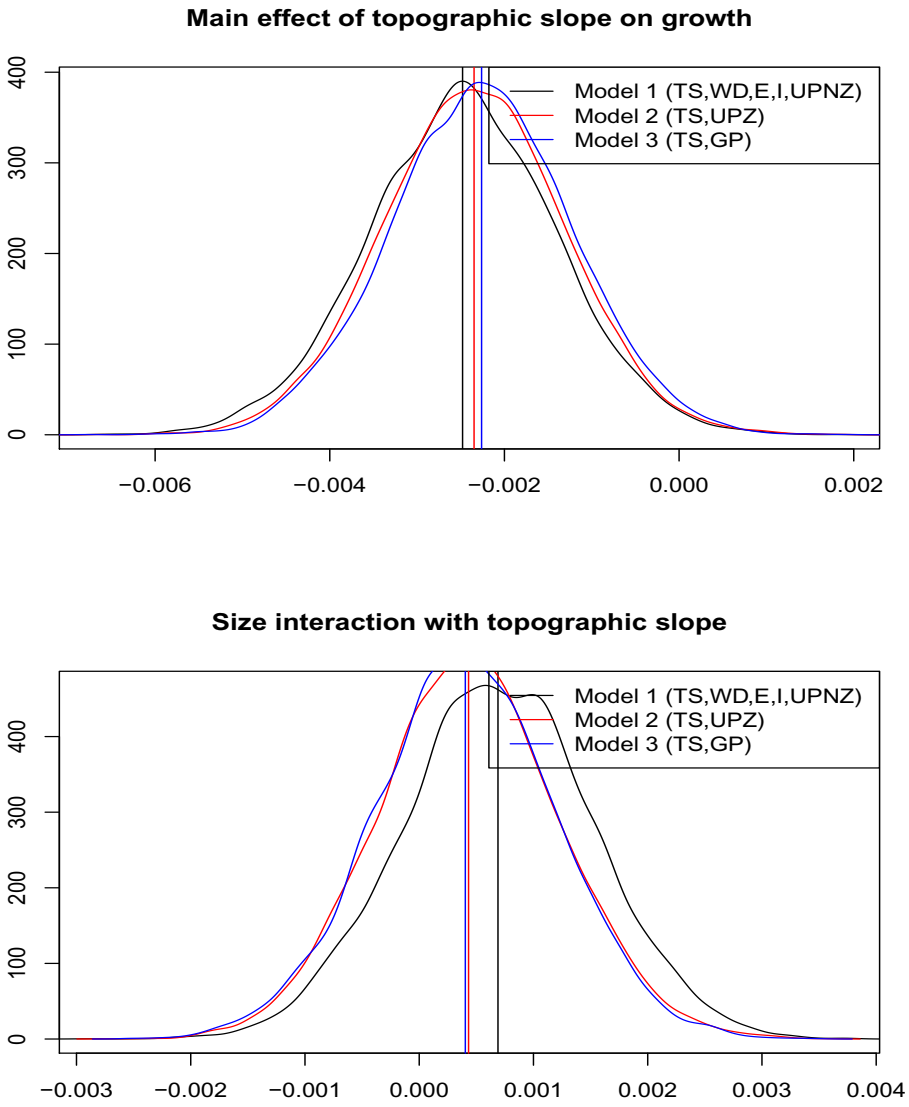

Figure G.4: For topographic slope, removal of other parameters does not change parameter estimates much. (Upper) The estimate of the main effect for topographic slope γ^{tslope} is quite robust both to priors used for observation error and to removal of annual deficit, insolation, and elevation. (Lower) The same is true for the interaction of topographic slope with size κ^{tslope}.

TS = topographic slope; WD = annual water deficit; E = elevation, I = Insolation; Observation Error Prior: UPNZ (uniform prior, no zero allowed — the prior used in Chapter 1), UPZ (uniform prior, zero allowed), GP (gamma prior with equal shape and rate).

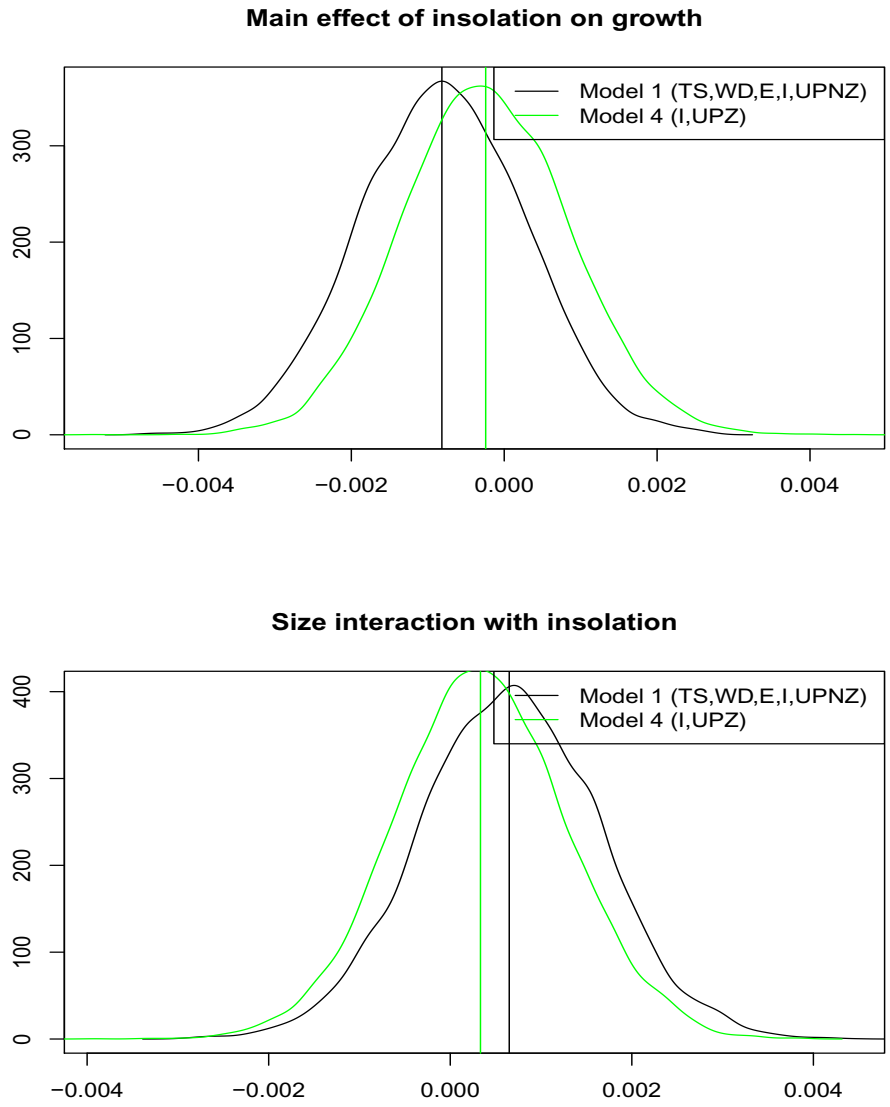

Figure G.5: The estimate for insolation parameters is robust to removing other covariates (topographic slope, elevation, annual water deficit); it does not become more significant (less overlap of zero) when it is in the model and topographic slope is not: it doesn't take on explanatory power. (Upper) Main effect of insolation on growth γ^{insol}; (Lower) Interaction of growth with size κ^{insol}.
TS = topographic slope; WD = annual water deficit; E = elevation, I = Insolation; Observation Error Prior: UPNZ (uniform prior, no zero allowed — the prior used in Chapter 1), UPZ (uniform prior, zero allowed), GP (gamma prior with equal shape and rate).

Appendix H

Comparison of Conditions at Blodgett Forest Research Station (BFRS) with Forest Inventory and Analysis (FIA) Data

Comparing slope, elevation, and insolation at BFRS with ranges derived from Forest Inventory and Analysis (FIA) plots from 37.87718 to 39.97781 in latitude (USDA Forest Service 2011) and restricted to "California mixed conifer" vegetation type, one can see that BFRS represents a small fraction of elevation (only approximately 8 %) and insolation (only approximately 6 %) conditions for white fir and represents a somewhat larger fraction (approximately 52% of the variation from the FIA plots) of topographic slope conditions, but still not by any means a representative set of conditions for the mixed conifer forest type in this latitudinal range in CA.

Annual water deficit values for the years of the inventories represent 82 % of the variation seen since 1985 from PRISM data (PRISM Climate Group 2011), and Blodgett is typically wetter than other sites. For example, annual water deficits averaged 176 mm at BFRS between 1976-2009, and a study by van Mantgem and Stephenson (2007) reported deficits at BFRS between 200 and 250 mm. Thus water supply may not be as strong a limiting factor for the trees in this study. In addition, as Clark et al. (2011) showed, tree fecundity (rather than growth or survival) was the most sensitive demographic parameter to climate.

Different basal area conditions, on the other hand, are well represented at BFRS, with the ranges of values comparable to those in the FIA plots.

Appendix I

Abundances of All Species in Inventory and Plot Basal Area and Water Deficit Analysis

Code	Species Name	Common Name	Trees	Plots	Years	Records
All	–	All Species	3319	109	34	9000
IC	*Calocedrus decurrens*	incense-cedar	1378	98	34	3733
WF	*Abies concolor*	white fir	846	93	34	2224
DF	*Pseudotsuga menziesii*	Douglas-fir	395	69	34	1110
PP	*Pinus ponderosa*	ponderosa pine	332	75	34	955
TO	*Notholithocarpus densiflorus*	tanoak	110	15	25	284
BO	*Quercus kelloggii*	black oak	104	29	34	292
SP	*Pinus lambertiana*	sugar pine	101	39	34	271
YW	*Taxus brevifolia*	Pacific yew	18	2	24	42
CH	*Chrysolepis spp.*	chinquapin	16	5	29	43
DW	*Cornus nuttallii*	Pacific dogwood	9	3	11	18
AL	*Alnus spp.*	alder	7	3	24	21
MD	*Arbutus menziesii*	Pacific madrone	3	2	25	7

Table I.1: Species abundances, number of records, and number of deaths. This table includes information on the subcanopy species as well as the canopy species included in the paper.

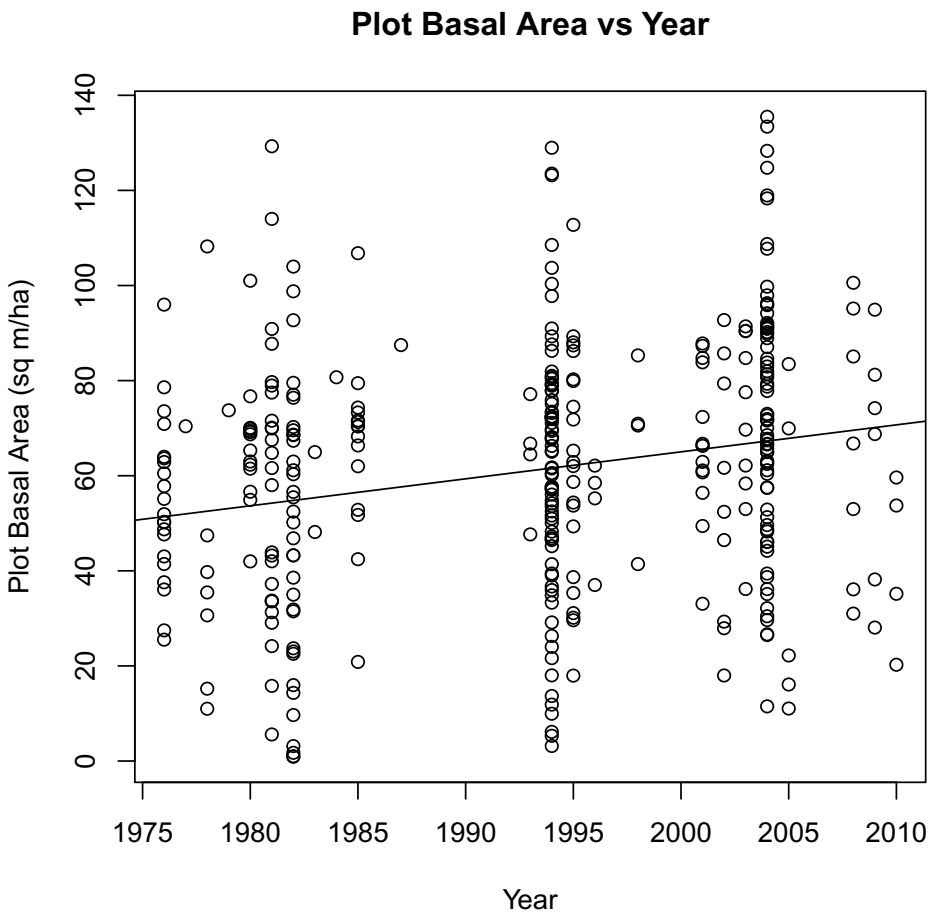

Figure I.1: Plot basal area for each plot at each time. Trend line is for a linear mixed model with plot as a random effect and year as a fixed effect.

I.1 Plot Basal Area and Annual Water Deficit Analysis

To test for a trend in plot basal area, I fit a linear mixed model with a fixed effect for time and a random effect for plot (using "`lme`" in package "`nlme`"; Figure I.1). The time trend was significant: 0.567 $m^2/ha/year$, $p =$ 1.85e-18.

To get a sense of the variation due to year as compared to the variation due to plot, I also fit a model with a random effect for year and a random effect for plot (using "`lmer`" in package "`lme4`", Bates and Maechler 2010). Plot variation is much higher than year variation for plot basal area: the standard deviation of year in this model was 7.9, and for plot was

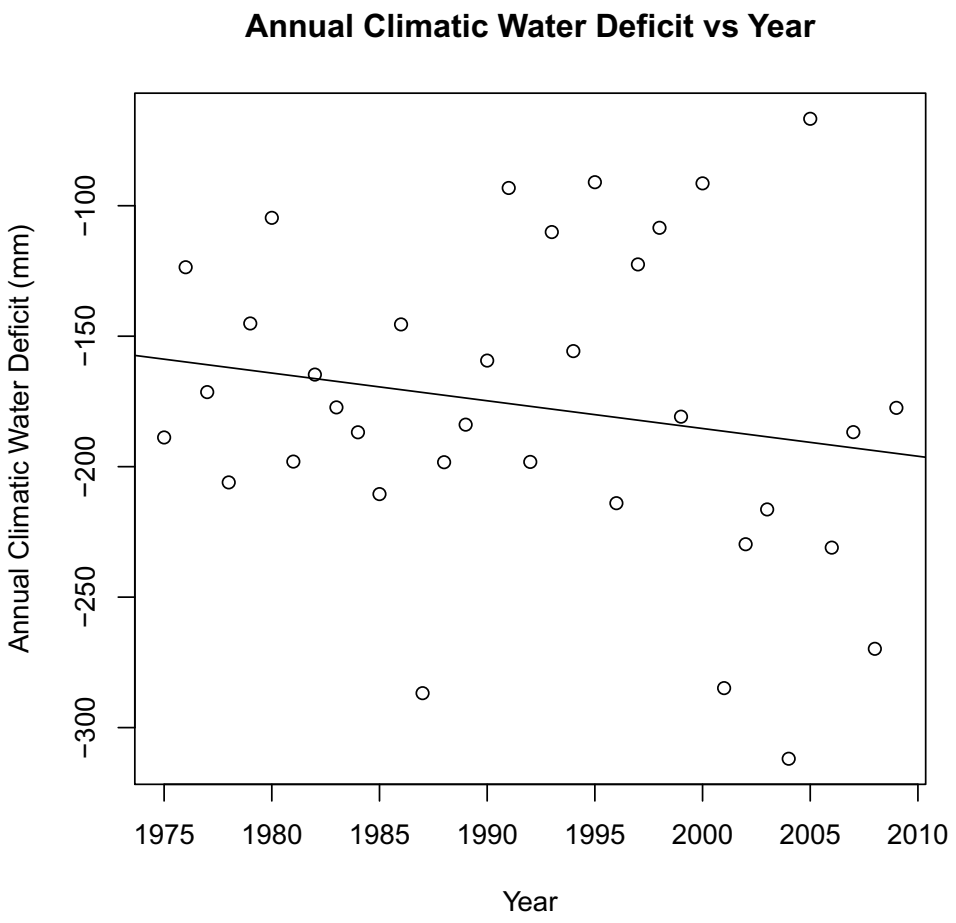

Figure I.2: Annual climatic water deficit by year. Trend line is for a linear model with year as predictor.

24.7. This helps to explain how an effect can be estimated for plot basal area, a random effect for plot, and a time trend, all in the same model: there is enough variation in each component to provide information about each of these effects.

The trend for annual climatic water deficit is small and not significant: -1.06 $mm/year$, p=0.29(Figure I.2). More negative deficits indicate a more stressful environment.

Appendix J

Estimating Missing Sizes

When the field crews encounter a dead tree, they have no way of knowing when it died in the intervening time since the last survey, and the size is recorded in the database as zero. Because tree size is interpolated to create a yearly individual-level variable for the survival models, sizes are needed for all trees in all years. When a tree is measured more than twice, the size at potential times of death can be interpolated from those other sizes, but when the tree is measured once and then dies, only one size is available to be interpolated. Therefore, I need a way to estimate that missing size in order to interpolate.

I tried three strategies to address these missing sizes: 1) exclude the trees; 2) use a naïve linear model for the trees which do have multiple sizes to estimate the growth as a function of size, averaging over census intervals, and predict the missing size; and 3) use the results from another state-space model for growth, which is designed to estimate missing sizes (Eitzel et al. 2013). Strategy three was explored only for white fir.

Note that the ideal solution is to jointly estimate the missing size **and** the missing survival status in each year, but such a joint estimation is beyond the scope of this project. In the case of modeling growth, where size at the previous time predicts size at the next time (Eitzel et al. 2013), interpolating or estimating missing sizes as a separate step would be much more statistically questionable than it is here. In this case, the size is a predictor of mortality, not both a predictor and response variable. Thus the estimation and interpolation introduce minimal error into the models.

Method 1 had disastrous consequences. Excluding these trees biases results (because only dead trees, and no live trees, were eliminated). Eliminating many dead trees left the annual mortality rates unrealistically low (less than a percent per year, when correct estimates range between 1.21 and 2.3 % (see Table L.1). Spurious climate trends emerged for many species and the species-specific stories described in Chapter 2 disappeared.

Methods 2 and 3 produced similar final sizes. Examination of posterior sensitivity showed that the methods did not change most parameter estimates at all; the only parameters which

were affected were, predictably, the linear and quadratic size factors. However, no choice for a missing size estimation method changed whether these two parameters were selected during the model selection procedure. Because the two methods gave similar results, and estimating a state-space model for missing sizes for each species is also beyond the scope of Chapter 2, I elected to use Method 2 for the models in this paper.

Appendix K

Model Details and Algebra For Standardization and Rescaling Parameter Estimates (Survival Model)

K.1 Model Details

Existing work on hierarchical models for tree mortality in second growth systems typically adopts an annual compounding framework to account for uneven census intervals (Hurst et al. 2011, Peng et al. 2011, Thorpe and Daniels 2012, Luo and Chen 2013, but see Yang and Huang 2013). This formulation makes incorporating yearly variables difficult. One solution has been to use a single average value for the predictor for a given census interval, but tree mortality responses may track yearly variation in variables and thus mortality trends may not be detected if the representative variable cannot be modeled at the yearly level. Hierarchical state-space models (Metcalf et al. 2009*b*, Clark et al. 2012, Csilléry et al. 2013) address these challenges by explicitly modeling the latent unmeasured survival status of the tree in each year; then a yearly climate variable as well as yearly stand development variables can be directly included.

The fullest model is specified as follows:

$$\begin{aligned} z_{ij}(t+1) &\sim Bern(\phi_{ij}(t) z_{ij}(t)) \\ logit(\phi_{ij}(t)) &= \beta_{0_j} + \sum_k \beta^k x_{ij}^k(t) \\ \beta_{0_j} &= b + p_j \\ p_j &\sim N(0, \sigma_p^2) \end{aligned} \tag{K.1}$$

where $z_{ij}(t+1)$ is the alive/dead status in year $t+1$ (0 is dead) and $\phi_{ij}(t)$ is the probability of survival for tree i in plot j from year t to year $t+1$. Multiplying by the status at the

previous time ensures that dead trees stay dead. β_{0_j} is an intercept, with a mean intercept b and an intercept random effect due to plot p_j, which are normally distributed with variance σ_p^2. k indexes one of the explanatory variables. Insolation (x_j^{insol}), topographic slope (x_j^{slope}), and elevation (x_j^{elev}) are all measured at plot level; plot basal area ($x_j^{ba}(t)$) is measured at plot and year level; tree size ($x_{ij}^{DBH}(t)$) is measured at the tree and year level; and annual water deficit is measured at year level ($x^{def}(t-1)$). In order to test for a time trend I also include the measurement year $x^t = t$ among the linear predictors in some models. The β^k are coefficients for each of these explanatory variables. To allow a more flexible mortality relationship with size, both a linear (β^{DBH}) and quadratic (β^{DBH^2}) term for tree diameter are included.

A note on estimation: in order to ensure that the prior for mean survival b is flat on the probability scale, it must be specified in the following way in BUGS:

```
>         i.overall<-logit(li.overall)
>         li.overall~dunif(0, 1)
```

Where `i.overall` is b, and `li.overall` is the inverse logit ("expit") of b. See Buoro et al. 2012 for more details on ensuring flat priors on the probability scale. I adopt their method only for the intercept b and not for the β^ks because doing so resulted in very slow mixing and did not appreciably change parameter estimates.

See Figure K.1 for a graphical representation of the model.

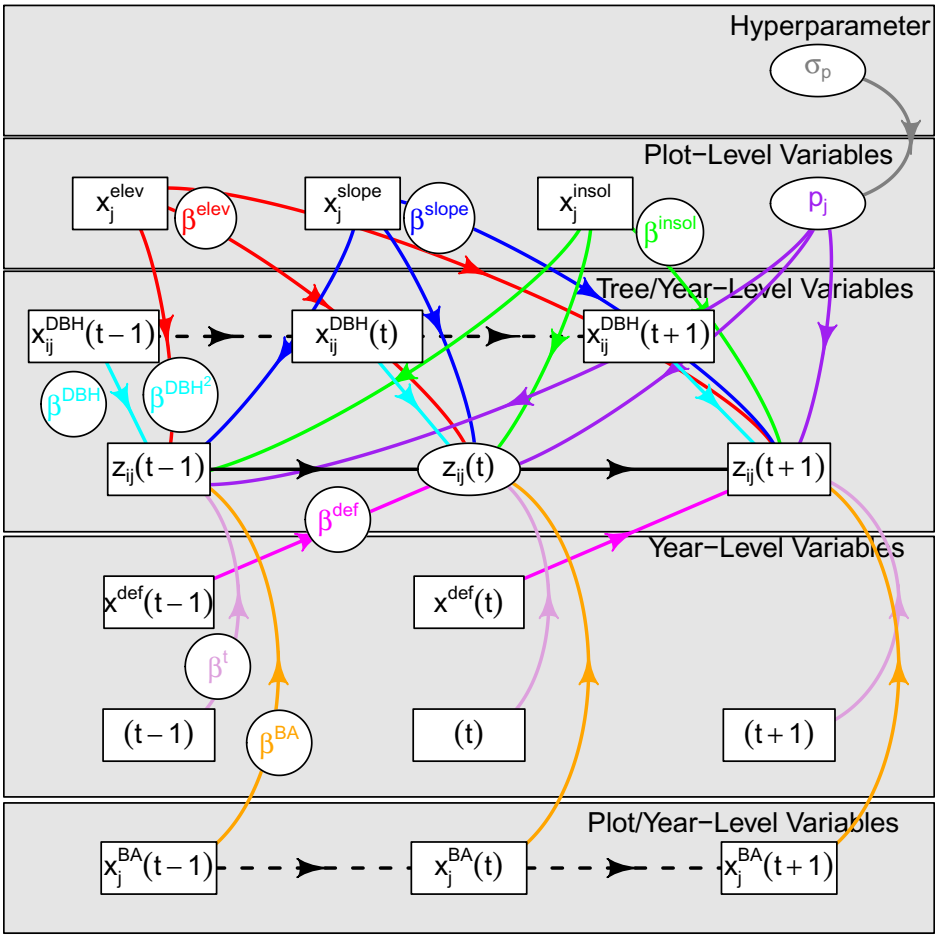

Figure K.1: Diagram of the state-space model. Data are indicated in rectangles, with parameters and latent states to be estimated indicated in ellipses. Colored arrows indicate multiplication of data (rectangle) by a parameter (typically β, labeled circles) to be summed in the linear predictors for z. The black arrows indicate the process portion of the model wherein the current survival status is conditional on the previous survival status: if a tree is dead at time $(t-1)$, it remains dead at time t. This diagram is for tree i in plot j; imagine another such diagram for every tree in the inventory. Each linear predictor coefficient β is only labeled once, but every arrow of the same color also indicates that parameter. Note that are depicted as $z_{ij}(t-1)$ and $z_{ij}(t+1)$ as known states — observations — and $z_{ij}(t)$ as an unknown latent state to be estimated. Dashed black lines indicate variables which have been interpolated.

K.2 Standardizing Explanatory Variables

Original variables are standardized by centering (subtracting the mean of the explanatory variable μ_{x^k}) and scaling (dividing by the standard deviation of the explanatory variable σ_{x^k}):

$$x^k(t)' = \frac{x^k(t) - \mu_{x^k}}{\sigma_{x^k}} \tag{K.2}$$

The estimated model is in terms of these standardized variables:

$$logit(\phi_{ij}(t)') = {\beta_{0_j}}' + \sum_k \beta^{k\prime} x_{ij}^k(t)' \tag{K.3}$$

Rewriting this in terms of the original variables:

$$logit(\phi_{ij}(t)') = {\beta_{0_j}}' + \sum_k \beta^{k\prime} \frac{x^k(t) - \mu_{x^k}}{\sigma_{x^k}} \tag{K.4}$$

To keep the variables centered but return them to the scale of the original variables, this equation should be written in terms of a centered variable:

$$x^k(t)'' = x^k(t) - \mu_{x^k} \tag{K.5}$$

Rewriting the model equation with respect to the double-primed variables:

$$logit(\phi_{ij}(t)'') = {\beta_{0_j}}'' + \sum_k \beta^{k\prime\prime} \frac{x^k(t)''}{\sigma_{x^k}} \tag{K.6}$$

One can see that the intercept ${\beta_{0_j}}'' = {\beta_{0_j}}' = \beta_{0_j}$ (and b and p_j along with it); but that the coefficient for $x^k(t)$ must be divided by σ_{x^k}:

$$\beta^{k\prime\prime} = \beta^{k\prime}/\sigma_{x^k} \tag{K.7}$$

The quadratic term for size, β^{DBH^2}, would need to be divided by $\sigma^2_{x^{DBH}}$.

Appendix L

Simple Mortality Calculations (m) to Validate Simple State-Space Models

A state-space model will ultimately be needed to incorporate the mosaic of measurement intervals, random effects, and explanatory variables. To validate the state space model, I also calculate a simple average annual mortality after Sheil et al. (1995). I compare this calculation with a simple version of the state-space model which uses a slightly different but equivalent model structure and a different inferential framework to estimate the same overall mortality parameter. Any subtle differences are due to the different estimation methods. Sheil et al. (1995) calculate an annual mortality rate in the following way: $N_1 = N_0(1-m)^t$, where N_1 is the current number of stems, N_0 is the initial number of stems, m is the mortality per year, and t is the measurement interval. Solving for m: $m = 1 - (N_1/N_0)^{1/t}$.

This calculation is a kind of cohort analysis, assuming that no individuals have been added to N_1 which were not in N_0. In these data, however, with every new census new individuals are included and therefore every unique measurement interval constitutes a different measurement cohort. In order to obtain a single estimate of m for each species, I maximize a binomial likelihood with N_0, N_1, and t for each cohort (defined as all trees sharing an inventory interval, aggregating over plots) as data and a single m for all intervals as the binomial probability parameter. Tables L.1 and L.2 summarize the results of these m calculations. I obtain 95% confidence intervals using a profile likelihood; the details of the calculation, including code, are given in this Appendix. Calculations of m assume a constant mortality rate and have no easy way to include year-level variables (including time trends and climate and stand development variables). Note that the mortality rates from maximum likelihood m estimations agree with the Bayesian estimates for the simple state-space model (Table L.2). Therefore, any differences between the estimates of m and the survival rates in the more complex models are likely due to the hierarchical structure (plot effects) and covariates, not to the Bayesian vs. frequentist estimation methods (de Valpine 2009).

L.1 Simple Visualization of m for Each Time Interval for Each Species

Figures L.1 and L.2 show m calculated for every interval in the inventory, pooled across plots sharing the same interval. The darker the line, the larger number of individuals used to calculate m.

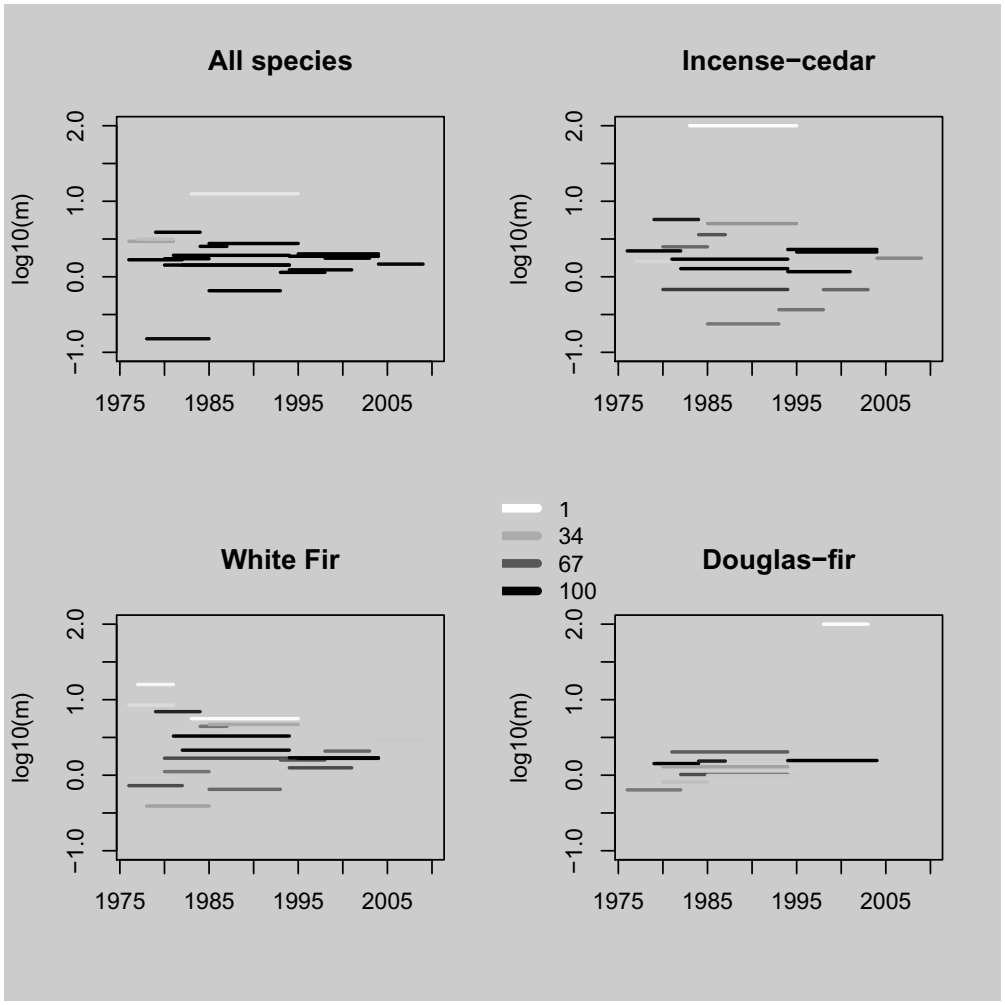

Figure L.1: m for each interval in the inventory. Plots sharing the same interval are combined. Shown on the log scale for clarity. Grayscale indicates how many trees are used for the calculation; darker indicates more individuals.

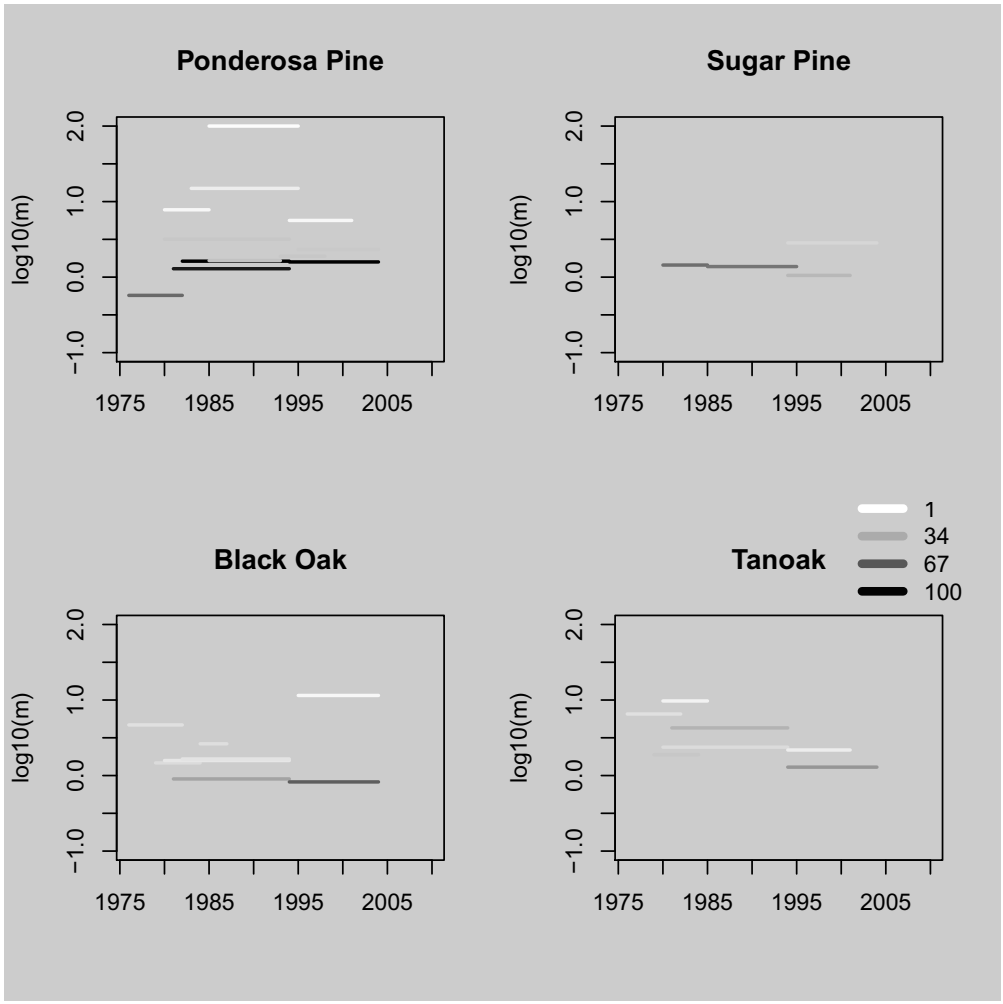

Figure L.2: m for each interval in the inventory. Plots sharing the same interval are combined. Shown on the log scale for clarity. Grayscale indicates how many trees are used for the calculation; darker indicates more individuals.

L.2 Maximum Likelihood Calculations and Code for Obtaining Single m Estimates by Species

For each interval, N_0 is the initial number of stems, N_1 is the final number of stems, and t is the number of years in that measurement interval. The annual probability of death, m, can be estimated based on the reduction in stems over these time periods. The likelihood is $L(N_0, N_1, t|m) = Bin(N_1, N_0, P_s)$, with the number of successes, or survived trees, N_1, out of the number of trials, or number of initial trees, N_0, given a survival probability determined by the number of years a tree has successfully survived: $P_s = (1 - m)^t$. In the simple case,

with one interval, $\hat{P}_s$ is just N_1/N_0. This is the same as Sheil et al. (1995)'s equation, which was derived based on discretized differential equations. Maximizing the likelihood of all the intervals together:

$$\sum_i^n log(L(N_{0_i}, N_{1_i}, t_i|m)) = \sum_i^n log\left(Bin(N_{1_i}, N_{0_i}, P_{s_i})\right)$$
$$P_{s_i} = (1-m)^{t_i} \tag{L.1}$$

This calculation will correctly weight plots with many or few trees, and those with no deaths or complete mortality. Here is the R code for the likelihood function:

```
> mylogLik<-function(m, dat) {
+          Psurv<-(1-m)^dat$t
+          sum(dbinom(dat$Nfin, dat$Nini, Psurv, log=TRUE))
+          }
```

And here is the code to maximize the likelihood using `optim` and get the profile likelihood for confidence intervals:

```
> get.profCI.m<-function(datf){
+         res<-optimize(mylogLik,c(0,1),datf,maximum=TRUE)
+         mgrid<-seq(0,.1, length.out=1000)
+         proflik<-numeric()
+         for(i in 1:length(mgrid))
+                 proflik[i]<-mylogLik(mgrid[i],datf)
+         liklim<-res$objective-qchisq(p=.95,df=1)
+         ms<-mgrid[which(proflik>=liklim)]
+         return(c(m=res$maximum,LCI=min(ms), UCI=max(ms)))
+ }
>
```

Table L.1 summarizes the results for m for all species with any deaths in the inventory. Table L.2 compares m calculations with the same result from my simple state-space model, for each species from Chapter 2 (including all species aggregated). The annual mortality probability, $1 - expit(b)$, for the simple state-space model is shown alongside m, estimated using Bayesian MCMC with credible intervals from highest posterior density calculations (Plummer et al. 2010). These two calculations should and do agree closely for each species.

The following tables also show the results for annual mortality probability for a model with a plot effect and no explanatory variables. The right-hand side of table L.2 shows mortality for high-mortality (low survival, $1-(expit(b-\sigma_p))$), mean mortality (1- $expit(b)$), and low-mortality (high survival, $1-(expit(b+\sigma_p))$) plots. Note that because random plot effects for sugar pine, black oak, and tanoak are poorly determined, the annual mortality estimates show a great deal of overlap between mean and high/low mortality plots. They also

have lower credible bounds of zero. See Appendix N for posterior densities of plot standard deviations for all species.

The large differences in high- and low-mortality plots highlight the potential difficulties in using a single number to represent a species' survival. Sheil and May (1996) indicate that annual mortality probabilities like m will be influenced by temporal variation in mortality processes and census interval lengths. This observation highlights the need to use a state-space model to handle the mosaic of census intervals. In addition, these simple models mask the lower survival rates of susceptible subsets of the population, as concluded in Chapter 2. Sheil and May (1996) additionally point out that "different species demonstrate varying vulnerabilities when exposed to different environmental phenomena or conditions" and the more heterogeneous the population, the worse the effects on m.

Table L.2 also includes the mortality rate associated with an average tree $100 * (1 - expit(b))$ in the final model for each species. Note that due to Jensen's Inequality, and the extreme nonlinearity of the logit near 0 and 1, I expect the intercept $expit(b)$ will be pushed closer and closer to 1 and the corresponding mortality closer to 0 as I add complexity to the model. Therefore the rising survival of the 'average' tree partly has to do with the model's complexity and partly to do with the nonlinearity of the logit/expit. The model's complexity, however, reflects the complexity in the real system.

Species	Deaths	Trees	(m)
All	833	3319	1.79 (1.62, 1.95)
IC	356	1378	1.81 (1.56, 2.08)
WF	231	846	2.14 (1.78, 2.54)
DF	78	395	1.39 (1.01, 1.87)
PP	90	332	1.62 (1.19, 2.12)
TO	17	110	1.28 (0.61, 2.32)
BO	20	104	1.21 (0.61, 2.1)
SP	31	101	2.3 (1.34, 3.61)
YW	3	18	1.27 (0.15, 4.44)
CH	6	16	2.75 (0.69, 7)
AL	1	7	0.64 (0.01, 4.23)

Table L.1: m results for all species.

	Maximum Likelihood	Bayesian State-Space Models				
Species	Simple m	$expit(b)$ only	With Plot Effect			Final Model
			Low	Mean	High	
All Species	1.79 (1.62, 1.95)	1.79 (1.66, 1.9)	0.7 (0.53, 0.87)	1.5 (1.27, 1.75)	3.12 (2.56, 3.77)	1.07 (0.83, 1.32)
Incense-cedar	1.81 (1.56, 2.08)	1.82 (1.64, 2.01)	0.52 (0.29, 0.73)	1.44 (1.06, 1.84)	3.92 (2.78, 5.16)	1.06 (0.71, 1.44)
White Fir	2.14 (1.78, 2.54)	2.14 (1.87, 2.41)	0.67 (0.41, 0.96)	1.69 (1.26, 2.15)	4.11 (3, 5.43)	2.13 (1.51, 2.73)
Douglas-fir	1.39 (1.01, 1.87)	1.41 (1.12, 1.73)	0.63 (0.24, 1.09)	1.13 (0.72, 1.55)	2.1 (1.39, 2.93)	0.99 (0.62, 1.39)
Ponderosa Pine	1.62 (1.19, 2.12)	1.63 (1.31, 1.98)	0.55 (0.22, 0.91)	1.43 (0.91, 1.94)	3.7 (2.15, 5.56)	0.72 (0.38, 1.08)
Tanoak	1.28 (0.61, 2.32)	1.37 (0.79, 2)	0.49 (0, 1.32)	2 (0.05, 4.09)	12.26 (0.67, 76.17)	0.9 (0.23, 1.65)
Black Oak	1.21 (0.61, 2.1)	1.25 (0.75, 1.8)	0.15 (0, 0.46)	1.17 (0.09, 2.49)	14.15 (0.97, 50.92)	0.46 (0.02, 1.2)
Sugar Pine	2.3 (1.34, 3.61)	2.37 (1.58, 3.18)	1.09 (0.11, 2.11)	2.2 (1.09, 3.31)	4.96 (1.89, 9.5)	1.53 (0.55, 2.51)

Table L.2: Comparison of m calculations, $1 - expit(b)$ calculations from the simple state-space model, mortality in high, mean, and low-mortality plots from the state-space model with plot effects, and $1 - expit(b)$ from the final model as selected for each species. Note that the m values and simple state-space model values agree closely and that the final model's $1 - expit(b)$ is typically smaller than either the mean plot or simple models.

Appendix M

Explanation of Posterior Tail Probability and Parameter Estimates for All Models for All Species

This appendix gives details on the model selection strategy, including how the posterior tail probability is calculated, and summarizes the parameter estimates for all models tested for each species. Code is included for calculating the posterior tail probability, which is based on `HPDinterval` (Plummer et al. 2010).

M.1 Explanation of the Posterior Tail Probability

Controversy has developed around model selection for Bayesian estimation methods (Celeux et al. 2006 and discussants), and in particular for hierarchical models and the deviance information criterion (DIC) (Celeux et al. 2006, Millar 2009). The DIC was derived "heuristically" for models without random effects (Spiegelhalter et al. 2002), based on asymptotic (large sample) results, specifically for a normal posterior. The caveats discussed by Plummer (2008) for spatial disease models also apply to my time-series models, which contain a large number of latent unmeasured states as well as random effects. The lack of a clear choice among possible random-effects versions of DIC (Celeux et al. 2006) has added to concerns about its behavior in real applications (Millar 2009, Kery and Schaub 2012). Spiegelhalter et al. (2014) remark that they were always aware of the limitations of the DIC and are surprised it has held up this well.

As described in Chapter 2, the posterior tail probability is a preferable way to rank potential variable additions and to determine when to stop adding variables. The posterior tail probability indicates how strong the support of the posterior is away from zero, complementary to the 95% credible interval which indicates how broad the posterior is in parameter space, i.e. many parameter values are highly supported. Figure M.1 shows a basic example,

and in Tables M.1 through M.8 I give the posterior mean, posterior tail probability, and 95 % credible interval for each parameter in each model for each species.

The posterior tail probability, practically speaking, is calculated using a version of "`HPDinterval`" (modified from "`coda`," Plummer et al. 2010) to check a grid of possible credible interval probabilities (including 95), and choosing the credible interval probability with one bound as close to zero as possible (see code at the end of this section). For example, the posterior tail probability for a parameter might be 0.20. This means that zero is at the edge of the 80 % most supported values (highest values of the posterior).

In this Bayesian context, a posterior tail probability of 0.001 would indicate that the credible interval with one bound at 0 is a 99.9% credible interval, containing the 99.9% most highly supported parameter values in the posterior. The reader may evaluate whether the posterior tail probability is high enough to warrant that variable's selection; compare a parameter with a posterior tail probability of 0.066 with a parameter with a posterior tail probability of 0.535. The latter parameter has only 46.5 % of the posterior within a credible interval with one bound at zero; one would say that zero has substantial support. The former parameter, while it has not met the 5 % threshold, may be judged to be important based on context (e.g. has a higher posterior tail probability in most other models).

This metric is justified for the case of comparing single variable additions to a model, as opposed to model selection with non-nested models. When comparing models differing by a single variable, the AIC difference and the p-value associated with a likelihood ratio test are monotonically related (Murtaugh 2014). For a Bayesian analysis where the likelihood dominates the prior (see Appendix K for details on flat priors for my models), the posterior tail area for a parameter that is away from zero will be very similar to the analogous p-value because of large-sample likelihood theory. Therefore, at each variable selection step, I select the variable with the highest posterior weight of differing from zero. Since this will be similar to the p-value, which varies monotonically with the AIC, the results should be similar to AIC model selection.

To examine whether the AIC-motivated approach would have differed from a DIC approach, I compared DIC values for models from the species with the fewest random effects: tanoak, which did not include plot effects. These models are closest to the case that the DIC was derived for. For the tanoak models, although not every comparison among specific variables would have had the same result, the final model did have the lowest DIC among all the models I calculated as part of the forward selection. Tables M.1 through M.8 list parameter estimates (posterior means), tail probabilities, and 95 % credible intervals for all models used in the forward selection strategy for each species.

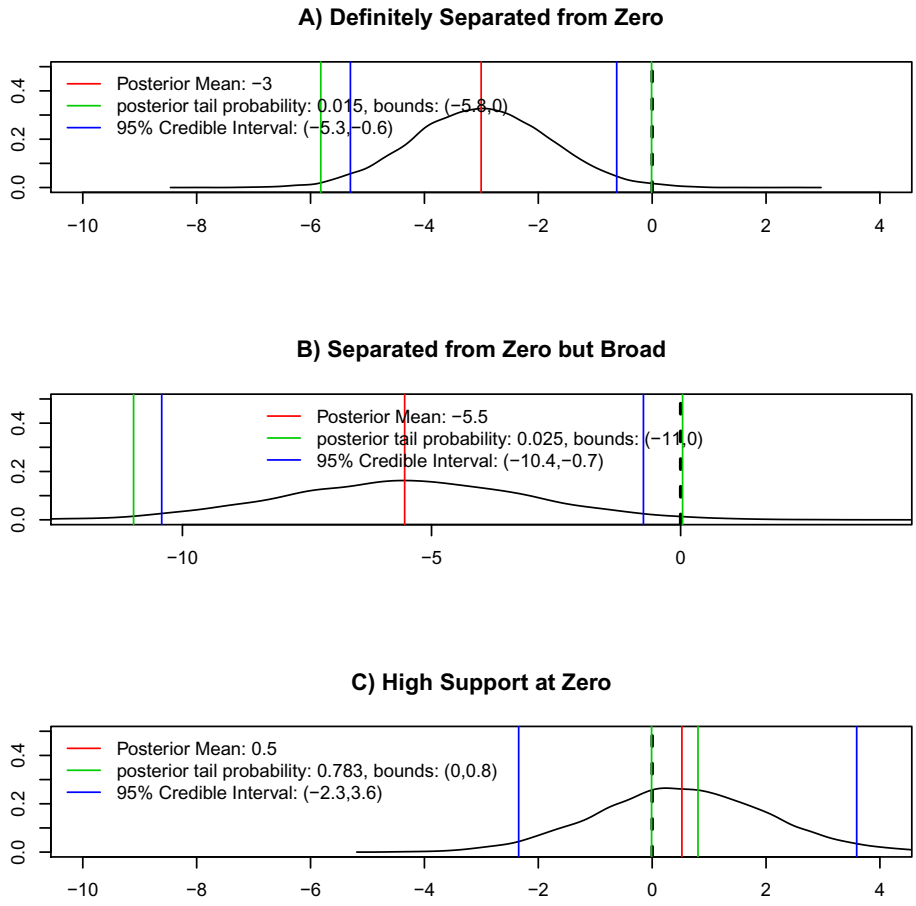

Figure M.1: Demonstration of posterior tail probability, 95% credible interval, and posterior mean. A) A case where the posterior is clearly separated from zero, with a posterior tail probability close to 0. B) A case with the same mean, but a wider posterior; the posterior tail probability is lower, but the 95 % credible interval helps to see just how much broader the posterior is and how uncertain the parameter estimate is. C) In this case, the posterior is close to zero; only the 78.3 % most probable values fall between 0 and 2.4.

M.2 Code to Obtain Posterior Tail Probability

```
> #workhorse piece of HPDinterval, coded separately so I can use apply
> evaluate.interval<-function(gap1,vals1){
+         nsamp1<-length(vals1)
+         temp1<-nsamp1-gap1
+         init1<-1:temp1
+         #starting point for each interval to evaluate
+         inds1<- which.min(vals1[init1 + gap1,drop=FALSE]
+         - vals1[init1,drop=FALSE])
+         #how long is the interval & pick smallest one
+         ans1 <- c(vals1[inds1],vals1[inds1 + gap1])
+         #return the endpoints of that interval
+         return(ans1)
+ }
> # a version of HPDinterval which can handle a vector of
> # probabilities, prob; obj is the vector of posterior samples
> HPDintervalMVE <- function(obj, prob){
+         nprob<-length(prob) # number of probabilities to do the calc for
+         vals <- sort(obj) # sort chain to get ecdf
+         nsamp <- length(vals) # number of samples in chain
+         gap <-round(nsamp * prob)
+         # length of credible interval in samples
+         gap[gap==nsamp]<-nsamp-1
+         # must have more than one gap to calculate
+         gap[gap==0]<-1 # cannot do the calc with a gap of 0
+         gap<-as.array(gap)
+         # for apply
+        # get the interval based on the gap
+         ans<-apply(gap,1,evaluate.interval,vals)
+         ans<-as.data.frame(t(ans)) #make it a dataframe
+         names(ans)<-c("lower","upper") #rename columns
+         ans$prob<-prob #add probability for convenience
+         return(ans)
+ }
> # function that calls HPDingervalMVE
> # nodechain is the chain for the node of interest, a
> # numeric variable containing all the posterior samples
> get.credible.level<-function(nodechain){
+         lgrid<-seq(0,1,length.out=10000)
+         #hard-code in how fine a grid to use of probs
+         CIs<-HPDintervalMVE(as.numeric(nodechain),lgrid)
+         #get the data frame of intervals
+         CIs$min<-apply(abs(CIs[,c("lower","upper")]),1,min)
```

```
+         #find the upper or lower bound closest to zero
+         res<-mean(CIs$prob[CIs$min==min(CIs$min)])
+         #pull out the prob corresponding to the closest to zero
+        #bound closest to zero
+         return(1-res)
+ }
>
```

M.3 Model Results for All Species, Variables, and Models

Tables M.1 through M.8 summarize model results for explanatory variables for each species. Variables included are: Mean survival, $expit(b)$; Plot standard deviation, σ_p ; Linear size, β^{DBH}; Quadratic size, β^{DBH^2}; Plot Basal Area, β^{ba}; Insolation, β^{insol}; Elevation, β^{elev}; Topographic Slope, β^{slope}; time trend, β^{t}; annual climatic water deficit, β^{def}. All variables are standardized (see Appendix K); therefore these parameters correspond to centered and scaled covariates. This means they can be directly compared across models and between variables, but not between species. Parameters which have tail probability less than 0.05% (alternatively, 95% credible interval does not overlap zero) are highlighted in **bold text**. For each variable and model, the table lists the posterior mean, the posterior tail probability (in parentheses), and the 95% credible interval (in parentheses: "lower, upper").

Table M.1: Survival model parameter estimates for all species aggregated.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.98, 0.99)	0.82 (0.213) (0, 1.07)	**0.72 (0) (0.53, 0.86)**							
0.99 (0) (0.99, 0.99)	0.94 (0) (0.76, 1.12)	**0.86 (0) (0.75, 0.97)**	**-0.15 (0) (-0.2, -0.1)**						
0.98 (0) (0.98, 0.99)	0.7 (0.219) (0, 0.93)		**0.12 (0) (0.07, 0.17)**						
0.98 (0) (0.98, 0.99)	0.73 (0.012) (0.58, 0.95)			0.16 (0.085) (-0.01, 0.34)					
0.98 (0) (0.98, 0.99)	0.72 (0.209) (0, 0.91)				-0.03 (0.75) (-0.19, 0.12)				
0.98 (0) (0.98, 0.99)	0.68 (0.309) (0, 0.89)					0.13 (0.077) (-0.02, 0.29)			
0.99 (0) (0.98, 0.99)	0.63 (0.202) (0, 0.92)						-0.46 (0.13) (-1.01, 0.11)		
0.99 (0.286) (0.98, 0.99)	0.77 (0) (0.61, 0.94)							**-0.38 (0) (-0.59, -0.18)**	
0.99 (0) (0.98, 0.99)	0.78 (0) (0.62, 0.95)								**-0.13 (0.035) (-0.25, -0.01)**
0.99 (0) (0.99, 0.99)	0.93 (0) (0.76, 1.12)	**0.86 (0) (0.74, 0.96)**	**-0.15 (0) (-0.19, -0.1)**	0.13 (0.216) (-0.07, 0.34)					
0.99 (0) (0.98, 0.99)	0.87 (0.344) (0, 1.11)	0.84 (0.375) (0.64, 0.98)	**-0.14 (0) (-0.19, -0.09)**		-0.05 (0.986) (-0.27, 0.14)				
0.99 (0) (0.98, 0.99)	0.72 (0.238) (0, 1.08)	**0.81 (0) (0.57, 0.97)**	**-0.14 (0) (-0.19, -0.08)**			0.13 (0.145) (-0.06, 0.35)			

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0.001) (0.98, 0.99)	0.77 (0.232) (0, 1.11)	**0.81 (0) (0.59, 0.97)**	**-0.14 (0) (-0.19, -0.08)**				-0.26 (0.76) (-0.93, 0.39)		
0.99 (0) (0.99, 0.99)	0.97 (0.001) (0.79, 1.17)	**0.85 (0) (0.74, 0.96)**	**-0.14 (0) (-0.19, -0.09)**					**-0.41 (0.001) (-0.66, -0.16)**	
0.99 (0) (0.99, 0.99)	0.97 (0.001) (0.79, 1.16)	**0.86 (0) (0.74, 0.97)**	**-0.14 (0) (-0.19, -0.09)**						**-0.16 (0.01) (-0.29, -0.04)**
0.99 (0) (0.99, 0.99)	0.9 (0.385) (0, 1.14)	**0.83 (0) (0.62, 0.98)**	**-0.14 (0) (-0.19, -0.09)**	0.08 (0.494) (-0.13, 0.28)				**-0.39 (0) (-0.61, -0.17)**	
0.99 (0) (0.99, 0.99)	0.97 (0) (0.79, 1.17)	**0.85 (0) (0.73, 0.95)**	**-0.14 (0) (-0.18, -0.08)**		0.02 (0.901) (-0.19, 0.26)			**-0.4 (0) (-0.65, -0.19)**	
0.99 (0) (0.99, 0.99)	0.97 (0) (0.78, 1.17)	**0.85 (0) (0.73, 0.95)**	**-0.14 (0) (-0.19, -0.09)**			0.11 (0.27) (-0.1, 0.29)		**-0.41 (0.002) (-0.66, -0.16)**	
0.99 (0) (0.99, 0.99)	0.97 (0) (0.79, 1.19)	0.85 (0.25) (0.75, 0.97)	**-0.14 (0) (-0.19, -0.09)**				-0.14 (0.947) (-0.68, 0.37)	**-0.4 (0.004) (-0.63, -0.14)**	
0.99 (0) (0.99, 0.99)	0.92 (0.023) (0, 1.15)	**0.84 (0) (0.69, 1)**	**-0.14 (0) (-0.19, -0.09)**					**-0.4 (0.003) (-0.67, -0.09)**	0 (0.892) (-0.17, 0.18)

Table M.2: Survival model parameter estimates for incense-cedar.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
1 (0) (0.99, 1)	1.35 (0.338) (0, 1.9)	**0.69 (0) (0.3, 1.02)**							
1 (0) (0.99, 1)	1.44 (0.006) (0.96, 2.08)	**0.7 (0.002) (0.32, 1.04)**	0.04 (0.938) (-0.13, 0.23)						
1 (0.4) (0.99, 1)	1.32 (0) (0.89, 1.79)		0.26 (0.155) (0.08, 0.46)						

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
1 (0) (0.99, 1)	1.25 (0.001) (0.85, 1.75)			0.41 (0.08) (-0.01, 0.87)					
1 (0) (0.99, 1)	1.1 (0.295) (0, 1.68)				-0.15 (0.532) (-0.52, 0.17)				
1 (0) (0.99, 1)	1.26 (0.002) (0.83, 1.79)					0.05 (0.876) (-0.35, 0.5)			
1 (0) (1, 1)	1.65 (0.01) (1.15, 2.37)						2.21 (0.4) (1.65, 2.85)		
1 (0.4) (1, 1)	1.82 (0) (1.28, 2.43)							**-2.65** **(0)** **(-3.28, -2.01)**	
1 (0) (1, 1)	1.9 (0.001) (1.34, 2.51)								**-1.28** **(0)** **(-1.62, -0.97)**
1 (0) (0.99, 1)	1.44 (0.012) (0.96, 2.15)	**0.71** **(0.001)** **(0.34, 1.11)**	0.03 (0.912) (-0.13, 0.21)	0.46 (0.112) (-0.07, 1.03)					
1 (0) (0.99, 1)	1.31 (0.321) (0, 1.93)	**0.65** **(0.007)** **(0.18, 1.03)**	0.05 (0.977) (-0.13, 0.25)		-0.19 (0.685) (-0.66, 0.18)				
1 (0) (0.99, 1)	1.18 (0.236) (0, 1.9)	**0.63** **(0.011)** **(0.12, 1.07)**	0.05 (0.873) (-0.13, 0.27)			-0.01 (0.941) (-0.41, 0.4)			
1 (0) (1, 1)	1.94 (0) (1.37, 2.53)	**0.86** **(0)** **(0.46, 1.21)**	0.02 (0.788) (-0.15, 0.24)				**2.38** **(0)** **(1.81, 3.05)**		
1 (0) (1, 1)	2.26 (0.001) (1.57, 2.96)	**0.86** **(0)** **(0.51, 1.22)**	0.06 (0.643) (-0.13, 0.27)					**-3.11** **(0)** **(-3.9, -2.31)**	
1 (0) (1, 1)	2.26 (0) (1.65, 2.98)	**0.88** **(0)** **(0.54, 1.23)**	0.04 (0.846) (-0.13, 0.25)						**-1.45** **(0)** **(-1.8, -1.13)**
1 (0)	2.3 (0)	**0.84** **(0)**	0.06 (0.963)	0.05 (0.902)				**-3.12** **(0)**	

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
(1, 1)	(1.61, 3.05)	**(0.49, 1.21)**	(-0.14, 0.27)	(-0.62, 0.75)				**(-3.88, -2.37)**	
1 (0) (1, 1)	2.3 (0) (1.58, 3.05)	**0.84** **(0)** **(0.49, 1.18)**	0.05 (0.814) (-0.13, 0.27)		0.02 (0.874) (-0.58, 0.58)			**-3.13** **(0)** **(-3.95, -2.33)**	
1 (0) (1, 1)	2.24 (0.278) (1.53, 3.23)	**0.84** **(0.005)** **(0.4, 1.3)**	0.05 (0.854) (-0.15, 0.27)			-0.08 (0.92) (-0.81, 0.6)		**-3.04** **(0.001)** **(-4.11, -2.19)**	
1 (0) (1, 1)	2.04 (0.001) (1.41, 2.7)	**0.9** **(0)** **(0.55, 1.3)**	0.07 (0.854) (-0.13, 0.31)				**1.92** **(0.01)** **(0.98, 2.91)**	**-2.25** **(0.002)** **(-2.88, -1.54)**	
1 (0) (1, 1)	2.06 (0) (1.48, 2.71)	**0.86** **(0)** **(0.51, 1.22)**	0.06 (0.7) (-0.13, 0.26)					**-1.75** **(0)** **(-2.75, -0.75)**	**-0.77** **(0.001)** **(-1.26, -0.25)**
1 (0) (1, 1)	1.92 (0.24) (0, 2.68)	**0.8** **(0.025)** **(0.11, 1.25)**	0.07 (0.764) (-0.12, 0.31)	0.26 (0.502) (-0.39, 0.98)				**-1.54** **(0)** **(-2.8, -0.61)**	**-0.8** **(0.003)** **(-1.35, -0.16)**
1 (0) (1, 1)	2.09 (0) (1.49, 2.78)	**0.86** **(0)** **(0.5, 1.23)**	0.05 (0.786) (-0.13, 0.25)		-0.09 (0.717) (-0.7, 0.47)			**-1.72** **(0)** **(-2.63, -0.72)**	**-0.78** **(0.001)** **(-1.3, -0.32)**
1 (0) (1, 1)	2.08 (0) (1.47, 2.73)	**0.85** **(0)** **(0.5, 1.19)**	0.06 (0.583) (-0.13, 0.27)			-0.01 (0.807) (-0.55, 0.54)		**-1.78** **(0)** **(-2.7, -0.91)**	**-0.75** **(0.001)** **(-1.2, -0.3)**
	2.06 (0) (1.45, 2.75)	**0.88** **(0)** **(0.51, 1.26)**	0.06 (0.803) (-0.13, 0.27)				0.73 (0.36) (-1.37, 2.63)	**-1.89** **(0)** **(-3, -0.84)**	-0.46 (0.271) (-1.23, 0.26)
0.99 (0) (0.99, 1)	1.28 (0) (0.97, 1.62)	**1.02** **(0)** **(0.81, 1.24)**	**-0.09** **(0.017)** **(-0.16, -0.02)**	0.21 (0.224) (-0.14, 0.55)			-0.19 (0.826) (-1.23, 0.81)	-0.41 (0.105) (-0.9, 0.07)	-0.23 (0.116) (-0.53, 0.05)
0.99 (0) (0.99, 1)	1.28 (0) (0.92, 1.59)	**1.02** **(0)** **(0.81, 1.23)**	**-0.1** **(0.018)** **(-0.16, -0.02)**		0.01 (0.967) (-0.32, 0.34)		-0.14 (0.957) (-1.2, 0.79)	**-0.48** **(0.037)** **(-0.94, -0.02)**	-0.2 (0.164) (-0.47, 0.08)
0.99 (0) (0.99, 1)	1.26 (0) (0.94, 1.6)	**1.01** **(0)** **(0.78, 1.22)**	**-0.09** **(0.023)** **(-0.16, -0.01)**			0.15 (0.367) (-0.17, 0.47)	-0.03 (0.672) (-1.12, 0.88)	**-0.51** **(0.021)** **(-0.94, -0.06)**	-0.17 (0.208) (-0.42, 0.11)

Table M.3: Survival model parameter estimates for white fir.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.98 (0) (0.98, 0.99)	0.94 (0) (0.67, 1.22)	**0.61** **(0)** **(0.42,** **0.81)**							
0.98 (0) (0.97, 0.99)	0.95 (0) (0.69, 1.23)	**0.57** **(0)** **(0.35,** **0.79)**	**0.29** **(0.008)** **(0.08,** **0.52)**						
0.98 (0) (0.97, 0.98)	0.84 (0.209) (0, 1.17)		**0.39** **(0)** **(0.2,** **0.58)**						
0.98 (0) (0.98, 0.99)	0.93 (0.003) (0.66, 1.24)			0 (0.862) (-0.28, 0.27)					
0.98 (0) (0.98, 0.99)	0.94 (0) (0.68, 1.21)				-0.06 (0.684) (-0.32, 0.17)				
0.98 (0) (0.98, 0.99)	0.91 (0) (0.64, 1.18)					0.18 (0.162) (-0.1, 0.43)			
0.98 (0) (0.98, 0.99)	0.93 (0.001) (0.67, 1.22)						0.15 (0.565) (-0.52, 0.82)		
0.98 (0) (0.98, 0.99)	0.94 (0.001) (0.68, 1.24)							-0.17 (0.318) (-0.49, 0.14)	
0.98 (0) (0.98, 0.99)	0.91 (0.004) (0.65, 1.21)								0.05 (0.706) (-0.19, 0.29)
0.98 (0) (0.97, 0.99)	0.95 (0) (0.7, 1.24)	**0.57** **(0)** **(0.36,** **0.8)**	**0.3** **(0.005)** **(0.08,** **0.54)**	-0.08 (0.513) (-0.37, 0.17)					
0.98 (0) (0.97, 0.99)	0.95 (0.002) (0.68, 1.28)	**0.56** **(0)** **(0.35,** **0.8)**	**0.3** **(0.003)** **(0.08,** **0.52)**		-0.04 (0.86) (-0.29, 0.21)				
0.98 (0) (0.97, 0.98)	0.58 (0.343) (0, 1.12)	**0.55** **(0)** **(0.34,** **0.77)**	**0.28** **(0.007)** **(0.06,** **0.51)**			**0.26** **(0.035)** **(0.02,** **0.49)**			

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.98 (0) (0.97, 0.99)	0.86 (0.368) (0, 1.2)	**0.56 (0) (0.35, 0.8)**	**0.29 (0.006) (0.08, 0.53)**				0.21 (0.469) (-0.48, 0.87)		
0.98 (0) (0.98, 0.99)	0.97 (0) (0.7, 1.25)	**0.56 (0) (0.34, 0.77)**	**0.31 (0.006) (0.08, 0.53)**					-0.22 (0.177) (-0.53, 0.1)	
0.98 (0) (0.97, 0.99)	0.95 (0) (0.67, 1.22)	**0.57 (0) (0.35, 0.8)**	**0.29 (0.008) (0.07, 0.53)**						0.04 (0.999) (-0.22, 0.28)
0.98 (0) (0.97, 0.99)	0.93 (0) (0.68, 1.22)	**0.58 (0) (0.36, 0.8)**	**0.31 (0.004) (0.09, 0.52)**	-0.01 (0.92) (-0.28, 0.27)		0.24 (0.091) (-0.01, 0.51)			
0.98 (0) (0.97, 0.99)	0.64 (0.287) (0, 1.15)	**0.55 (0) (0.34, 0.77)**	**0.29 (0.008) (0.07, 0.54)**		-0.09 (0.55) (-0.37, 0.12)	**0.28 (0.026) (0.04, 0.54)**			
0.98 (0) (0.97, 0.99)	0.93 (0) (0.68, 1.21)	**0.58 (0) (0.37, 0.81)**	**0.31 (0.002) (0.09, 0.55)**			0.27 (0.061) (-0.02, 0.55)	0.34 (0.269) (-0.43, 0.97)		
0.98 (0) (0.97, 0.99)	0.82 (0.227) (0, 1.18)	**0.56 (0) (0.34, 0.79)**	**0.32 (0.003) (0.1, 0.56)**			0.25 (0.055) (0, 0.52)		-0.22 (0.185) (-0.54, 0.1)	
0.98 (0) (0.97, 0.99)	0.9 (0.002) (0.62, 1.2)	**0.59 (0) (0.37, 0.82)**	**0.29 (0.007) (0.06, 0.53)**			**0.27 (0.049) (0, 0.52)**			0.08 (0.539) (-0.16, 0.35)

Table M.4: Survival model parameter estimates for Douglas-fir.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.99, 0.99)	0.62 (0.005) (0.13, 1.04)	**0.83 (0) (0.48, 1.15)**							
0.99 (0) (0.99, 0.99)	0.65 (0.001) (0.26, 1.06)	**0.85 (0) (0.52, 1.17)**	-0.04 (0.375) (-0.22, 0.17)						
0.99 (0) (0.98, 0.99)	0.37 (0.333) (0, 0.91)		0.09 (0.376) (-0.05, 0.26)						

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.99, 0.99)	0.53 (0.037) (0, 0.97)			0.29 (0.105) (-0.06, 0.64)					
0.99 (0) (0.98, 0.99)	0.68 (0.001) (0.26, 1.08)				0.13 (0.489) (-0.22, 0.47)				
0.99 (0) (0.99, 0.99)	0.51 (0.376) (0, 0.95)					**-0.39 (0.043) (-0.77, -0.01)**			
0.99 (0) (0.99, 1)	0.68 (0.003) (0.27, 1.15)						0.32 (0.434) (-1.12, 1.55)		
0.99 (0) (0.99, 0.99)	0.74 (0) (0.35, 1.16)							-0.42 (0.121) (-0.98, 0.06)	
0.99 (0) (0.98, 0.99)	0.67 (0.012) (0.15, 1.2)								-0.23 (0.239) (-0.59, 0.15)
0.99 (0) (0.99, 0.99)	0.56 (0.002) (0.09, 0.95)	**0.84 (0) (0.51, 1.15)**	-0.04 (0.456) (-0.21, 0.17)	0.29 (0.107) (-0.06, 0.65)					
0.99 (0) (0.99, 0.99)	0.61 (0.084) (0, 1.02)	**0.84 (0) (0.54, 1.19)**	-0.04 (0.453) (-0.2, 0.16)		0.11 (0.49) (-0.23, 0.44)				
0.99 (0) (0.99, 1)	0.59 (0.014) (0.1, 1.06)	**0.83 (0.001) (0.52, 1.18)**	-0.04 (0.373) (-0.2, 0.16)			-0.33 (0.091) (-0.74, 0.03)			
0.99 (0) (0.99, 1)	0.66 (0.001) (0.24, 1.08)	**0.87 (0) (0.53, 1.21)**	-0.04 (0.441) (-0.21, 0.16)				0.59 (0.211) (-0.84, 1.75)		
0.99 (0) (0.99, 1)	0.74 (0.023) (0, 1.16)	**0.86 (0.001) (0.53, 1.2)**	-0.04 (0.389) (-0.21, 0.17)					-0.53 (0.064) (-1.08, 0.07)	
0.99 (0) (0.99, 1)	0.65 (0.298) (0, 1.11)	**0.86 (0) (0.54, 1.19)**	-0.03 (0.448) (-0.21, 0.19)						-0.32 (0.104) (-0.71, 0.06)
0.99 (0)	0.71 (0.016)	**0.87 (0)**	-0.04 (0.369)	0.22 (0.256)				-0.48 (0.139)	

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
(0.99, 1)	(0.18, 1.28)	**(0.53, 1.21)**	(-0.21, 0.17)	(-0.16, 0.62)				(-1.11, 0.09)	
0.99 (0) (0.99, 1)	0.8 (0.003) (0.39, 1.28)	**0.87 (0) (0.55, 1.21)**	-0.04 (0.363) (-0.21, 0.17)		0.24 (0.28) (-0.17, 0.67)			**-0.63 (0.029) (-1.17, -0.05)**	
0.99 (0) (0.99, 1)	0.76 (0.002) (0.25, 1.29)	**0.86 (0) (0.53, 1.19)**	-0.04 (0.403) (-0.22, 0.18)			-0.3 (0.167) (-0.74, 0.11)		-0.52 (0.082) (-1.11, 0.06)	
0.99 (0) (0.99, 1)	0.73 (0.009) (0.21, 1.33)	**0.9 (0) (0.56, 1.21)**	-0.03 (0.429) (-0.19, 0.18)				0.85 (0.127) (-0.62, 1.97)	**-0.56 (0.027) (-1.1, -0.03)**	
0.99 (0) (0.99, 1)	0.74 (0.105) (0, 1.16)	**0.88 (0) (0.56, 1.21)**	-0.03 (0.472) (-0.21, 0.2)					-0.44 (0.115) (-0.96, 0.08)	-0.27 (0.17) (-0.67, 0.12)
0.99 (0) (0.99, 1)	0.66 (0.007) (0.13, 1.16)	**0.88 (0.001) (0.53, 1.22)**	-0.03 (0.439) (-0.2, 0.21)	0.26 (0.19) (-0.09, 0.67)			0.86 (0.143) (-0.79, 1.9)	-0.45 (0.109) (-0.97, 0.12)	
0.99 (0) (0.99, 1)	0.76 (0.002) (0.31, 1.25)	**0.9 (0) (0.54, 1.25)**	-0.04 (0.425) (-0.22, 0.19)		0.21 (0.285) (-0.13, 0.59)		0.82 (0.129) (-0.47, 1.87)	**-0.63 (0.029) (-1.16, -0.08)**	
0.99 (0) (0.99, 1)	0.73 (0.003) (0.2, 1.25)	**0.86 (0) (0.54, 1.21)**	-0.05 (0.405) (-0.22, 0.15)			-0.24 (0.233) (-0.71, 0.18)	0.74 (0.177) (-0.6, 1.82)	**-0.56 (0.047) (-1.14, -0.01)**	
0.99 (0) (0.99, 1)	0.74 (0.011) (0.22, 1.32)	**0.88 (0) (0.55, 1.21)**	-0.04 (0.472) (-0.21, 0.16)				0.61 (0.333) (-0.69, 1.95)	-0.52 (0.071) (-1.11, 0.02)	-0.09 (0.624) (-0.56, 0.38)

Table M.5: Survival model parameter estimates for ponderosa pine.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.98, 0.99)	1.05 (0.001) (0.56, 1.57)	**0.86 (0) (0.58, 1.14)**							
0.99 (0) (0.99, 1)	0.93 (0.331) (0, 1.44)	0.84 (0.286) (0.57, 1.1)	**-0.43 (0) (-0.61, -0.23)**						
0.99 (0)	0.86 (0.004)		**-0.25 (0.004)**						

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
(0.98, 0.99)	(0.34, 1.35)		**(-0.41, -0.1)**						
0.99 (0) (0.98, 0.99)	0.89 (0.001) (0.37, 1.41)			0.29 (0.07) (-0.04, 0.6)					
0.99 (0) (0.98, 0.99)	1.01 (0) (0.57, 1.51)				-0.01 (0.999) (-0.35, 0.36)				
0.99 (0) (0.98, 0.99)	0.99 (0) (0.54, 1.49)					0.14 (0.423) (-0.24, 0.5)			
0.99 (0) (0.98, 1)	1.02 (0) (0.56, 1.51)						0.28 (0.56) (-1.22, 1.63)		
0.99 (0) (0.98, 0.99)	1.05 (0.001) (0.58, 1.56)							-0.27 (0.287) (-0.75, 0.18)	
0.99 (0) (0.98, 0.99)	1 (0) (0.56, 1.48)								-0.26 (0.161) (-0.61, 0.09)
0.99 (0) (0.99, 1)	0.87 (0.003) (0.32, 1.43)	**0.85 (0) (0.58, 1.1)**	**-0.43 (0) (-0.62, -0.24)**	**0.37 (0.01) (0.08, 0.67)**					
0.99 (0) (0.99, 1)	1.07 (0.001) (0.6, 1.61)	**0.86 (0) (0.61, 1.1)**	**-0.43 (0) (-0.62, -0.24)**		0.03 (0.835) (-0.34, 0.38)				
0.99 (0) (0.99, 1)	1.01 (0.001) (0.51, 1.56)	**0.86 (0) (0.6, 1.11)**	**-0.43 (0) (-0.61, -0.24)**			0.26 (0.213) (-0.12, 0.69)			
1 (0) (0.99, 1)	1.1 (0.085) (0, 1.69)	**0.9 (0) (0.63, 1.24)**	**-0.44 (0) (-0.63, -0.23)**				0.29 (0.459) (-2.01, 2.29)		
0.99 (0) (0.99, 1)	1.07 (0.199) (0, 1.68)	**0.92 (0) (0.62, 1.21)**	**-0.45 (0) (-0.64, -0.25)**					**-0.65 (0.008) (-1.21, -0.12)**	
0.99 (0) (0.99, 1)	1.23 (0) (0.67, 1.79)	**0.94 (0) (0.67, 1.24)**	**-0.43 (0) (-0.62, -0.24)**						**-0.5 (0.01) (-0.91, -0.12)**

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.99, 1)	0.92 (0.001) (0.4, 1.47)	**0.85 (0) (0.59, 1.1)**	**-0.43 (0) (-0.62, -0.25)**	**0.37 (0.034) (0.05, 0.72)**	0 (0.968) (-0.33, 0.33)				
0.99 (0) (0.99, 1)	0.67 (0.247) (0, 1.28)	**0.83 (0) (0.59, 1.11)**	**-0.42 (0) (-0.6, -0.24)**	**0.35 (0.02) (0.05, 0.65)**		0.24 (0.112) (-0.08, 0.56)			
1 (0) (0.99, 1)	0.85 (0.327) (0, 1.46)	**0.89 (0) (0.62, 1.19)**	**-0.43 (0) (-0.63, -0.2)**	**0.4 (0.007) (0.11, 0.71)**			0.57 (0.267) (-2.39, 2.32)		
0.99 (0) (0.99, 1)	1.07 (0) (0.51, 1.7)	**0.93 (0) (0.66, 1.19)**	**-0.45 (0) (-0.63, -0.25)**	0.26 (0.168) (-0.13, 0.65)				**-0.51 (0.04) (-1.09, -0.03)**	
0.99 (0) (0.99, 1)	1.06 (0.001) (0.51, 1.66)	**0.94 (0) (0.68, 1.2)**	**-0.42 (0) (-0.6, -0.24)**	**0.43 (0.019) (0.08, 0.81)**					**-0.5 (0.015) (-0.9, -0.09)**
0.99 (0) (0.99, 1)	1.12 (0.001) (0.6, 1.73)	**0.94 (0) (0.66, 1.2)**	**-0.43 (0) (-0.62, -0.23)**	**0.41 (0.039) (0.03, 0.79)**	-0.02 (0.904) (-0.38, 0.41)				**-0.52 (0.007) (-0.94, -0.12)**
0.99 (0) (0.99, 1)	1.07 (0.001) (0.57, 1.64)	**0.95 (0) (0.67, 1.21)**	**-0.42 (0) (-0.6, -0.24)**	**0.4 (0.024) (0.06, 0.75)**		0.24 (0.219) (-0.13, 0.65)			**-0.5 (0.015) (-0.9, -0.12)**
0.99 (0) (0.99, 1)	1.06 (0.016) (0.42, 1.78)	**0.94 (0) (0.65, 1.21)**	**-0.43 (0) (-0.62, -0.23)**	**0.43 (0.012) (0.1, 0.77)**			0.05 (0.876) (-1.85, 1.62)		-0.4 (0.052) (-0.83, 0.01)
0.99 (0) (0.99, 1)	1.11 (0) (0.57, 1.69)	**0.96 (0) (0.68, 1.24)**	**-0.44 (0) (-0.61, -0.24)**	0.36 (0.077) (-0.05, 0.75)				-0.29 (0.346) (-0.85, 0.19)	-0.4 (0.077) (-0.86, 0.04)

Table M.6: Survival model parameter estimates for sugar pine.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.98 (0) (0.97, 0.99)	1 (0.05) (0, 1.98)	**0.75 (0.002) (0.22, 1.3)**							
0.98 (0) (0.97, 1)	1.07 (0.022) (0.04, 2.16)	**0.8 (0.002) (0.25, 1.41)**	-0.07 (0.999) (-0.54, 0.39)						

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.97 (0) (0.96, 0.99)	1.02 (0.033) (0.01, 2.02)		0.15 (0.637) (-0.21, 0.56)						
0.98 (0) (0.97, 0.99)	0.92 (0.01) (0.04, 1.82)			0.17 (0.649) (-0.41, 0.78)					
0.98 (0) (0.97, 0.99)	0.83 (0.018) (0.03, 1.73)				0.11 (0.613) (-0.41, 0.62)				
0.98 (0) (0.97, 0.99)	0.88 (0.007) (0.05, 1.73)					0.13 (0.612) (-0.46, 0.69)			
0.98 (0) (0.97, 0.99)	0.83 (0.005) (0.05, 1.66)						-0.19 (0.992) (-1.2, 0.88)		
0.98 (0) (0.97, 0.99)	0.88 (0.056) (0.02, 1.76)							0.12 (0.725) (-0.57, 0.87)	
0.98 (0) (0.97, 0.99)	0.85 (0.014) (0.04, 1.6)								**0.81** **(0.015)** **(0.15,** **1.5)**
0.98 (0) (0.97, 1)	1.1 (0.011) (0.03, 2.35)	**0.84** **(0.001)** **(0.28,** **1.44)**	-0.1 (0.511) (-0.55, 0.39)	0.31 (0.338) (-0.4, 1)					
0.98 (0) (0.97, 0.99)	0.92 (0.48) (0, 2.12)	**0.81** **(0.001)** **(0.29,** **1.4)**	-0.09 (0.662) (-0.54, 0.36)		0.27 (0.257) (-0.28, 0.83)				
0.98 (0) (0.97, 1)	1.17 (0.14) (0, 2.33)	**0.82** **(0.003)** **(0.29,** **1.45)**	-0.07 (0.684) (-0.51, 0.42)			0.12 (0.75) (-0.53, 0.76)			
0.99 (0) (0.97, 1)	1.09 (0.021) (0.04, 2.22)	**0.81** **(0.004)** **(0.25,** **1.44)**	-0.06 (0.688) (-0.53, 0.39)				-0.21 (0.776) (-1.36, 0.93)		
0.98 (0) (0.97, 1)	1.15 (0.007) (0.07, 2.26)	**0.85** **(0.003)** **(0.25,** **1.48)**	-0.08 (0.784) (-0.55, 0.4)					-0.2 (0.559) (-1.05, 0.6)	
0.99 (0)	0.84 (0.012)	**0.7** **(0.003)**	-0.13 (0.464)						**0.71** **(0.043)**

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
(0.98, 1)	(0.03, 1.8)	**(0.17, 1.23)**	(-0.59, 0.3)						**(0.02, 1.42)**
0.99 (0) (0.98, 1)	1.02 (0.014) (0.05, 2.04)	**0.78 (0.002) (0.24, 1.37)**	-0.15 (0.49) (-0.63, 0.31)	0.2 (0.554) (-0.48, 0.86)					0.67 (0.065) (-0.06, 1.35)
0.99 (0) (0.97, 1)	0.83 (0.049) (0.01, 1.86)	**0.76 (0.002) (0.25, 1.32)**	-0.13 (0.497) (-0.56, 0.37)		0.27 (0.29) (-0.3, 0.81)				**0.69 (0.049) (0, 1.37)**
0.99 (0) (0.98, 1)	1.02 (0.008) (0.08, 2.01)	**0.72 (0.004) (0.17, 1.24)**	-0.13 (0.566) (-0.57, 0.33)			0.11 (0.672) (-0.53, 0.83)			**0.7 (0.047) (0.02, 1.46)**
0.99 (0) (0.98, 1)	0.98 (0.246) (0, 1.99)	**0.72 (0.006) (0.22, 1.27)**	-0.12 (0.51) (-0.6, 0.33)				-0.64 (0.558) (-2.31, 1.02)		**0.86 (0.022) (0.07, 1.62)**
0.99 (0) (0.98, 1)	1.02 (0.343) (0, 2.07)	**0.83 (0.002) (0.26, 1.45)**	-0.19 (0.407) (-0.64, 0.34)					-0.53 (0.369) (-1.54, 0.36)	**0.88 (0.026) (0.1, 1.7)**

Table M.7: Survival model parameter estimates for black oak.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.97, 1)	2.97 (0.003) (1.09, 5.52)	0.65 (0.105) (-0.17, 1.44)							
1 (0) (0.99, 1)	3.08 (0.291) (0, 5.98)	**1.58 (0.017) (0.16, 2.82)**	**-0.86 (0.005) (-1.65, -0.22)**						
0.99 (0) (0.98, 1)	2.72 (0.005) (0.65, 5.24)		-0.29 (0.205) (-0.72, 0.08)						
0.99 (0) (0.98, 1)	2.22 (0.025) (0.02, 4.35)			1.06 (0.999) (-0.17, 2.39)					
0.99 (0) (0.97, 1)	2.88 (0.002) (0.93, 5.6)				-0.23 (0.81) (-1.58, 1.26)				
0.99 (0)	2.79 (0.007)					0.49 (0.477)			

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
(0.97, 1)	(0.64, 5.66)					(-0.95, 1.98)			
0.99 (0) (0.98, 1)	2.6 (0.002) (0.76, 4.97)						-0.04 (0.832) (-1.9, 1.71)		
0.99 (0) (0.97, 1)	2.79 (0.004) (0.81, 5.4)							-0.4 (0.533) (-1.64, 0.94)	
0.99 (0) (0.98, 1)	2.55 (0.005) (0.75, 4.85)								0.03 (0.922) (-0.84, 0.84)
1 (0) (0.99, 1)	3.1 (0.002) (0.87, 5.91)	**1.76 (0) (0.57, 2.94)**	**-0.81 (0.002) (-1.46, -0.26)**	1.52 (0.397) (-0.17, 3.38)					
0.99 (0) (0.98, 1)	3.96 (0.004) (1.3, 7.59)	**1.78 (0.002) (0.56, 3.05)**	**-0.93 (0.001) (-1.63, -0.28)**		-0.39 (0.646) (-2.17, 1.37)				
1 (0) (0.99, 1)	3.73 (0.002) (1.26, 7.19)	**1.76 (0) (0.64, 2.96)**	**-0.94 (0.002) (-1.65, -0.35)**			0.82 (0.381) (-1.13, 2.83)			
1 (0) (0.99, 1)	3.47 (0.2) (0, 6.59)	**1.67 (0.005) (0.4, 3.21)**	**-0.91 (0.004) (-1.79, -0.18)**				-0.14 (0.901) (-2.18, 1.52)		
1 (0) (0.99, 1)	3.52 (0.002) (1.21, 6.48)	**1.87 (0.001) (0.66, 3.33)**	**-0.96 (0.002) (-1.7, -0.3)**					-0.96 (0.206) (-2.46, 0.61)	
1 (0) (0.99, 1)	3.6 (0.002) (1.27, 6.64)	**1.8 (0.001) (0.6, 3.09)**	**-0.95 (0.001) (-1.67, -0.26)**						-0.1 (0.855) (-1.05, 0.88)
1 (0) (0.98, 1)	4.37 (0.006) (0.96, 10.17)	**2.01 (0.001) (0.67, 3.28)**	**-0.94 (0) (-1.72, -0.3)**	2 (0.132) (-0.63, 5.07)	-1.11 (0.55) (-4.74, 1.8)				
1 (0) (0.99, 1)	3.27 (0.002) (0.9, 6.5)	**1.9 (0) (0.76, 3.28)**	**-0.92 (0.001) (-1.63, -0.26)**	1.61 (0.068) (-0.13, 3.55)		1.03 (0.305) (-0.43, 2.98)			
1 (0) (0.99, 1)	2.4 (0.293) (0, 5.43)	**1.64 (0.001) (0.41, 3.05)**	**-0.78 (0.001) (-1.48, -0.22)**	1.5 (0.064) (-0.08, 3.28)			0.02 (0.677) (-1.98, 1.56)		

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
1 (0) (0.99, 1)	3.55 (0.006) (1, 6.96)	**1.99** **(0.001)** **(0.76,** **3.39)**	**-0.9** **(0.001)** **(-1.66,** **-0.27)**	1.51 (0.114) (-0.32, 3.41)				-0.67 (0.422) (-2.2, 1.03)	
1 (0) (0.99, 1)	3.44 (0.004) (0.94, 6.46)	**1.92** **(0)** **(0.7,** **3.19)**	**-0.92** **(0.001)** **(-1.61,** **-0.27)**	1.61 (0.089) (-0.24, 3.55)					-0.3 (0.628) (-1.25, 0.59)

Table M.8: Survival model parameter estimates for tanoak.

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.99 (0) (0.98, 0.99)		0.45 (0.059) (-0.02, 0.94)							
0.99 (0) (0.98, 1)		0.38 (0.127) (-0.08, 0.88)	-0.12 (0.4) (-0.44, 0.23)						
0.99 (0) (0.98, 1)			-0.2 (0.163) (-0.5, 0.1)						
0.99 (0) (0.98, 1)				-0.58 (0.284) (-1.53, 0.33)					
0.98 (0) (0.97, 0.99)					**-0.64** **(0.033)** **(-1.2, -** **0.07)**				
0.99 (0) (0.98, 0.99)						-0.34 (0.547) (-1.13, 0.36)			
0.99 (0) (0.98, 1)							-0.25 (0.94) (-2.01, 1.24)		
0.99 (0) (0.98, 0.99)								-0.28 (0.566) (-1.36, 0.69)	
0.99 (0) (0.98, 0.99)									-0.22 (0.609) (-1.02, 0.58)

$expit(b)$	σ_p	β^{DBH}	β^{DBH^2}	β^{slope}	β^{elev}	β^{insol}	β^{def}	β^{ba}	β^{t}
0.98 (0) (0.98, 0.99)		0.35 (0.135) (-0.09, 0.81)			-0.54 (0.079) (-1.14, 0.08)				
0.98 (0) (0.97, 0.99)		0.37 (0.136) (-0.11, 0.88)	0.07 (0.88) (-0.36, 0.52)		-0.62 (0.097) (-1.43, 0.1)				
0.99 (0) (0.98, 1)				-0.39 (0.384) (-1.18, 0.37)	-0.59 (0.061) (-1.17, 0.02)				
0.98 (0) (0.98, 0.99)					**-0.61 (0.047) (-1.17, -0.01)**	-0.18 (0.999) (-0.89, 0.46)			
0.99 (0) (0.98, 1)					**-0.64 (0.034) (-1.16, -0.02)**		-0.19 (0.914) (-1.75, 1.27)		
0.98 (0) (0.97, 0.99)					**-0.63 (0.043) (-1.21, -0.04)**			-0.04 (0.901) (-1.12, 0.96)	
0.98 (0) (0.97, 0.99)					**-0.63 (0.042) (-1.22, -0.05)**				-0.05 (0.882) (-0.86, 0.78)
0.99 (0) (0.98, 1)		**0.61 (0.015) (0.1, 1.08)**		**-0.89 (0.032) (-1.77, -0.07)**	-0.54 (0.099) (-1.17, 0.15)				
0.99 (0) (0.98, 0.99)		0.35 (0.14) (-0.1, 0.8)			-0.52 (0.098) (-1.15, 0.07)	-0.18 (0.918) (-0.86, 0.45)			
0.99 (0) (0.98, 1)		0.33 (0.16) (-0.11, 0.79)			-0.55 (0.075) (-1.17, 0.07)		-0.09 (0.872) (-1.66, 1.31)		
0.99 (0) (0.98, 0.99)		0.4 (0.119) (-0.08, 0.91)			-0.6 (0.071) (-1.24, 0.06)			0.32 (0.754) (-0.88, 1.62)	
0.99 (0) (0.98, 0.99)		0.38 (0.146) (-0.09, 0.88)			-0.58 (0.072) (-1.2, 0.05)				0.18 (0.641) (-0.72, 1.06)

Appendix N

Plot Standard Deviation Model Selection and Posteriors (Survival Model)

This appendix contains a description of the model selection strategy for random plot effects in the survival models, and shows posteriors for plot standard deviation σ_p for all species aggregated, incense-cedar, white fir, Douglas-fir, ponderosa pine, sugar pine, black oak, and tanoak.

N.1 Selection Strategy for Random Effects

The plot random effect standard deviation cannot overlap zero and one therefore cannot use a 95% credible interval overlap criterion. Instead, one can consider the shape and location of the posterior, interpreting it as not meaningful when zero has substantial support. In addition, random effects cannot be treated equivalently to fixed effects in the selection strategy because they are variance components. The forward-selection strategy is not optimal for selecting random effects because one prefers to test the random effects with a "beyond optimal" model which includes all possible fixed effects (Zuur et al. 2009), which is not feasible here. Therefore I use the following procedure to select random plot effects:

1) Fit models with only plot effects. This model should demonstrate which random effects should definitely not be selected, because if there are no other variables and the plot standard deviation posterior has high support at zero, there is little evidence in the data that any variation is accounted for with this effect. Therefore, if plot effect standard deviations have support at zero, I do not include plot effects in future models for fixed effect selection.

2) Otherwise, include plot effects which could potentially be significant and proceed with fixed effect model selection as described in Chapter 2.

3) After selecting fixed effects, check the plot random effect standard deviation again in the final models to confirm that it has remained distinct and separated from zero.

N.2 Results of Selection for Survival Models

Figure N.1 shows posteriors from the model with random plot effects and no explanatory variables (from step 1). Tanoak is the only species in the models which clearly has little support for a plot random effect, so I do not include it in further model selection. All other species have plot random effect posteriors which are separated from zero and therefore plot effects are included for those species throughout the model selection process.

For these species, after selecting fixed effects, I check the plot random effect standard deviation again to see if the variation initially attributed to plot has now been attributed to fixed effects (from step 3, Figure N.2). For all species aggregated, incense-cedar, Douglas-fir, and ponderosa pine, the plot standard deviation posteriors are still clearly separated from zero. Sugar pine's plot random effect has become more ambiguous. Black oak and white fir now have bimodal plot standard deviation posteriors (the MCMC consistently alternates between two possible values). For black oak, 'pairs' plots in R of the MCMC chains show that the plot random effect is compensating for the size effect — when the plot random effect node is close to zero, then the size parameters β^{DBH} and β^{DBH^2} also have smaller (close to zero) values. However, the more favored mode (with the higher density) is the one with a nonzero plot effect and strong size dependence. For white fir, no tradeoffs are obvious in the 'pairs' plot and it is possible that the plot random effect may not be important for this species.

Note that variation in incense-cedar is higher than all species aggregated, Douglas-fir, and ponderosa pine. One possible reason for incense-cedar's higher variation could be that it is the most abundant species at BFRS. It is found on almost the same number of plots as white fir, which is next most abundant, and yet has greater plot variation. It may be that much of the recruitment and gap dynamics are occurring within this most abundant species. Note also that spatial influences on mortality can be more elegantly handled in mapped plots, where the hierarchical model can directly incorporate spatial autocorrelation (Finley et al. 2011).

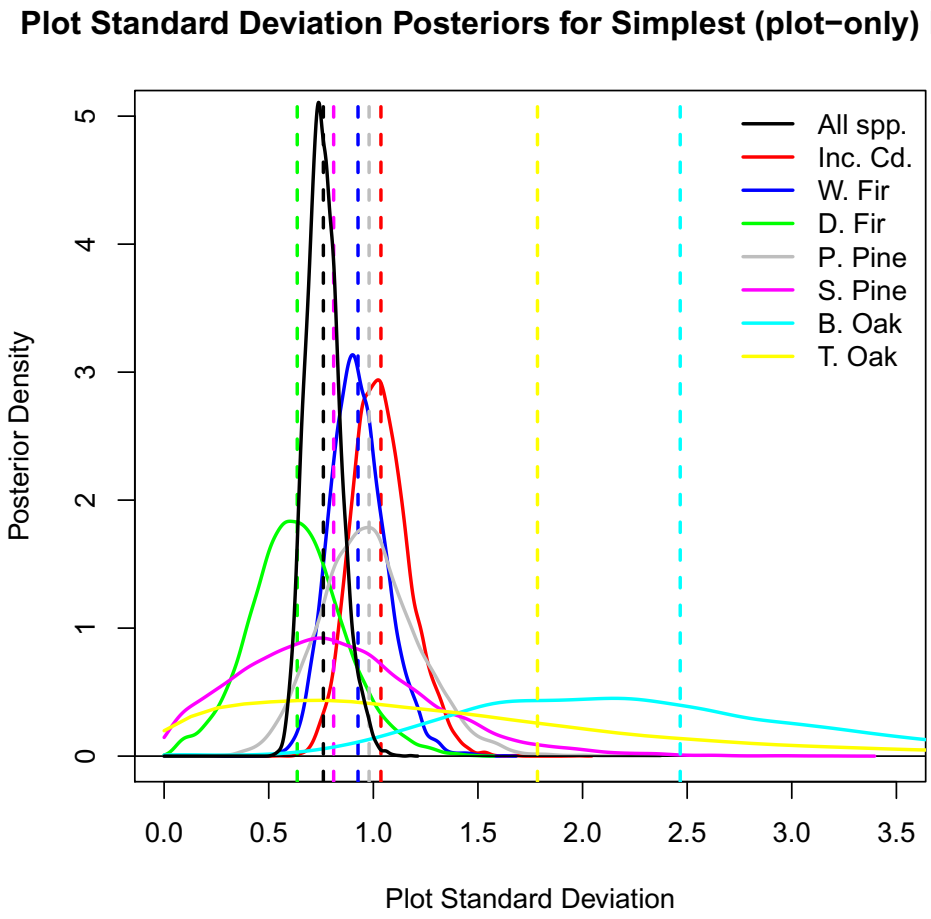

Figure N.1: Plot standard deviation posteriors for plot-only models. Means are shown as a vertical dashed line.

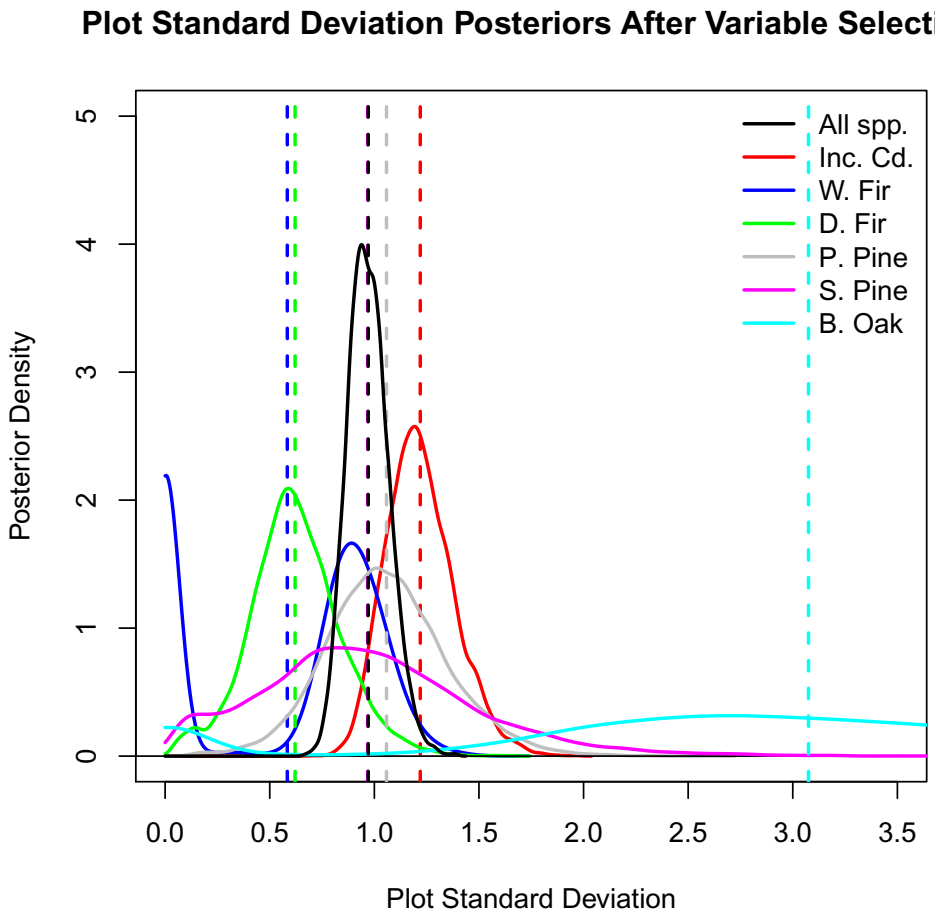

Figure N.2: Plot standard deviation posteriors for models with model-selected fixed effects. Means are shown as a vertical dashed line.

Appendix O

Additional Figures for Significant Variables on Probability Scale for Survival Models

Figures O.1 through O.7 show additional probability/size relationships for species-variable combinations which were not included in Figure 2.3 but were selected by the model selection procedure. Two borderline cases are also shown here: Douglas-fir with plot basal area, and incense-cedar and the secular time trend. Note that black oak only selected linear and quadratic size and therefore no other variable is shown in Figure O.3.

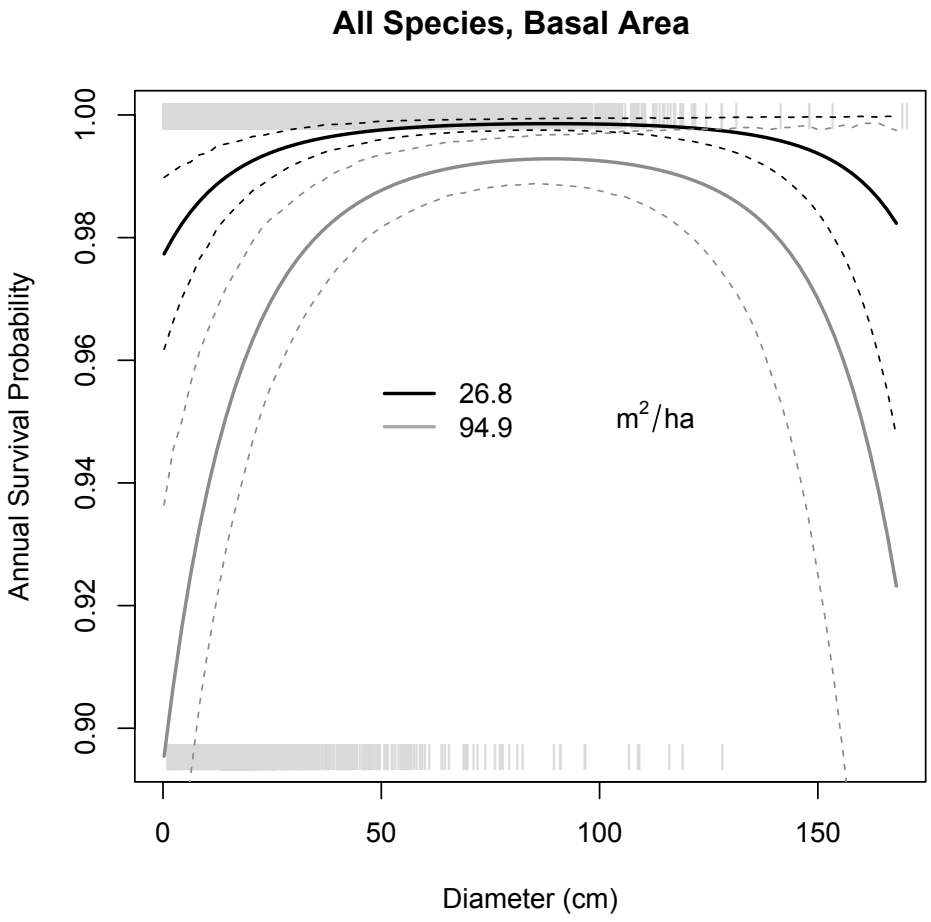

Figure O.1: All species aggregated, plot basal area probability plots.

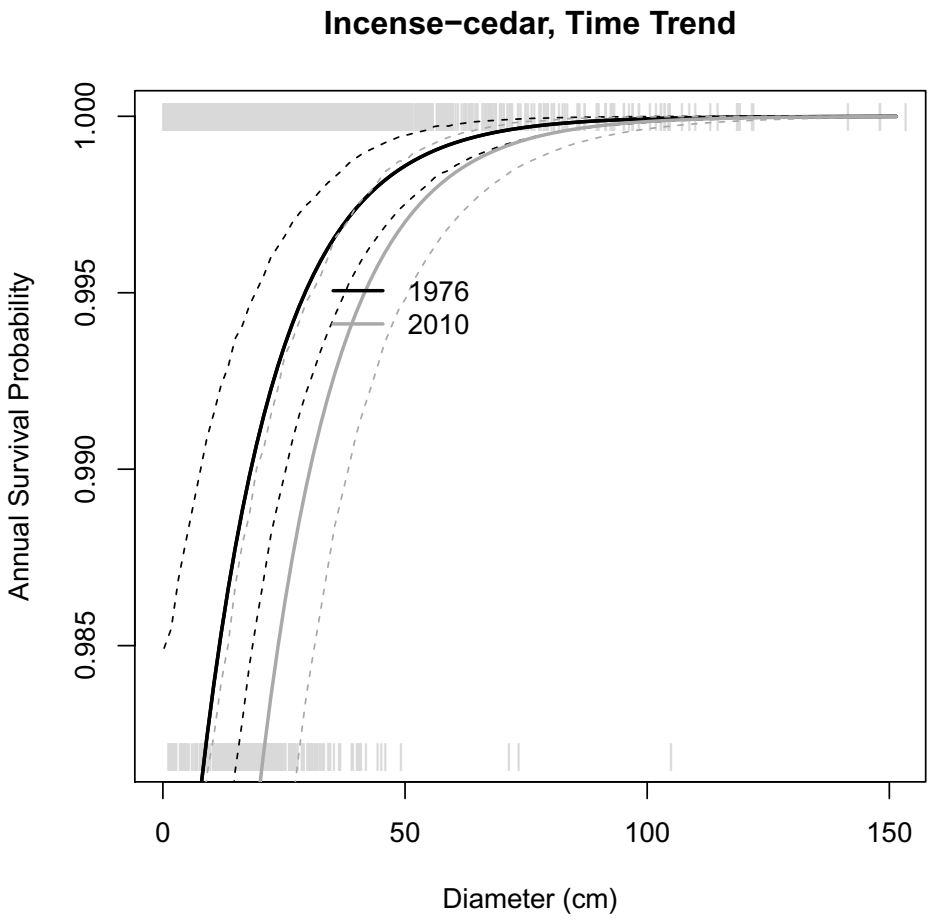

Figure O.2: Incense-cedar, secular time trend probability plots.

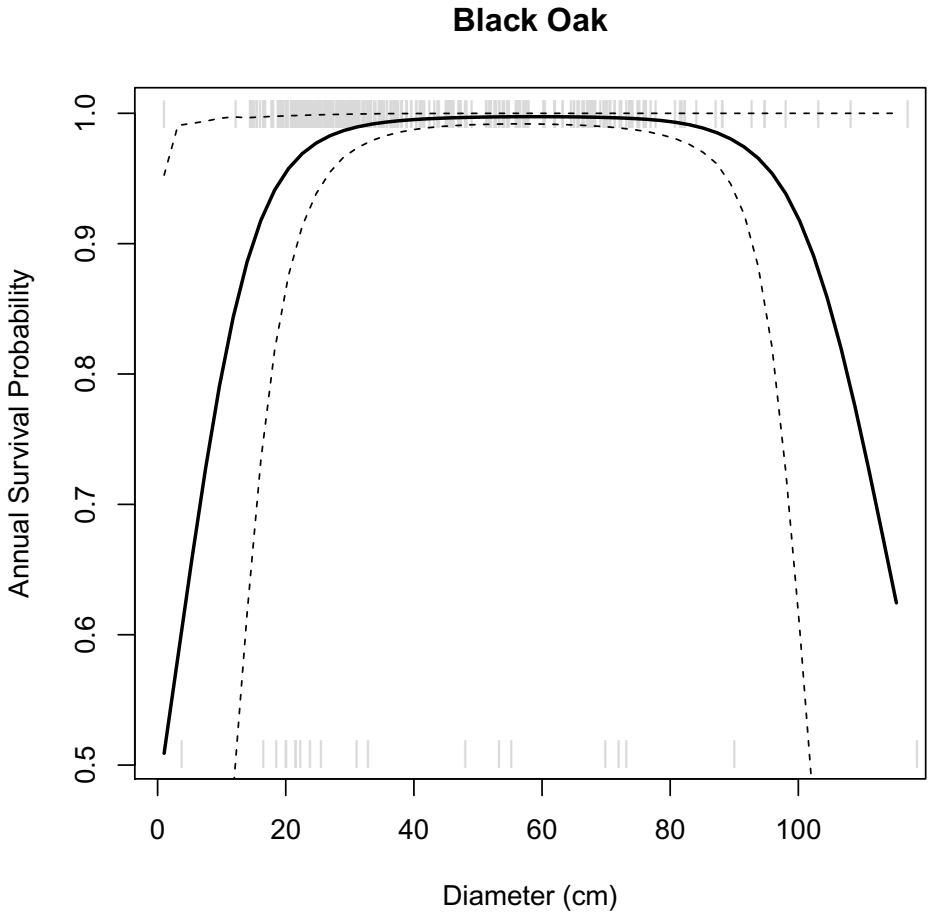

Figure O.3: Black oak, diameter probability plots.

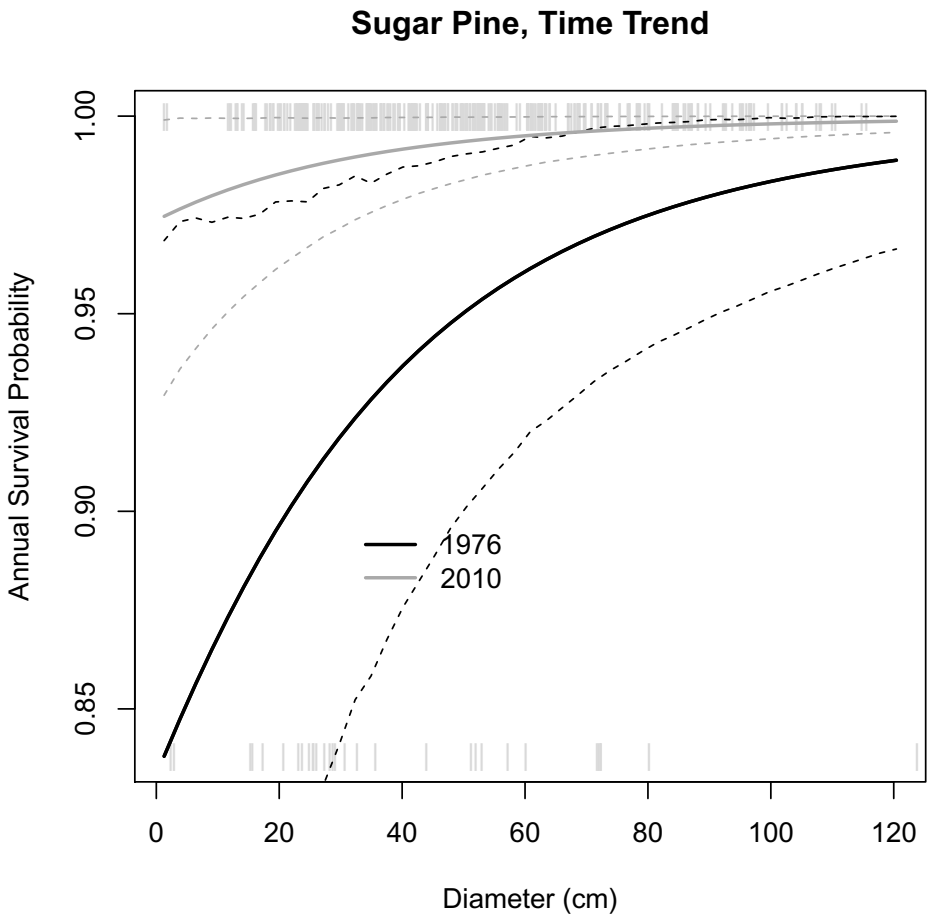

Figure O.4: Sugar pine, secular time trend probability plots.

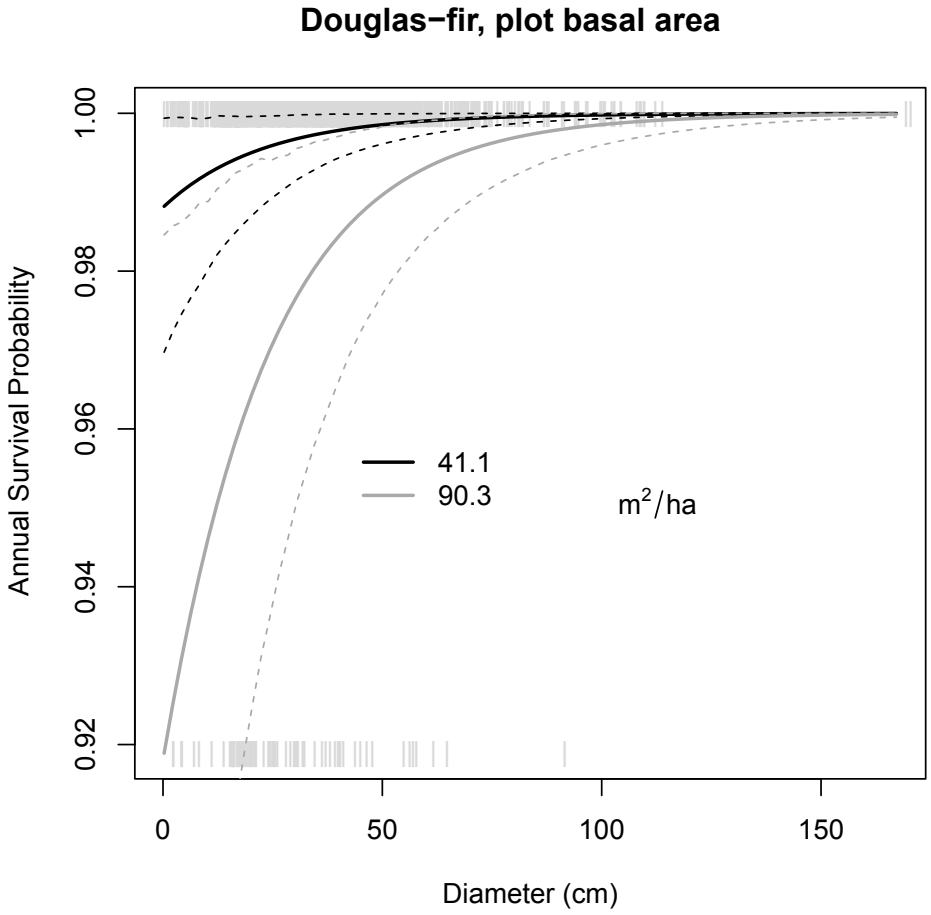

Figure O.5: Douglas-fir, plot basal area probability plots.

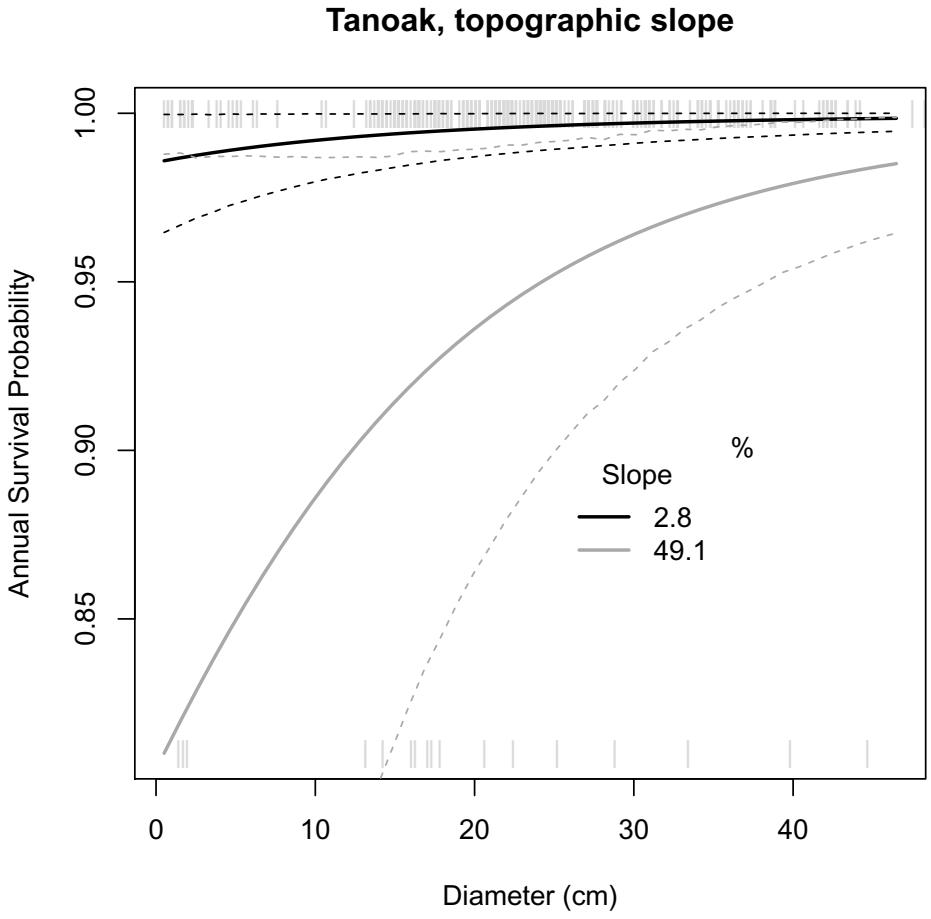

Figure O.6: Tanoak, topographic slope probability plots.

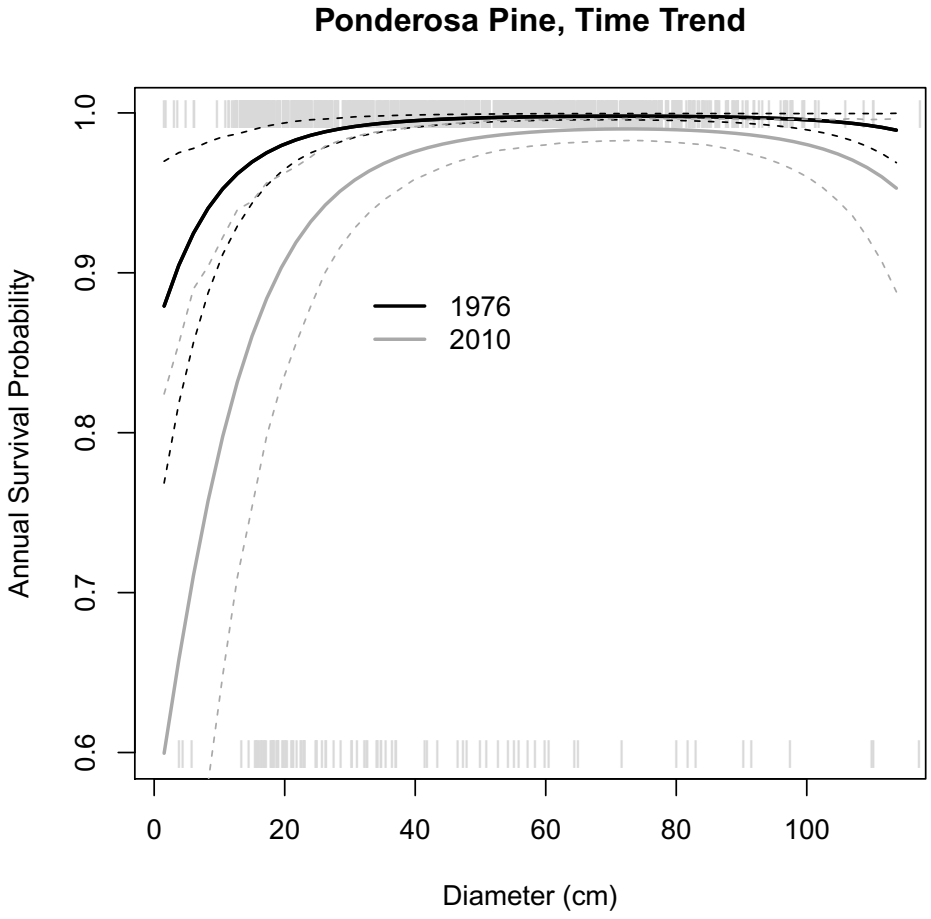

Figure O.7: Ponderosa Pine, secular time trend probability plots.

Appendix P

Exploring Different Smoothing Functions

Figure P.1 shows, for Iaqua Buttes, the comparison between two different generalized additive model smoothing function bases, thin plate regression splines and Markovian random fields, as well as a slightly different method of calculating the polygon centroid to incorporate geographical location into the statistical model. Solid lines are associated with parameter estimates and dashed lines show 95% confidence intervals; parameter estimates which are significant at $p < 0.05$ have solid circles and those that are not have open circles. The Y-axis is a standardized regression parameter estimate β, and the X-axis is the cell size in meters. There are no appreciable differences between these methods. Similarly, the two methods of locating a polygon centroid (to extract the geographical coordinates for the statistical model), "inside" vs. "centroid" do not produce appreciably different results.

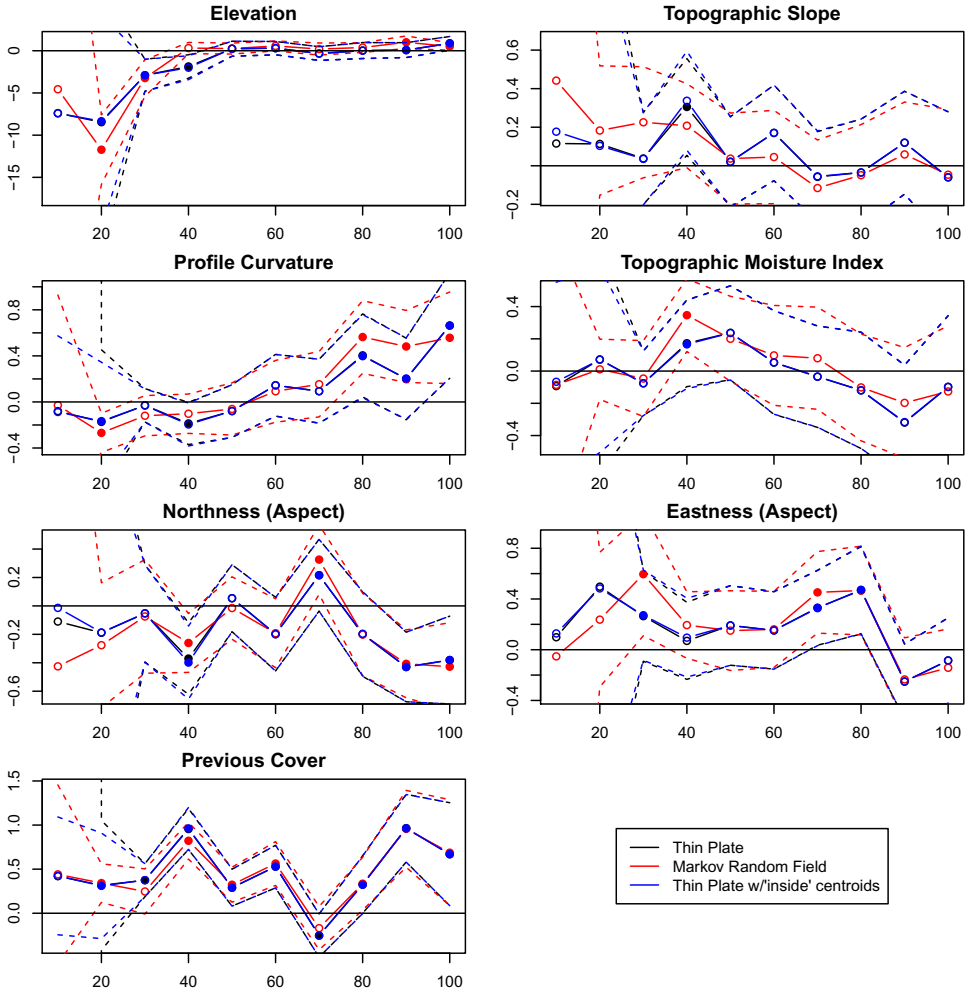

Figure P.1: Thin plate regression spline and Markovian random field bases give similar results for Iaqua Buttes. Also shown: different choice of centroid location ("inside" vs. "centroid") for geographical coordinates, which also gives similar results.

Appendix Q

Imagery and Classifications for Other Sites

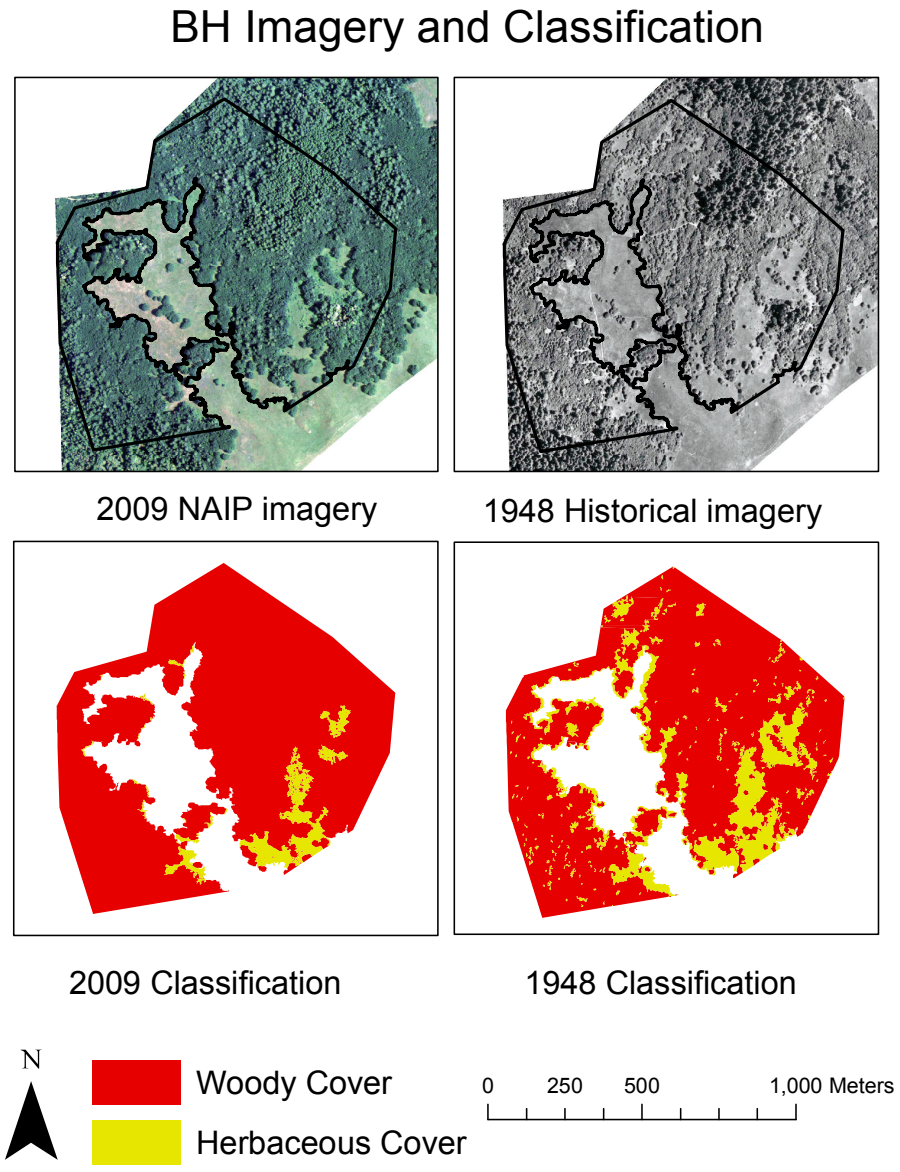

Figure Q.1: Imagery and classification for the Bald Hills (BH) site.

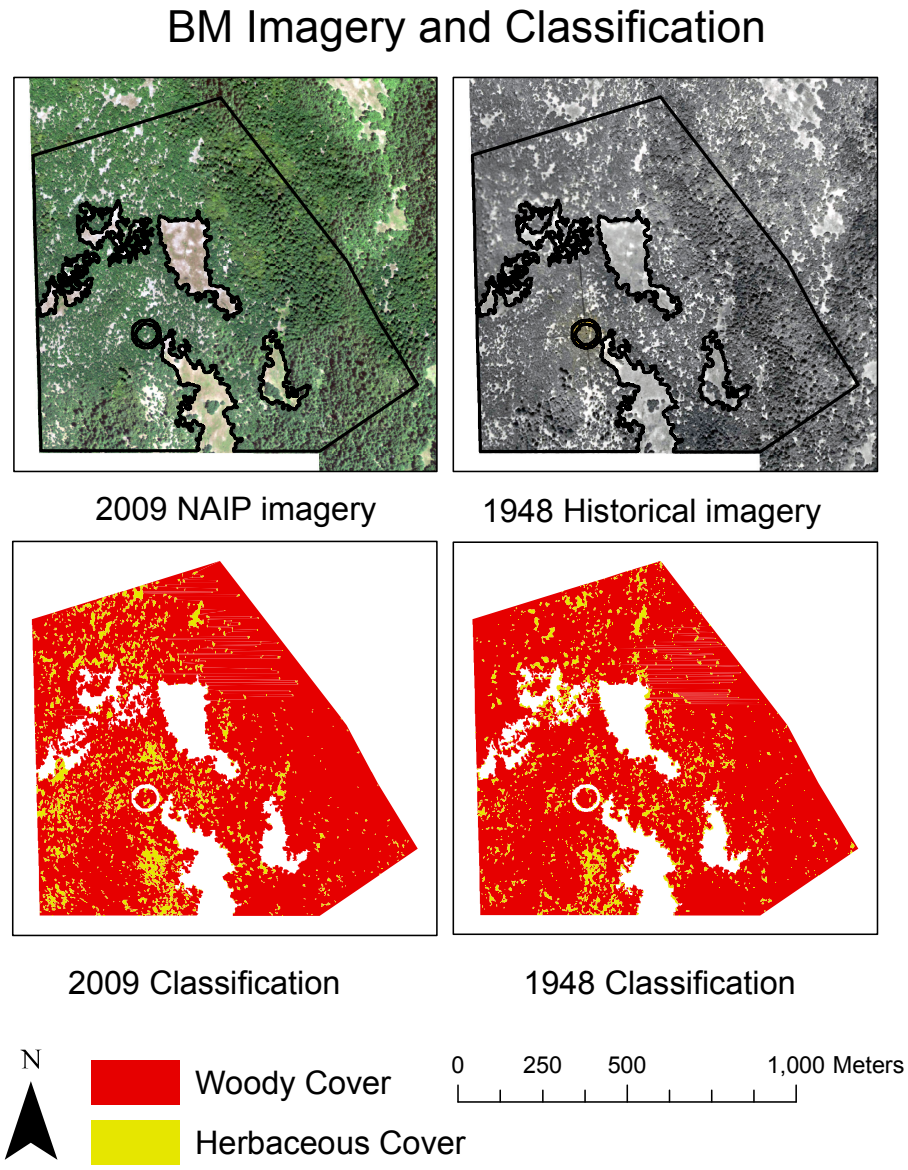

Figure Q.2: Imagery and classification for the Blake Mountain (BM) site.

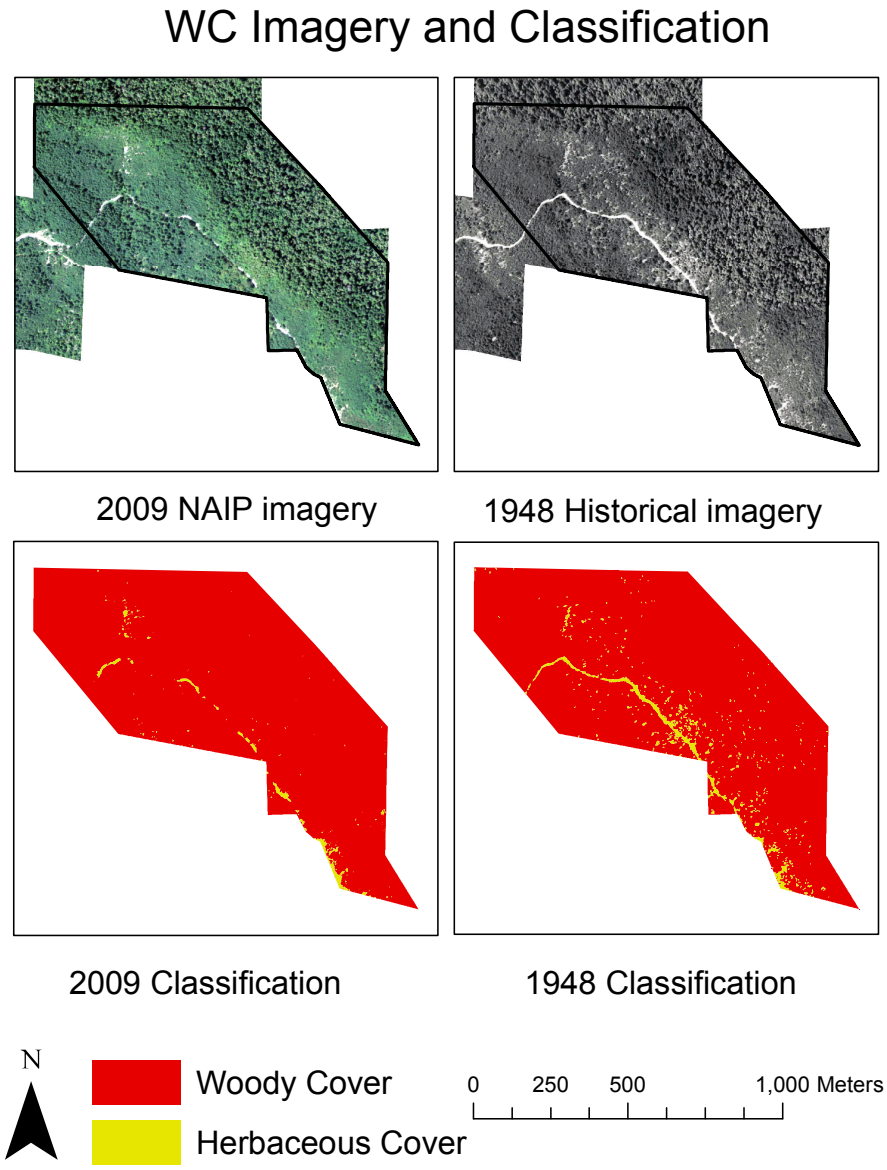

Figure Q.3: Imagery and classification for the Willow Creek (WC) site.

Appendix R

Scaling Relationships with Non-Standardized Variables

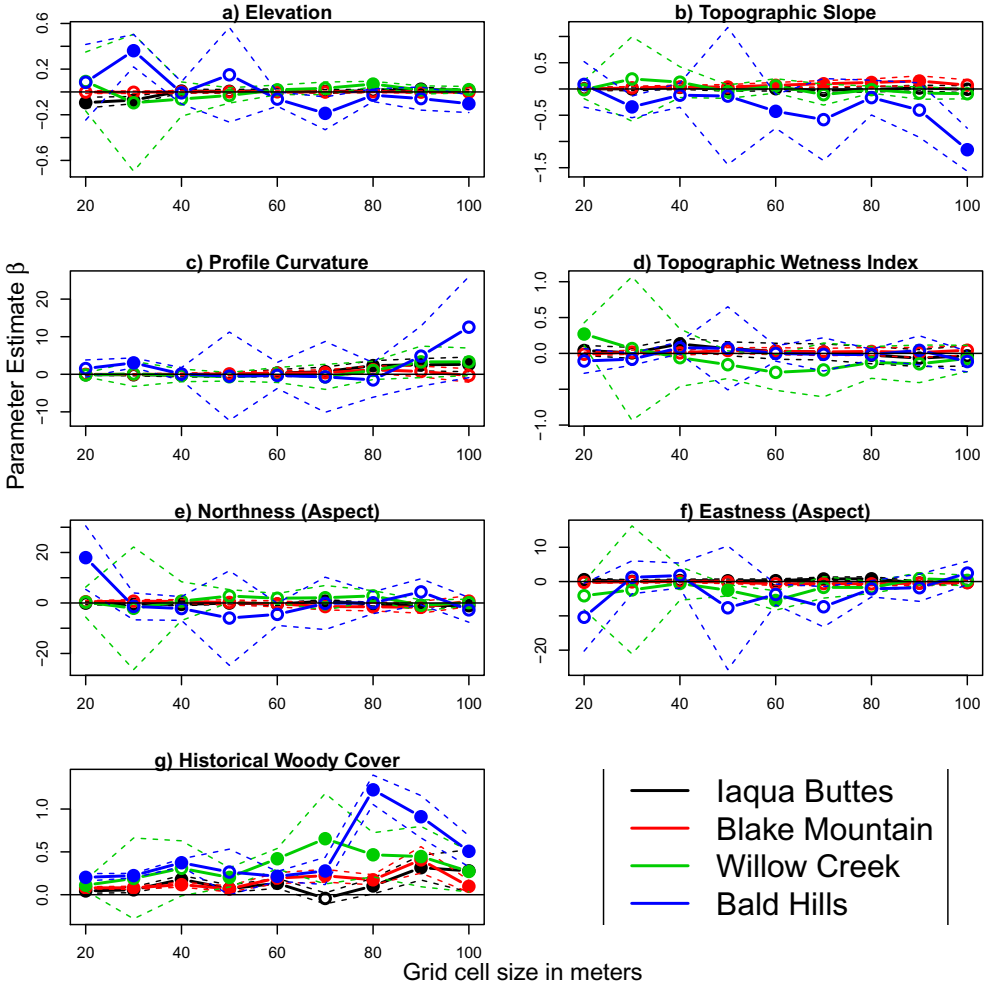

Figure R.1: Results of scaling from 20 m grid cells to 100 m grid cells. Each color represents one site, with solid lines associated with parameter estimates and dashed lines showing 95% confidence intervals; parameter estimates which are significant at p <0.05 have solid circles and those that are not have open circles. Parameters are for non-standardized variables so the magnitude of the effects can be directly compared between sites and scales; the results are generally unchanged from the standardized case, therefore comparisons in Figure 3.5 are legitimate.

Appendix S

Error Matrices for Image Classifications

Tables S.1 and S.2 show error matrices for each classified image and site (IB – Iaqua Buttes, BH – Bald Hills, BM – Blake Mountain, WC – Willow Creek). The tables include the number of samples correctly classified, as well as omission and commission errors (as Producer's and User's accuracies). Accuracy for Iaqua Buttes in 1948 was 92% (with a kappa statistic of 0.78); for the 2009 image, accuracy was 98% (kappa = 0.89). For Bald Hills, the 1948 image had a classification accuracy of 96% (kappa = 0.86); for the 2009 image, classification accuracy was 98% (kappa = 0.89). For Blake Mountain, the 1948 image had a classification accuracy of 94% (kappa = 0.74); for the 2009 image, classification accuracy was 99% (kappa = 0.95). For Willow Creek, the 1948 image had a classification accuracy of 99% (kappa = 0.94); for the 2009 image, classification accuracy was 99% (kappa = 0.94). In six cases out of 800, a sample fell into a clearly mixed class between woody and herbaceous cover and was thrown out.

	Assessed Classes				
Assigned Classes	IB		Woody	Herbaceous	User Acc.
		Woody	71	4	94.7%
		Herbaceous	4	20	83.3%
		Prod. Acc	94.7%	83.3%	91.9%
	BH		Woody	Herbaceous	User Acc.
		Woody	79	3	96.3%
		Herbaceous	1	16	94.1%
		Prod. Acc	98.8%	84.2%	96.0%
	BM		Woody	Herbaceous	User Acc.
		Woody	83	6	93.3%
		Herbaceous	0	10	100.0%
		Prod. Acc	100.0%	62.5%	93.9%
	WC		Woody	Herbaceous	User Acc.
		Woody	90	0	100.0%
		Herbaceous	1	9	90.0%
		Prod. Acc	98.9%	100.0%	99.0%

Table S.1: Error matrices for each classified 1948 image

	Assessed Classes				
Assigned Classes	IB		Woody	Herbaceous	User Acc.
		Woody	88	1	98.9%
		Herbaceous	1	9	90.0%
		Prod. Acc.	98.9%	90.0%	98.0%
	BH		Woody	Herbaceous	User Acc.
		Woody	88	1	98.9%
		Herbaceous	1	9	90.0%
		Prod. Acc.	98.9%	90.0%	98.0%
	BM		Woody	Herbaceous	User Acc.
		Woody	88	1	98.9%
		Herbaceous	0	10	100.0%
		Prod. Acc.	100.0%	90.9%	99.0%
	WC		Woody	Herbaceous	User Acc.
		Woody	90	0	100.0%
		Herbaceous	1	9	90.0%
		Prod. Acc.	98.9%	100.0%	99.0%

Table S.2: Error matrices for each classified 2009 image

Appendix T

Model Diagnostics for All Sites and Scales

Table T.1 gives the model diagnostics for each site and scale. The values of the basis dimension k used for the smoothing spline are all greater than the default of 10. The Moran test for autocorrelation (in which significant p-values indicate the presence of autocorrelation) shows no autocorrelation for any site or scale. All models were run with large enough k to allow the smoother to be flexible enough to account for spatial autocorrelation. The estimated dispersion parameters for the quasibinomial are all much greater than one (binomial distribution with no dispersion would have a dispersion parameter of one), indicating that the quasibinomial was necessary. The spatial smoothing term is usually significant, but not always. Finally, the table gives the run time for each model.

Site	Grid cell size	N	Basis dimension k	Moran Test p.value	Estimated dispersion	Spatial term p.value	Run time (seconds)
BH	100	130	13	0.77	4.89	2.22E-008	0.10
	90	154	15	1.00	6.44	6.77E-010	0.13
	80	187	19	0.99	2.22	2.23E-015	0.15
	70	246	25	1.00	22.31	9.01E-002	0.25
	60	323	32	0.98	6.16	2.81E-013	0.74
	50	452	45	0.70	152.78	1.00E+000	1.82
	40	664	66	0.99	4.54	1.26E-065	3.94
	30	1154	115	1.00	4.13	1.98E-096	73.36
	20	2453	245	1.00	21.10	3.76E-010	695.18
BM	100	137	14	0.85	3.37	3.10E-019	0.10
	90	162	32	1.00	3.47	1.12E-012	0.24
	80	201	40	0.99	3.19	9.48E-030	0.13
	70	250	50	0.90	3.75	3.87E-025	0.20
	60	332	66	1.00	4.34	2.15E-038	0.41
	50	473	94	1.00	4.61	8.60E-057	0.99
	40	722	144	1.00	4.22	3.71E-120	3.32
	30	1215	244	1.00	4.97	3.43E-226	16.15
	20	2678	268	0.97	9.29	9.42E-266	148.29
IB	100	99	10	0.76	7.11	2.06E-001	0.08
	90	117	12	0.84	4.76	3.84E-005	0.06
	80	145	14	0.93	5.88	7.12E-009	0.08
	70	184	18	0.95	5.26	2.57E-009	0.11
	60	242	24	0.98	6.99	4.89E-013	0.14
	50	333	33	0.99	7.87	2.98E-014	0.56
	40	490	49	1.00	7.71	6.63E-034	1.48
	30	842	84	0.97	9.05	2.35E-042	5.00
	20	1800	180	1.00	11.74	2.31E-071	45.50
WC	100	106	11	0.98	1.38	1.19E-007	0.09
	90	127	13	0.98	2.20	3.44E-005	0.10
	80	151	15	0.99	1.40	7.85E-022	0.07
	70	198	20	0.97	7.01	1.48E-001	0.14
	60	263	26	1.00	2.59	3.13E-021	0.36
	50	361	36	0.94	1.87	1.38E-028	4.57
	40	553	55	1.00	14.67	1.42E-002	2.36
	30	967	97	1.00	159.28	1.00E+000	46.26
	20	2064	206	1.00	9.78	6.93E-010	111.86

Table T.1: Model diagnostics for each site and scale (IB – Iaqua Buttes, BH – Bald Hills, BM – Blake Mountain, WC – Willow Creek).