Landscape- and regional-level correlative modeling techniques for the prediction of ecological processes

Nicholas A. Povak

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Jerry F. Franklin, Chair

Paul F. Hessburg

Don McKenzie

Program Authorized to Offer Degree:

School of Environmental and Forest Sciences

University of Washington

Abstract

Landscape- and regional-level correlative modeling techniques for the prediction of ecological processes

Nicholas A. Povak

Chair of Supervisory Committee:
Jerry F. Franklin, Professor
School of Environmental and Forest Sciences

Quantifying landscape patterns and relating them to key biophysical drivers is often of primary interest in ecological research. Knowledge of such relationships can increase our understanding of pattern-process linkages and aid in making predictions across other spatial or temporal extents. With the recent influx of freely available, high resolution and spatially explicit data sources landscape- and regional-level models have gained in popularity. However, certain challenges arise when modeling ecological process at these scales, which are often overlooked. Processes in ecological systems occur at several spatial scales and interactions among processes are many, complex, and often non-linear and can result in highly heterogeneous conditions across spatial scales. The current research represents three separate applications of landscape- and regional- level correlative modeling. The objectives of each application were to develop correlative models to identify key environmental drivers of ecological process of interest, and to predict the process across a large spatial extent. For each application, I used a

multi-model approach where a variety of statistical regression and machine learning methods were tested to objectively identify the model or models that best captured the complexity of an ecological process. The first two studies were located in the southern Appalachian mountain region where stream industrial emissions have acidified stream waters for more than a century. The objectives of these studies were to identify main drivers of stream water acid neutralizing capacity (ANC; a metric related to the ability of a stream to buffer against acidic inputs), and base cation weathering (BC_w; the level of base cations supplied to the stream by catchment soils) and to predict ANC and BC_w across the study region. The third study was located in the Methow Valley region of eastern Washington, which was aimed at identifying the landscape vegetation components most limiting to cavity-nesting bird populations in the area. Each study had different modeling objectives and different associated nuisances related to the data design, the spatial extent of the study region, and the correlative structure of the modeled process in relation to the environmental drivers.

ACKNOWELDGEMENTS

I would like to thank Paul Hessburg for his unwavering support over the past five years. Paul has been integral in my development as an ecologist, and he has gone above and beyond as a boss, a colleague, and a friend to encourage me over the course of this degree. Thank you to Jerry Franklin for graciously accepting me as a student and bestowing his wisdom through lectures, field trips, lab meetings, and closed-door therapy sessions.

To my lab mates, James Freund, Derek Churchill, Keala Hagmann, Van Kane, Linda Winter, and Lauren Urgenson for their conversations about ecology (or otherwise) in the lab (or otherwise); I thank them for all of their constructive feedback, encouragement, and camaraderie. Brion Salter has been invaluable over the past five years in focusing my ideas, believing in me, and supplying me with innumerable GIS maps. Needless to say, I would not have made it without him.

To my friends and family here in Seattle, I thank them for their support and much needed distractions, and for making this city feel like home. My parents, who have instilled in me a sense of pride for all that I accomplish, without them this would not have been possible.

And to Erin, with whom I have enjoyed countless adventures, being with you has made this one more incredible trip.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1 - PREDICTING RESPONSES FROM MULTI-SCALE LANDSCAPE PATTERNS AND PROCESSES: EXPLORING OPTIONS FOR CORRELATIVE MODELING

"In the face of complexity, ecologists often strive to identify models that capture the essence of a system, explaining the observed distribution and perhaps ultimately permitting prediction." *(Whittingham et al. 2006)*

Ecological patterns observed on the landscape are an emergent property of the many interactions between biotic and abiotic pattern and process interactions over space and time (Wu and David 2002). Quantifying landscape and regional patterns and relating them to key biophysical drivers is often of primary interest in ecological research (Guisan et al. 2002, Araújo and Guisan 2006, Elith et al. 2006, Maggini et al. 2006). Knowledge of these relationships can increase our understanding of pattern-process linkages and can be used to make future projections across other spatial or temporal extents (Franklin and Miller 2009). To this end, correlative statistical models are used to relate environmental predictors (e.g. climate, soils, geology, or land surface forms) to a response (e.g., species presence/absence). With the recent influx of large-scale high-resolution data, ecologists are increasingly incorporating correlative landscape-level analyses in their research (Miller et al. 2004), and are faced with many challenges associated with analyzing large-scale ecological datasets.

A variety of correlative modeling frameworks are used in landscape-level analysis to quantify relations between environmental variables and an ecological process (Guisan et al. 2002, Elith and Leathwick 2009). Parametric models, such as linear regression and generalized linear models, estimate model parameters based on a predetermined functional model form (i.e., Poisson, linear, logistic). Interactions among predictors and polynomial terms to model non-linear and multiplicative responses can be included, but need to be explicitly stated (Guisan et al. 2002, Elith and Leathwick 2009). Variable selection is often conducted using automated stepwise procedures, despite known limitations (Whittingham et al. 2006). Regardless, parametric regression analyses have been successfully implemented in many applications, particularly in species distribution modeling (Guisan et al. 2002, Maggini et al. 2006).

Alternatively, when information regarding the functional model form is unknown, exploratory data analysis using non-parametric or semi-parametric models has been used. Such models include classification and regression trees (and derivatives thereof), neural networks, support vector machines, and Maximum Entropy models (Franklin and Miller 2009). These models are referred to as machine learning algorithms (Hastie et al. 2005). Correlative regression or classification models represent supervised learning problems, where predictor variables are mapped to an observed response, and rules are developed to predict new observations. These models are useful for (1) selecting key predictors from among a potentially large pool of variables, (2) modeling linear or non-linear responses, and (3) identifying variable interactions without explicitly incorporating them into the model. However, because these models are non-parametric, model interpretations are restricted to graphical representations, and methods are often considered "black box." Also, because these models are data driven, issues with overfitting are well known. Despite these limitations machine learning is gaining popularity in ecological

literature, and where correctly applied, has proven to be highly predictive and capable of identifying important correlates of ecological processes (Prasad et al. 2006, Cutler et al. 2007).

Regardless of the approach, other practical issues related to correlative landscape-level modeling include the scale and quality of the environmental data layers (Elith and Leathwick 2009), biased or imbalanced sampling designs (Loiselle et al. 2008, He and Garcia 2009), small sample sizes (Raudys and Jain 1991), and poor representation of the sampled environmental space. Each of these factors contributes to model error, and may bias results. For example, many ecological data sets are collected without consideration of a probabilistic sampling design (e.g., opportunistic data), which needs to be addressed during model development (Edwards et al. 2006, Loiselle et al. 2008). Appropriate model evaluation is required to quantify the validity of model predictions and identify the influence of data deficiencies on model predictive ability across both geographic and data space.

In total, an appropriate modeling framework for correlative landscape-level modeling should include theory that outlines research goals, (2) information on data requirements where data deficiencies and limitations are identified, (3) a statistical model appropriate for the theory and modeling objectives, and (4) a model evaluation scheme that uses model selection and performance measures to adequately assess model stability, goodness-of-fit, and the robustness of predictions (Austin 2002).

The chapters that follow present three novel applications of correlative landscape- and regional-level modeling. Each application specifically addresses nuances of the ecological process being modeled, and the data used in analysis. The studies represent a variety of spatial scales, and ecological processes, and each uses an assortment of statistical and machine learning

models, data resampling, and model evaluation techniques appropriate to the model objectives.

Chapter 2 – Landscape modeling to predict the biogeochemical and climatic environment of acidic streams in the Appalachian Mountains, USA

The first chapter was aimed at identifying environmental cofactors of streamwater acidity across the southern Appalachian Mountain region. This region has a long history of industrially generated sulfur (S) deposition that has acidified streams and led to losses in macroinvertebrate and fish biodiversity. Efforts aimed at reducing future acidification in streams across the region hinge on identifying those most susceptible to atmospheric S deposition and those that are least capable of naturally buffering acidity. In a previous modeling effort, Sullivan et al. (2007) used logistic regression, and a restricted set of predictors related to catchment lithology and terrain to predict catchment acid neutralizing capacity (ANC). Model error was assessed using predictions to the training data.

I built upon the models of Sullivan et al. (2007), but incorporated a larger study area and additional water chemistry sample points. I used a variety of modeling approaches to address known biases in the sample design, identify additional environmental correlates of ANC, and predicted continuous values of ANC across the study area. I used a combination of data resampling techniques and hurdle modeling to adjust for the biased sample. Linear regression, random forests, and boosted regression trees were used in combination to identify the model with the lowest error-rate. These models were used within a hurdle modeling framework, to address an imbalanced data design. Predictor variables included not only soils, lithology, and terrain variables, but also aspects of the climate regime, levels of wet and dry atmospheric sulfur deposition, and environmental setting. Model errors were subsequently mapped across the region

to identify areas for additional future sampling.

Chapter 3 - Landscape modeling to predict catchment-level base cation weathering rates across the southern Appalachian Mountains region, USA

The acid-base chemistry of a stream is largely determined by its natural resupply rate of base cations. Most base cation inputs to a stream result from the physical and chemical weathering of upslope catchment soils. Estimates of these rates are a main component of Critical Loads (CL) calculations, which define resource susceptibility to continued acidic depositions. CLs are used to guide policy decisions regarding acceptable pollution levels. The base cation weathering (BC_w) parameter is the single most important input parameter to CL calculations, yet is fraught with the highest estimation error (McNulty and Boggs 2010). BC_w is estimated using either process or empirical models that relate soils and geology to BC_w. Process-based models, such as PROFILE (Sverdrup and Warfvinge 1993) or MAGIC (Cosby et al. 2001), provide accurate BC_w estimates, but due to large data and processing requirements, estimates are restricted to individual catchments. Empirical models, such as the Steady State Water Chemistry Model (SSWC; Henriksen et al. 2002), use empirical models to estimate BC_w across large spatial scales, but these models were largely developed for European soils, and may not be applicable in North America (Li and McNulty 2007).

I used a hybrid modeling approach, where correlative models were used to extrapolate 140 catchment-level MAGIC modeled estimates of BC_w across the southern Appalachian study region. A similar set of variables were used as described in the previous ANC assessment. I compared the utility of five statistical and machine-learning methods for predicting a continuous BC_w surface, and all models were evaluated using 10-fold cross-validation. Models predicted

BC_w for the entire study area and also for the three main physiographic regions within the study area to determine if relationships between environmental drivers and BC_w were consistent across the study region.

Chapter 4 - Nest site selection preferences and niche diversification for cavity-nesting birds in the eastern Cascade Region of Washington State, USA

Cavity nesting birds (CNB) play a keystone role in forested ecosystems. Excavations made by primary CNBs are important habitat for other non-excavating secondary CNBs as well as a variety of other vertebrate species (Short 1979, Thomas et al. 1979, Bednarz et al. 2004, Virkkala 2006), yet populations of some species are declining (Newton 1994, Eadie et al. 1998). Specific management needed to ensure stable populations over time is largely unknown. Recent research has focused on snag characteristics and demographics (Zarnowitz and Manuwal 1985, Brawn and Balda 1988, Newton 1994). However, nest-site selection is inherently a multi-scaled process (Kotliar and Wiens 1990, Cushman and McGarigal 2004). Thus, multi-level models may be required to elucidate co-factors associated with site selection (Lawler and Edwards 2006). Multi-scale model development requires identifying ecologically relevant scales of potential covariates, and quantifying variance explained at several spatial data levels and shared across levels (i.e., cross-correlations). Furthermore, social interactions among birds are rarely included in models, yet these interactions can influence nest-use patterns as individuals may cluster or disperse nests in response to other birds (Ahlering and Faaborg 2006).

To this end, I studied CNB nest-site selection preferences for four species within a dry forest of eastern Washington. CNB nests were georeferenced, and available habitat was

quantified using a complete census of snags across seven forested drainages covering a total of 2,100 ha. Multi-scaled vegetation data were quantified using a raster-based vegetation classification and photo-interpreted patches. I used multi-level habitat models and niche overlap modeling to identify nest selection preferences of four CNBs, and niche overlap among birds, respectively. Habitat models were built with and without inter-nest distance predictors to quantify the effect of social interactions among birds. To reduce imbalances in the ratio of nests to non-nests (>100:1), I used two resampling techniques: oversampling nest sites (SMOTE; He and Garcia 2009), and under-sampling of the non-nest snags at different sampling intensities. Variance decomposition was used to quantify the amount of pure variance explained at each individual spatial level, and the amount of shared variance among levels. Resulting models predicted CNB habitat suitability across the study area.

REFERENCES

Ahlering, M. A. and J. Faaborg. 2006. Avian habitat management meets conspecific attraction: If you build it, will they come? The Auk **123**:301-312.

Araújo, M. B. and A. Guisan. 2006. Five (or so) challenges for species distribution modelling. Journal of Biogeography **33**:1677-1688.

Austin, M. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling **157**:101-118.

Bednarz, J. C., D. Ripper, and P. M. Radley. 2004. Emerging concepts and research directions in the study of cavity-nesting birds: Keystone ecological processes. The Condor **106**:1-4.

Brawn, J. D. and R. P. Balda. 1988. Population biology of cavity nesters in northern Arizona: do nest sites limit breeding densities? Condor **90**:61-71.

Cosby, B., R. Ferrier, A. Jenkins, and R. Wright. 2001. Modelling the effects of acid deposition: refinements, adjustments and inclusion of nitrogen dynamics in the MAGIC model. Hydrology and Earth System Sciences **5**:499-518.

Cushman, S. A. and K. McGarigal. 2004. Hierarchical analysis of forest bird species-environment relationships in the Oregon Coast Range. Ecological Applications **14**:1090-1105.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology **88**:2783-2792.

Eadie, J., P. Sherman, and B. Semel, editors. 1998. Conspecific brood parasitism, population dynamics, and the conservation of cavity-nesting birds. Oxford Univeristy Press, New York, New York.

Edwards, T. C., D. R. Cutler, N. E. Zimmermann, L. Geiser, and G. G. Moisen. 2006. Effects of sample survey design on the accuracy of classification tree models in species distribution models. Ecological Modelling **199**:132-141.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, and A. Lehmann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography **29**:128-151.

Elith, J. and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics **40**:677-697.

Franklin, J. and J. A. Miller. 2009. Mapping species distributions: spatial inference and prediction. Cambridge University Press, New York.

Guisan, A., T. C. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling **157**:89-100.

Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin. 2005. The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer **27**:83-85.

He, H. and E. A. Garcia. 2009. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering **21**:1263-1284.

Henriksen, A., P. Dillon, and J. Aherne. 2002. Critical loads of acidity for surface waters in south-central Ontario, Canada: regional application of the Steady-State Water Chemistry (SSWC) model. Canadian Journal of Fisheries and Aquatic Sciences **59**:1287-1295.

Kotliar, N. B. and J. A. Wiens. 1990. Multiple scales of patchiness and patch structure: A hierarchical framework for the study of heterogeneity. Oikos **59**:253-260.

Lawler, J. J. and T. C. Edwards. 2006. A variance-decomposition approach to investigating multiscale habitat associations. The Condor **108**:47-58.

Li, H. and S. G. McNulty. 2007. Uncertainty analysis on simple mass balance model to calculate critical loads for soil acidity. Environmental Pollution **149**:315-326.

Loiselle, B. A., P. M. Jørgensen, T. Consiglio, I. Jiménez, J. G. Blake, L. G. Lohmann, and O. M. Montiel. 2008. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? Journal of Biogeography **35**:105-116.

Maggini, R., A. Lehmann, N. E. Zimmermann, and A. Guisan. 2006. Improving generalized regression analysis for the spatial prediction of forest communities. Journal of Biogeography **33**:1729-1749.

McNulty, S. G. and J. L. Boggs. 2010. A conceptual framework: Redefining forest soil's critical acid loads under a changing climate. Environmental Pollution **158**:2053-2058.

Miller, J. R., M. G. Turner, E. A. H. Smith, C.L. Dent, and E.H. Stanley. 2004. Spatial extrapolation: the science of predicting ecological patterns and processes. BioScience **54**:310-320.

Newton, I. 1994. The role of nest sites in limiting the numbers of hole-nesting birds: A review. Biological Conservation **70**:265-276.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems **9**:181-199.

Raudys, S. J. and A. K. Jain. 1991. Small sample size effects in statistical pattern recognition: Recommendations for practitioners. IEEE Transactions on pattern analysis and machine intelligence **13**:252-264.

Short, L. L. 1979. Burdens of the picid hole-excavating habit. The Wilson Bulletin **91**:16-28.

Sullivan, T., B. Cosby, A. Herlihy, C. Driscoll, I. Fernandez, T. C. McDonnell, C. W. Boylen, S. Nierzwicki-Bauer, and K. Snyder. 2007. Assessment of the extent to which intensively studied lakes are representative of the Adirondack Mountain region. New York State Energy Research and Development Authority.

Sverdrup, H. and P. Warfvinge. 1993. Calculating field weathering rates using a mechanistic geochemical model PROFILE. Applied Geochemistry **8**:273-283.

Thomas, J., Anderson, R., Maser, C., Bull, E.L., 1979. Snags. In: Thomas, J.W. (Ed.), Wildlife habitats in managed forests the Blue Mountains of Oregon and Washington, Washington, D.C., pp. 60-77

Virkkala, R. 2006. Why study woodpeckers? The significance of woodpeckers in forest ecosystems. Annales Zoologici Fennici **43**:82-85.

Whittingham, M. J., P. A. Stephens, R. B. Bradbury, and R. P. Freckleton. 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology **75**:1182-1189.

Wu, J. and J. L. David. 2002. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. Ecological Modelling **153**:7-26.

Zarnowitz, J. E. and D. A. Manuwal. 1985. The effects of forest management on cavity-nesting birds in northwestern Washington. The Journal of Wildlife Management **49**: 255-263.

CHAPTER 2 - LANDSCAPE MODELING TO PREDICT THE BIOGEOCHEMICAL AND CLIMATIC ENVIRONMENT OF ACIDIC STREAMS IN THE APPALACHIAN MOUNTAINS, USA

ABSTRACT

In many industrialized regions of the world, atmospherically deposited sulfur reduces stream water quality, and resulting acidic conditions threaten aquatic resources. Accurate maps of predicted stream water acidity are an essential aid to managers who must identify acid-sensitive streams, potentially affected biota, and resource protection strategies. In this study, we developed correlative models to predict the acid neutralizing capacity (ANC) of streams across the southern Appalachian Mountain region, USA. Models were developed using stream water chemistry data from 933 sampled locations and continuous maps of pertinent environmental and climatic predictors. Environmental predictors were averaged across the upslope contributing area for each sampled stream location and submitted to both statistical and machine learning regression models. Predictor variables represented key aspects of the contributing geology, soils, climate, topography, and acidic deposition. To reduce model error rates, we employed hurdle modeling to screen out well buffered sites and predict continuous ANC for the remainder of the stream network. Models predicted acid-sensitive streams in forested watersheds with small contributing areas, siliceous lithologies, cool and moist environments, low clay content soils, and moderate or higher dry sulfur deposition. Our results confirmed findings from other studies, but also identified several influential climatic variables and variable interactions. Model predictions indicated that one-quarter of the total stream network was sensitive to additional sulfur inputs (i.e., ANC <100 $\mu eq\ L^{-1}$), while <10% displayed much lower ANC (<50 $\mu eq\ L^{-1}$). These

methods may be readily adapted in other regions to assess stream water quality and potential biotic sensitivity to acidic inputs.

1. INTRODUCTION

Since the industrial revolution, atmospherically deposited sulfur (S) has acidified streams across the eastern United States (US) (U.S. EPA 2009), Europe (Schöpp et al. 2003), China (Galloway et al. 1987), southeast Asia, and other industrialized regions (Galloway 2001, Kuylenstierna et al. 2001, Menz and Seip 2004). Within these areas, increased acidification has depleted base cations from soils through sulfate (SO_4^{2-}) leaching, mobilized of inorganic aluminum (Al_i) to streams (Sullivan 2000), and reduced richness of fish and aquatic invertebrates (Rago and Wiener 1986, Guerold et al. 2000, Sullivan et al. 2007a, U.S. EPA 2009).

Recent reductions in industrial emissions, particularly in the US and Europe, have reduced atmospheric S, but lagged effects of prior deposition are still apparent, particularly in geologically sensitive headwater streams (Guerold et al. 2000, Driscoll et al. 2003, U.S. EPA 2009). Results from recent eastern US lake evaluations suggest that recovery from chronic exposure to atmospheric S may take decades (Driscoll et al. 2003).

Acid neutralizing capacity (ANC) is one measure of stream water acid-base status, which is well correlated with biological health and species richness in acid-sensitive systems (Lien et al. 1992, Sullivan et al. 2007a, U.S. EPA 2009). ANC is the sum of the concentrations of all base cations, minus concentrations of anionic sulfate (SO_4^{2-}), nitrate (NO_3^-), and chloride (Cl^-), reported in $\mu eq\ L^{-1}$. As rates of acidic deposition increase in watersheds, especially those with shallow acid-sensitive soils, surface water ANC generally decreases, but in proportion to the natural re-supply of base cations. Consequently, re-balancing long-term acid-base chemistry in acid-impacted watersheds partially depends on reducing atmospheric S to levels below the

natural re-supply of base cations (Sullivan 2000). Base cation resupply comes from soil mineral base cation weathering (BC_w) or exogenous inputs (Cosby et al. 1985, Henriksen and Posch 2001, McDonnell et al. 2010).

Various ANC thresholds are associated with biological effects (U.S. EPA 2009). Negative effects on macroinvertebrate and fish species richness have been associated with ANC concentrations between ~50-100 $\mu eq \cdot L^{-1}$ (Cosby et al. 2006, Sullivan et al. 2007a), and more substantial effects are observed at lower levels (Cosby et al. 2006, Sullivan et al. 2007a, U.S. EPA 2009). Accurate identification of low ANC streams that are subjected to high acidic deposition levels can focus efforts to reduce key pollution sources.

When monitoring stream water acid-base status, ANC is often chosen over other metrics, such as pH, due to its relative insensitivity to changes in concentrations of CO_2, aluminum reactions, and presence of organic acids (Neal et al. 1999). ANC is also widely used in conjunction with regional estimates of steady-state critical loads (CL) (Henriksen et al. 1995, Duan et al. 2000, U.S. EPA 2009, Clark et al. 2012) to identify acid-sensitive stream reaches (McDonnell et al. in review). The CL is a quantitative estimate of the level of sustained S deposition above which harmful ecosystem effects are likely (Nilsson and Grennfelt 1988).

Taken together, accurately estimated ANC and CL within individual stream reaches can inform decisions about where best to mitigate S deposition in aquatic habitats. ANC modeling reported here was used in conjunction with BC_w and CL estimation for the study region (McDonnell et al. in review). All three estimates are used in a decision support modeling framework to guide resource management and policy decisions regarding S emissions (Reynolds et al. 2012).

The southern Appalachian Mountain region has a long history of atmospheric S deposition, and contains threatened aquatic resources. The region exhibits complex land-use patterns superimposed on steep climatic and topographic gradients. As such, it is well-suited to broad-scale ANC estimation.

Developing regression models that explain the contributions of biogeochemical and climatic variables to acid neutralization can be a difficult process, particularly when modeling these relations at large spatial scales. Difficulties arise from inherent geographic variability in interactions among biological, geochemical, and climatic variables that can influence the susceptibility of a stream to S deposition (Turner 1989, Levin 1992). These interactions can be non-linear and temporally non-stationary, and therefore are poorly addressed by traditional modeling frameworks (Elith et al. (2008)).

In an effort to address these factors, machine-learning techniques have recently been introduced to mainstream ecological research. Machine learning methods (a) are reportedly robust against multi-collinearity and outliers; (b) include methods to reduce model overfitting; (c) able to identify identify important predictor variables, non-linear relationships, and complex interactions among predictors; (d) are unaffected by data transformations; and (e) can incorporate categorical, ordinal, or continuous numeric predictors (Elith et al. 2008, Olden et al. 2008, Franklin and Miller 2009). A disadvantage of machine learning techniques is that most are non-parametric and do not produce model coefficients associated with traditional statistical models. Example machine learning models include ensemble decision trees, neural networks, support vector machines, and Bayesian belief networks (Hastie et al. 2005).

Here, we employ machine learning and statistical regression models to predict the biogeochemical, climatic, vegetative, and acidic deposition that are associated with low-ANC streams in the southern Appalachian Mountain region. To accomplish this, we:

1) gathered available stream water ANC datasets within the region,
2) used available remotely sensed, surveyed or process modeled climate, land cover, atmospheric deposition, geologic, edaphic, and topographic data,
3) compared traditional and machine learning approaches to select the best performing model, and
4) used hurdle modeling and data resampling to address data imbalances.

Our objectives were to develop and validate models that best explained observed ANC, identify key explanatory variables, predict ANC for a continuous stream network, and identify acid-sensitive streams and focal areas for future sampling, monitoring, and potential remediation.

This research advances the work of Sullivan et al. (2007b) who used logistic regression over a portion of the same region. Here, we evaluate the effects of imbalance in the ANC data distribution, compare machine learning and statistical regression model performance, evaluate a much broader set of potential predictors, average predictors to the upslope contributing area, and provide continuous estimates of ANC across a larger region.

2. 2. METHODS

2.1. Study area and background

The study area was located in the southern Appalachian Mountain region ($14.3 \cdot 10^6$ ha), which extends from northern Georgia to southern Pennsylvania, and from eastern Kentucky to central Virginia (Figure 2.1). The region is primarily composed of the Blue Ridge, Ridge and Valley,

and Central Appalachian ecoregions (Omernik 1987). The Blue Ridge ecoregion is dominated by metamorphic and igneous parent materials, whereas the Ridge and Valley and Central Appalachian ecoregions are primarily sedimentary, with northeast to southwest trending sandstone ridges and limestone valleys (Figure 2.1). Elevations range from about 300 - 2000 m. The dominant land cover of the area consists of oak, hickory, pine, spruce and hemlock forests, interspersed with crop and pasture lands, and urban areas.

2.2. Acid neutralizing capacity data

Water chemistry data were used to calculate stream water ANC. Data were obtained from national and regional databases, including the National Stream Survey, Environmental Monitoring and Assessment Program, Virginia Trout Stream Sensitivity Study (VTSSS), and others (see also Sullivan et al. 2004, Sullivan et al. 2007b). A total of 933 sampled sites were included in this study.

Water chemistry data were collected mainly during the spring season between 1986 and 2009, with 43% of the data collected during the 2000 VTSSS survey. All water chemistry samples were georeferenced to a synthetic stream network created using a hydrologically conditioned 30-m digital elevation model (DEM; U.S. EPA and U.S.G.S. 2005) within a geographical information system (GIS).

ANC was calculated as the sum of the charge balance of Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Cl^{-}, NO_3^{-}, NH_4^{+}, and SO_4^{2-}. ANC values across in the database ranged from -109 to 3,889 $\mu eq \cdot L^{-1}$ (mean=188$\pm$414 (standard deviation, SD) $\mu eq \cdot L^{-1}$, median = 72), and the data were right-skewed (Figure 2.2). We address this below in Section 2.6 and 2.7.

2.3. Predictor variables

An initial set of environmental predictors was chosen to represent broad- to fine-scale climatic, lithologic, geomorphic, topographic, edaphic, vegetation, land ownership, and S deposition conditions that were potentially influential to ANC (see Appendix 2.1 for complete details). All data layers were resampled to 30-m raster grids (Appendix 2.1), and data values were averaged across the upslope contributing area of each 30-m grid cell, using the methods described in McDonnell et al. (2012). The equation for upslope averaging follows,

$$\bar{P} = \frac{P_i + \sum_1^j P_j}{(N+1)} \quad \text{Eq. 1}$$

where $\bar{P}$ is the upslope averaged value for the candidate cell (P_i), $\sum_j^N P_j$ is the summation of all cell values upslope of P_i, and N is the total number of upslope cells. Upslope-averaging enabled me to attribute the average of each predictor variable across the landscape draining into each individual cell. Environmental data were obtained from several sources, as described below.

Eighteen climate variables representing 1961-1990 climate normals were taken from a 1-km resolution Ameriflux dataset developed by Hargrove and Hoffman (2004) for the conterminous United States (CONUS). Climate variables represented various aspects of the temperature, precipitation, and insolation regimes, each conditioned by local growing and non-growing seasons. Five soil, one topography, two vegetation, and ten productivity variables were also provided by this dataset (Appendix 2.1).

The National Land Cover Data set (NLCD; Homer et al. 2007) provided data on the percentage areal cover of major land cover types, including coniferous, hardwood, and all forest types combined. Two additional classes were derived that combined mixed coniferous and hardwood forest by weighted average. A landownership variable was also included to represent the percentage of catchment area in federal vs. non-federal ownership (National Atlas of the

United States 2006). Due to different management histories on federal versus privately owned lands in the eastern US (more intensive logging on non-federal lands, in general), this layer was intended to provide proxy information on degree of past logging disturbance – a known influence on ANC (Sullivan et al. 1999).

Using a 30-m digital elevation model (DEM), we derived three topographic variables – a steady-state topographic wetness index (TWI; Wood et al. 1990), surface area ratio (SAR), and flow accumulation (FAC). TWI was computed as the log of the catchment size divided by the catchment slope (radians), to represent the propensity of each grid cell to accumulate water (Moore et al. 1993). SAR measured terrain roughness as the ratio of sloped to flat surface area of a grid cell (Jenness 2004). FAC represented the total area contributing overland flow to a grid cell (Jenson and Domingue 1988). These variables represented watershed characteristics and were not upslope averaged.

Additional soils data were obtained from the Soil Survey Geographic (NRCS Soil Survey Staff 2010a) and the U.S. General Soil Map (NRCS Soil Survey Staff 2010b) databases. Soil variables included percentage clay, soil pH, and soil depth. A broad-scale lithology classification provided by Sullivan et al. (2007b) was used to capture the percentage composition of parent materials across the study area. Classes included siliceous, argillic, felsic, mafic, and carbonate substrates. Mapped surface lithologies were composites from State geologic maps (USGS 2005b, a).

Total wet and dry S deposition were calculated based on three-year averages of the NADP (National Atmospheric Deposition Program, (Grimm and Lynch 2004)) interpolated wet (375m resolution) and CMAQ (Community Multiscale Air Quality, (Byun and Schere 2006)) modeled dry deposition (12km resolution), centered on the 2002 weather year.

The initial set of predictors included 57 variables; only those with Pearson's correlation scores <0.7 were retained, leading to a modeling set of 33 variables. Among correlated variables, those with the highest Pearson's correlation with ANC were retained. Other correlation cutoff values were considered and evaluated, but none improved the final models.

2.4. Model development

To identify best modeling approaches, we compared traditional regression and machine learning models using a common set of performance metrics (described below). Statistical models included linear models (LM) and logistic regression (logR), random forest (RF), and boosted classification and regression trees (BCT, BRT).

Exploratory ecological analyses like this one are opportunistic, involving combined datasets, which can yield imbalanced sampling designs and skewed data distributions (Chawla et al. 2002, Barandela et al. 2003). While most statistical models, including advanced machine learning algorithms (Elith et al. 2008, Elith and Leathwick 2009), assume balanced data designs, some produce robust predictions with even moderate imbalances, but few perform well with large imbalances. We explored a variety of techniques to ameliorate the influence of imbalanced data in our study and increase the overall accuracy of model predictions.

We began by predicting ANC as a continuous response using LM, RF, and BRT models. Next, we used two-stage hurdle modeling, in combination with the regression techniques, to screen out sites with a high probability of being well-buffered, and then predicted a continuous ANC value for the remainder. Finally, we tested two data resampling techniques that imposed a balanced data distribution within the hurdle modeling framework.

2.5. Machine learning algorithms

We compared the performance of boosted classification tree (BCT), BRT, (*gbm* package in R v2.12.2, (Ridgeway 2006)) and RF (*randomForest* package in R v2.12.2, (Liaw and Wiener 2002)) models against LM, logistic regression (logR) models (*stats* package in R, (R Development Core Team 2011)) to determine the best approach, considering the available data and the prediction goals. Models were validated using a 75/25% training/testing set, over 50 iterations to ensure accuracy of predicted model error-rates. Model error metrics are discussed below in Section 2.6.

Machine learning algorithms are relatively new to ecological research (Hastie et al. 2005, Olden et al. 2008). Many are data driven, meaning that models do not produce a parameterized statistical model, but identify patterns in the data with few assumptions regarding the underlying probability distribution of the training data (Breiman 2001b). Accordingly, the main advantage of these algorithms is their potential to model complex and non-linear relationships without having a priori knowledge of the key predictor variables and the functional form of their relationships to the dependent variable (Gahegan 2003, Olden et al. 2008, Franklin and Miller 2009).

Boosted and RF methods represent ensemble versions of traditional classification and regression tree (CART) analysis that output a majority vote or mean value from a series of trees (De'ath and Fabricius 2000, Prasad et al. 2006, Elith et al. 2008). In lieu of a parameterized statistical model, CART splits the data into successively smaller and more homogenous groups until some stopping criterion is met. This is done by iteratively sorting each predictor variable, and then splitting the data into two mutually exclusive groups at each iteration. This is repeated for every value of the predictor and for all predictors individually. The predictor and splitting

value are chosen for each split that minimizes the within-group heterogeneity of the response variable (Breiman 1984, De'ath and Fabricius 2000).

RF and boosted methods build tens to thousands of CART trees; predictions made from the models are based on a majority vote (classification) or averaged value (regression). Boosting works by using the entire set of predictors and data to build each individual tree, and the algorithm is sequential, using information about model residuals of past trees to guide development of subsequent trees (De'ath 2007, Elith et al. 2008). RF models use a bagging algorithm, in which individual trees are built using random samples of the predictors and of the data, each in a predefined proportion. Model performance is assessed using an independent test set, n-fold cross-validation (BCT, BRT), or out-of-bag estimates (RF). Out-of-bag samples refer to the training data left while building individual regression trees. Mean out-of-bag error rates are estimated by models for each out-of-bag sample by entering them into the tree they were omitted from, calculating error estimates for each tree, and averaging error estimates across all trees.

Variable importance measures are also calculated by these algorithms. For BRT/BCT, variable importance is based on the number of times the variable was chosen as a predictor in the individual trees, and weighted by the deviance the variable explained across all trees (Elith et al. 2008). For RF, variable importance is calculated as the difference between the error-rate of an individual tree and the error rate of the tree calculated using randomly assigned values for the specific predictor, averaged across all trees in the RF model (Breiman 2001a).

2.6. Hurdle modeling

Initially, we found that single regression models exhibited high error rates, particularly for streams with ANC values >150 $\mu eq \cdot L^{-1}$. A two-stage hurdle model was used to minimize RMSE

of predicted ANC values (Figure 2.3). In a first stage, a binomial model (e.g., logR, BCT, RF) predicted the probability that ANC values for any 30-m grid cell were below a specified threshold (e.g., ANC <200 $\mu eq \cdot L^{-1}$). If a cell exhibited a high probability (e.g., probability >0.5) of a low ANC value, that cell was entered into a second regression model (e.g., LM, BRT, or RF), where continuous ANC values were predicted. If a cell exhibited a low probability (e.g., <0.5) of a low-ANC value, it was considered well-buffered, assigned an arbitrarily high ANC value, and not considered further by the continuous model. Tested ANC threshold values were 150, 200, 250, and 300 $\mu eq \cdot L^{-1}$; tested probability cut-off values were (0.4, 0.5, 0.6, and 0.7).

Threshold and continuous models were trained separately to identify the optimal statistical model, predictors, and parameters. Models were constructed using 3, 5, 7, 10, 15, 20, 25, and 33 of the most influential predictors. Performance was evaluated for each of these models to optimize model parsimony and prediction accuracy. Performance metrics were calculated for threshold and continuous models using a random 25% draw on the dataset. Threshold models were compared using misclassification rate, Kappa statistic (Maclure and Willet 1987), G mean, and area under the receiver operating curve (AUC) (He and Garcia 2009). Continuous models were compared using model RMSE (lower numbers indicate higher prediction accuracy) and coefficient of determination (R^2).

Within the hurdle model, data resampling techniques were used to help reduce the influence of imbalanced data on model performance. These included both (1) randomly oversampling the "high ANC" stream sample sites (those with values greater than the specified threshold, see *Hurdle modeling* above) until the sample number equaled that of the low-ANC sites, and (2) randomly undersampling the low-ANC sites until the sample sizes were equal.

2.7. Validating the complete hurdle model

Once both the threshold and continuous models were trained and the optimal models identified, the complete hurdle model (threshold + continuous models) was trained and validated using a randomly drawn training set consisting of 75% of the sampled water chemistry sites. During model training, three parameters were adjusted to optimize hurdle model parameterization: the threshold ANC value (150, 200, 250, or 300 $\mu eq \cdot L^{-1}$), the probability cutoff value (0.4, 0.5, 0.6, or 0.7), and the data resampling technique (no resample, oversample high-ANC, undersample low-ANC) – resulting in 48 unique model parameterizations. All combinations of these parameterizations were run 50 times each, using unique random draws on the data to create the training and testing sets, and model statistics were averaged to identify optimal hurdle model parameterization. All model parameterizations were compared by plotting mean ($\pm$SD) RMSEs for the continuous models against mean ($\pm$SD) false positive rates for the threshold models. The false positive rate, as opposed to overall model error-rate, was used to minimize instances in which a truly low ANC site was erroneously predicted as high-ANC. The "best" model parameterization exhibited the lowest values for the two error-rates.

2.8. Spatial autocorrelation

Spatial autocorrelation in the final ANC model was assessed with a Moran's I correlogram on model residuals using the *spdep* package within R (Bivand et al. 2011). This statistic varies between -1 (perfect negative correlation) and 1 (perfect positive correlation); values approaching 0 indicate complete spatial randomness. The correlogram also displays the results of null hypothesis testing at varying geographic distances.

3. RESULTS

3.1. LM regression and machine learning

We initially predicted continuous surfaces of ANC using RF, BRT, and LM. RMSE of predictions (based on independent test sets withheld from model training over 50 iterations) were 7-14% lower for machine learning algorithms (RF: 258.1 $\pm$ 36.2, BRT: 277.2 $\pm$ 38.7) compared with LM regression (298.3 $\pm$ 37.4), but error-rates were relatively high for all models in relation to the scale of the ANC values. Error-rates for all models were highest for sites with ANC > 150 $\mu eq \cdot L^{-1}$ (Figure 2.4).

3.2. Hurdle modeling training

3.2.1. Threshold and continuous model selection

Compared to the initial continuous models, hurdle models reduced overall RMSE rates and identified key predictors contributing to low ANC. Within the hurdle modeling framework, RF models displayed the lowest error rates of all tested (Figure 2.5). Across all models, performance declined when fewer than 10 predictors were entered into the model (Figure 2.5).

RF threshold models exhibited AUC scores >0.9, Kappa scores >0.7, and error-rates < 8% when models included between 7-20 predictors (Figure 2.5). Over this range of predictors, LM models exhibited ~10% error-rates, AUC scores <0.9, and Kappa scores ~0.5. Similar results were found for the continuous models where R^2 for RF models were $\geq$0.5 compared to 0.40-0.48 for LM, and RMSE rates were consistently high for RF.

The final hurdle model included RF threshold and continuous models. Both models included 10 predictors, and were derived from 1,000 individual regression trees (*ntree* parameter) with a bagging fraction of 30% (*mtry* parameter).

3.2.2. Variable importance

Threshold model – From the threshold model, low ANC occurred in small forested catchment areas with non-carbonate lithology, cool and moist climates, with relatively low soil pH and clay levels (Table 2.1 and Figure 2.6). With few exceptions, predictors influenced ANC in a non-linear manner, and non-linearities occurred at the ends of the response curves. For example, the probability of a low value of ANC increased non-linearly with average percent forested cover, but the majority of samples (Figure 2.6, panel 6, hash marks on the x-axis) had between 90-100% forested cover.

Significant interactions among variables were identified by the RF model for percentage calcareous, percentage public land, the maximum number of very wet days (i.e., vapor pressure deficit <1000 pa, VWDAYMAX), and maximum number of continuous very dry days (i.e., vapor pressure deficit >750 pa; VDCONTDAY). Areas with <10% calcareous lithology in fairly moist environments had the highest probabilities of being low-ANC; probabilities decreased most with slight increase in calcareous lithology and in the driest environments (Appendix 2.2). Other important interactions included the amount of catchment area in public lands, and the highest probabilities of low-ANC occurred on public lands, and in moist environments. However, the interaction between public lands and carbonaceous lithology suggested that probabilities were highest outside of public lands and on non-calcareous lithology (Appendix 2.2). Slight increases in public lands reduced the probability of low-ANC, but probabilities increased slightly with increased amounts of public land; increases in calcareous lithology reduced the probability of low-ANC only for catchments with negligible amounts of public land (Appendix 2.2).

Continuous model – From the continuous model, low ANC occurred in areas with siliceous lithologies, a relatively moist, cool, and short growing season, in conditions with low soil pH and clay levels, and in small, forested catchment areas (Table 2.1, Figure 2.7). Like the threshold model, most variables had non-linear response functions, and non-linearities generally occurred at the extremes of the response curves.

Important interactions were identified for the amount of siliceous lithology, precipitation, temperature, forest cover, and topography variables (Appendix 2.3). Topographic wetness index had a highly non-linear influence on ANC levels (Figure 2.7), but when interacting with the level of siliceous lithology this was less apparent and ANC appeared most influenced by lithology. Interactions among climate variables clearly showed that low-ANC was associated with low temperatures and high precipitation. Catchments with warm climates appeared to reduce the influence of siliceous lithology on ANC; catchments with 80% of their area in siliceous lithology and >10 days with temperatures >32.2^{o} C had a predicted ANC of ~140 $\mu eq \cdot L^{-1}$ compared to only ~50 $\mu eq \cdot L^{-1}$ in the coldest environments. The amount of forested area had a similar influence, and catchments with <90% forest cover and <80% siliceous lithology appeared well buffered.

To better understand differences in important drivers of ANC at high compared to low elevations, we stratified low (<500 m, 1st quartile) from high elevation (>850 m, 3rd quartile) sites and developed continuous RF models for observations with ANC $\leq$300 $\mu eq \cdot L^{-1}$ separately for each elevation setting. Drivers of low-ANC for the low elevation model were similar to those of the full model (Appendix 2.4), but in the high elevation model, low ANC was found in areas with high coniferous and low deciduous cover (Appendix 2.5). Wet S deposition was also a

leading predictor in the high elevation model, but not in the low-elevation model or hurdle model, and ANC was lowest for intermediate levels of deposition.

3.2.3. Hurdle model performance

Based on model performance results presented in Figure 2.8, the final model parameterization included RF threshold and continuous models, imbalanced data (no resampling), and a 0.7 probability cutoff, with a 300 $\mu eq \cdot L^{-1}$ ANC threshold (Model 46, Figure 2.8, Appendix 2.6). This model displayed a 5.6% omission error-rate, 9.5% overall error-rate of classification, and an RMSE of 107.5 $\mu eq \cdot L^{-1}$ based on 50 model iterations. Overall, hurdle modeling reduced model RMSE by 140-177%. Model RMSE was lowest for streams segments with ANC <150 $\mu eq \cdot L^{-1}$ (Figure 2.4); but this model consistently under-predicted ANC for values >150 $\mu eq \cdot L^{-1}$. Under-prediction was likely associated with few sampled water chemistry data sites in streams with ANC >150 $\mu eq \cdot L^{-1}$. This can be remedied with additional sampling. When only sites with ANC < 150 $\mu eq \cdot L^{-1}$ were included in the model, RMSE was 36.6 $\mu eq \cdot L^{-1}$.

The majority of the RMSE for well-buffered sites was related to threshold model misclassifications of high-ANC sites as low; a behavior consistent across all models tested. When misclassifications were removed from the error analysis, the resultant RMSE was 49.4 $\mu eq \cdot L^{-1}$ for the full model.

Results of the spatial autocorrelation analysis indicated that spatial autocorrelation among model residuals was negligible (Table 2.2).

3.2.4. ANC model predictions

Stream water chemistry often associated with biological harm (e.g., ANC <100 $\mu eq\ L^{-1}$) occurred throughout the study area (Figure 2.9). Approximately 8% of the stream network displayed ANC values <50 $\mu eq\ L^{-1}$, and 24.5% with values <100 $\mu eq \cdot L^{-1}$ (Figure 2.10). Lithologic patterns

clearly influenced ANC predictions, with high-ANC areas occurring mainly in limestone valley bottoms, and low ANC occurring on siliceous bedrock.

We compared predictions from RF hurdle models to those from a linear regression model (LM). Differences in the patterns of predictions made by the LM and RF hurdle models were apparent. The LM model predicted >84% of the stream network (as measured by length) had ANC values >150 $\mu eq \cdot L^{-1}$, compared to only 58% for the RF hurdle model (Figure 2.10). The RF model predicted 35% of streams had ANC between values 50-150 $\mu eq \cdot L^{-1}$ compared to only 8.9% predicted by LM (Figure 2.10). For the most acidic class (i.e., ANC <0 $\mu eq \cdot L^{-1}$), the LM model predicted a slightly higher percentage of streams (3.9% LM *versus* 0.9% RF hurdle).

LM tended to predict higher ANC levels particularly in the central Appalachian region (Appendix 2.7, 2.8). However, in the southern Blue Ridge, and in pockets of eastern and central West Virginia, where the RF hurdle model predicted low ANC rates (e.g., ANC <50 $\mu eq \cdot L^{-1}$), the LM model generally predicted very low ANC rates (Appendix 2.7, 2.8). Low ANC predictions from the LM model were fairly isolated geographically and generally corresponded to areas of highest sampling intensity, potentially indicating that the LM model had poor predictive ability outside the limits of the training data.

Spatial patterning of RF model uncertainty was also apparent across the study area. Model uncertainty was expressed as the standard deviation of continuous predictions made among the ensemble of individual trees comprising the RF model. As might be expected, model uncertainty was highest in the western portion of the study domain where few samples occurred (Figure 2.9) and lowest in the areas with the most samples.

4. DISCUSSION

4.1. Modeling low-ANC

Biogeochemical and climatic influences on ANC almost certainly involve many interactions among environmental and climatic processes that together influence the ultimate susceptibility of a stream to the acidifying effects of S deposition. Some of these factors are as yet unknown, and others are known, but complex, non-linear, non-stationary over space and time, and difficult to measure (Driscoll et al. 1987, U.S. EPA 2009). Known processes include nutrient uptake by plants and transport of litter fall and delivery to neighboring streams, mobilization of Al^{3+}, erosion and sedimentation of cations and anions from upslope catchments, deposition of S from exogenous sources, and physical and chemical weathering of parent materials proximal and distal to the stream channel (Driscoll et al. 1987, Reuss et al. 1987, Christophersen and Neal 1990, Sullivan 2000, Johnson and Host 2010). Modeling the susceptibility of aquatic systems to acidic deposition at a regional scale requires an understanding of these processes, access to adequately resolved data layers that accurately portray the complexity of environmental conditions, and reliable statistical modeling.

Our aim was to predict how integrated biological, geological, chemical, and climatic processes within catchments can be used to predict stream water chemistry, which has linear flow properties that may concentrate or dilute local buffering capacity within stream networks, and which result in a unique correlation structure among observed responses and the environments driving the response. Results from this study indicated that predictions of ANC within a region can be made using a combination of remotely sensed, spatially explicit environmental data, a representative sample of water chemistry across the region, and a modeling framework appropriate for the analysis.

We addressed several aspects of the environmental and stream water chemistry during the modeling process including the spatial scale of the environmental predictors, the timespan over which water chemistry was sampled, and the spatial distribution of water sample sites.

Stream water chemistry is governed by biological and physical processes that are both endogenous and exogenous to the contributing watersheds (Sullivan et al. 2007b). Water chemistry may be sampled at points along streams, but the environmental and edaphic influences on that chemistry must be related to the contributing space assignable to those points (Johnson and Host 2010, Steel et al. 2010). We accounted for this by using an upslope-averaging technique (McDonnell et al. 2012) that generated values for predictor variables that reflected conditions over the contributing area to each point along the stream network. Point estimates of the biogeoclimatic setting at the location of water chemistry sample points would not adequately represent catchment level influences on water chemistry.

The water chemistry data used in this study were obtained from a series of sampling efforts undertaken intermittently over a span of years (1986-2009). These data were used because they were readily available, sampled a fairly large spatial extent, were collected at fairly high density, and used a consistent method for calculating stream water acid-base status. All sample data used for modeling were collected during spring flow. As such, the model represents a description of the environment expected to influence spring season ANC concentrations, which are often the lowest that occur in these streams (Herlihy et al. 1993). Although S deposition declined substantially between 1986 and 2009, stream water ANC has shown little recovery, mainly because of base cation depletion and continued S adsorption on soils (Sullivan et al. 2004, Sullivan et al. 2008, Sullivan et al. 2011). As a result, we expect that data collected over a period of two decades generally reflect ambient stream acid-base chemistry.

The lack of an *a priori* sampling design necessitated an assessment of potential sampling bias. The effects of data resampling on model prediction were mixed. Models that undersampled low-ANC sites (the majority sample) reduced RMSE, but increased threshold model error-rates. Models that oversampled high-ANC sites (the minority sample) produced negligible differences in model performance (Figure 2.8). Hence, rebalancing the sample was unnecessary, and RF models performed reasonably well, even with unbalanced sampling. Our results suggest that the multivariate signature of low ANC was unique enough to be readily detected against the background noise of other ANC levels. Moreover, the RF and boosted models consistently outperformed traditional regression approaches, and two-stage hurdle modeling adequately minimized effects of sampling bias on predicted ANC.

Comparison of mean RMSE values for RF models with and without hurdle modeling showed that the hurdle model provided a more than two-fold decrease in model RMSE (258.1, RF-only, and 107.5, hurdle model). The RMSE for ANC predictions <150 $\mu eq \cdot L^{-1}$ was 36.6 $\mu eq\ L^{-1}$. This is an important observation--we can accept higher error rates in predictions of high ANC, but require more accurate predictions where ANC values are within the range of known ecological degradation. Model performance was substantially improved for the ANC conditions that are most biologically relevant ($\leq 100\ \mu eq\ L^{-1}$).

A major aim of ecological models is often to predict various phenomena or responses over large geographic areas. When models produce sufficiently accurate predictions, they provide valuable insight into pattern and process interactions, and variability of those interactions. Our model allowed us to (1) identify areas of potential concern for ecological remediation and management, (2) identify discrete areas where exogenous factors such as the broad-scale climatic setting may trump local factors such as topography or soils (e.g., the

western portion of the region where ANC predictions were low despite deeply dissected topography; presumably due to precipitation and S deposition patterns), (3) identify areas where spatial clustering of environmental degradation is predicted, and (4) locate areas of high uncertainty, which may indicate inadequate sampling or variable interactions among the predictors.

Knowledge of the uncertainty in model predictions is essential to evaluating model performance across geographic areas. Where model uncertainty is high, there may be deficiencies in sampling, unique environments not captured by the training data, lack of phenomenological understanding, or errors in model parameterization. Model validation results indicated that higher model uncertainty occurred in areas where observed ANC values were between 100-300 $\mu eq\ L^{-1}$; mapped SD values (Figure 2.9) confirmed this observation. These areas would benefit from additional sampling and monitoring.

Predictions from our model indicate that streams with ANC <100 $\mu eq\ L^{-1}$ make up approximately one-quarter of the total stream length in the study area. ANC levels <100 $\mu eq\ L^{-1}$ are associated with potential reductions in fish and macroinvertebrate taxonomic richness (Cosby et al. 2006). Approximately 10% of the total stream length in the study area was predicted to exhibit ANC < 50 $\mu eq\ L^{-1}$. Streams with ANC <50 $\mu eq\ L^{-1}$ are sensitive to episodic acidification, and fish species richness is greatly reduced (Schindler 1988). Low-ANC streams were generally located in the mountainous regions of the Central Appalachian and Blue Ridge Mountain provinces and in the central Ridge and Valley Province (Figure 2.9). These areas are characterized by rugged topography, relatively steep temperature and precipitation gradients, high percentage of forest cover, and a variety of geologic parent materials.

In lieu of using a statistical modeling approach, previous estimates of regional stream water ANC in our study area have been made using stratified random sampling designed to evaluate the proportion of surveyed streams contained within specified ANC classes (Herlihy et al. 1993). Estimates from these surveys generally corroborate our findings when individual ecoregions are assessed, although direct comparison is not possible. The authors concluded that about 13-25% of the sample streams displayed ANC values <50 $\mu eq\ L^{-1}$; our area prediction was lower (<10%).

4.2. Endogenous and exogenous drivers of stream water acidity

Results from the ANC modeling suggested low ANC of streams in the southern Appalachian Mountain Region derived from lithologies, land-surface forms, temperature and precipitation regimes, exogenous S deposition, and endogenous patterns of physiognomies, soils, and topographies within catchments.

Streams situated on siliceous lithologies had relatively low ANC, and this variable was the strongest predictor in the continuous model. Other studies have also found strong empirical relationships between lithology and stream water ANC (Puckett and Bricker 1992, Herlihy et al. 1993, Sullivan et al. 2007b). Soils created from siliceous materials tend to be shallow, acidic (Krug and Frink 1983), and generally exhibit low productivity, conductivity and base cation weathering (Herlihy et al. 1993). Despite the known importance of lithology on governing stream water quality and other ecological processes, region-wide data layers describing mineralogy (NRCS Soil Survey Staff 2010a) are of inconsistent quality and resolution (Sullivan et al. 2007b). For instance, areas of West Virginia and Georgia delineate mineralogy at different resolutions, which likely contributes to reduced model performance in these areas (Figure 2.9). Because lithology and soil characteristics were important predictors in both the threshold and

continuous models (Table 2.1), we suspect that improved quality, resolution, and consistency in soil and geologic data would improve model prediction and spatial accuracy.

Low ANC was also associated with a high percentage of forest cover (Figure 2.6, 2.7). Sullivan et al. (2007b) and Herlihy et al. (1998) found a similar relationship between low ANC and high forest cover for streams within our study area. Afforested areas in Europe have also been found to consistently incite high levels of acidity in soils and stream water (Miles 1986, Whitehead et al. 1988, Gee and Stoner 1989, Jenkins et al. 1990). Forests can have a variable influence on water chemistry depending on the prevailing climate, underlying geology, catchment size, successional stage of forest development, forest composition, and legacy influences of past disturbances, among others (Harriman and Morrison 1982, Sullivan 2000). Forests generally scavenge wet and dry deposition of atmospheric pollutants (including, but not limited to S) and translate these pollutants to watershed soils and the associated stream reach. Furthermore, forests can take up base cations, and reduce the capacity of soils to buffer against acidic inputs. However, forested areas in the Appalachian region are generally located at higher elevations and on ridges where catchments are small and reside on siliceous bedrocks (Herlihy et al. 1993). Therefore, the geographic locations of forested areas may be coincident with areas predisposed to having low ANC due to the edaphic and lithologic setting.

When high elevation sites were modeled alone, percentage coniferous forest cover was the leading variable in the model, followed by deciduous forest cover. Moreover, these cover types displayed opposite influences on ANC (i.e., high coniferous cover was associated with low ANC and high deciduous cover was related to high ANC; Appendix 2.5). This observation is generally consistent with other work (Nihlgård 1970, Cronan et al. 1978). Indeed, high-elevation catchments dominated by deciduous forest ($\geq$60%; threshold value, Appendix 2.5) had mean

ANC of 83.8 μeq L^{-1}, compared to -19.8 μeq L^{-1} for catchments with a lesser deciduous component. At high elevation, deciduous forests were generally located in larger catchments, with a lower percentage of siliceous parent material, less dry S deposition, and warmer and drier climates than coniferous forests. Whether coniferous forest types contributed to low ANC at higher elevations through acidic foliage deposition, pollutant sequestration, and subsequent acidic throughfall and stemflow (Miles 1986, Dunford et al. 2012), or if the environmental setting of coniferous forests influenced low ANC remains unclear. It is plausible that the influences of catchment size, mineral substrate, weathering rates, and dominant vegetation are highly interactive and change across the study region. The continuous model in our study identified important interactions among forest cover and lithology, and ANC was higher in catchments dominated by siliceous lithology when forest cover was low, compared to well-forested catchments with similar lithologies (Appendix 2.3).

Although elevation is generally considered a surrogate for climate at broad-scales, elevation was not explicitly included as a predictor variable because remotely sensed climate data are now regionally available at a scale and accuracy sufficient for our application. Predictive modeling is most robust when predictor variables expected to be directly linked to the modeled response are used in place of those that may have indirect effects (Austin 2002). For example, elevation is often included as a predictor in species distribution modeling (Austin 2002), but rarely do species respond to elevation alone. Rather, more direct predictors such as temperature and precipitation, which often are correlated with elevation, are more directly related to a species' distributional limits (Austin 2002, Elith and Leathwick 2009).

Elevation is often used as a predictor in the ANC literature, where increases in elevation are often associated with (1) increased S deposition (Lawrence et al. 1999), (2) higher

precipitation (Sullivan et al. 1999), (3) cooler temperatures, (4) higher percentage siliceous parent material, (5) thinner and coarser soils with lower cation exchange capacity, (6) smaller catchments with steep slopes, and (7) higher percentage forest cover (Lynch and Dise 1985, Herlihy et al. 1993, Herlihy et al. 1998, Sullivan et al. 2007b), all of which lead to decreased surface water ANC (Herlihy et al. 1993, Sullivan et al. 2007b, Sullivan et al. 2011). However, it is not clear how these factors rank in importance in driving ANC at regional scales.

Our modeling results indicated that elevation was either uncorrelated or loosely correlated with stream water ANC (r=-0.21), total S deposition (r=0.11), percentage siliceous lithology (r=-0.04), soil depth (r=0.06), soil clay (r=-0.26), catchment size (r=-0.09), and forest cover (r=-0.01). However, elevation was correlated with some, but not all, climate variables. These variables included number of days >32^{o}C (AB90GROW, r=-0.70), non-growing season precipitation levels (r=0.63), number of days without precipitation (NPDAYMAX, r=-0.63), number of humid days (VWDAYMAX, r=0.70), non-growing season solar insolation (SOLARTOTNG, r=0.61), and non-growing season gross primary productivity (GPPNG, r=0.60). Three of these climate variables (NPDAYMAX, VWDAYMAX, AB90GROW) were among the top five variables in the continuous model in our study, and indicated that low-ANC streams coincided with areas that experience few precipitation-free days, few dry days, and few days >32^{o}C. These variables had higher variable importance scores than other edaphic, topographic, vegetation, and sulfur deposition variables in the model, suggesting that climate may drive regional stream water acidity.

The importance of identifying climatic correlates of stream water ANC with our models has implications for monitoring and continued modeling of future stream water acid-base chemistry (Evans 2005, Wright et al. 2006, Benčoková et al. 2011). Climate projections for the

southern Appalachian Mountain region indicate both warmer and drier (south) and warmer and wetter (north) futures (year 2100 projections) (Solomon 2007, Hayhoe et al. 2008, Karl et al. 2009). ANC modeling results reported here indicate that areas with relatively low precipitation and warm temperatures have relatively high stream water ANC. A warmer climate would increase BC_w and therefore stream ANC. Drier conditions would decrease wet S deposition and leaching losses of base cations from soils. Interactions from the continuous model (Appendix 2.3) suggest that higher temperatures and lower precipitation can moderate the influence of siliceous lithology on reducing stream water ANC levels. Monitoring ANC over time will be necessary to better understand these interactions.

Ecosystem responses to longer-term climatic change are largely unknown due to the potential complexity of interactions among the changes. Accompanying changes in temperature, precipitation, and insolation regimes, changes in overall productivity, vegetative communities, plant cover, carbon and nutrient uptake by plants, and influences of disturbances may all affect the acid-base chemistry of streams within the study domain. Thus, it will be important to continue to monitor and model ANC. Moreover, new combinations or changes in the importance of predictors will likely necessitate water chemistry sampling that more evenly samples the variability of conditions represented by the predictors.

Since enactment of the Clean Air Act Amendments of 1990, levels of S deposition in the eastern US have decreased by more than 40% (Baumgardner et al. 2002). Ongoing monitoring suggests that reduced deposition improved stream and lake acid-base chemistry at some locations (U.S. EPA 2009). In the current assessment, the amount of S deposited was an important predictor of ANC, but S deposition was only included in one model (the continuous model) and was not included in any other of the models in the analysis. Instead, other variables related to the

surrounding geology, soils, climate, and vegetation were generally more influential than the amount of S deposition in the catchment.

Furthermore, the continuous model did not indicate a monotonic decrease in ANC with increasing levels of S deposition; rather, non-linearities in this relation were apparent. Modeled ANC was lowest for intermediate levels of dry S deposition, pointing to interactions among S and other predictors. For example, intermediate levels of dry S deposition typically occurred in montane environments, where wet deposition was highest, catchment areas were small, and climate was cool and moist with frequent cloud cover.

Non-linearities may also derive from the rather coarse resolution of CMAQ modeled dry and wet S deposition or from CMAQ prediction errors. Considerable uncertainty exists in these layers, but it is difficult to quantify the extent of these uncertainties as no "true estimates" of deposition exist to test model errors (U.S. EPA 2009).

Another possible explanation for the predicted non-linear relationship between S deposition and ANC may be related to the temporal discontinuity between a portion of the observed ANC data and the modeled CMAQ data. Because the water chemistry data were taken over a 24 year period and the CMAQ data used to represent dry S deposition were taken from 2002, the modeled dry S deposition may not accurately represent conditions at the time of the water chemistry sampling. However, CMAQ data for 2002 likely provide the relative extent to which each catchment has been exposed to elevated S deposition. Furthermore, base cation depletion of soils has limited the ANC recovery of the most acid-sensitive streams (Sullivan et al. 2004, Sullivan et al. 2011). Nevertheless, results here indicate that reductions in S inputs alone may not lead directly to increases in stream water ANC because interactions with others

driving variables, such as temperature and precipitation, confound simple monotonic interpretation of relations between acidic S inputs and stream water ANC.

5. CONCLUSION

Models developed here predict stream water ANC across the southern Appalachian Mountains. Results suggest that aquatic biota may be at risk from the deleterious effects of low ANC (ANC <100 ueq L-1; Sullivan 2000) over approximately one-fourth of the stream network. Acid-sensitive areas have siliceous lithology, cool and moist climates, and forests with low soil clay and pH levels, and relatively small contributing areas. Our findings suggest that predicting future ANC will require incorporation of data from ongoing stream water chemistry monitoring into a modeling framework capable of quantifying the inherently non-linear interactions among relevant biogeochemical and climatic variables. Continued water chemistry monitoring should include previously sampled locations and be expanded to include undersampled geographic areas and environments (e.g., areas with high model uncertainty). Although correlation does not imply causation, the identification of several climate variables as key drivers of ANC suggests that future climate change may influence future stream acid-base chemistry in unique and complex ways. The analysis approach taken here can be readily applied in other regions where adequate coverage of high-resolution, spatially explicit environmental data are available, and where stream water quality surveys have been conducted at a reasonable sampling intensity.

6. REFERENCES

Austin, M. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling **157**:101-118.

Barandela, R., E. Rangel, J. Sánchez, and F. Ferri. 2003. Restricted decontamination for the imbalanced training sample problem. Progress in Pattern Recognition, Speech and Image Analysis:424-431.

Baumgardner, R. E., T. F. Lavery, C. M. Rogers, and S. S. Isil. 2002. Estimates of the atmospheric deposition of sulfur and nitrogen species: Clean Air Status and Trends Network, 1990-2000. Environmental Science & Technology **36**:2614-2629.

Benčoková, A., J. Hruška, and P. Krám. 2011. Modeling anticipated climate change impact on biogeochemical cycles of an acidified headwater catchment. Applied Geochemistry **26, Supplement**:S6-S8.

Bivand, R., L. Anselin, O. Berke, A. Bernat, M. Carvalho, Y. Chun, C. Dormann, S. Dray, R. Halbersma, and N. Lewin-Koh. 2011. spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-31, http://cran.r-project.org/web/packages/spdep/index.html.

Breiman, L. 1984. Classification and regression trees. Chapman & Hall/CRC.

Breiman, L. 2001a. Random forests. Machine learning **45**:5-32.

Breiman, L. 2001b. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical Science **16**:199-231.

Byun, D. and K. L. Schere. 2006. Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. Applied Mechanics Reviews **59**:51-77.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research **16**:321-357.

Christophersen, N. and C. Neal. 1990. Linking hydrological, geochemical, and soil chemical processes on the catchment scale: An interplay between modeling and field work. Water Resources Research **26**:3077-3086.

Clark, J. M., A. Heinemeyer, P. Martin, and S. H. Bottrell. 2012. Processes controlling DOC in pore water during simulated drought cycles in six different UK peats. Biogeochemistry **109**:253-270.

Cosby, B., G. Hornberger, J. Galloway, and R. Wright. 1985. Modeling the effects of acid deposition: Assessment of a lumped parameter model of soil water and streamwater chemistry. Water Resources Research **21**:51-63.

Cosby, B., J. Webb, J. Galloway, and F. Deviney. 2006. Acidic deposition impacts on natural resources in Shenandoah National Park Technical report NPS/NER/NRTR-2006/066, Philadelphia.

Cronan, C. S., W. A. Reiners, R. C. Reynolds, and G. E. Lang. 1978. Forest floor leaching: contributions from mineral, organic, and carbonic acids in New Hampshire subalpine forests. Science **200**:309.

De'ath, G. 2007. Boosted trees for ecological modeling and prediction. Ecology **88**:243-251.

De'ath, G. and K. E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology **81**:3178-3192.

Driscoll, C. T., K. M. Driscoll, K. M. Roy, and M. J. Mitchell. 2003. Chemical response of lakes in the Adirondack region of New York to declines in acidic deposition. Environmental Science & Technology **37**:2036-2042.

Driscoll, C. T., B. J. Wyskowski, C. C. Cosentini, and M. E. Smith. 1987. Processes regulating temporal and longitudinal variations in the chemistry of a low-order woodland stream in the Adirondack Region of New York. Biogeochemistry **3**:225-241.

Duan, L., J. Hao, S. Xie, and K. Du. 2000. Critical loads of acidity for surface waters in China. Science of the Total Environment **246**:1-10.

Dunford, R. W., D. N. M. Donoghue, and T. P. Burt. 2012. Forest land cover continues to exacerbate freshwater acidification despite decline in sulphate emissions. Environmental Pollution **167**:58-69.

Elith, J., J. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology **77**:802-813.

Elith, J. and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics **40**:677-697.

Evans, C. D. 2005. Modelling the effects of climate change on an acidic upland stream. Biogeochemistry **74**:21-46.

Franklin, J. and J. A. Miller. 2009. Mapping species distributions: spatial inference and prediction. Cambridge University Press

Gahegan, M. 2003. Is inductive machine learning just another wild goose (or might it lay the golden egg)? International Journal of Geographical Information Science **17**:69-92.

Galloway, J. 2001. Acidification of the world: Natural and anthropogenic. Water, Air, & Soil Pollution **130**:17-24.

Galloway, J. N., Z. Dianwu, X. Jiling, and G. E. Likens. 1987. Acid Rain: China, United States, and a Remote Area. Science **236**:1559-1562.

Gee, A. and J. Stoner. 1989. A review of the causes and effects of acidification of surface waters in Wales and potential mitigation techniques. Archives of Environmental Contamination and Toxicology **18**:121-130.

Grimm, J. W. and J. A. Lynch. 2004. Enhanced wet deposition estimates using modeled precipitation inputs. Environmental Monitoring and Assessment **90**:243-268.

Guerold, F., J.-P. Boudot, G. Jacquemin, D. Vein, D. Merlet, and J. Rouiller. 2000. Macroinvertebrate community loss as a result of headwater stream acidification in the Vosges Mountains (N-E France). Biodiversity and Conservation **9**:767-783.

Hargrove, W. and F. Hoffman. 2004. A flux atlas for representativeness and statistical extrapolation of the Ameriflux network. Oak Ridge National Laboratory Technical Memorandum **ORNL-TM-2004**:1-152.

Harriman, R. and B. Morrison. 1982. Ecology of streams draining forested and non-forested catchments in an area of central Scotland subject to acid precipitation. Hydrobiologia **88**:251-263.

Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin. 2005. The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer **27**:83-85.

Hayhoe, K., C. Wake, B. Anderson, X. Z. Liang, E. Maurer, J. Zhu, J. Bradbury, A. DeGaetano, A. M. Stoner, and D. Wuebbles. 2008. Regional climate change projections for the Northeast USA. Mitigation and Adaptation Strategies for Global Change **13**:425-436.

He, H. and E. A. Garcia. 2009. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering **21**:1263-1284.

Henriksen, A. and M. Posch. 2001. Steady-state models for calculating critical loads of acidity for surface waters. Water, Air, & Soil Pollution: Focus **1**:375-398.

Henriksen, A., M. Posch, H. Hultberg, and L. Lien. 1995. Critical loads of acidity for surface waters: Can the ANC limit be considered variable? Water, Air, & Soil Pollution **85**:2419-2424.

Herlihy, A. T., P. R. Kaufmann, M. R. Church, P. J. Wigington, Jr., J. R. Webb, and M. J. Sale. 1993. The effects of acidic deposition on streams in the Appalachian Mountain and Piedmont Region of the Mid-Atlantic United States. Water Resources Research **29**:2687-2703.

Herlihy, A. T., J. L. Stoddard, and C. B. Johnson. 1998. The relationship between stream chemistry and watershed land cover data in the mid-Atlantic region, US. Water, Air, & Soil Pollution **105**:377-386.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. VanDriel, and J. Wickham. 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. Photogrammetric Engineering and Remote Sensing **73**:337-341.

Jenkins, A., B. J. Cosby, R. C. Ferrier, T. A. B. Walker, and J. D. Miller. 1990. Modelling stream acidification in afforested catchments: An assessment of the relative effects of acid deposition and afforestation. Journal of Hydrology **120**:163-181.

Jenness, J. S. 2004. Calculating landscape surface area from digital elevation models. Wildlife Society Bulletin **32**:829-839.

Jenson, S. and J. Domingue. 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. Photogrammetric Engineering and Remote Sensing **54**:1593-1600.

Johnson, L. B. and G. E. Host. 2010. Recent developments in landscape approaches for the study of aquatic ecosystems. Journal of the North American Benthological Society **29**:41-66.

Karl, T. R., J. M. Melillo, and T. C. Peterson. 2009. Global climate change impacts in the United States. Cambridge Univ Pr.

Krug, E. C. and C. R. Frink. 1983. Acid rain on acid Soil: A new perspective. Science **221**:520-525.

Kuylenstierna, J. C. I., H. Rodhe, S. Cinderby, and K. Hicks. 2001. Acidification in developing countries: Ecosystem sensitivity and the critical load approach on a global scale. Ambio **30**:20-28.

Lawrence, G. B., M. B. David, G. M. Lovett, P. S. Murdoch, D. A. Burns, J. L. Stoddard, B. P. Baldigo, J. H. Porter, and A. W. Thompson. 1999. Soil calcium status and the response of stream chemistry to changing acidic deposition rates. Ecological Applications **9**:1059-1072.

Levin, S. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology **73**:1943-1967.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R News **2**:18-22.

Lien, L., G. Raddum, and A. Fjellheim. 1992. Critical loads of acidity to freshwater fish and invertebrates. Naturens Talegrenser no **23**.

Lynch, D. D. and N. B. Dise. 1985. Sensitivity of stream basins in Shenandoah National Park to acid deposition. US Dept. of the Interior, Geological Survey, Richmond, VA: US Geological Survey.

Maclure, M. and W. C. Willet. 1987. Misinterpretation and misuse of the Kappa statistic. American Journal of Epidemiology **126**:161-169.

McDonnell, T., B. Cosby, T. Sullivan, S. McNulty, and E. Cohen. 2010. Comparison among model estimates of critical loads of acidic deposition using different sources and scales of input data. Environmental Pollution **158**:2934-2939.

McDonnell, T. C., B. J. Cosby, and T. J. Sullivan. 2012. Regionalization of soil base cation weathering for evaluating stream water acidification in the Appalachian Mountains, USA. Environmental Pollution **162**:338-344.

McDonnell, T. C., T. J. Sullivan, P. F. Hessburg, K. Reynolds, N. A. Povak, B. J. Cosby, W. Jackson, and R. B. Salter. in review. Critical Loads of sulfur deposition for aquatic resource protection in the southern Appalachian Mountains. Water Resources Research.

Menz, F. C. and H. M. Seip. 2004. Acid rain in Europe and the United States: An update. Environmental Science and Policy **7**:253-265.

Miles, J. 1986. What are the effects of trees on soils? Pages 55-62 *in* D. Jenkins, editor. Trees and wildlife in the Scottish uplands. NERC/ITE.

Moore, I. D., P. E. Gessler, G. A. Nielsen, and G. A. Peterson. 1993. Soil Attribute Prediction Using Terrain Analysis. Soil Science Society America Journal **57**:NP-NP.

National Atlas of the United States. 2006. Federal lands of the United States: National atlas of the United States, Reston, VA. National Atlas of the United States.

Neal, C., B. Reynolds, and A. J. Robson. 1999. Acid neutralisation capacity measurements within natural waters: towards a standardised approach. Science of the Total Environment **243–244**:233-241.

Nihlgård, B. 1970. Precipitation, its chemical composition and effect on soil water in a beech and a spruce forest in south Sweden. Oikos:208-217.

Nilsson, J. and P. Grennfelt. 1988. Critical loads for sulfur and nitrogen. Nordic Council of Ministers, Copenhagen, Denmark:31.

NRCS Soil Survey Staff. 2010a. Soil Survey Geographic (SSURGO) database for southern Appalachian Region.

NRCS Soil Survey Staff. 2010b. U.S. General Soil Map State Soil Geographic (STATSGO) database.

Olden, J. D., J. J. Lawler, and N. Poff. 2008. Machine learning methods without tears: A primer for ecologists. Quarterly Review of Biology **83**:171-194.

Omernik, J. M. 1987. Ecoregions of the Conterminous United States. Annals of the Association of American Geographers **77**:118 - 125.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems **9**:181-199.

Puckett, L. J. and O. P. Bricker. 1992. Factors controlling the major ion chemistry of streams in the blue ridge and valley and ridge physiographic provinces of Virginia and Maryland. Hydrological Processes **6**:79-97.

R Development Core Team. 2011. R: A language and environment for statistical computing. Vienna, Austria.

Rago, P. J. and J. G. Wiener. 1986. Does pH affect fish species richness when lake area is considered? Transactions of the American Fisheries Society **115**:438-447.

Reuss, J. O., B. J. Cosby, and R. F. Wright. 1987. Chemical processes governing soil and water acidification. Nature **329**:27-32.

Reynolds, K. M., P. F. Hessburg, T. J. Sullivan, N. A. Povak, T. C. McDonnell, B. J. Cosby, and W. Jackson. 2012. Spatial decision support for assessing impacts of atmospheric sulfur deposition on aquatic ecosystems in the southern Appalachian Region. Pages 1197-1206 *in* Proceedings of the 45th annual Hawaii International Conference on System Sciences. IEEE, Wailea, Maui, Hawaii.

Ridgeway, G. 2006. Generalized boosted regression models. Documentation on the R Package 'gbm', version 1· 5 **7**.

Schindler, D. W. 1988. Effects of acid rain on freshwater ecosystems. Science **239**:149-157.

Schöpp, W., M. Posch, S. Mylona, and M. Johansson. 2003. Long-term development of acid deposition (1880–2030) in sensitive freshwater regions in Europe. Hydrol. Earth Syst. Sci. **7**:436-446.

Solomon, S. 2007. Climate Change 2007: the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ Pr.

Steel, E. A., R. M. Hughes, A. H. Fullerton, S. Schmutz, J. A. Young, M. Fukushima, S. Muhar, M. Poppe, B. E. Feist, and C. Trautwein. 2010. Are we meeting the challenges of landscape-scale riverine research? A review. Living Review of Landscape Research **4**.

Sullivan, T., B. Cosby, A. Herlihy, C. Driscoll, I. Fernandez, T. C. McDonnell, C. W. Boylen, S. Nierzwicki-Bauer, and K. Snyder. 2007a. Assessment of the extent to which intensively studied lakes are representative of the Adirondack Mountain region. New York State Energy Research and Development Authority.

Sullivan, T., J. Webb, K. Snyder, A. Herlihy, and B. Cosby. 2007b. Spatial distribution of acid-sensitive and acid-impacted streams in relation to watershed features in the southern Appalachian Mountains. Water, Air, & Soil Pollution **182**:57-71.

Sullivan, T. J. 2000. Aquatic effects of acidic deposition. CRC Press, Boca Raton, FL.

Sullivan, T. J., D. F. Charles, J. A. Bernert, B. McMartin, K. B. Vaché, and J. Zehr. 1999. Relationship between landscape characteristics, history, and lakewater acidification in the Adirondack Mountains, New York. Water, Air, & Soil Pollution **112**:407-427.

Sullivan, T. J., B. J. Cosby, A. T. Herlihy, J. R. Webb, A. J. Bulger, K. U. Snyder, P. F. Brewer, E. H. Gilbert, and D. L. Moore. 2004. Regional model projections of future effects of sulfur and nitrogen deposition on streams in the southern Appalachian Mountains. Water Resources Research **40**:W02101.

Sullivan, T. J., B. J. Cosby, W. A. Jackson, K. U. Snyder, and A. T. Herlihy. 2011. Acidification and prognosis for future recovery of acid-sensitive streams in the southern Blue Ridge Province. Water, Air, & Soil Pollution **219**:11-26.

Sullivan, T. J., B. J. Cosby, J. R. Webb, R. L. Dennis, A. J. Bulger, and F. A. Deviney. 2008. Streamwater acid-base chemistry and critical loads of atmospheric sulfur deposition in Shenandoah National Park, Virginia. Environmental Monitoring and Assessment **137**:85-99.

Turner, M. 1989. Landscape ecology: the effect of pattern on process. Annual Review of Ecology and Systematics **20**:171-197.

U.S. EPA. 2009. Risk and exposure assessment for review of the secondary National Ambient Air Quality Standards for Oxides of Nitrogen and Oxides of Sulfur: Final.*in* U. S. EPA, editor. Center for Environmental Assessment, Office of Research and Development, Research Triangle Park, NC.

U.S. EPA and U.S.G.S. 2005. National Hydrography Dataset Plus - NHDPlus Version 1.0.

USGS. 2005a. Preliminary integrated geologic map databases for the United States : Delaware, Maryland, New York, Pennsylvania, and Virginia. Open-File Report **2005**.

USGS. 2005b. Preliminary integrated geologic map databases for the United States : Kentucky, Ohio, Tennessee, and West Virginia. Open-File Report **2005**.

Whitehead, P. G., B. Reynolds, M. Hornung, C. Neal, J. Cosby, and P. Paricos. 1988. Modelling long term stream acidification trends in upland wales at plynlimon. Hydrological Processes **2**:357-368.

Wood, E. F., M. Sivapalan, and K. Beven. 1990. Similarity and scale in catchment storm response. Reviews of Geophysics **28**:1-18.

Wright, R. F., J. Aherne, K. Bishop, L. Camarero, B. J. Cosby, M. Erlandsson, C. D. Evans, M. Forsius, D. W. Hardekopf, R. Helliwell, J. Hruška, A. Jenkins, J. Kopáček, F. Moldan, M. Posch, and M. Rogora. 2006. Modelling the effect of climate change on recovery of acidified freshwaters: Relative sensitivity of individual processes in the MAGIC model. Science of the Total Environment **365**:154-166.

TABLES

Table 2.1. Predictor variables included in the threshold model (left) and continuous model (right) within the hurdle modeling framework. GS = growing season. Relative importance for a single predictor variable is based on the mean increase in mean squared error (regression) or model accuracy (classification) for each decision tree when the values of the predictor variable are randomized during model calibration. The values have been standardized to sum to 100%. See Appendix 2.1 for a description of the predictor variables.

Threshold model variables	**Short name**	**Relative Importance**	**Continuous model variables**	**Short name**	**Relative Importance**
Percent carbonate lithology	LITH_CAR	20.23	Percent siliceous lithology	LITH_SIL	18.64
Soil pH	SOIL_PH	12.20	Mean penultimate maximum days without precipitation while $\geq 10^{o}C$	NPDAYMAX	11.98
Mean penultimate maximum days VPD <1000 pa while $\geq 10^{o}C$	VWDAYMAX	12.08	Mean number of GS days above $32.2^{o}C$	AB90GROW	11.51
Mean 95% of maximum GS temperature difference	DIFF95GR	10.68	Mean penultimate maximum days vapor pressure deficit<1000 pa while $\geq 10^{o}C$	VWDAYMAX	9.24
Percent land in public ownership	PUBLIC	10.60	Dry sulfur deposition	S_DRY	9.13
Percent forest cover	FOREST	8.86	Percent forest cover	FOREST	8.79
Percent argillic lithology	LITH_ARG	7.14	Percent soil clay	SOIL_CLAY	8.42
Mean precipitation sum during the local non-GS	PRECIPNG	6.83	Soil pH	SOIL_PH	7.99
Mean penultimate maximum consecutive days VPD > 750 pa while $\geq 10^{o}C$	VDCONTDAY	6.17	Topographic wetness index	TWI	7.38
Mean non-GS gross primary productivity	GPPNG	5.20	Flow accumulation	FAC	6.92

Table 2.2. Results from a Moran's I correlogram on model residuals from the hurdle model. Values represented in bold indicate significant autocorrelation.

Lag	Estimate	Expected	Variance	SD	P
1	-0.001	-0.001	0.000	-0.014	0.989
2	0.004	-0.001	0.000	0.815	0.415
3	-0.007	-0.001	0.000	-1.058	0.290
4	-0.002	-0.001	0.000	-0.130	0.897
5	**-0.016**	**-0.001**	**0.000**	**-2.116**	**0.034**
6	-0.001	-0.001	0.000	0.099	0.921

FIGURES

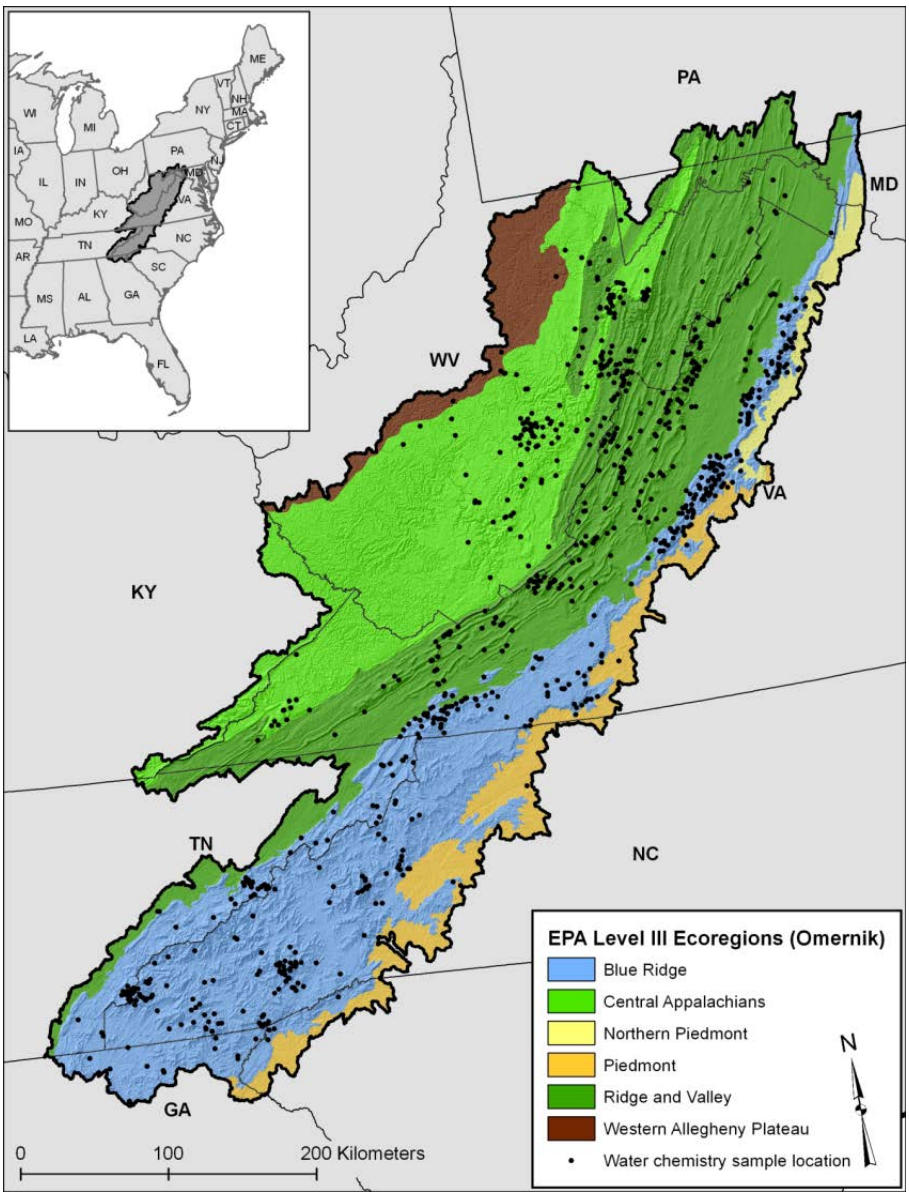


Figure 2.1. Distribution of water chemistry samples within Omernik (1987) ecoregions for the southern Appalachian Mountain region.

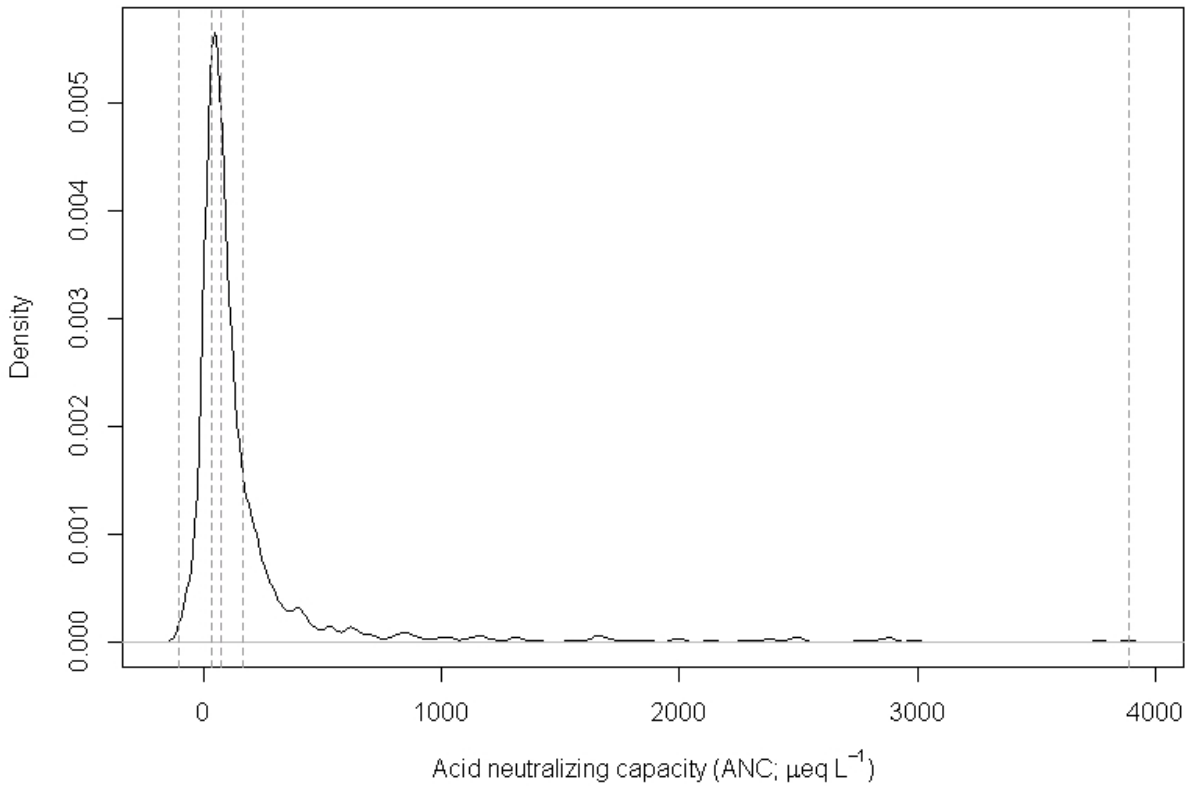


Figure 2.2. A plot of the kernel density function estimated for the 933 sampled ANC values across the southern Appalachian Mountain region. Grey vertical lines indicate minimum, first quantile, median, second quantile and maximum ANC values.

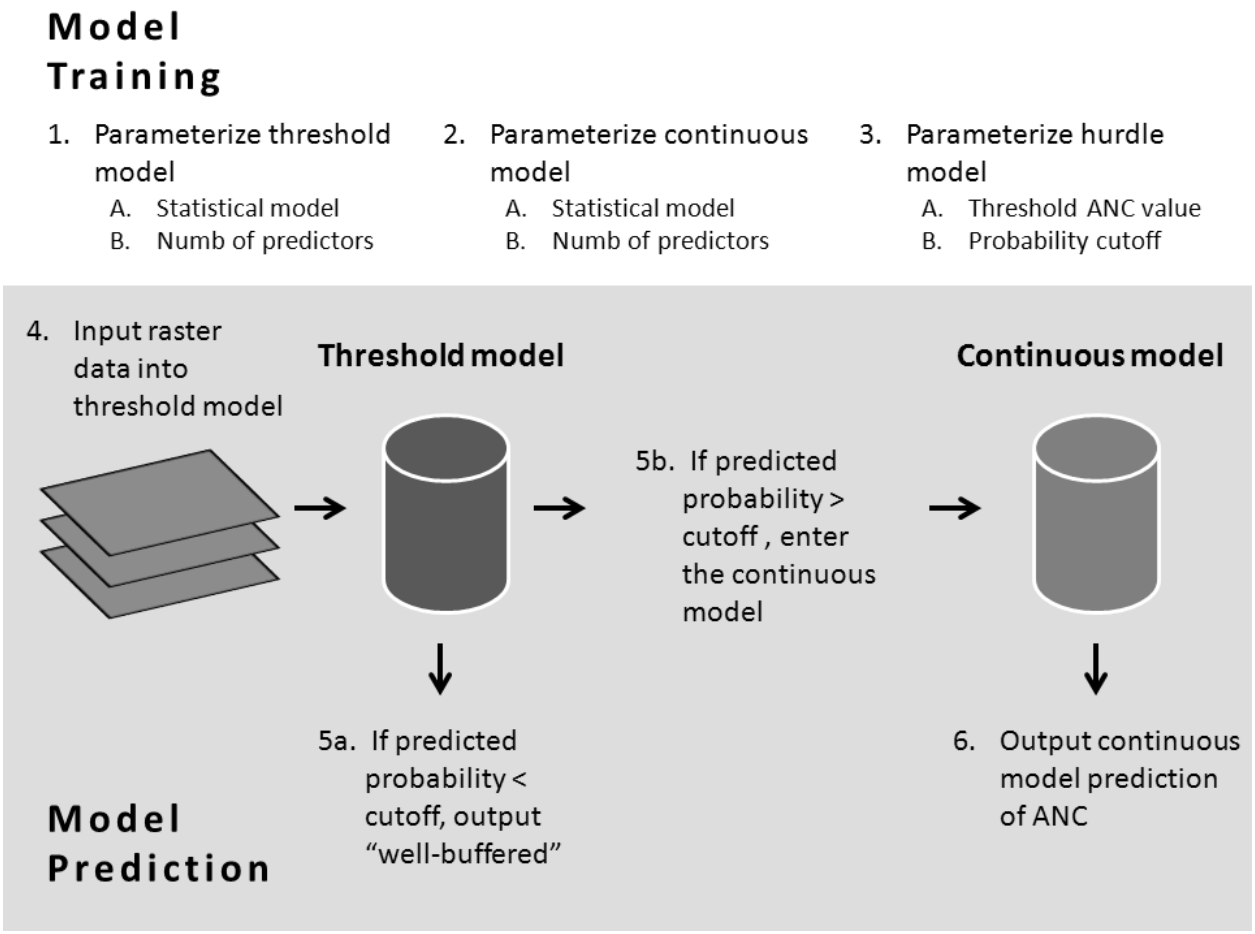


Figure 2.3. Conceptual diagram of the hurdle modeling framework for predicting ANC values in the southern Appalachian Mountain region.

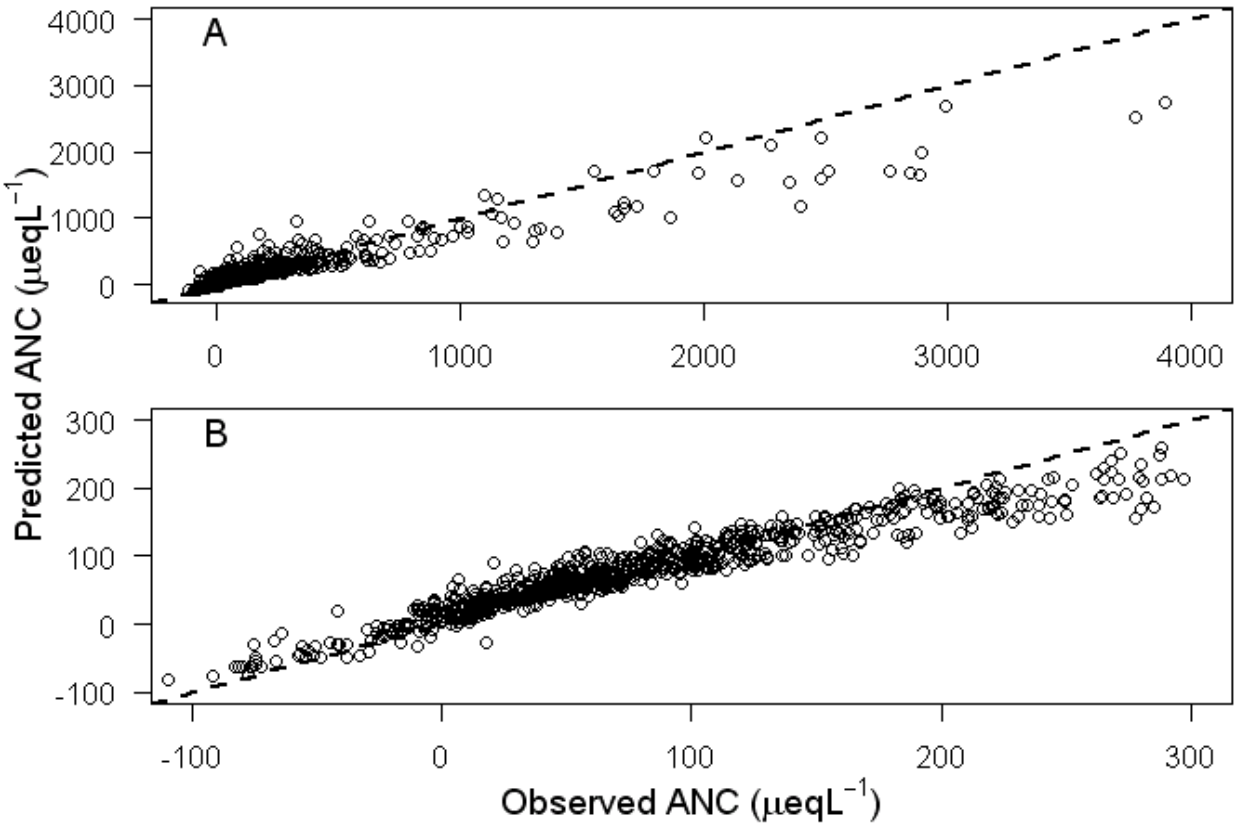


Figure 2.4. Predicted versus observed ANC values resulting from (A) the random forest-only model (i.e., single random forest regression model) and (B) the final hurdle model. Unlike Figure 2.8, where models were developed using a testing/training set, models here were trained and predicted using all 933 sample sites. Note the difference in x- and y-axis scaling between the top and bottom panels.

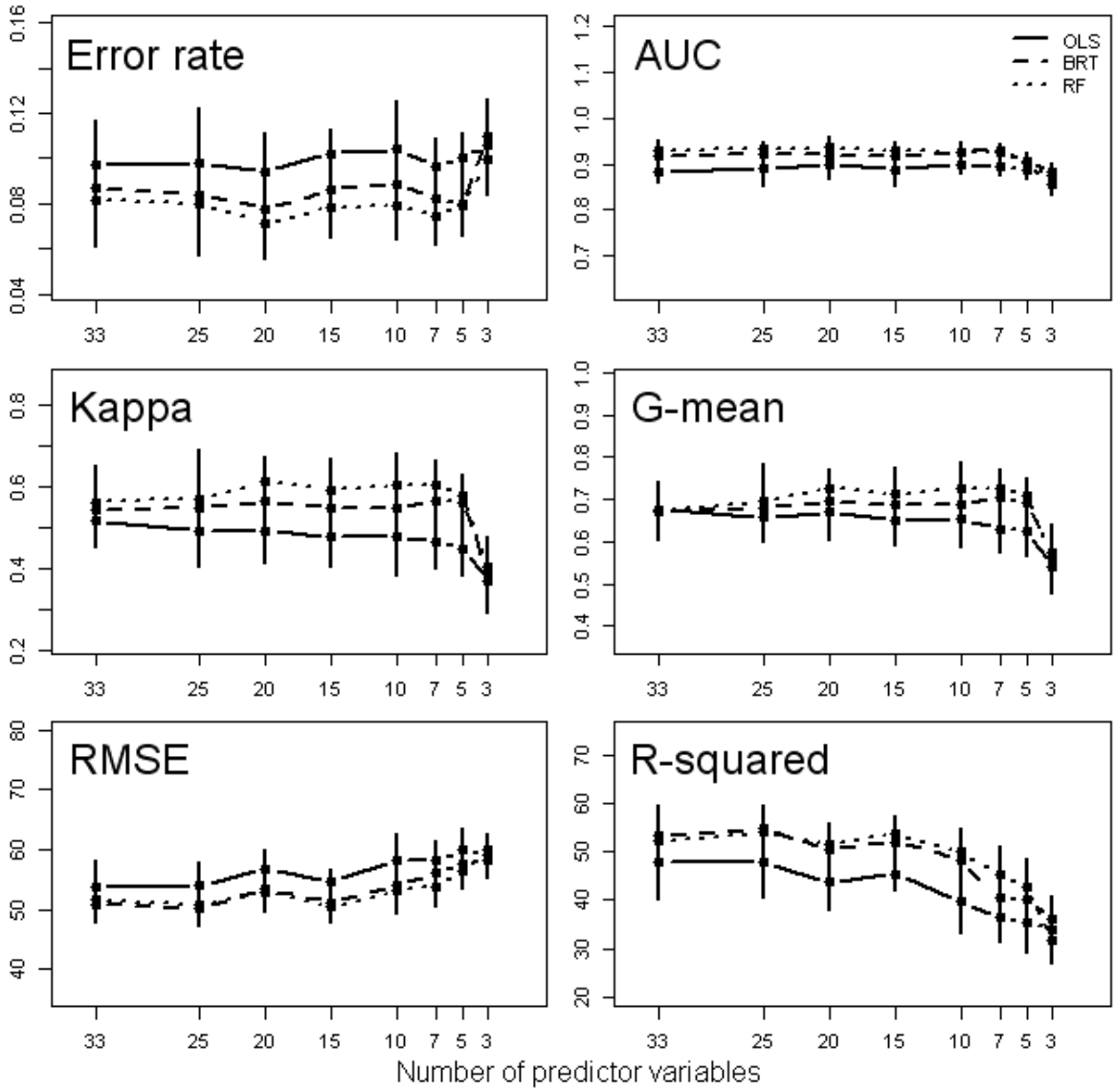


Figure 2.5. Validation results for threshold (top four panels) and continuous models (lower two panels) of the hurdle modeling framework. Comparisons are shown for the ordinary least squares (LM; continuous model only), boosted regression tree (BRT), and random forest (RF) models. Abbreviations are: area under the receiver operator curve (AUC) and root mean squared error (RMSE). Each unique combination of models was run 15 times to obtain estimates of variability due to the collection of training data unique to each run. Error bars represent the (+/-) standard deviation of the error estimates. Variables included in each model were chosen using the top predictor variables based on the average variable importance measures from the BRT/BCT and RF models.

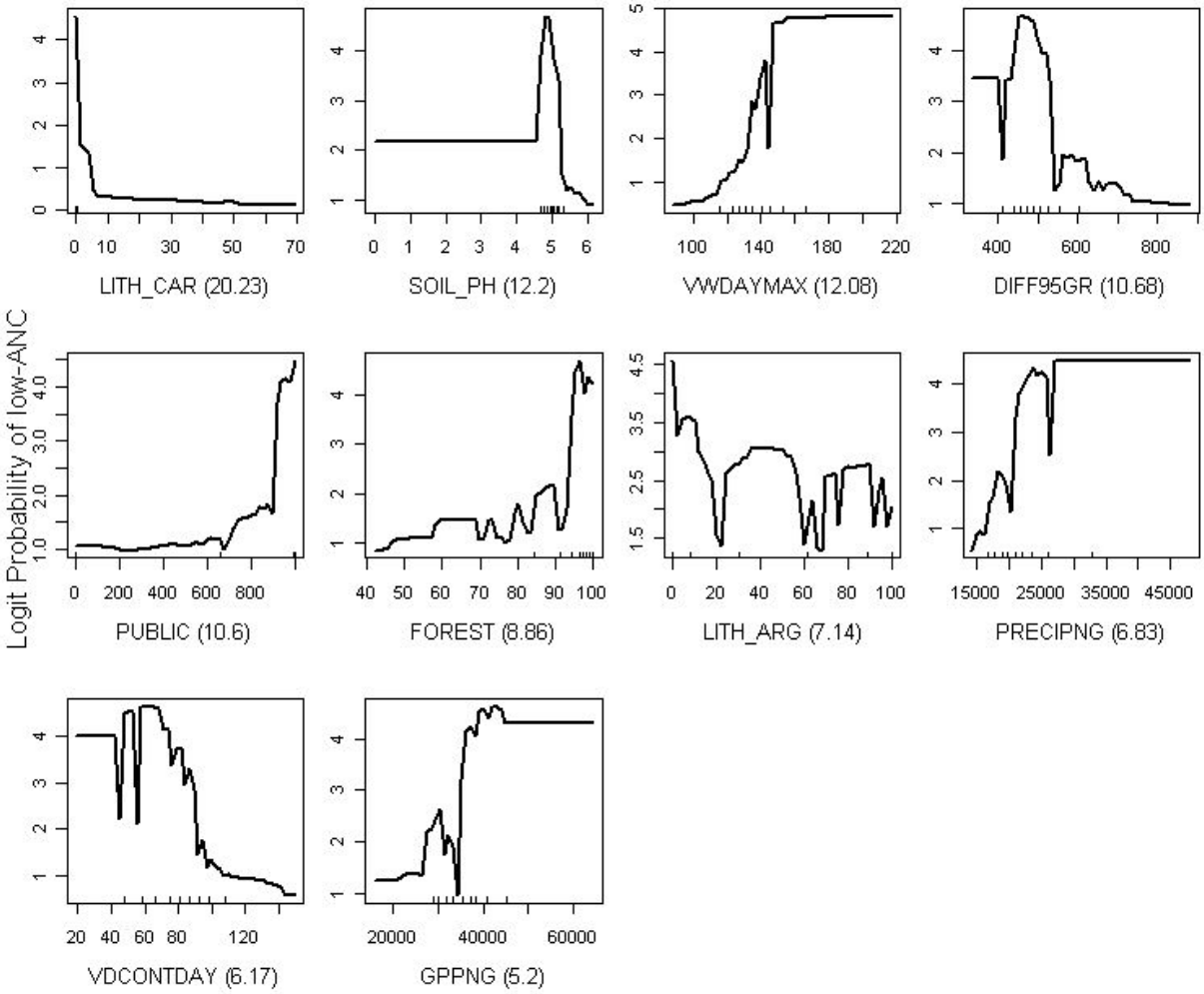


Figure 2.6. Response curves showing relations between the predicted ANC and individual predictor variables included in the threshold model within the hurdle modeling framework. Black tick marks on the x-axis indicate decile classes for the predictors. The y-axis indicates the relative effect of the predictor on ANC on a logit scale. In general, higher y-axis values indicate a higher probability of predicting a low-ANC value. See Appendix 2.1 for a description of the predictors.

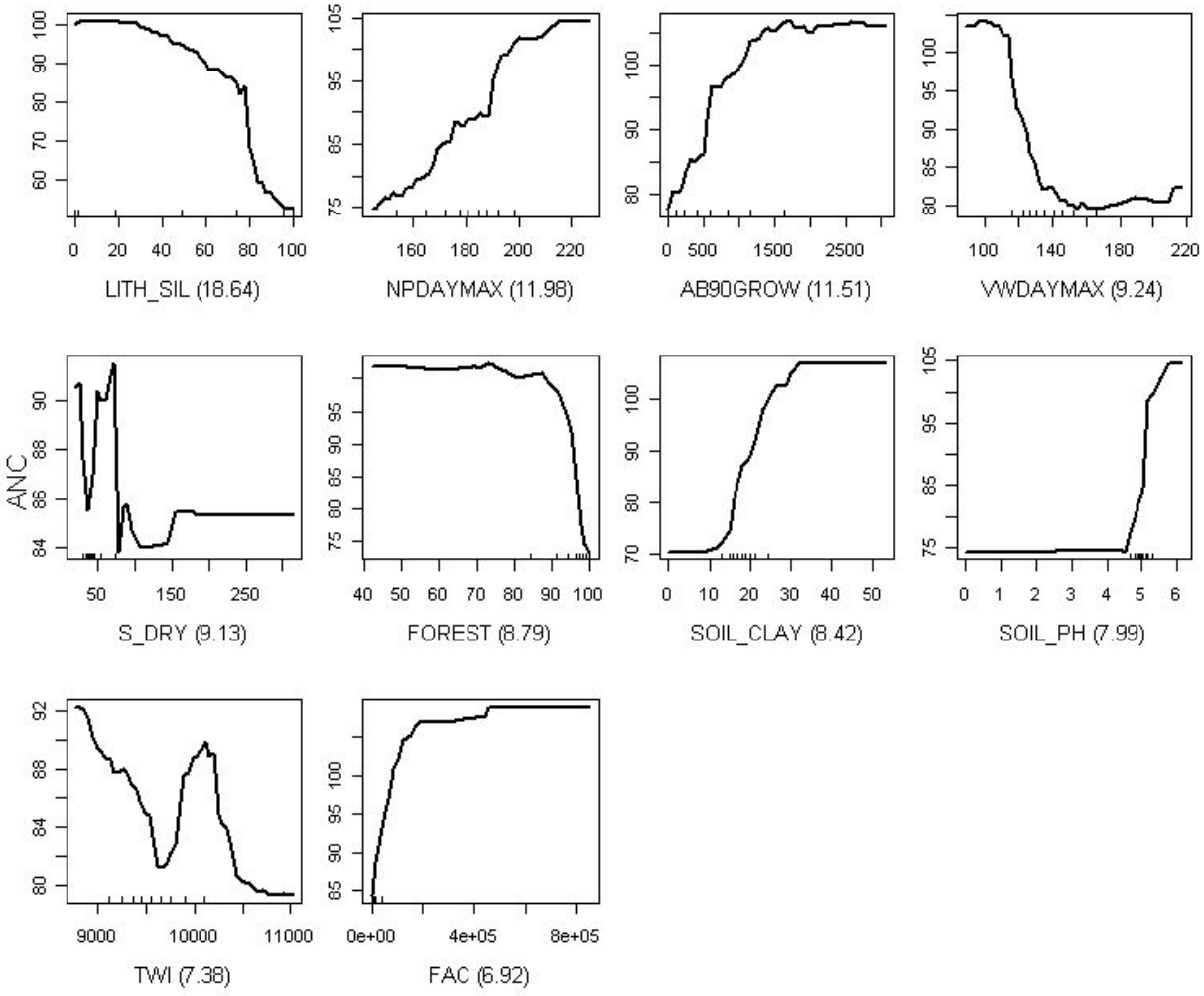


Figure 2.7. Response curves showing relations between the predicted ANC and individual predictor variables included in the continuous model within the hurdle modeling framework. Black tick marks on x-axis indicate decile classes for the predictors. The y-axis indicates the relative effect of the predictors on ANC. In general, lower y-axis values indicate lower ANC values. Note that the majority of climate variables have been multiplied by a constant, which had no effect on predictions or model performance. See Appendix 2.1 for a description of the predictor variables.

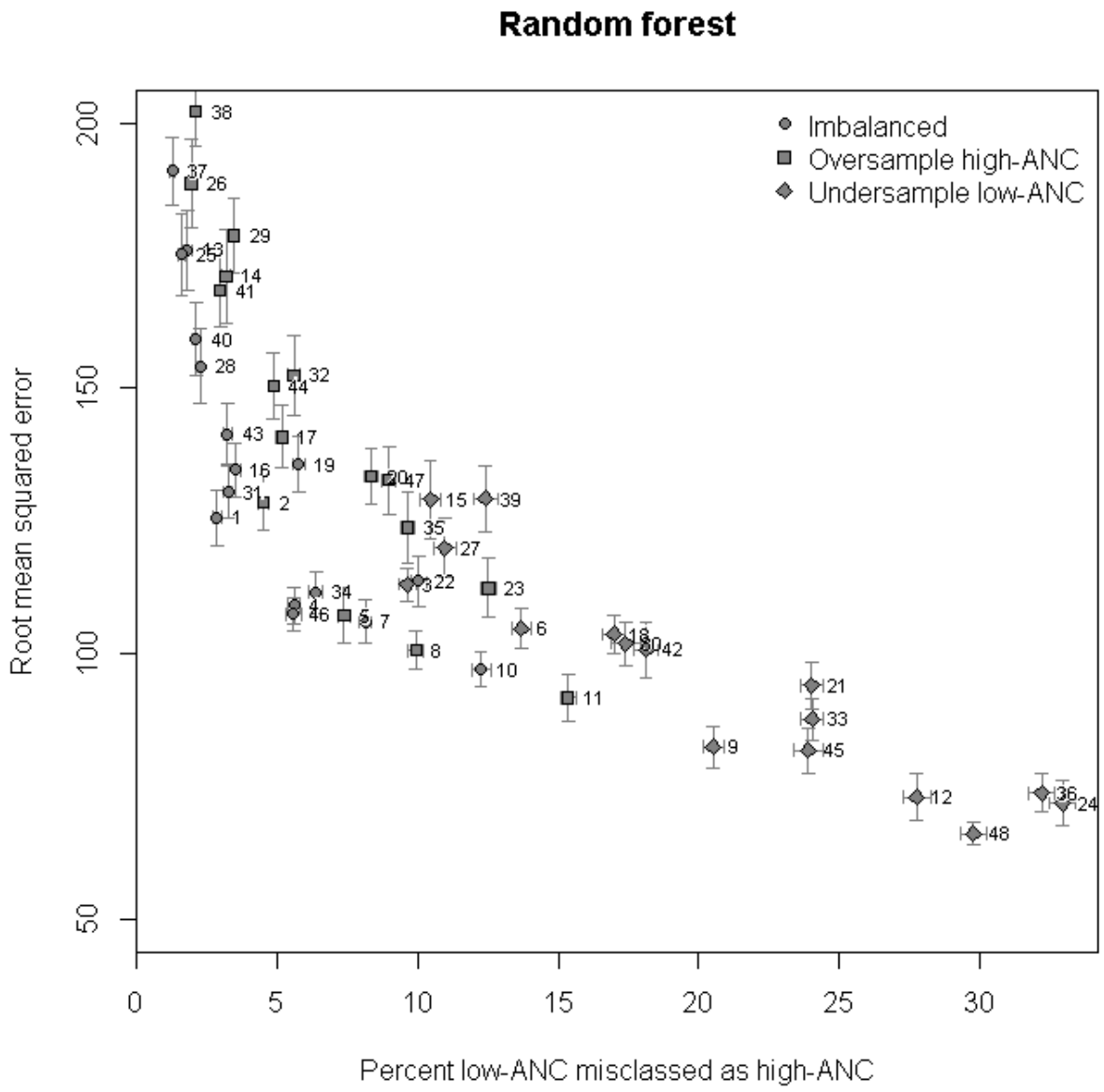


Figure 2.8. Scatterplot of the performance of 48 random forest continuous models varying by combination of data resampling method (3), ANC threshold (4), and probability threshold (4). Validation statistics are based on predictions to a randomly drawn 25% subset of the data. A complete list of model results is included in Appendix 2.6. Model number 46, which showed the lowest RMSE and percent misclassified, was used as the final model.

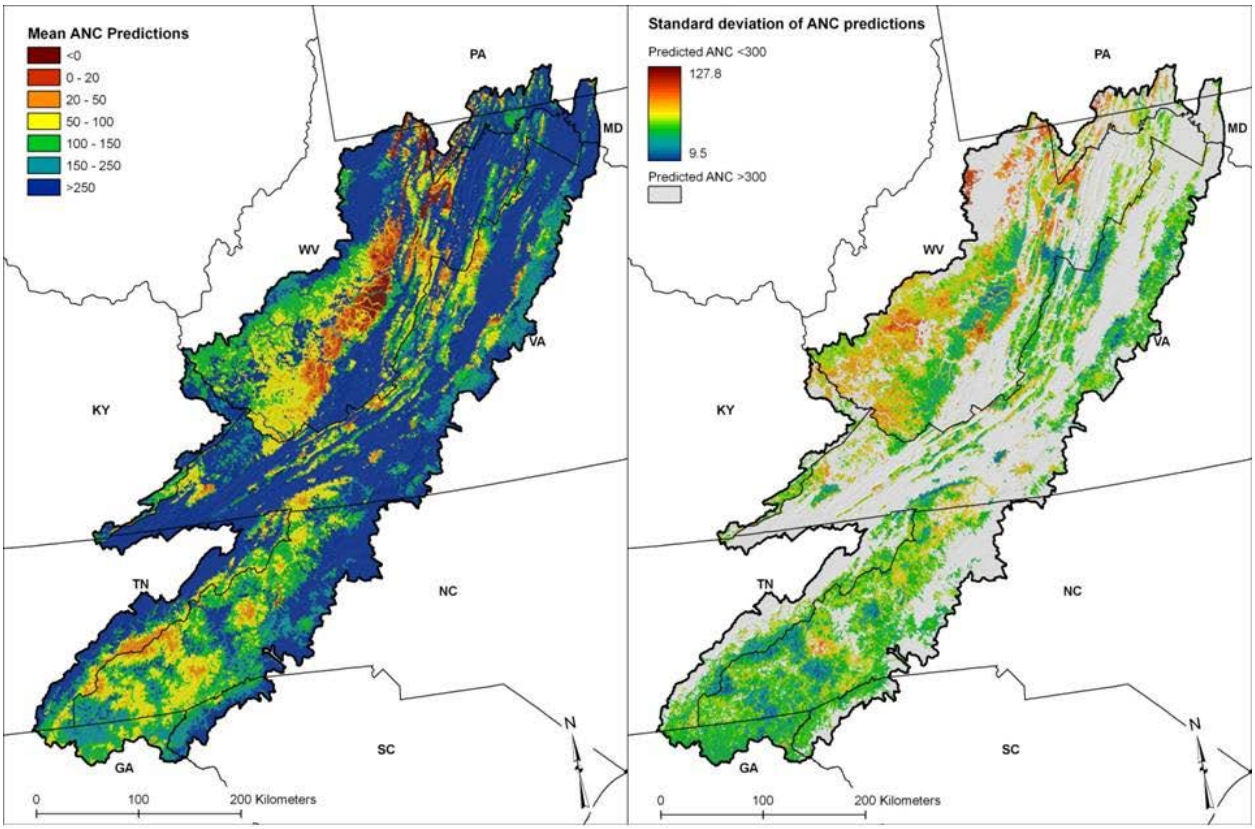


Figure 2.9. ANC predictions from the best hurdle model, number 46 (left panel, see also Figure 2.8), and the standard deviation of predictions made by the continuous model within the final hurdle model (right panel). Standard deviations were calculated from the predictions made from the ensemble of 1000 individual regression trees that made up the continuous random forest model. Areas shown in white (filtered out by the threshold model) were predicted to exhibit ANC values >300 $\mu eq \cdot L^{-1}$ and were not submitted to the continuous model).

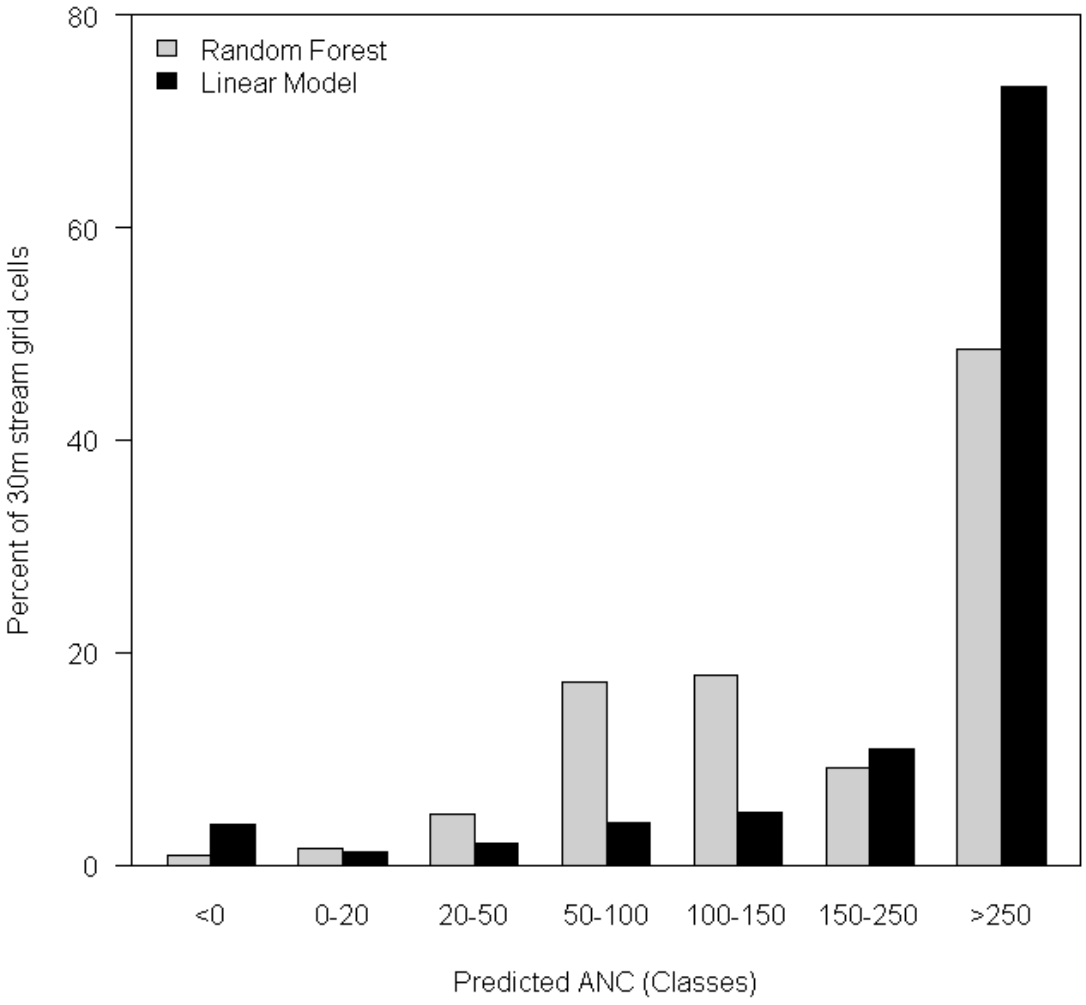


Figure 2.10. Histogram of the percentage of 30-m stream network grid cells predicted to be within seven ANC (μeq L^{-1}) classes. Dark shaded bars represent the predictions made from the linear regression model, and the lighter bars are predictions made from the random forest hurdle model.

APPENDICES

Appendix 2.1. The complete list of potential predictor variables used in the modeling of stream water acid neutralizing capacity in the southern Appalachian region, US.

Variable Type	Variable shortname	Description	Multiplic-ative constant	Units	Citation
Climate	AB90GROW	Mean number of days above 90°F during the local growing season	0.01	days	Hargrove and Hoffman (2004)
Climate	BE32INTNG	Mean number of days below 32°F during the local non-growing season	0.01	days	Hargrove and Hoffman (2004)
Climate	DDAYHMAX	Mean penultimate maximum degrees x days heating > 18°C	none	Degrees Celsius	Hargrove and Hoffman (2004)
Climate	DIFF95GR	Mean 95th percentile of maximum diurnal surface temperature difference during the local growing season	0.02	Degrees Kelvin	Hargrove and Hoffman (2004)
Climate	DIFF95NG	Mean 95th percentile of maximum diurnal surface temperature difference during the local non-growing season	0.02	Degrees Kelvin	Hargrove and Hoffman (2004)
Climate	LSTGROW	Mean degree-days heat sum above 42°F from daytime land surface temperature during the local growing season	0.02	Degrees Kelvin x days	Hargrove and Hoffman (2004)
Climate	LSTPMNG	Mean degree-days cold sum below 42°F from nighttime land surface temperature	0.02	Degrees Kelvin	Hargrove and Hoffman (2004)

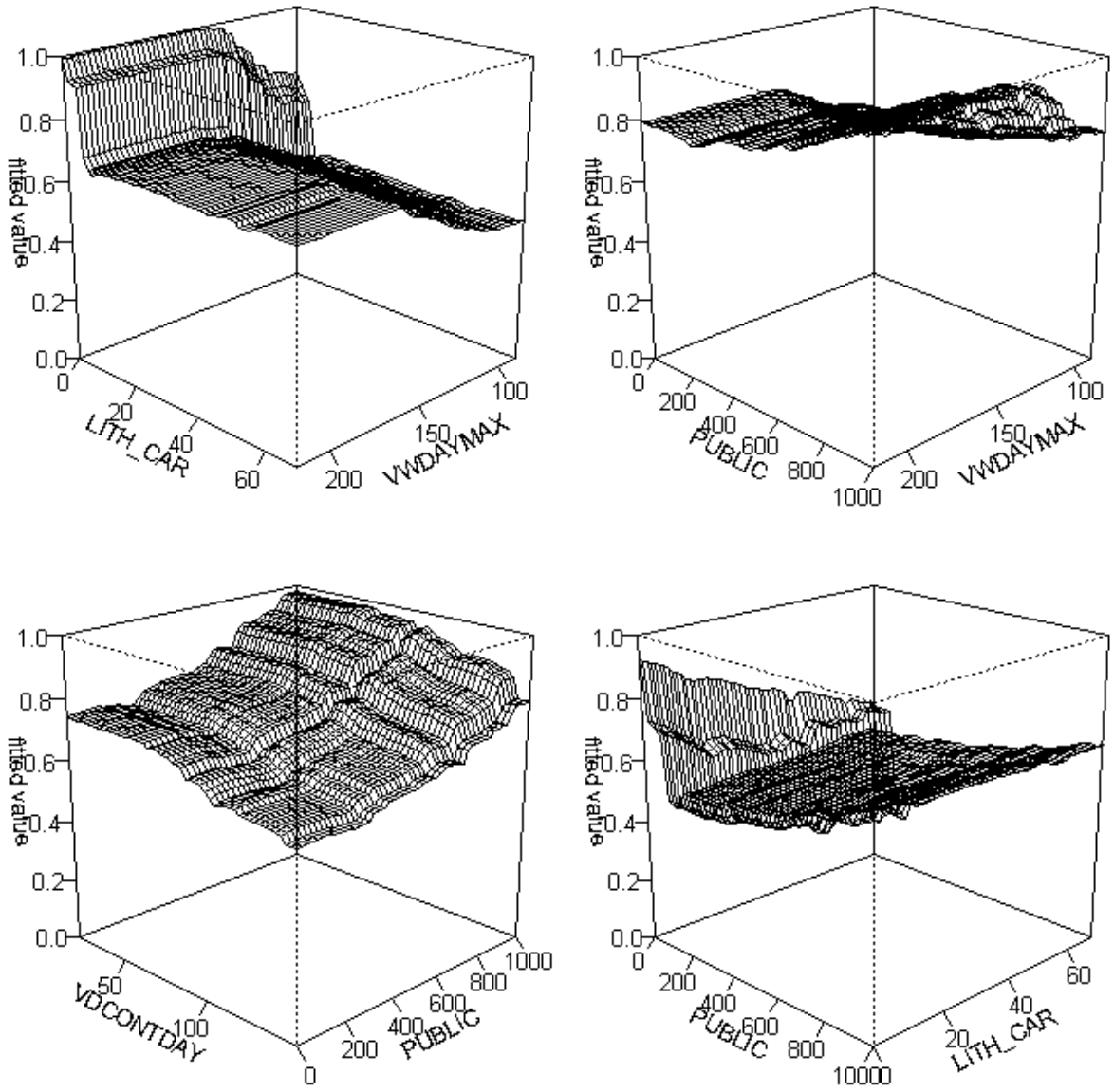
		during the local non-growing season		x days	
Climate	NPCONTDAY	Mean penultimate maximum consecutive days without precipitation (<0.3cm) while > 10°C	none	days	Hargrove and Hoffman (2004)
Climate	NPDAYMAX	Mean penultimate maximum days without precipitation (<0.3cm) while > 10°C	none	days	Hargrove and Hoffman (2004)
Climate	PDAYMAX	Mean penultimate maximum days with precipitation (>0.3cm) while > 10°C	none	days	Hargrove and Hoffman (2004)
Climate	PRECIPNG	Mean precipitation sum during the local non-growing season	0.01	millimeters	Hargrove and Hoffman (2004)
Climate	PSUMMAX	Mean penultimate maximum precipitation sum while ≥ 10°C	0.01	millimeters	Hargrove and Hoffman (2004)
Climate	SOLARTOTGR	Mean total solar insulation during the local growing season, including clouds, aerosols, slope and aspect physiography	none	$kWh \cdot m^{-2}$	Hargrove and Hoffman (2004)
Climate	SOLARTOTNG	Mean total solar insulation during the local non-growing season, including clouds, aerosols, slope and aspect physiography	none	$kWh \cdot m^{-2}$	Hargrove and Hoffman (2004)
Climate	TDAYMEAN	Mean mean daytime temperature	none	Degrees Celsius	Hargrove and Hoffman (2004)
Climate	VDCONTDAY	Mean penultimate maximum consecutive days	none	days	Hargrove and Hoffman

		vpd > 750 pa while > 10°C			(2004)
Climate	VDDAYMAX	Mean penultimate maximum days vpd > 750 pa while > 10C	none	days	Hargrove and Hoffman (2004)
Climate	VWDAYMAX	Mean penultimate maximum days vpd < 1000 pa while > 10°C	none	days	Hargrove and Hoffman (2004)
Deposition	S_DRY	Mean dry sulfur deposition	none	meq/ m^2/yr	Byun and Schere (2006; CMAQ)
Deposition	S_WET	Mean wet sulfur deposition	none	meq/ m^2/yr	Byun and Schere (2006; CMAQ)
Deposition	S_TOT	Mean total sulfur deposition	none	meq·m^{-2}·yr	Byun and Schere (2006; CMAQ)
Geology	LITH_ARG	Percentage contributing area in argillic lithology	none	percent	Sullivan et al. (2007)
Geology	LITH_CAR	Percentage contributing area in carbonaceous lithology	none	percent	Sullivan et al. (2007)
Geology	LITH_FEL	Percentage contributing area in felsic lithology	none	percent	Sullivan et al. (2007)
Geology	LITH_MAF	Percentage contributing area in mafic lithology	none	percent	Sullivan et al. (2007)
Geology	LITH_SIL	Percentage contributing area in siliceous lithology	none	percent	Sullivan et al. (2007)
Productivity	EVIINTGR	Mean enhanced vegetation index (EVI) integrated over the local growing season	0.0001	Unitles s x days	Hargrove and Hoffman (2004)
Productivity	EVIINTNGOW	Mean enhanced vegetation index (EVI) integrated over the local non-growing	0.0001	Unitles s x days	Hargrove and Hoffman (2004)

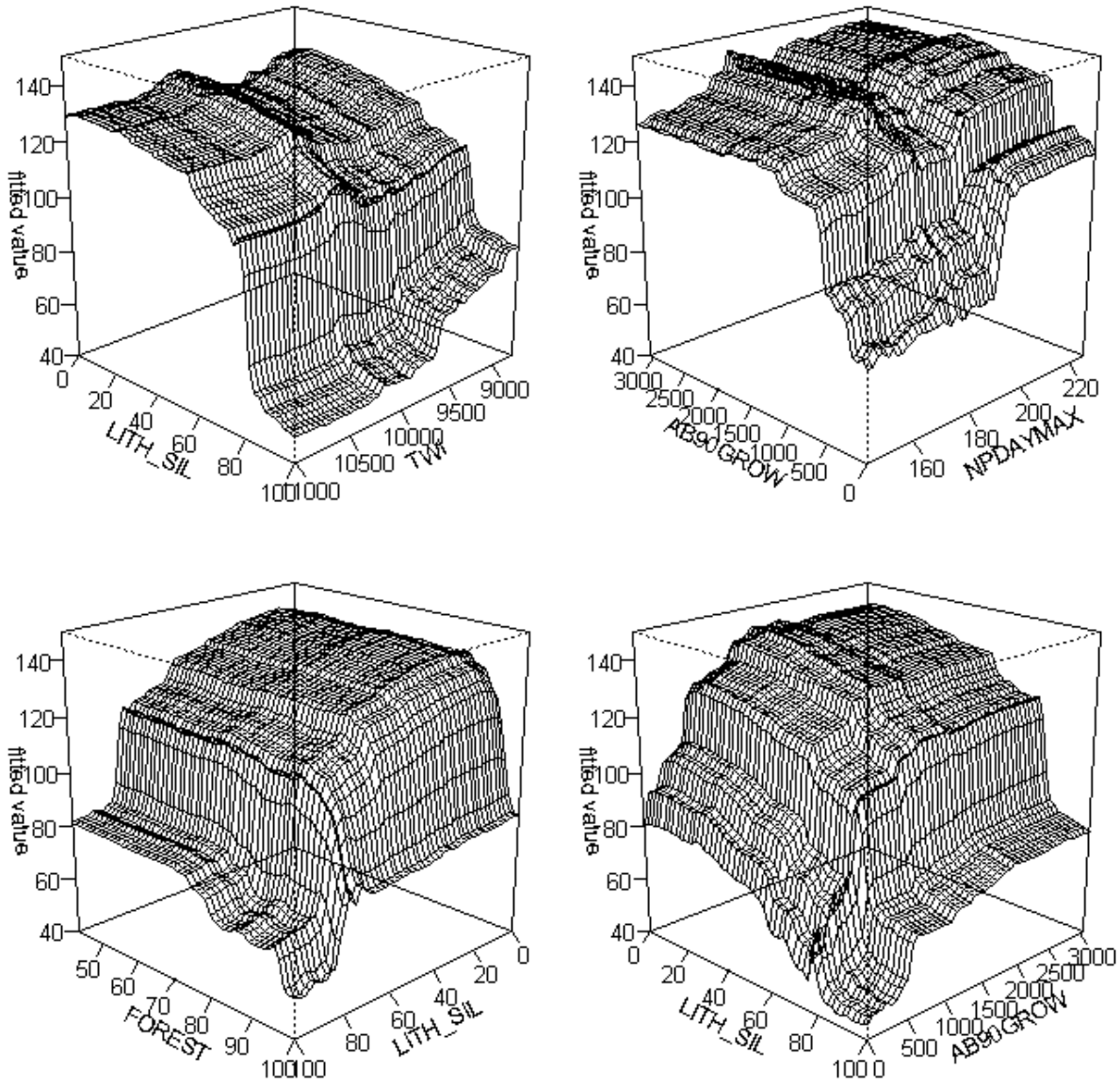
		season			
Productivity	FPARGROW	Mean fraction of photosynthetically active radiation (FPAR) absorbed by vegetation integrated over the local growing season	0.01	Percent x days	Hargrove and Hoffman (2004)
Productivity	FPARNG	Mean fraction of photosynthetically active radiation (FPAR) absorbed by vegetation integrated over the local non-growing season	0.01	Percent x days	Hargrove and Hoffman (2004)
Productivity	GPPGROW	Mean gross primary production (GPP) integrated over the local growing season	0.0001	(kg C/m^2/8 days) x days	Hargrove and Hoffman (2004)
Productivity	GPPNG	Mean gross primary production (GPP) integrated over the local non-growing season	0.0001	(kg C/m^2/8 days) x days	Hargrove and Hoffman (2004)
Productivity	LAIGROW	Mean leaf area index (LAI) integrated over the local growing season	0.1	(m^2/m^2) x days	Hargrove and Hoffman (2004)
Productivity	LAING	Mean leaf area index (LAI) integrated over the local non-growing season	0.1	(m^2/m^2) x days	Hargrove and Hoffman (2004)
Productivity	RIGROW	Mean respiration index (RI) integrated over the local growing season	0.0001	(kg C/m^2/8 days) x days	Hargrove and Hoffman (2004)
Productivity	RING	Mean respiration index (RI) integrated over the local non-growing season	0.0001	(kg C/ m^2/8 days) x days	Hargrove and Hoffman (2004)

Soils	AWCNEW	Mean soil plant-available water holding capacity to 1.5 m	none	percent by weight	Hargrove and Hoffman (2004)
Soils	MDEPTHNEW	Mean depth of mineral soil	none	cm	Hargrove and Hoffman (2004)
Soils	NITRONEW	Mean soil Kjeldahl nitrogen to 50 cm depth	none	kg N/ha	Hargrove and Hoffman (2004)
Soils	OMNEW	Mean soil organic matter to 50 cm depth	0.01	kg OM/ha	Hargrove and Hoffman (2004)
Soils	WATERD	Mean depth to water table	0.1	feet	Hargrove and Hoffman (2004)
Soils	SOIL_CLAY	Mean soil percentage clay	none	%	Soil Survey Staff NRCS (2010a)
Soils	SOIL_PH	Mean soil pH	none	pH	Soil Survey Staff NRCS (2010a)
Soils	SOIL_DEP	Mean soil depth	none	cm	Soil Survey Staff NRCS (2010a)
Topography	NEWCTI	Mean compound topographic index (CTI)	0.01	unitless	Hargrove and Hoffman (2004)
Topography	FAC	Flow accumulation	none	number of pixels	30-m DEM
Topography	ELEVRAW	Elevation at the sample point	none	meters	30-m DEM
Topography	ELEV	Mean elevation	none	meters	30-m DEM

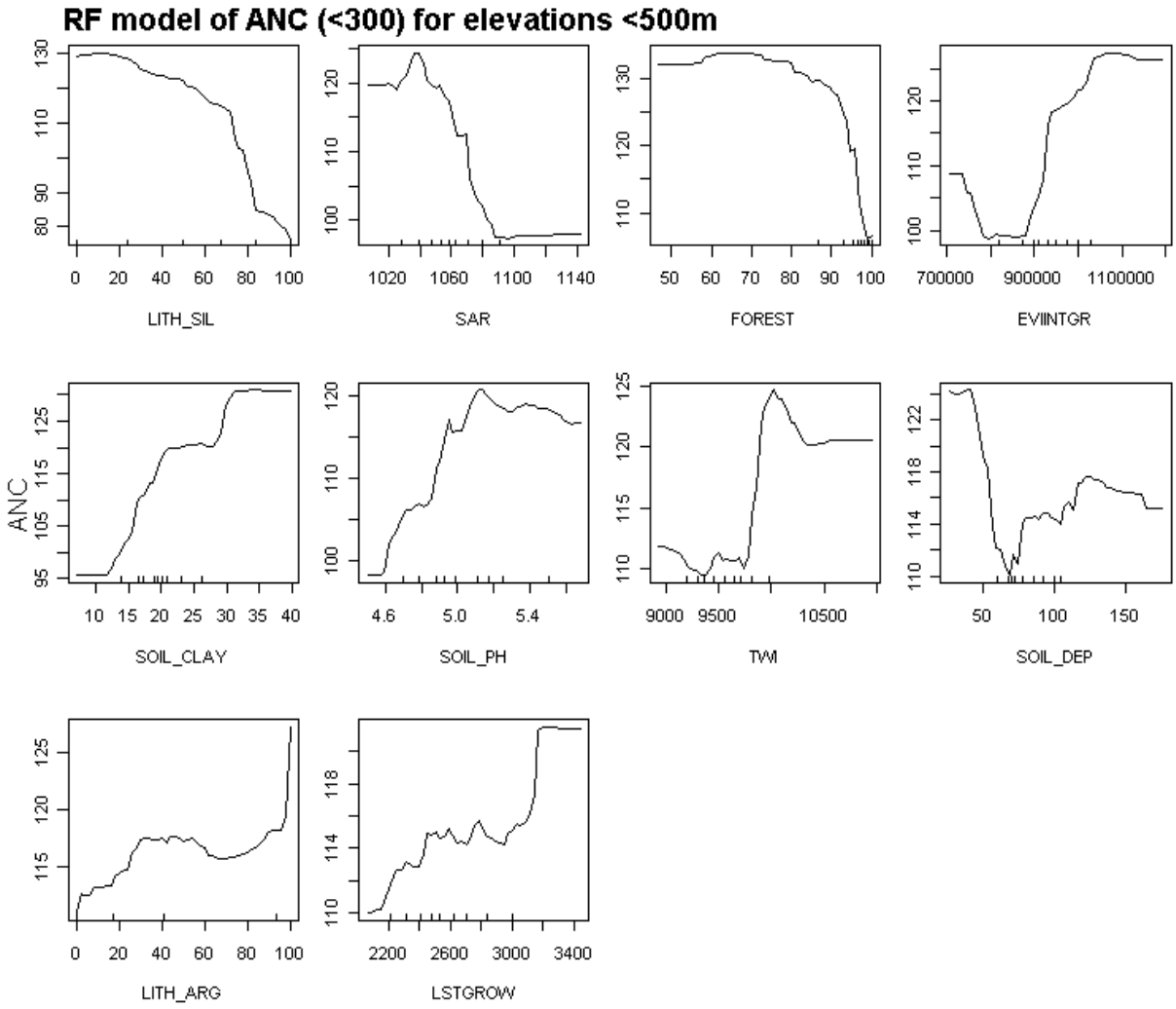
Topography	SAR	Surface area ratio (measure of terrain roughness)	none	ratio	30-m DEM
Topography	TWI	Topographic wetness index	none	none	30-m DEM (Wood et al. 1990)
Vegetation	PCTBARE	Mean percentage bare cover	none	percent	Hargrove and Hoffman (2004)
Vegetation	PCTTREE	Mean percentage tree cover	none	percent	Hargrove and Hoffman (2004)
Vegetation	CON42	Percentage contributing area in conifer cover	none	percent	Homer et al. (2007; NLCD)
Vegetation	CONMXD	Percentage contributing area in conifer (1) + mixed (0.5) cover (weighted by # in parentheses)	none	percent	Homer et al. (2007; NLCD)
Vegetation	DECID41	Percentage contributing area in deciduous cover	none	percent	Homer et al. (2007; NLCD)
Vegetation	DECIDMX	Percentage contributing area in deciduous (1) + mixed (0.5) cover (weighted by # in parentheses)	none	percent	Homer et al. (2007; NLCD)
Vegetation	FOREST	Percentage contributing area in forested cover	none	percent	Homer et al. (2007; NLCD)



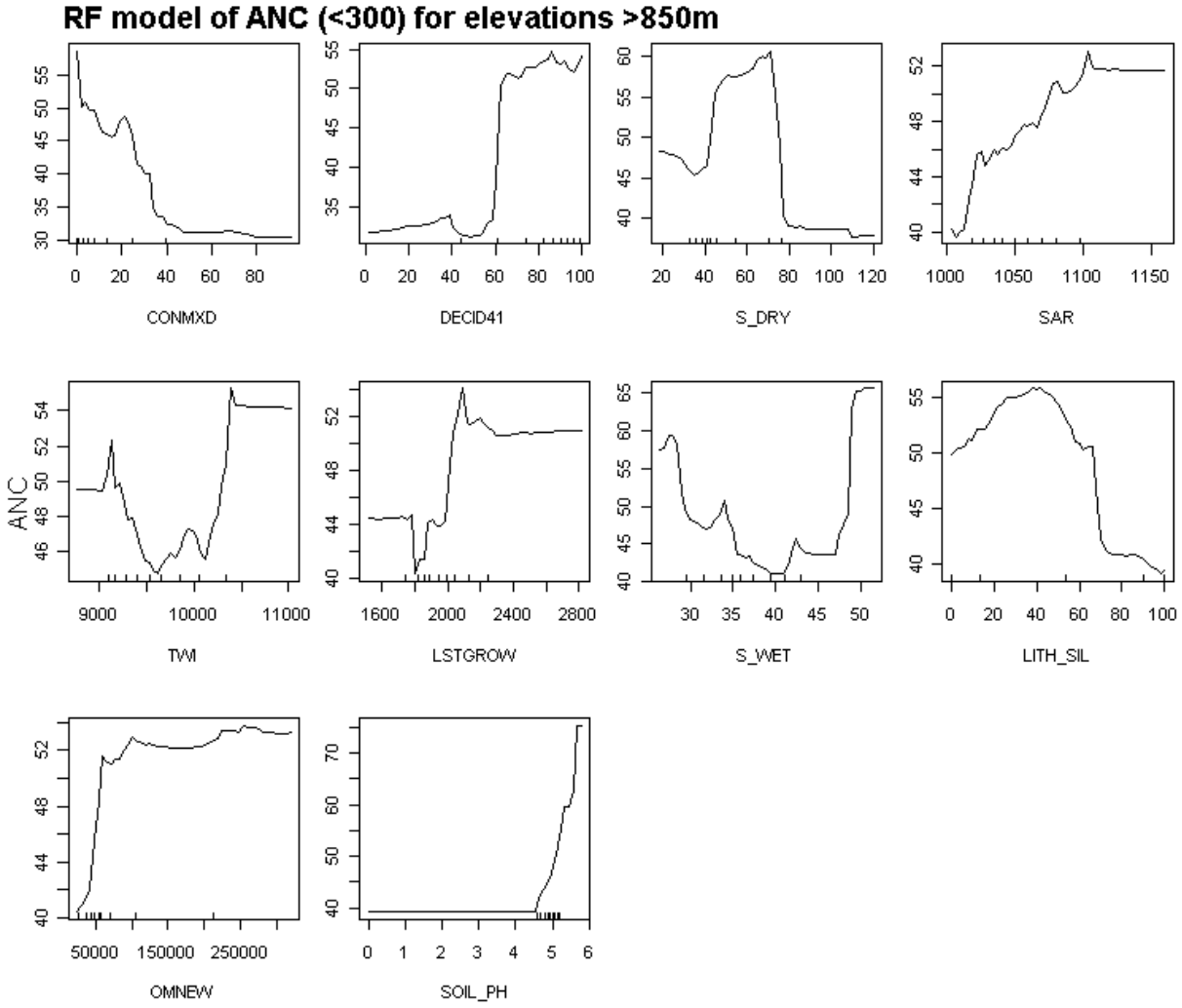
Appendix 2.2. Interaction plots from the threshold random forest model. Variables are, VWDAYMAX, maximum number of very wet days (i.e., vapor pressure deficit <1000 pa); PUBLIC, percent public land cover; LITH_CAR, percent calcareous lithology; FOREST, percent forest cover; SOIL_PH, soil pH; LITH_ARG, percentage argillic lithology.



Appendix 2.3. Interaction plots from the continuous random forest model. Variables are, LITH_SIL, percentage siliceous lithology; TWI, topographic wetness index; AB90GROW, number of days >32.2^{o}C; NPDAYMAX, number of days without precipitation; FOREST, percentage forest cover.



Appendix 2.4. Modeled response curves for ANC. Models were developed using a single random forest regression model for all sites with ANC <300 $\mu eq \cdot L^{-1}$ at **elevations <500 m**. Variables are LITH_SIL, percentage siliceous lithology; SAR, surface area roughness; FOREST, percentage forest cover; EVIINTGR, Mean enhanced vegetation index integrated over the local growing season; SOIL_CLAY, percentage soil clay; SOIL_PH, soil pH; TWI, topographic wetness index; SOIL_DEP, soil depth; LITH_ARG, percentage argillic lithology; LSTGROW, mean degree-days heat sum above 5.6^{o}C surface temperature during the local growing season.

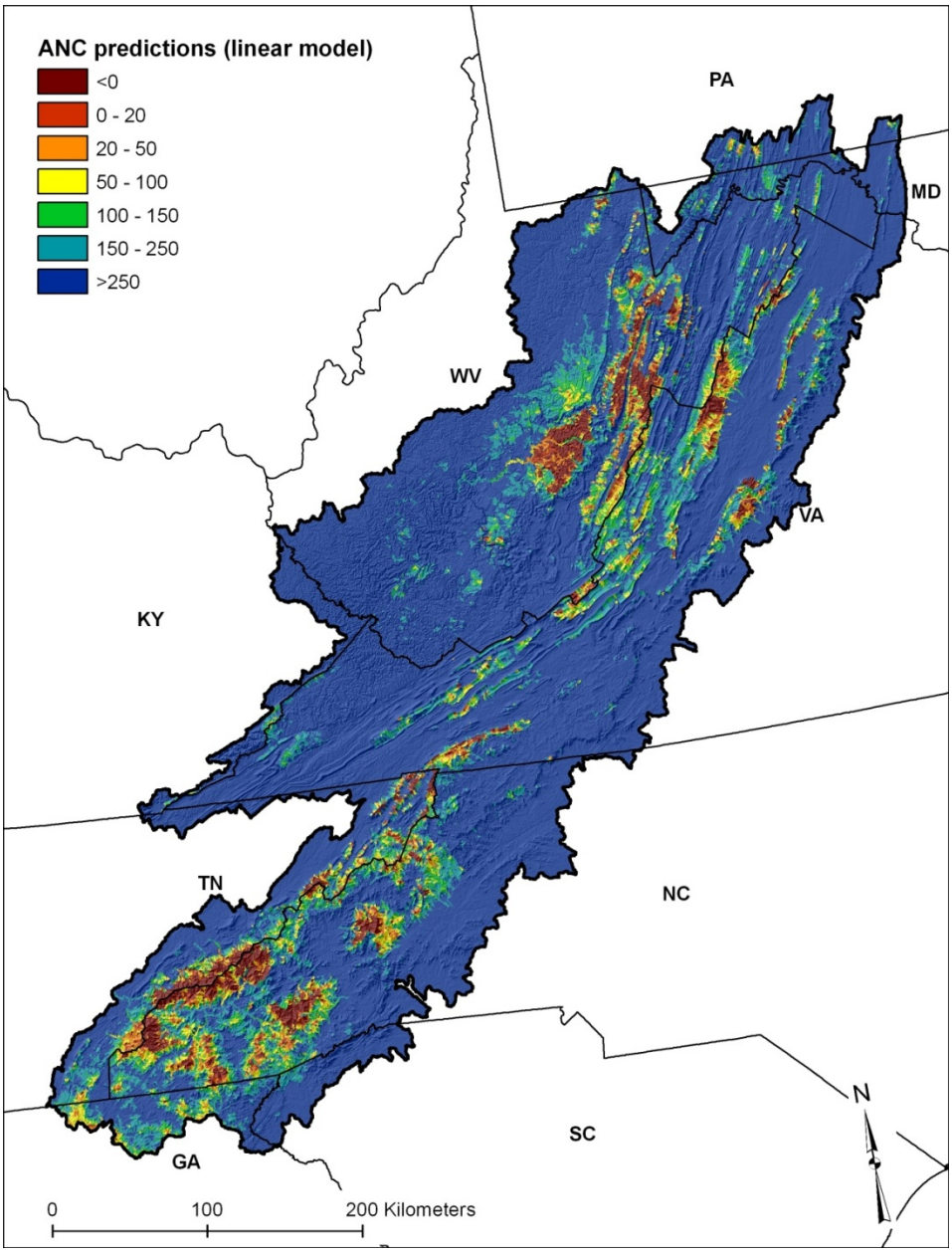


Appendix 2.5. Modeled response curves for ANC. Models were developed using a single random forest regression model for all sites with ANC <300 $\mu eq \cdot L^{-1}$ at **elevations >850 m**. Variables are, CONMXD, percentage mixed-conifer forest cover; DECID41, percentage deciduous forest cover, S_DRY, amount of dry sulfur deposition; SAR, surface area roughness; TWI, topographic wetness index; LSTGROW, mean degree-days heat sum above 5.6^{o}C surface temperature during the local growing season; S_WET, amount of wet sulfur deposition; LITH_SIL, percentage siliceous lithology; OMNEW, amount of soil organic matter; SOIL_PH, soil pH.

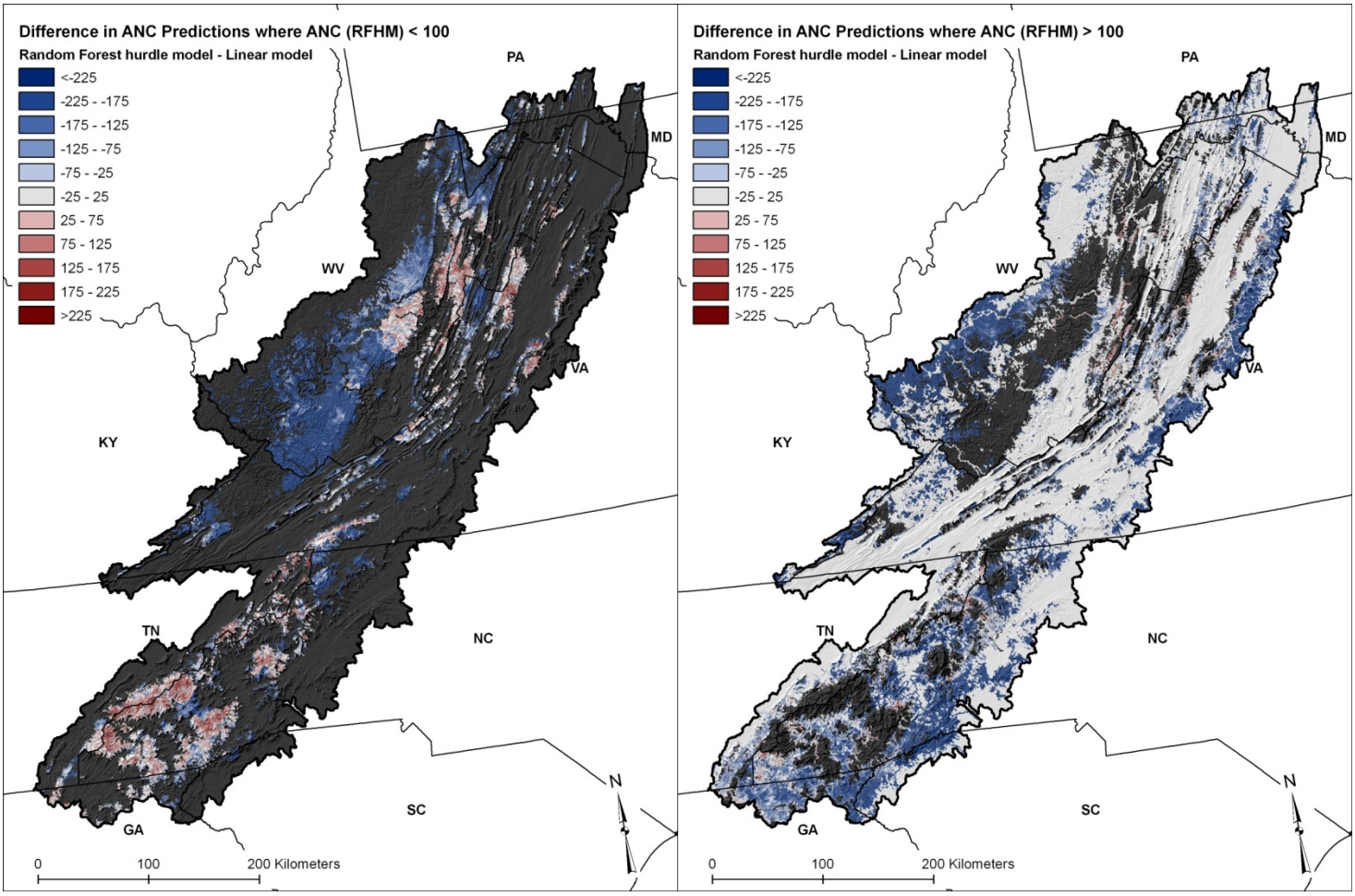
Appendix 2.6. Individual gatekeeper model performance results, for the random forests continuous model, for the combinations of model parameters plotted in Figure 2.8. ANC cutoff refers to the observed ANC value in the data below which sites are considered to exhibit "low-ANC" for the gatekeeper analysis. Probability cutoff refers to the probability threshold output by the threshold model, above which sites are classified as having a "low-ANC" and thereby entered into the continuous model within the gatekeeper modeling framework. Misclassification rate refers to the percentage of times the threshold model resulted in a misclassification of high/low ANC. All model validation statistics were computed using an independent test set withheld from model building. RMSE = root mean squared error.

ID	Sampling	Prob. Cutoff	ANC cutoff	RMSE	Misclass (%)	ID	Sampling	Prob. Cutoff	ANC cutoff	RMSE	Misclass (%)
1	Imbalanced	0.4	150	125.7	2.8	**25**	Imbalanced	0.4	250	175.2	1.6
2	Oversample	0.4	150	128.5	4.5	**26**	Oversample	0.4	250	188.5	1.9
3	Undersample	0.4	150	113.2	9.6	**27**	Undersample	0.4	250	120	11
4	Imbalanced	0.5	150	109.1	5.6	**28**	Imbalanced	0.5	250	154.1	2.2
5	Oversample	0.5	150	107.2	7.4	**29**	Oversample	0.5	250	178.8	3.4
6	Undersample	0.5	150	104.8	13.7	**30**	Undersample	0.5	250	102.1	17.4
7	Imbalanced	0.6	150	106	8.1	**31**	Imbalanced	0.6	250	130.5	3.2
8	Oversample	0.6	150	100.7	9.9	**32**	Oversample	0.6	250	152.5	5.6
9	Undersample	0.6	150	82.5	20.5	**33**	Undersample	0.6	250	87.8	24
10	Imbalanced	0.7	150	97.1	12.2	**34**	Imbalanced	0.7	250	111.7	6.4
11	Oversample	0.7	150	91.8	15.3	**35**	Oversample	0.7	250	123.9	9.6
12	Undersample	0.7	150	73.1	27.8	**36**	Undersample	0.7	250	73.9	32.2
13	Imbalanced	0.4	200	176	1.8	**37**	Imbalanced	0.4	300	191	1.3
14	Oversample	0.4	200	171.1	3.2	**38**	Oversample	0.4	300	202.2	2.1
15	Undersample	0.4	200	129.1	10.4	**39**	Undersample	0.4	300	129.2	12.4
16	Imbalanced	0.5	200	134.7	3.5	**40**	Imbalanced	0.5	300	159.3	2.1
17	Oversample	0.5	200	140.9	5.2	**41**	Oversample	0.5	300	168.4	3
18	Undersample	0.5	200	103.7	17	**42**	Undersample	0.5	300	100.8	18.1
19	Imbalanced	0.6	200	135.8	5.8	**43**	Imbalanced	0.6	300	141.4	3.2
20	Oversample	0.6	200	133.5	8.3	**44**	Oversample	0.6	300	150.5	4.9
21	Undersample	0.6	200	94.1	24	**45**	Undersample	0.6	300	81.8	23.9
22	Imbalanced	0.7	200	113.8	10	**46***	Imbalanced	0.7	300	107.5	5.6
23	Oversample	0.7	200	112.4	12.5	**47**	Oversample	0.7	300	132.7	9
24	Undersample	0.7	200	71.9	32.9	**48**	Undersample	0.7	300	66.3	29.8

* Indicates the model chosen for predicting ANC to the entire study domain



Appendix 2.7. Linear model predictions of ANC across the southern Appalachians.



Appendix 2.8. Differences in ANC predictions made by the random forest hurdle model and the linear model (RF-LM) for (left) areas where the hurdle model predicted ANC <100 $\mu eq \cdot L^{-1}$, and (right) areas where the hurdle model predicted ANC >100 $\mu eq \cdot L^{-1}$. Blue areas indicate where the linear model predicted high ANC compared to the predictions made by the hurdle model. Conversely, red areas indicate where the RF model predicted higher ANC.

CHAPTER 3 - PREDICTING CATCHMENT-LEVEL BASE CATION WEATHERING RATES ACROSS THE SOUTHERN APPALACHIAN MOUNTAIN REGION, USA

ABSTRACT

A key component of the steady-state critical load estimation is determining the natural resupply rate of base cations to surface waters through weathering (BC_w). Here, we used modeled catchment-level BC_w across the southern Appalachian Mountain region as the response in a correlative modeling framework that identified key environmental correlates, and predicted a continuous map of BC_w. More than 50 initial candidate predictors were submitted to a variety of conventional and machine learning regression models. Predictors included aspects of the underlying geology, soils, geomorphology, climate, topographic context, and acidic deposition rates. Low BC_w rates were predicted in catchments with low precipitation, siliceous lithologies, low soil clay, nitrogen (N) and organic matter contents, and relatively high levels of canopy cover in mixed-coniferous forest types. Results will be used to support regional sulfur critical loads modeling to identify areas impacted by industrially derived atmospheric S inputs.

1. INTRODUCTION

Atmospheric sulfur (S) inputs to southeastern United States (US) streams have reduced native fish and invertebrate populations, and altered aquatic community structure in favor of acid tolerant species (U.S. EPA 2009). Deposited S largely derives from industrial coal and oil combustion (National Acid Precipitation Assessment Program (NAPAP) 1991), and contributes to deteriorated streamwater acid-base status through base cation leaching, increased sulfate

concentrations, reduced pH and acid neutralizing capacity (ANC), and toxic aluminum (Al^{3+}) mobilization from soil to stream waters (Driscoll et al. 2001, U.S. EPA 2009).

The U.S. Environmental Protection Agency (EPA) and federal land management agencies are charged with assessing the ecological effects of S emissions and implementing restorative management to areas most susceptible to chronic deposition. The standard practice for assessing the levels of continued emission and deposition that are detrimental to terrestrial and aquatic environments and to identify biological effects associated with soil and water chemistry is to develop modeled estimates of critical loads (CL; Henriksen and Posch 2001). The CL is defined as the level of sustained atmospheric deposition of strong acids (here, S) above which harmful effects are expected (Nilsson and Grennfelt 1988). Several studies have used aquatic critical load calculations to identify exceedances for aquatic habitat (Henriksen et al. 1995, Henriksen et al. 2002, Sullivan et al. 2005, McDonnell et al. 2012, Reynolds et al. 2012). The critical load is calculated as:

$$CL(S) = BC_{dep} + BC_w - Bc_u - Cl_{dep} - ANC_{le(crit)} \qquad \text{Eq. 1,}$$

where BC_{dep} and Cl_{dep} represent the total atmospherically deposited major base cations Ca, K, Mg, Na, and chloride, repectively; BC_w is the calculated weathering rate for base cations, Bc_u represents the uptake of nutrient base cations by trees, and $ANC_{le(crit)}$ is the critical acid neutralizing capacity (ANC) leaching limit.

This research focuses on estimating BC_w; considered by some to be the most critical (Hodson and Langan 1999) and often the most poorly estimated term of the CL calculation (Skeffington 2006, Li and McNulty 2007, McDonnell et al. 2010) . Maintaining ecosystem health is contingent upon maintaining acidic deposition levels below the natural resupply of base cations. BC_w is generally calculated as an intermediate output in models that estimate CLs.

Models typically are (1) process-based or (2) empirical steady-state models (Hodson and Langan 1999, Hall et al. 2001).

Process-based models such as PROFILE (Warfvinge and Sverdrup 1992, Sverdrup and Warfvinge 1993) and MAGIC (Model of Goundwater Acidification in Catchments; Cosby et al. 1985) use intensive field-collected soils and stream chemistry data for model calibration, and produce accurate BC_w estimates at an individual catchment level. Policy makers and land managers often require estimates made at landscape and regional scales, and data and processing limitations prevent the use of process-based models at these scales. Empirical steady-state models such as the Simple Mass Balance method (SMB; de Vries 1991, Sverdrup and De Vries 1994) are widely used to predict CL at larger spatial scales. These models develop empirical relationships between BC_w and underlying soil texture and mineralogy of catchment soils, and assume (1) steady-state conditions with a single soil layer, (2) soils and geology data are well represented by available large-scale data sets, and (3) that these factors alone drive weathering rates.

However, the SMB model suffers from high error rates related to variability in biogeoclimatic conditions at regional scales, low spatial accuracy or inconsistent scaling of soils and geological data, and high dissimilarity between regions where models were developed and applied (Skeffington 2006, Li and McNulty 2007, McDonnell et al. 2010). Furthermore, BC_w modeled error rates are often unreported or unknown. Thus, the validity of associated CL estimates is uncertain. For example, Li and McNulty (2007) used the SMB model and indicated that BC_w estimates contributed 49% of the error in estimated CL for the United States. The authors concluded that improving regional estimates of BC_w should be the focus of future research.

The goal of this research is to improve estimates of BC_w across a large region of the Southern Appalachian Mountains (Sullivan et al. 2008). We take an objective statistical modeling approach to correlate biogeochemical and climatic variables with BC_w estimates made using a dynamic process-based model for catchment-level soil acidification. We use these relationships to make continuous BC_w predictions across the study area. Improved knowledge of the main environmental correlates of BC_w can confirm known and elucidate previously unknown relationships. These relationships can then be used within a correlative modeling framework to extrapolate estimates of catchment-level BC_w to larger geographic regions.

The work presented here builds on previous modeling efforts in Virginia and West Virginia (McDonnell et al. 2012). Advances made in the current assessment include (1) the addition of numerous MAGIC modeled BC_w catchment locations (the response variable), (2) the inclusion of new climate, productivity, soils, topography, and sulfur deposition predictor variables, (3) incorporation of machine learning algorithms capable of modeling non-linear responses, (4) assessment of model uncertainty, and (5) extrapolation to a larger study area. Modeled estimates of BCw developed here were incorporated in a decision support system model whose aim was to identify sensitive aquatic habitat across the same region (Reynolds et al. 2012).

2. METHODS

2.1. Study area description

The study was conducted across a broad geographic area ($14.3 \cdot 10^6$ ha) encompassing much of the southern Appalachian Mountain region of the eastern United States, from northern Georgia to southern Pennsylvania, and from eastern Kentucky and Tennessee to central Virginia and western North Carolina (Figure 2.1). The study area includes several Omernik (1987) Level III

ecoregions, including the Blue Ridge, Ridge and Valley, and Central Appalachian regions. Elevations in the study region range from 300 to 2000 m. Forests are primarily composed of oak (*Quercus* spp.), hickory (*Carya* spp.), and pine (*Pinus* spp.). Spruce-fir and northern hardwood forests occupy the highest elevation sites with cool, moist climates. Forests are interspersed with agricultural and urban lands, which generally occupy lowland valleys.

2.2. MAGIC model

MAGIC is a process model used to estimate effects of acidic deposition on soils and surface waters in catchments (Cosby et al. 1985). Within the model, physical and chemical processes associated with the acid-base status of soil and stream water chemistry are averaged across a catchment to predict its potential response to acidic inputs over time. MAGIC simulates the soil solution and surface water chemistry over a period of approximately 150 years to predict average annual and monthly concentrations of major ions. Model calibration ensures that modeled estimates of soil and stream water chemistry agree with observed levels. Because the pool of exchangeable base cations in soil changes in response to atmospheric deposition and biomass export, the chemical equilibrium in soil may shift over time, thereby driving changes in base cation concentrations in the soil solution. For this reason, MAGIC is typically calibrated within a catchment using observed stream, soil, and atmospheric deposition data, which provides the model with upper and lower limits on soil BC concentrations.

The BC_w estimates presented here represent the catchment average net supply of base cations draining from a predefined catchment to surface waters (McDonnell et al. 2010); and represent a major component of aquatic critical load calculations. This term is identical to the BC_w term in the Steady State Water Chemistry model (SSWC) often used to calculate regional scale aquatic critical load estimates. We note that MAGIC estimates of BCw do not include

averages within the rooting zone of terrestrial plants, which are required for terrestrial CL calculations, except where catchment soils are shallow.

In the current study, MAGIC was calibrated for 140 catchments, using methods described by McDonnell et al. (2012). Many study catchments occurred on public lands, and all were located in forested areas. The catchment BC_w rate was extracted from the MAGIC model, and these values were used as training data to model BC_w rates across the region (the response variable, McDonnell et al., (2012)).

To assess the representativeness of the environmental settings of the modeled catchments compared to those present across the study domain, we compared the estimated kernel density function for each predictor from the 140 sites to those constructed from 4,000 randomly selected catchments (Table 3.1). The proportion of overlap among the density functions is one indication of correspondence among the distributions. Comparisons of the interquartile range and median values, and the density function overlap analysis, indicated fairly high concordance between the sampled locations and those present across the study region (mean overlap = 0.74).

2.3. Predictor variables

An initial set of 55 environmental predictor variables was chosen to represent broad- to fine-scale climatic, lithologic, geomorphic, topographic, edaphic, vegetative, land ownership, and atmospheric S deposition variables that were potentially influential to explaining geographical patterns of BC_w. Variables are briefly described below, but see Appendix 2.1 for a more complete description of the predictor variables.

Climate variables (1-km resolution) included estimates of the growing season and non-growing season precipitation, temperature, and insolation regime, and measures of overall productivity (Hargrove and Hoffman 2004). Vegetation data (30-m) included percentage

composition of conifer, deciduous, and all forest types (Homer et al. 2007). Soil variables (1-km) included percentage clay, soil pH, and soil depth (NRCS Soil Survey Staff 2010a, b), and amount of organic matter, mean Kjeldahl N content, and mean soil plant-available water (Hargrove and Hoffman 2004). A broad-scale lithologic classification provided by Sullivan et al. (2007) was used to capture the percentage composition of parent materials across the study area. Classes included siliceous, argillic, felsic, mafic, and carbonate substrates. Topographic setting was characterized using a 30-m digital elevation model (DEM) in terms of surface area roughness, topographic wetness index, and flow accumulation (catchment size). Levels of wet and dry S deposition were also included (Grimm and Lynch 2004, Byun and Schere 2006).

The raster layer of each predictor variable was resampled to 30-m resolution and the values were averaged across the contributing upslope area for each grid cell (see Appendix 2.1 for native raster resolutions). Upslope-averaging incorporated the catchment influence of the predictors on BC_w as calibrated to the stream sampling location, or pour-point, at the base of each catchment (McDonnell et al. 2012). The equation for upslope averaging is,

$$\bar{P}_i = \frac{(P_i + \sum_{j=1}^{N} P_j)}{(N+1)} \qquad \text{Eq. 1,}$$

where $\bar{P}_i$ is the upslope averaged value for the candidate cell (P_i), $\sum_{j=1}^{N} P_j$ is the summation of values for all cells upslope from P_i, and N is the total number of upslope cells. For each cell within a predefined catchment, the sum of all upslope cell values was added to the value of each cell. Upslope-averaging enabled us to incorporate the average of each predictor variable across the landscape draining into each individual cell.

Also included in a subset of the models were the nine water chemistry variables sampled at each of the 140 MAGIC calibration sites. These included strong mineral acid anions SO_4^{2-}, Cl^-, and NO_3^- and base cations Ca^{2+}, Mg^{2+}, Na^+, K^+, and NH_4^+, as well as two acidity measures, pH

and ANC. Chemistry data were included to increase the accuracy of BC_w predictions for a sample of 933 stream water chemistry sites distributed throughout the study region where data had been collected. Hereafter, models that include water chemistry predictors are referred to as landscape+chemistry models, and those that include only environmental variables are referred to as landscape-only models. We incorporated the landscape+chemistry models in our analysis specifically to assess model improvements with the addition of water chemistry variables, and these models were not used in final model predictions to the study area. Demonstrating high predictive ability using these models would suggest that intensive MAGIC modeling may not be necessary to predict BCw at the catchment scale; water chemistry and environmental data layers are inexpensive and easily accessible.

2.4. Statistical modeling

We explored a variety of statistical models, from linear regression to more complex machine learning algorithms, in order to objectify the process of identifying the best statistical model for the dataset, and to identify the model parameterization with the lowest error for developing the continuous BC_w surface.

Statistical models tested included: linear regression (LM, stats package in R v 2.12.2, R Development Core Team 2011), boosted regression trees (BRT, gbm package in R v2.12.2, Ridgeway 2010), random forest (RF, randomForest package in R v2.12.2, Liaw and Wiener 2002), multivariate adaptive regression splines (MARS, mda package in R v2.12.2, Hastie and Tibshirani 2011), and geographically weighted regression (Bivand et al. 2010). Each statistical model has been used in previous landscape modeling assessments (e.g., Prasad et al. 2006). Model performance often depends on unique aspects of the data and relations among the patterns being modeled and the processes driving the patterns (Olden et al. 2008, Franklin and Miller

2009).

Machine learning algorithms are relatively new to ecological research (Hastie et al. 2005, Olden et al. 2008). These algorithms are non-parametric and data driven, meaning that models do not produce a parameterized statistical model, but identify patterns in the data with no assumptions regarding the underlying probability distribution of the training data (Breiman 2001b). RF (Breiman 2001a) and BRT (Friedman 2002) models are ensemble versions of traditional regression tree analyses, in which hundreds to thousands of individual regression trees are built within each algorithm and final predictions are averaged across the individual trees. In RF, individual regression trees are built using a random subset of the response and predictor datasets; resulting trees are not pruned. Model error-rates can be assessed using out-of-bag predictions (OOB), which are those made using the individual regression trees on the subset of data that was randomly withheld at each model iteration.

BRT uses the entire set of predictors and data to build each individual tree, and the algorithm is sequential, with higher weights given to poorly classified observations and final predictions are weighted means of each tree (De'ath 2007, Elith et al. 2008). MARS is a non-parametric extension of traditional linear regression models (Friedman 1991), which consist of a series piecewise linear functions that allow for non-linearities in a response. GWR is another extension of linear regression developed by Fotheringham et al (2002), which allows the parameters in a global regression model to be estimated locally at every point by giving higher weights to geographically proximal data points.

2.5. Model building and validation

Multi-collinearity in the initial set of predictor variables (landscape-only model and landscape+chemistry model) was reduced by retaining only those with Pearson's correlation

scores <0.7, leading to a modeling set of 33 and 38 predictors for landscape-only, and landscape+chemistry models, respectively. Among correlated variables, those with the highest Pearson's correlation with BC_w were retained. High and lower correlation cutoff values were tested, but none improved the final models. RF models were then run using the reduced predictor set for each iteration of the modeling with successively fewer predictor variables. At each iteration, the highest scoring variables from the prior run were re-entered into each of the statistical models tested. RF models were used in variable selection to (1) avoid potential bias associated with traditional stepwise reduction procedures (Whittingham et al. 2006), and (2) allow for direct comparison of model performance (e.g., linear, non-linear) with the same model predictors.

Separate statistical models were built using the reduced set of predictors (after reducing multi-collinearity) for the landscape-only and landscape+chemistry models. Goodness-of-fit statistics, including root mean squared error (RMSE) and R-squared, were calculated based on predictions on the training data (hereafter, training error). Model validation statistics were then calculated for each model using sequentially fewer predictors (25, 20, 15, 10, 5, and 3) to determine the number of predictors that balanced model parsimony with prediction accuracy. Model error was also evaluated using 10-fold cross-validation (hereafter, $CV_{10\text{-fold}}$) to reveal model overfitting (Cutler et al. 2007, Elith et al. 2008).

In subsequent analyses, we also built ecoregion-specific models using the same methods as outlined above, for three main Omernik ecoregions within the study domain that appeared to have an adequate number of BC_w samples. These models were developed to identify important differences in BC_w covariates that were related to physiography. Individual ecoregion model performance was compared to the single global model to identify changes in model error rates.

The Northern Piedmont, Piedmont and Western Allegheny Plateau ecoregions each included fewer than five BC_w sample locations and were not considered further in ecoregional modeling. These lesser ecoregions made up 12.5% of the study area.

3. RESULTS AND DISCUSSION

3.1. BC_w summary

MAGIC modeled BC_w rates ranged from 3.1 to 257.1 $meq \cdot m^{-2} \cdot yr^{-1}$, with a median of 66.4 $meq \cdot m^{-2} \cdot yr^{-1}$. The estimated kernel density function for the data was right-skewed with 20% of all catchment-level weathering rates >100 $meq \cdot m^{-2} \cdot yr^{-1}$. Median elevation for study catchments was 753 m (Inter-quartile range (IQR): 585, 937 m), and catchments tended to be located at slightly higher elevations, in cooler and moister climates, with higher percent forest cover, higher percent siliceous lithology and lower soil clay levels compared to the rest of the study area (Table 3.1).

3.2. Model calibration and validation

Machine learning algorithms (RF, BRT, and MARS) consistently outperformed linear regression and GWR models based on error rates (Figure 3.1). Most models, with the exception of RF, showed a steep increase in model error rate (or reduction in variance explained) as the number of predictor variables decreased (Figure 3.1). This was most pronounced for landscape-only models.

RF was least influenced by the number of predictors included in the model (Figure 3.1), and showed the best overall performance among the models tested (Figure 3.2). Training model RMSE averaged 16.5 ± 1.2 $meq \cdot m^{-2} \cdot yr^{-1}$ across all landscape-only models and 12.2 ± 0.8 $meq \cdot m^{-2} \cdot yr^{-1}$ for landscape+chemistry models (Figure 3.2). BRT and MARS models both showed

inconsistent behavior in the landscape-only models; performance was high when many predictors were included, but model error rates increased with fewer predictors (Figure 3.1).

LM and GWR models recorded the poorest overall mean training error rates when averaged across all models tested (28.8±3.9 and 25.7±4.7 $meq \cdot m^{-2} \cdot yr^{-1}$ landscape-only, and 20.5±3.1 and 19.3±2.2 $meq \cdot m^{-2} \cdot yr^{-1}$ landscape+chemistry for LM and GWR, respectively; Figure 3.2). The bandwidth selected by GWR model cross-validation was ~250 km, a distance that incorporated a high degree of environmental heterogeneity, and likely precluded local coefficient estimates from improving model performance over the LM model.

When water chemistry variables were included (landscape+chemistry), performance of all models was less influenced by the number of predictors (Figure 3.1) and training model error rates were generally lower than landscape-only models (Figure 3.1, and Figure 3.2). In general, the addition of the water chemistry predictors reduced training model error by more than 25%. This reduction in model error indicated that knowledge of downslope stream water chemistry improved predictions of upslope catchment BC_w rates.

Cross-validation estimates of RMSE for all models ranged from 30-35 $meq \cdot m^{-2} \cdot yr^{-1}$ for landscape-only models and 21-24 $meq \cdot m^{-2} \cdot yr^{-1}$ for landscape+chemistry models, and RF models reported the lowest RMSE rates of the models tested. $CV_{10\text{-fold}}$ error rates of the LM and GWR models were similar to those based on estimates for the training dataset, but this was not the case for the machine learning algorithms. Among the landscape-only models, $CV_{10\text{-fold}}$ error rates increased by <8% over training model error rates for the LMs, compared to >60% for the machine learning algorithms (Table 3.2). These results indicate model instability exists, likely due to under-sampled BCw within the geographical range of the predictors.

3.3. Ecoregional models

Individual ecoregion models did not improve overall model fit compared to using a single global model. RMSE training rates were lowest for Blue Ridge (BR) and Ridge and Valley (RV) ecoregions (~15 meq·m^{-2}·yr^{-1}) and highest for the Central Appalachians (CA; 24 meq·m^{-2}·yr^{-1}) based on RF model predictions (Table 3.3). The poor performance of the CA model is likely due to the low sample size compared to the size of the ecoregion. Cumulatively, total training model error was only slightly reduced by <1 meq·m^{-2}·yr^{-1} when BC_w was modeled using separate ecoregional models compared to a single global model. When the model calibrated for all ecoregions was used to predict for individual ecoregions, results indicated that the global model performed as well or slightly better than the individual ecoregion models (Table 3.4).

3.4. Model selection

The final model chosen for predicting for the study domain was the single global, landscape-only RF model with 10 predictor variables, 1000 individual regression trees, and three predictor variables included in the development of individual regression trees (Table 3.2). This model included only landscape predictors, could be used to predict BC_w rates across the entire study region, and had the lowest error rates compared to other models tested. The model had a training error rate of 16.1 meq·m^{-2}·yr^{-1}, and an R-squared of 0.86. The OOB error rate was 31.4 meq·m^{-2}·yr^{-1} (R^2_{OOB}= 0.452; data not shown), similar to the 10-fold cross-validation results (Table 3.2). The final RF model performed best when the observed BC_w was <125 meq·m^{-2}·yr^{-1} (Figure 3.3). Model errors were inflated for BC_w estimates above this value, presumably due to the sparseness of data points in relatively high BC_w environments.

3.5. Biophysical predictors of BC_w

Key predictor variables in the landscape-only models represented various aspects of the local

climatic regime, atmospheric S inputs, underlying lithology, and soils characteristics of the upslope catchment. Low BC_w rates were predicted for catchments that received low precipitation, in areas underlain by siliceous lithologies, with low soil clay, low Kjeldahl nitrogen, low organic matter content, and relatively high levels of canopy cover of mixed-coniferous forest (Figure 3.4). When water-chemistry variables were incorporated into models (landscape+chemistry models), low BC_w rates occurred in streams having low concentrations of streamwater Ca^{2+}, Na^{+}, and SO_4^{2-}, low streamwater pH, and low levels of precipitation during the non-growing season. Response functions for the variables included in the landscape-only and landscape+chemistry models were generally non-linear over the range of the predictor variables, but distinct directional patterns were obvious for most variables (Figure 3.4).

Siliceous lithology was identified as the most influential predictor of BC_w and was included in all global landscape-only models. Low BC_w was correlated with the presence of siliceous bedrock, which is known to be low in base cations (Duan et al. 2002, Posch et al. 2003). This result is also consistent with independent geological sensitivity mapping conducted throughout this region by Sullivan et al. (2007). A significant interaction between percent siliceous lithology and precipitation amount was identified, and suggested that areas with the highest observed precipitation amounts had the highest predicted BC_w rates regardless of lithology. A similar result was identified for the interaction between percent siliceous lithology and wet sulfur deposition.

Other variables related to soil properties were positively correlated with BC_w rates (Figure 3.4). Some evidence suggests that organic matter accelerates weathering rates of rocks (Drever and Stillings 1997), although relationships between organic matter and weathering are thought to be complex and influenced by pH and parent material composition (Drever 1994, Kump et al.

2000). A significant interaction was found between organic matter and precipitation (Figure 3.5), with the lowest BC_w rates occurring in areas with low to moderate levels of precipitation and low soil organic matter.

Besides soils and lithology, climate variables reflecting temperature and precipitation regimes are known to influence weathering rates (Peltier 1950, Johnson et al. 1994, Kump et al. 2000), and several climate variables were retained in the models. Specifically, our models indicated a positive correlation between BC_w rates and precipitation (total consecutive days with precipitation (PDAYMAX), and non-growing season accumulation (PRECIPNG); Figure 3.4). Precipitation likely influences rates of base cation transport within the soil solution and alters chemical concentrations in soil water over time (Drever 1994). Temperature influences chemical reaction and respiration rates in soils, and influences overall weathering rates (Lloyd and Taylor 1994, Winkler et al. 1996, Wright et al. 2006). The inclusion of the variable DIFF95NG (Figure 3.4) indicated that large diurnal differences in surface temperatures led to higher BC_w rates, which may be related to physical weathering by non-growing season freeze-thaw cycles (Peltier 1950).

Vegetation influences BC_w through physical processes that fracture bedrock, and chemical processes that control the deposition of organic matter, uptake of water and nutrients, and release of carbon dioxide into soil water (Hultberg 1985, Miles 1986, Drever 1994, Wright et al. 2006). Conifer forests, in general, develop humus layers high in organic acids, which can lead to increased levels of base cation leaching, lower soil pH, and increased chemical weathering rates (Hornung 1985, Johnson et al. 1994, Augusto et al. 2002), which in turn may influence BC_w rates. Accordingly, our models identified a negative relationship between the mixed-conifer cover and BC_w rates (Figure 3.4).

Sulfur deposition was positively correlated with BC_w, and wet S deposition was included as a main predictor variable in all ecoregion models. The role of S deposition in influencing BC_w, is not clear (Johnson et al. 1994). Because, the research presented here represents a proof-of-concept in predicting BC_w rates using the methods detailed here, S deposition variables were retained in the model. Sulfur deposition levels are strongly influenced by orography (Byun and Schere 2006), and this variable may have "coded" for a variety of topographic and climatic variables, thereby contributing to a high variable importance score. Excluding S variables from the model led to a minimal 5.8% reduction in overall RMSE. Variables that replaced S deposition, when it was excluded from the models, included catchment soil pH, compound topographic index, non-growing season solar insolation, and flow accumulation. This suggests that including S deposition predictors may not be necessary to successfully model BC_w.

Consistent with McDonnell et al. (2012) our models identified siliceous lithology (%), soil clay (%), and soil pH as leading variables; however, they also identified other soils variables including soil nitrogen and organic matter, climate variables including seasonal and annual precipitation, temperature, and S deposition levels that were predictive of BC_w rates across the study area.

3.6. Model predictions

Model predictions ranged between 25.2 and 174.2 $meq \cdot m^{-2} \cdot yr^{-1}$, and 72% of the southern Appalachians had predicted BC_w rates <100 $meq \cdot m^{-2} \cdot yr^{-1}$ (Table 3.5). Areas with high BC_w were generally located to the northwest where high levels of precipitation during the non-growing season may have contributed to relatively high modeled weathering rates (Figure 3.6). Patterns of siliceous lithologies are also apparent, particularly throughout the RV ecoregion, where some of the lowest predicted BC_w rates were found on pervasive NE/SW trending parallel sandstone

ridges (Figure 3.6). Such patterns were more apparent when S variables were excluded from the model (Appendix 3.1).

Both McDonnell et al. (2012) and our analysis used MAGIC model estimates to quantify site-specific BC_w rates. We compare predictions of the final regional models to highlight differences in the studies. McDonnell et al. (2012) used three separate ecoregional models to predict to their study domain (~70% of our study domain). The authors concluded that less than one-third of their study region had BC_w rates <100 $meq \cdot m^{-2} \cdot yr^{-1}$, compared to our finding of more than three-quarters of the study region (Table 3.5). Furthermore, the authors reported nearly one-quarter of the landscape with BC_w rates >200 $meq \cdot m^{-2} \cdot yr^{-1}$, compared to our finding of <2%. There are several likely explanations for these differences in findings. The McDonnell et al. (2012) study used (1) ecoregion specific models to predict to the study region, (2) linear models with stepwise variable selection, (3) fewer data points (n=92) used in model training, and (4) a more restricted geographic extent compared to that used in the current study. As a comparison, we used a linear model with 10 variables (see Appendix 3.2) for prediction, which led to a somewhat more even distribution of predictions among BC_w classes (Table 3.5), but predictions were still concentrated in the 50-100 and 100-150 $meq \cdot m^{-2} \cdot yr^{-1}$ classes. It is likely that MAGIC model estimates of BC_w at additional locations within the study region will be necessary to improve model error rates and model accuracy.

4. CONCLUSIONS

Correlative landscape models were used to identify relationships between environmental predictors and individual catchment-level BC_w estimates to predict BC_w rates across the southern Appalachian Mountains. Results from this analysis suggest that several broad- to fine-scale factors are related to weathering rates, including precipitation and temperature regimes,

lithology, soil properties, and vegetation. Predictions from an RF model indicate that the lowest BC_w rates generally occur in dry areas underlain by siliceous lithologies, and in catchments containing low levels of soil pH, N and organic matter. Models predicted more than three-quarters of the study region had BC_w rates <100 $meq \cdot m^{-2} \cdot yr^{-1}$, suggesting that weathering rates across the region are inherently low. MAGIC model estimates should be made at new sample locations throughout the study area, particularly in currently under-sampled geographic areas, to further enhance our understanding of the broad-scale biophysical drivers that are not directly modeled in the process-based calculations of BC_w and to make more refined predictions of weathering rates across the landscape.

5. REFERENCES

Augusto, L., J. Ranger, D. Binkley, and A. Rothe. 2002. Impact of several common tree species of European temperate forests on soil fertility. Ann. For. Sci. **59**:233-253.

Bivand, R., D. Yu, T. Nakaya, and M. Garcia-Lopez. 2010. spgwr: Geographically weighted regression.

Breiman, L. 2001a. Random forests. Machine learning **45**:5-32.

Breiman, L. 2001b. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical Science **16**:199-231.

Byun, D. and K. L. Schere. 2006. Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. Applied Mechanics Reviews **59**:51-77.

Cosby, B., G. Hornberger, J. Galloway, and R. Wright. 1985. Modeling the effects of acid deposition: Assessment of a lumped parameter model of soil water and streamwater chemistry. Water Resources Research **21**:51-63.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology **88**:2783-2792.

De'ath, G. 2007. Boosted trees for ecological modeling and prediction. Ecology **88**:243-251.

de Vries, W. 1991. Methodologies for the assessment and mapping of critical loads and of the impact of abatement strategies on forest soils.

Drever, J. I. 1994. The effect of land plants on weathering rates of silicate minerals. Geochimica et Cosmochimica Acta **58**:2325-2332.

Drever, J. I. and L. L. Stillings. 1997. The role of organic acids in mineral weathering. Colloids and Surfaces A: Physicochemical and Engineering Aspects **120**:167-181.

Driscoll, C. T., G. B. Lawrence, A. J. Bulger, T. J. Butler, C. S. Cronan, C. Eagar, K. F. Lambert, G. E. Likens, J. L. Stoddard, and K. C. Weathers. 2001. Acidic deposition in the northeastern United States: sources and inputs, ecosystem effects, and management strategies. BioScience **51**:180-198.

Duan, L., J. Hao, S. Xie, Z. Zhou, and X. Ye. 2002. Determining weathering rates of soils in China. Geoderma **110**:205-225.

Elith, J., J. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology **77**:802-813.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons Inc.

Franklin, J. and J. A. Miller. 2009. Mapping species distributions: spatial inference and prediction. Cambridge University Press

Friedman, J. H. 1991. Multivariate adaptive regression splines. The annals of statistics:1-67.

Friedman, J. H. 2002. Stochastic gradient boosting. Computational Statistics & Data Analysis **38**:367-378.

Grimm, J. W. and J. A. Lynch. 2004. Enhanced wet deposition estimates using modeled precipitation inputs. Environmental Monitoring and Assessment **90**:243-268.

Hall, J., B. Reynolds, S. Langan, M. Hornung, F. Kennedy, and J. Aherne. 2001. Investigating the Uncertainties in the Simple Mass Balance Equation for Acidity Critical Loads for Terrestrial Ecosystems in the United Kingdom. Water, Air, & Soil Pollution: Focus **1**:43-56.

Hargrove, W. and F. Hoffman. 2004. A flux atlas for representativeness and statistical extrapolation of the Ameriflux network. Oak Ridge National Laboratory Technical Memorandum **ORNL-TM-2004**:1-152.

Hastie, T. and R. Tibshirani. 2011. mda: Mixture and flexible discriminant analysis.

Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin. 2005. The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer **27**:83-85.

Henriksen, A., P. Dillon, and J. Aherne. 2002. Critical loads of acidity for surface waters in south-central Ontario, Canada: regional application of the Steady-State Water Chemistry (SSWC) model. Canadian Journal of Fisheries and Aquatic Sciences **59**:1287-1295.

Henriksen, A. and M. Posch. 2001. Steady-state models for calculating critical loads of acidity for surface waters. Water, Air, & Soil Pollution: Focus **1**:375-398.

Henriksen, A., M. Posch, H. Hultberg, and L. Lien. 1995. Critical loads of acidity for surface waters: Can the ANC limit be considered variable? Water, Air, & Soil Pollution **85**:2419-2424.

Hodson, M. and S. Langan. 1999. Considerations of uncertainty in setting critical loads of acidity of soils: the role of weathering rate determination. Environmental Pollution **106**:73-81.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. VanDriel, and J. Wickham. 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. Photogrammetric Engineering and Remote Sensing **73**:337-341.

Hornung, M. 1985. Acidification of soils by trees and forests. Soil Use and Management **1**:24-27.

Hultberg, H. 1985. Budgets of Base Cations, Chloride, Nitrogen and Sulphur in the Acid Lake Gårdsjön Catchment, SW Sweden. Ecological Bulletins:133-157.

Johnson, C. E., M. I. Litaor, M. F. Billett, and O. P. Bricker. 1994. Chapter 14: Chemical weathering in small catchments: Climatic and anthropogenic influences. Pages 323-341 *in* B. Moldan and J. Cerny, editors. Chemical Weathering in Small Catchments: Climatic and Anthropogenic Influences. John Wiley & Sons Ltd.

Kump, L. R., S. L. Brantley, and M. A. Arthur. 2000. Chemical weathering, atmospheric CO2, and climate. Annual Review of Earth and Planetary Sciences **28**:611-667.

Li, H. and S. G. McNulty. 2007. Uncertainty analysis on simple mass balance model to calculate critical loads for soil acidity. Environmental Pollution **149**:315-326.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R News **2**:18-22.

Lloyd, J. and J. A. Taylor. 1994. On the Temperature Dependence of Soil Respiration. Functional Ecology **8**:315-323.

McDonnell, T., B. Cosby, T. Sullivan, S. McNulty, and E. Cohen. 2010. Comparison among model estimates of critical loads of acidic deposition using different sources and scales of input data. Environmental Pollution **158**:2934-2939.

McDonnell, T. C., B. J. Cosby, and T. J. Sullivan. 2012. Regionalization of soil base cation weathering for evaluating stream water acidification in the Appalachian Mountains, USA. Environmental Pollution **162**:338-344.

Miles, J. 1986. What are the effects of trees on soils? Pages 55-62 *in* D. Jenkins, editor. Trees and wildlife in the Scottish uplands. NERC/ITE.

National Acid Precipitation Assessment Program (NAPAP). 1991. 1990 Integrated Assessment Report. Office of the Director, Washington DC.

Nilsson, J. and P. Grennfelt. 1988. Critical loads for sulfur and nitrogen. Nordic Council of Ministers, Copenhagen, Denmark:31.

NRCS Soil Survey Staff. 2010a. Soil Survey Geographic (SSURGO) database for southern Appalachian Region.

NRCS Soil Survey Staff. 2010b. U.S. General Soil Map State Soil Geographic (STATSGO) database.

Olden, J. D., J. J. Lawler, and N. Poff. 2008. Machine learning methods without tears: A primer for ecologists. Quarterly Review of Biology **83**:171-194.

Omernik, J. M. 1987. Ecoregions of the Conterminous United States. Annals of the Association of American Geographers **77**:118 - 125.

Peltier, L. C. 1950. The geographic cycle in periglacial regions as it is related to climatic geomorphology. Annals of the Association of American Geographers **40**:214-236.

Posch, M., J. Hettelingh, and J. Slootweg. 2003. Manual for dynamic modelling of soil response to atmospheric deposition.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems **9**:181-199.

R Development Core Team. 2011. R: A language and environment for statistical computing. Vienna, Austria.

Reynolds, K. M., P. F. Hessburg, T. J. Sullivan, N. A. Povak, T. C. McDonnell, B. J. Cosby, and W. Jackson. 2012. Spatial decision support for assessing impacts of atmospheric sulfur deposition on aquatic ecosystems in the southern Appalachian Region. Pages 1197-1206 *in* Proceedings of the 45th annual Hawaii International Conference on System Sciences. IEEE, Wailea, Maui, Hawaii.

Ridgeway, G. 2010. gbm: Generalized Boosted Regression Models.

Skeffington, R. 2006. Quantifying uncertainty in critical loads:(A) literature review. Water, Air, & Soil Pollution **169**:3-24.

Sullivan, T., B. Cosby, K. Tonnessen, and D. Clow. 2005. Surface water acidification responses and critical loads of sulfur and nitrogen deposition in Loch Vale watershed, Colorado. Water Resources Research **41**:W01021.

Sullivan, T., J. Webb, K. Snyder, A. Herlihy, and B. Cosby. 2007. Spatial distribution of acid-sensitive and acid-impacted streams in relation to watershed features in the southern Appalachian Mountains. Water, Air, & Soil Pollution **182**:57-71.

Sullivan, T. J., B. J. Cosby, J. R. Webb, R. L. Dennis, A. J. Bulger, and F. A. Deviney. 2008. Streamwater acid-base chemistry and critical loads of atmospheric sulfur deposition in Shenandoah National Park, Virginia. Environmental Monitoring and Assessment **137**:85-99.

Sverdrup, H. and W. De Vries. 1994. Calculating critical loads for acidity with the simple mass balance method. Water, Air, & Soil Pollution **72**:143-162.

Sverdrup, H. and P. Warfvinge. 1993. Calculating field weathering rates using a mechanistic geochemical model PROFILE. Applied Geochemistry **8**:273-283.

U.S. EPA. 2009. Risk and exposure assessment for review of the secondary National Ambient Air Quality Standards for Oxides of Nitrogen and Oxides of Sulfur: Final.*in* U. S. EPA, editor. Center for Environmental Assessment, Office of Research and Development, Research Triangle Park, NC.

Warfvinge, P. and H. Sverdrup. 1992. Calculating critical loads of acid deposition with PROFILE—a steady-state soil chemistry model. Water, Air, & Soil Pollution **63**:119-143.

Whittingham, M. J., P. A. Stephens, R. B. Bradbury, and R. P. Freckleton. 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology **75**:1182-1189.

Winkler, J. P., R. S. Cherry, and W. H. Schlesinger. 1996. The Q10 relationship of microbial respiration in a temperate forest soil. Soil Biology and Biochemistry **28**:1067-1072.

Wright, R. F., J. Aherne, K. Bishop, L. Camarero, B. J. Cosby, M. Erlandsson, C. D. Evans, M. Forsius, D. W. Hardekopf, R. Helliwell, J. Hruška, A. Jenkins, J. Kopáček, F. Moldan, M. Posch, and M. Rogora. 2006. Modelling the effect of climate change on recovery of acidified freshwaters: Relative sensitivity of individual processes in the MAGIC model. Science of the Total Environment **365**:154-166.

TABLES

Table 3.1. Environmental covariates used in modeling BC_w rates (n=140) compared to conditions at over 4,000 randomly select catchments. Proportion overlap of the two kernel density functions estimated for each data set and each environmental variable was calculated. High overlap indicates good concordance between the two distributions.

Variable shortname	**Description**	**Proportion overlap**
AB90GROW	Mean number of days above 90°F during the local growing season	0.72
DDAYHMAX	Mean penultimate maximum degrees x days heating > 18°C	0.63
TDAYMEAN	Mean mean daytime temperature	0.62
PRECIPNG	Mean precipitation sum during the local non-growing season	0.68
PDAYMAX	Mean penultimate maximum days with precipitation (>0.3cm) while > 10°C	0.90
DIFF95GR	Mean 95th percentile of maximum diurnal surface temperature difference during the local growing season	0.61
GPPGROW	Mean gross primary production (GPP) integrated over the local growing season	0.80
GPPNG	Mean gross primary production (GPP) integrated over the local non-growing season	0.70
CON42	Percent contributing area in conifer cover	0.65
DECID41	Percent contributing area in conifer (1) + mixed (0.5) cover (weighted by # in parentheses)	0.80
FAC	Flow accumulation	0.79
LITH_CAR	Percent contributing area in carbonaceous lithology	0.81
LITH_SIL	Percent contributing area in siliceous lithology	0.64
SOIL_DEP	Mean soil depth	0.89
SOIL_CLAY	Mean soil clay (%)	0.65
SOIL_PH	Mean soil pH	0.70
OMNEW	Mean organic matter content	0.83
NITRONEW	Mean soil Kjeldahl nitrogen to 50 cm depth	0.66
S_WET	Mean wet sulfur deposition	0.87
S_DRY	Mean dry sulfur deposition	0.85

Table 3.2. Model validation statistics calculated for linear regression (LM), geographically weighted regression (GWR), multivariate adaptive regression splines (MARS), boosted regression trees (BRT) and random forest (RF) models. RMSE is root mean squared error, AIC is Akaike information criteria. The final column is the mean (± 1 SD) RMSE resulting from 10-fold cross-validation ($CV_{10\text{-fold}}$). All models included 10 predictor variables.

Landscape-only models (10-variable models)			
Model	R-squared	RMSE	$CV_{10\text{-fold}}$ RMSE
LM	51.2	29.6	30.6±9.1
GWR	56.2	28.1	30.2±9.7
MARS	73.8	21.7	34.9±12.3
BRT	81.8	18.1	31.8±11.2
RF	85.6	16.1	29.5±10.2

Landscape+chemistry models (10-variable models)			
Model	R-squared	RMSE	$CV_{10\text{-fold}}$ RMSE
LM	74.9	21.3	21.5±6.6
GWR	76.2	20.7	21.7±6.7
MARS	89.7	13.6	24.4±7.9
BRT	90.1	13.4	23.2±9.1
RF	91.5	12.4	21.3±8.3

Table 3.3. Model validation and variable importance measures for ecoregional models used to predict BC_w rates across the southern Appalachian study region. Numbers indicate the ranking of each variable in the model by its variable importance score within the random forest model. The (+/-) signs indicate the approximate positive or negative effect of the predictor on BC_w. Ecoregions are: Blue Ridge Mountains (BR), Central Appalachian (CA), and Ridge and Valley (RV). RMSE is the root mean squared error for models based on predictions made on the training data. Models were built using the top 7 predictor variables, including the ALL ecoregion model, which is shown for comparison.

Variable	**Description**	**BR**	**CA**	**RV**	**ALL**
S_WET	Wet sulfur deposition	7 (+)	7 (+)	2 (+)	2 (+)
LITH_SIL	Percent siliceous lithology	4 (-)		1 (-)	1 (-)
PRECIPNG	NGS precipitation	3 (+)		4 (+)	5 (+)
PDAYMAX	Maximum days with precipitation (when $\geq 10^oC$)	1 (+)			4 (+)
OMNEW	Amount of soil organic matter	6 (+)		5 (+)	3 (+)
SOIL_CLAY	Percent soil clay				7 (+)
SOIL_PH	Soil pH	2 (+)			
NITRONEW	Soil Kjeldahl nitrogen			3 (+)	6 (+)
NEWCTI	Compound topographic index		6 (+)		
LITH_FEL	Percent felsic lithology	5 (-)			
LSTGROW	Number of GS days $\geq 5.6^o$ C		2 (+)		
PSUMMAX	Maximum annual precipitation amount (when $\geq 10^oC$)		3 (+)		
AB90GROW	Number of days $\geq 32.2^o$ C		4 (+)		
GPPNG	NGS gross primary productivity			6 (-)	
RIGROW	Growing season respiration index			7 (-)	
CON42	Percent conifer forest cover		1 (-)		
FOREST	Percent forest cover		5 (-)		
Sample size		79	21	38	140
RMSE		15.7	23.8	15.2	16.8

Table 3.4. Root mean squared error (RMSE) rates for random forest models built with 7 predictor variables for each individual ecoregion and all ecoregions combined. Each of the resulting four models was then tested in turn on each ecoregions and all ecoregions combined. RMSE scores on the diagonal represent estimates of model error based on estimates made on the training data for that ecoregion.

	Training data			
Testing data	Blue Ridge	Central Appalachian	Ridge and Valley	All
Blue Ridge	15.7	46.5	31.5	14.9
Central Appalachian	54.1	23.8	51.5	26.4
Ridge and Valley	44.1	46.9	15.2	13.9
All	33.2	44.4	32.0	16.8

Table 3.5. Comparisons among random forest (RF) and linear regression (LM) model predictions across the study region. Values represent the percentage of 30-m grid within predefined BC_w classes.

BC_w classes	**Observed**	**RF**	**LM***
<20	4.3	0.0	2.2
20 – 50	27.1	4.0	4.8
50 – 100	48.6	67.6	47.7
100 – 150	14.3	26.8	42.1
>150	5.7	1.6	3.2
Sum	100.0	100.0	100.0

* See Appendix 3.3 for the associated map of BC_w predicted values using this model.

FIGURES

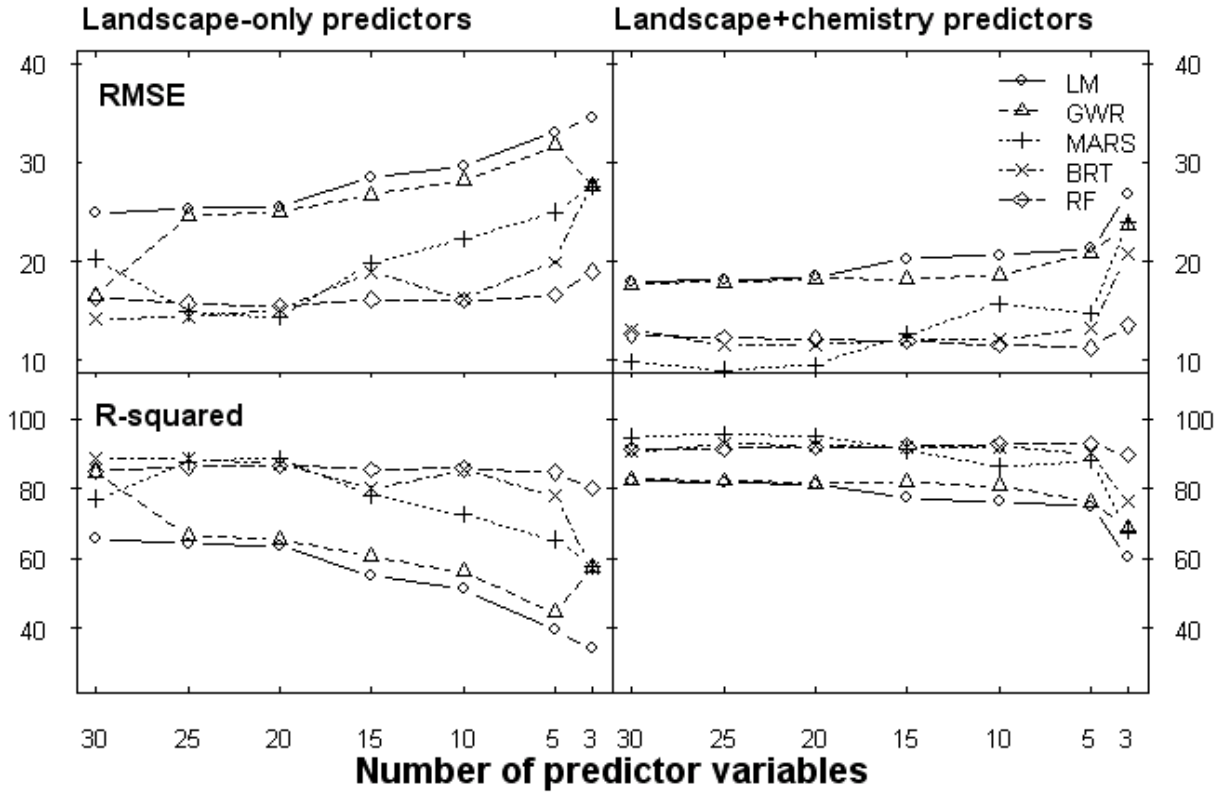


Figure 3.1. Error rates for five statistical models tested (LM = linear regression, GWR=geographically weighted regression, MARS=multivariate adaptive regression splines, BRT=boosted regression trees, RF=random forest).

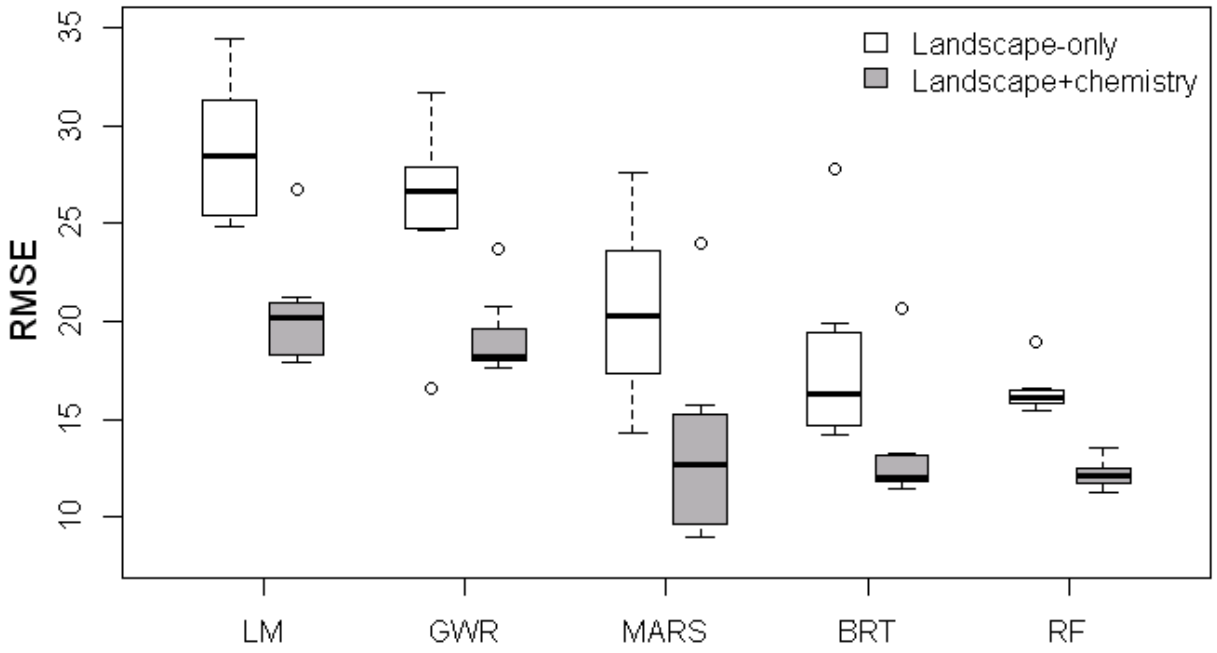


Figure 3.2. Boxplot of error rates averaged across all combinations of number of variables (i.e., from 33 to 3 for landscape-only models) entered into each statistical model. Abbreviations are the same as Figure 3.1. Grey boxes indicate water chemistry variables were included in the model and white boxes indicate landscape-only models.

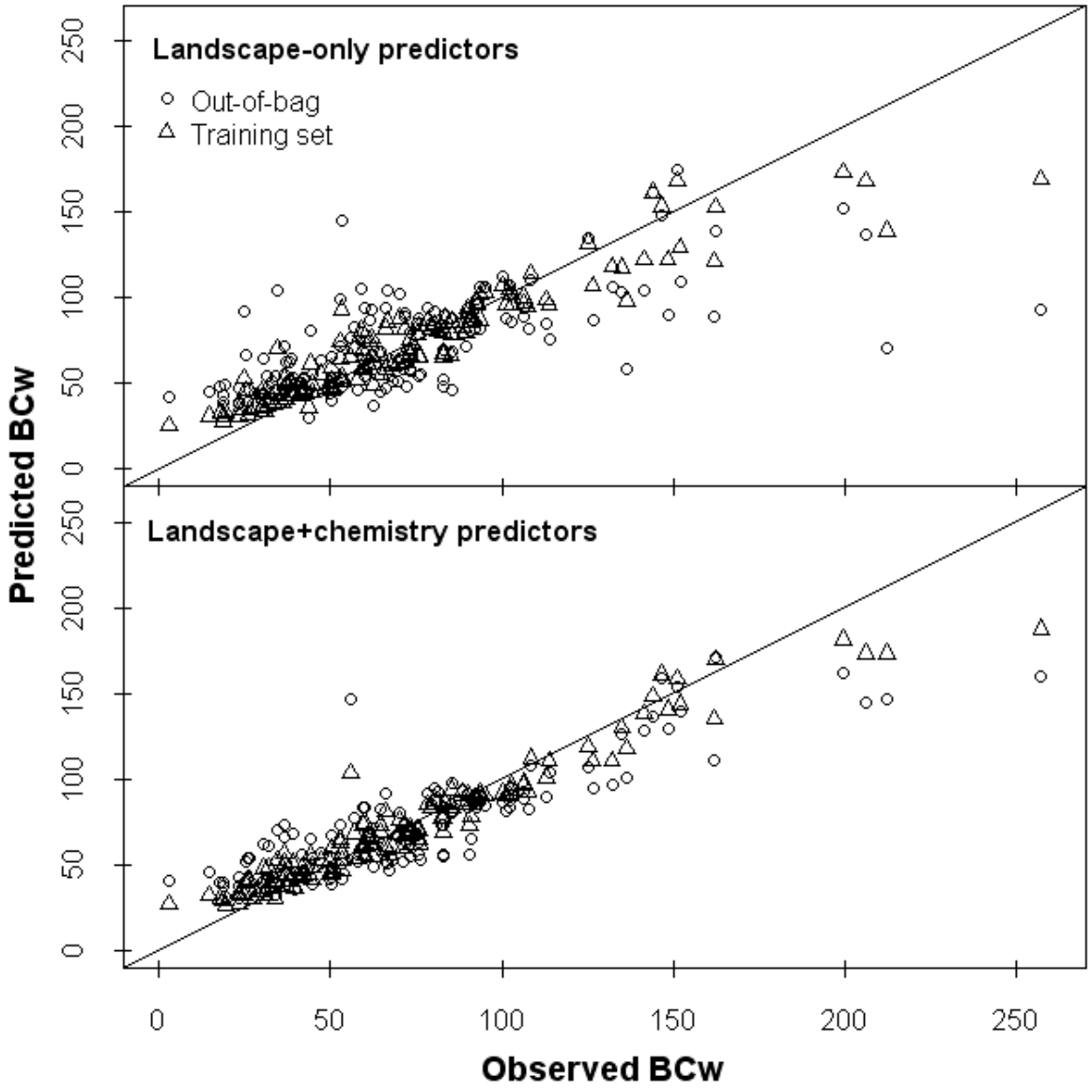


Figure 3.3. Final random forest model predictions of BC_w plotted against observed (MAGIC calibrated) BC_w. Black circles indicate out-of-bag predictions (similar to cross-validation, but see text) and black triangles indicate predictions made on the original training set of data.

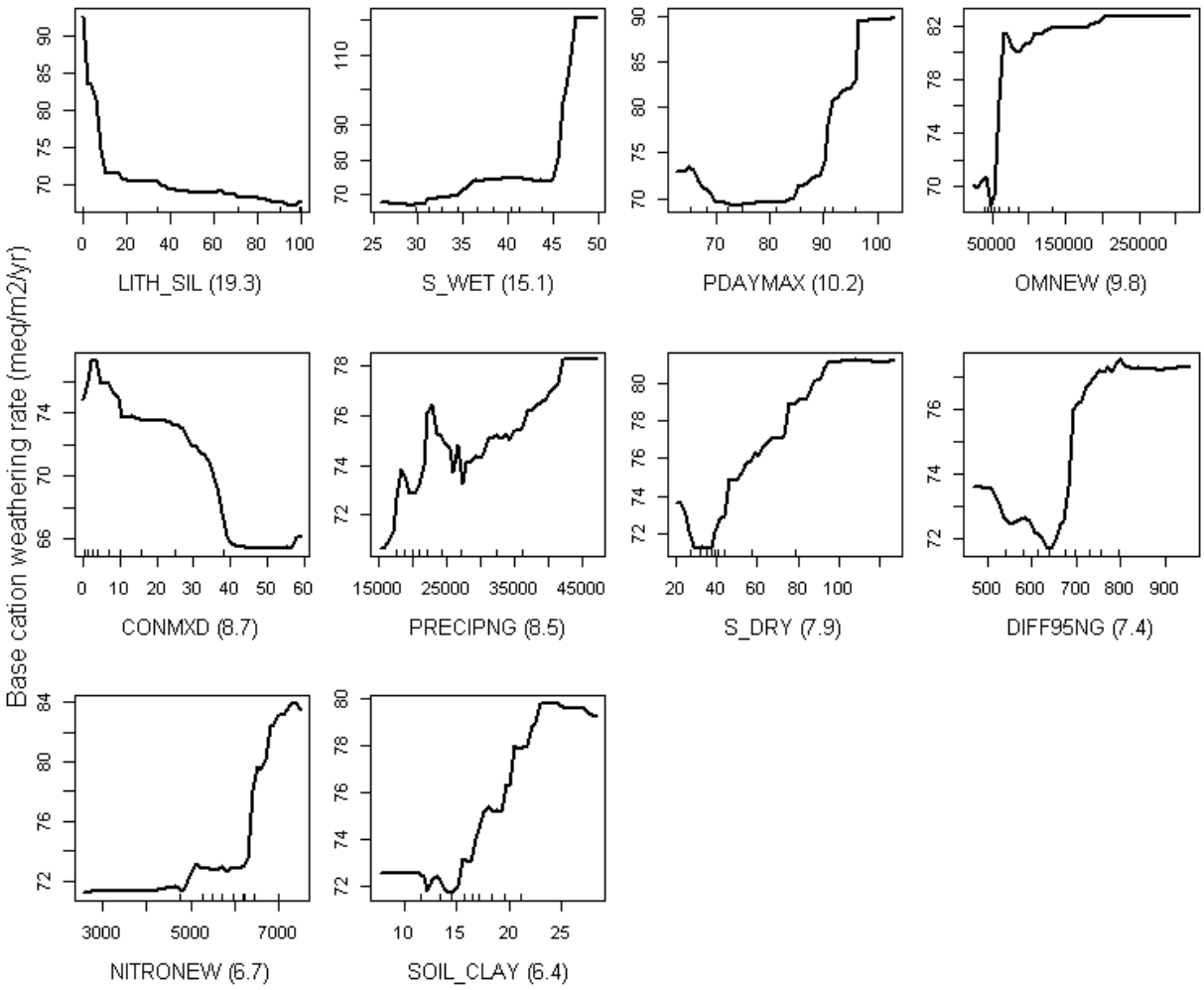


Figure 3.4. Response curves showing relations between predicted BC_w and individual predictor variables included in the final landscape-only RF model. Black tick marks (rug plot) on the x-axis indicate decile classes for the predictor variables. The y-axis indicates the relative effect of the predictor on BC_w. In general, higher y-axis values indicate higher predicted BC_w. Numbers in parentheses indicate the variable importance measure (i.e., mean decrease in squared error) scaled to sum to 100% over all variables included in the model. Variables are: S_WET, amount of wet sulfur deposition; LITH_SIL, percent of catchment in siliceous lithologies; PRECIPNG, amount of precipitation in the non-growing season; PDAYMAX, mean penultimate maximum days with precipitation; S_DRY, amount of dry sulfur deposition; CONMXD, percent of catchment in mixed-conifer; SOIL_CLAY, mean percent clay of catchment soils; OMNEW, mean soil organic matter; DIFF95NG, Mean 95th percentile of maximum diurnal surface temperature difference during the local non-growing season; NITRONEW, mean soil Kjeldahl

nitrogen to 50 cm depth. Note that the majority of climate variables have been multiplied by a constant, which had no effect on predictions or model performance.

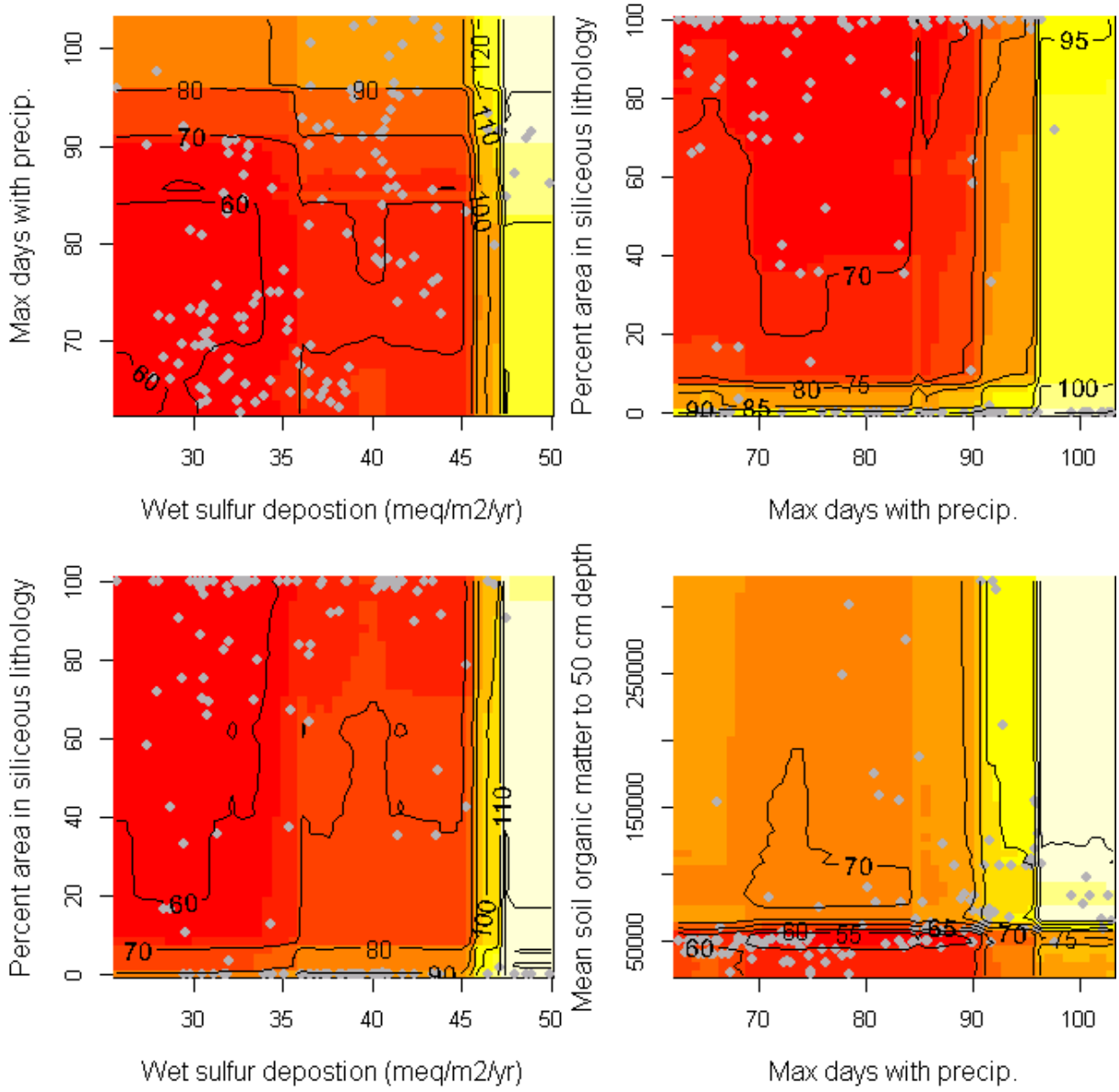


Figure 3.5. Surface plots displaying interactions between two continuous predictor variables used in modeling BC_w rates with landscape predictors only included in the model. Top interactions were calculated for final RF models using methods similar to those of Elith et al. (2008). Contour labels are predicted BC_w rates. Grey points represent the environmental conditions of the MAGIC modeled BC_w sampled catchments.

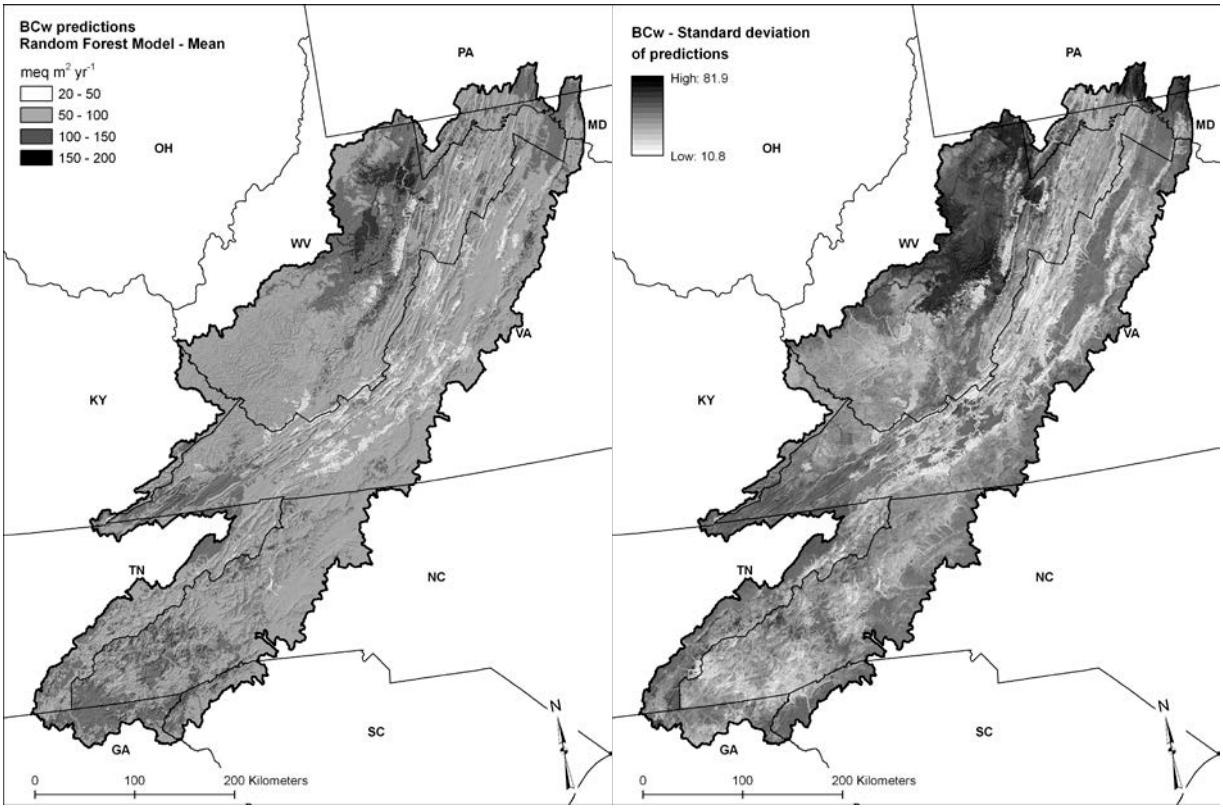
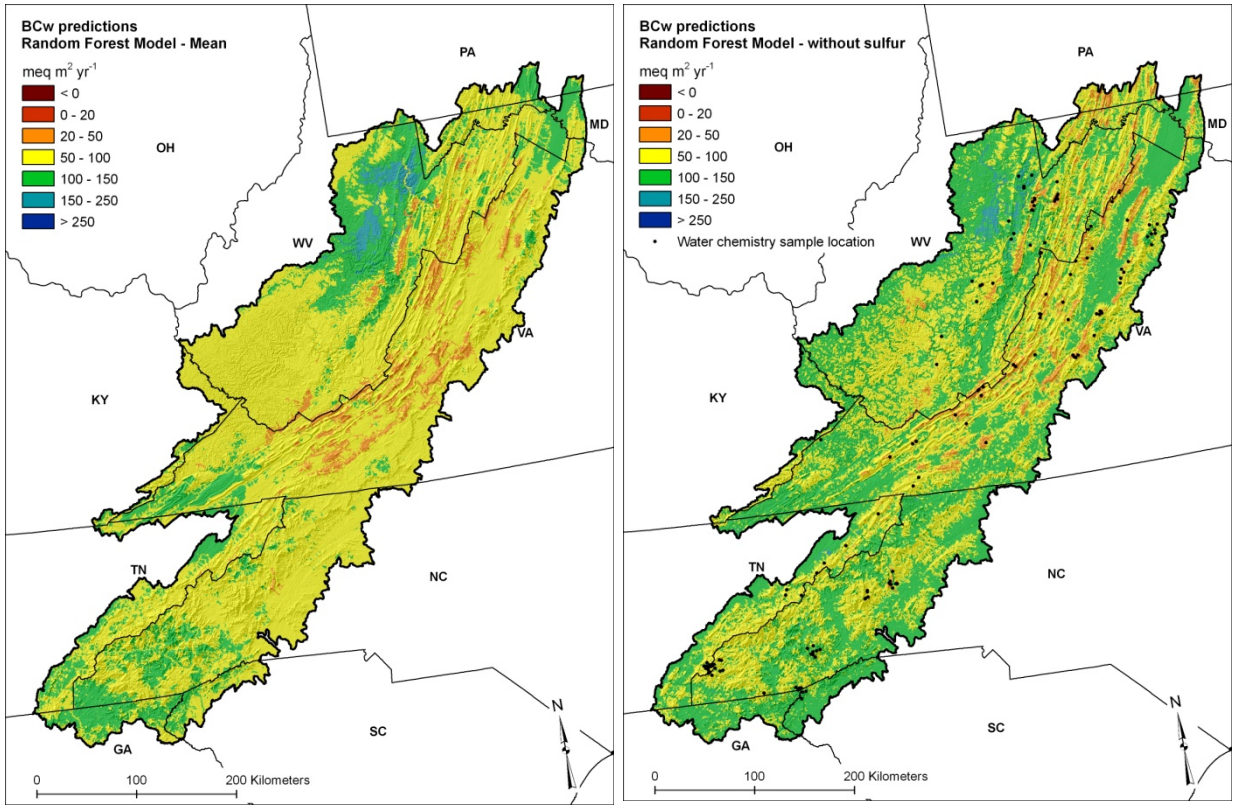
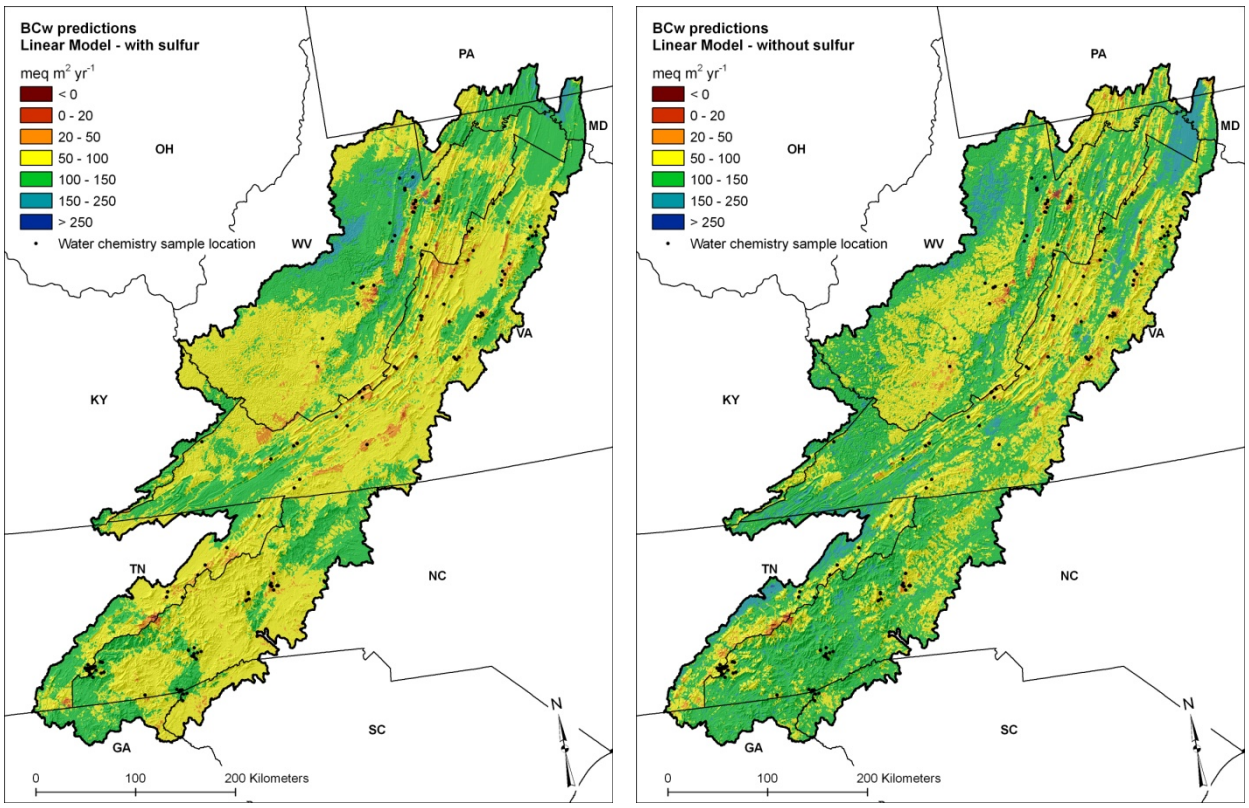


Figure 3.6. Continuous surface of predicted BC_w (left panel) and standard deviation of predictions (right panel) from the final landscape-only model. Standard deviations were calculated from the predictions made from the 1000 individual regression trees within the RF algorithm.

APPENDICES



Appendix 3.1. Geographic distribution of BCw predictions made using **random forest models** (left) including sulfur deposition predictors, and (right) without S deposition predictors. The scales on the distribution of values are consistent across all maps.



Appendix 3.2. Geographic distribution of BCw predictions made using **linear regression models** (left) including sulfur deposition predictors, and (right) without S deposition predictors. The scales on the distribution of values are consistent across all maps.

CHAPTER 4 - NEST SITE SELECTION PREFERENCES AND NICHE DIVERSIFICATION AMONG FOUR CAVITY-NESTING BIRDS IN DRY MIXED-CONIFER FORESTS

ABSTRACT

Cavity-nesting birds (CNB) are keystone species of forested ecosystems that excavate cavities in trees, which provide habitat for a variety of other species. Snags are necessary components of CNB habitat, and nest sites are selected based on characteristics of the nest snag, proximity to other resources, and levels of predation and competition. Specific requirements in nest-site selection for CNBs are largely unknown and likely vary by region. We used multi-scale habitat models, and niche-overlap analysis to identify nest-site selection patterns for four CNB species within a dry mixed-conifer forest in eastern Washington that had not experienced recent wildfire. Habitat was quantified using a georeferenced census of >12,000 snags, and remotely sensed raster- and patch-based vegetation data summarized to two spatial scales around each snag. Distance to nearest CNB nest was calculated for every snag and used as a predictor of habitat selection. Habitat modeling results indicated that most of the variance in nest-site selection was explained at the snag scale, and selection was less influenced by multi-scaled vegetation patterns. Nests were generally located in live ponderosa pine trees with little decay, and with past excavations. When nearest-nest distances were included as predictors, total variance explained increased by >18% for all models. Most CNB species nested in significantly different habitat from one another, and selection preferences likely varied with foraging and nesting requirements.

Results suggest that nests were spatially distributed to minimize both inter-bird interactions and energy expenditures related to foraging, nest building, and maintenance.

1. INTRODUCTION

Ecologists are often interested in modeling species resource-use patterns to better understand the patterns and processes driving habitat selection (Whitham 1980, Manly 2002). New insights can identify important selection criteria, mechanisms for the cohabitation of species, and key patterns and processes that create suitable habitat. Ultimately, these insights can guide restoration practices required to increase the amount and arrangement of limiting resources (Bull et al. 1997, Bunnell et al. 2002).

Understanding species-habitat relations of cavity-nesting birds (CNB) is especially relevant as these species play an important keystone role in many forested systems (Daily et al. 1993, Bednarz et al. 2004, Virkkala 2006), and populations of some species are declining (Newton 1994, Eadie et al. 1998). The excavations made by CNBs provide habitat for an estimated 25-30% of vertebrate species in the Pacific Northwest (Bunnell et al. 1999), and CNBs contribute significantly to the overall species richness and abundance within a community (Short 1979, Martin et al. 2004, Virkkala 2006). Recent declines in some CNB species have been documented, for example pileated (*Dryocopus pileatus*) and black-backed woodpeckers (*Picoides arcticus*) are listed species of concern (Sauer et al. 2011). CNBs are particularly sensitive to changes in forest structure and composition following harvests or natural disturbances (Klenner and Huggard 1998, Wightman and Germaine 2006); current declines are attributed to alterations in CNB habitat from past management (Dobkin et al. 1995, Eadie et al. 1998, Blewett and Marzluff 2005, Guinan et al. 2008). Identifying key species-habitat relations can help better incorporate fine-filter management approaches in the context of broader

management goals (Bull et al. 2005, Knapp et al. 2005, Stephens and Moghaddas 2005, Lehmkuhl et al. 2007, Ritchie et al. 2013).

Research into specific habitat preferences for CNBs has focused largely on snag (i.e., standing dead trees or trees with dead portions, sensu Thomas et al. 1979) characteristics and demographics, as these structures are often assumed to be most limiting in managed forests (Zarnowitz and Manuwal 1985, Brawn and Balda 1988, Newton 1994). Snags are a necessary component of CNB nesting, roosting, and foraging habitat (Conner et al. 1975, Thomas et al. 1979, Hunter 1990). While most managers acknowledge the important functional roles snags play in sustaining CNB populations, many prescriptions rely on simple density guidelines with little attention given to protecting and conserving snags that satisfy specific CNB habitat requirements (Thomas et al. 1979).

For instance, preferences for specific snag species, sizes, and decay level appear to influence CNB nesting and foraging patterns (Scott 1978, Short 1979, Newton 1994, Ganey and Vojta 2004). In some areas, a lack of preferred nest trees may limit CNB populations even where snags are otherwise abundant (Raphael and White 1984, Cody 1985, Brawn and Balda 1988, Newton 1994). Furthermore, individual species preferences are not well understood, and may change across regions, forest types, or with the level of competition for limiting resources (Jackson et al. 2002, Guinan et al. 2008, Wiebe and Moore 2008, Gyug et al. 2012). Identifying specific mortality or decay agents related to preferred CNB nest sites may help improve our understanding of the mechanisms associated with resource selection, and lead to more effective management (Walter and Maguire 2005).

The spatial context of snags may also influence nest site selection (Gutzwiller and Anderson 1987, Kotliar and Wiens 1990). In choosing a nest location, a CNB attempts to satisfy multiple resource requirements at once, given the prevailing environmental conditions, levels of intra- and inter-specific competition, and predation (Holt and Martin 1997, Saab 1999, Rodewald et al. 2001, Cushman and McGarigal 2002). Birds must consider not only the suitability of the potential nest snag itself, but also its environmental setting, and proximity to reliable foraging habitat (Gutzwiller and Anderson 1987, Li and Martin 1991, Weikel and Hayes 1999). Understanding nest selection, therefore, requires information not only about the nest snag itself but also of the surrounding neighborhood (Cushman and McGarigal 2004, Meyer 2007). Multi-scale habitat models are increasingly being used to identify resource selection patterns (Bednarz et al. 2004, Battin and Lawler 2006). Diagnostic tests, such as variance partitioning, can be used to decompose the contribution of each scales' conditions to resource selection (Cushman and McGarigal 2002), which can lead to a better understanding of the features and spatial scales most influential to habitat selection.

Social interactions among neighboring individuals are also influential to observed resource use (Slagsvold 1980, Ahlering and Faaborg 2006). Interactions can be positive (e.g., conspecific or heterospecific attraction), or negative (e.g., competitive exclusion, territorial), and can relegate certain individuals to sub-optimal environments (Van Horne 1983). For instance, a CNB may select a habitat near other conspecifics to reduce the search time required to locate adequate habitat (Ahlering and Faaborg 2006). Competitive interactions may cause an individual to inhabit low quality habitat or search elsewhere for a nest site – particularly where significant overlap exists in resource use among species (Fretwell and Lucas 1969). Information on neighboring competitors is often excluded from habitat models (Johnson 2007), yet predictions

from these models can overestimate suitable habitat, and lead to potentially spurious conclusions about nesting preferences (Van Horne 1983, Johnson 2007).

Here, we examined the influence of (1) snag size, fungal decay, and other abiotic damages, (2) multi-scaled vegetation patterns, and (3) social interactions on nest-site selection preferences for four CNB species in a dry mixed-conifer forest that had not experienced recent wildfire. We censused and georeferenced all potential nest snags within seven forested drainages, and summarized vegetation conditions at two spatial scales around each censused snag. These data were used to develop multi-scale habitat models for each CNB species and all species combined. We included both habitat features and inter-nest distances (i.e., a surrogate for social interactions among CNBs) as predictors to directly test their contribution to nest-site selection. Due to imbalances in the ratio of non-nest to nest snags (>100:1), we tested the influence of two data resampling techniques and a machine learning algorithm (random forest) on model performance. In a separate analysis, a niche overlap analysis was used to identify significant differences in nesting habitats among the four CNB species.

2. METHODS

2.1. Study area description

This study was conducted in the eastern Cascade region of Washington State, USA (48.556, -120.096; Figure 4.1). Elevations range from 307-1052-m, and the climate is continental with a mean annual temperature of 7.2^{o} C. Mean monthly range is -6.3^{o} C (January) to 19.6^{o} C (July) (PRISM climate data, Daly et al. 2000). Annual precipitation averages 37.6 cm, with about two-thirds falling as snow between November and March.

Seven study units, ranging in size from 150-400 ha (cumulative area 2,107 ha), were located within a 51.5 km long and 29 km wide area (Figure 4.1). Study sites were chosen from a pool of 13 potential units based on size (>150 ha), accessibility, and forest type. Forest type was dry-mixed conifer with shrub and grass understories. Units were dominated by tree cover composed primarily of ponderosa pine (*Pinus ponderosa*) and Douglas-fir (*Pseudotsuga menziesii*). Dominant understory species were snowberry (*Symphoricarpos albus*), spirea (*Spirea* spp.), serviceberry (*Amelanchier alnifolia*), and chokecherry (*Prunus virginiana*).

2.2. Data description

2.2.1. Cavity nesting bird surveys

A systematic survey of CNB nests was conducted throughout two spring breeding seasons (May -June of 2002 and 2003). Variable width belt transects were established every 200-m within each unit. Observations of eggs, young, or parental incubation exchanges made using cavity viewers or auditory cues were considered evidence of nesting. Nests were monitored every three to four days to determine the fledgling success. Each annual survey was conducted independently for each unit, and the data were combined for all analyses unless stated otherwise: any nest tree

occupied in both years, or doubly occupied in a single year was recorded as two separate observations. Each nest tree was georeferenced using a Global Positioning System (GPS). Nest-specific data recorded were CNB species, nest height, cavity orientation and the original excavator of the cavity (if known).

CNB species monitored across the two breeding seasons were the woodpeckers Lewis's (*Melanerpes lewis*), black-backed (*Picoides arcticus*), three-toed (*P. tridactylus*), white-headed (*P. albolarvatus*), hairy (*P. villosus;* HAWO), downy (*P. pubescens*), and pileated (*Dryocopus pileatus*); Williamson's (*Sphyrapicus thyroideus;* WISA) and red-naped sapsuckers (*S. nuchalis*); northern flicker (*Colaptes auratus;* NOFL), and western (*Sialia mexicana;* WEBL) and mountain (*S. currucoides*) bluebirds. Other CNB species were not included in the bird survey as they were not primary cavity nesters or species of conservation concern.

Individual species analyses were conducted for the hairy woodpecker (HAWO, N=20), northern flicker (NOFL, N=34), Williamson's sapsucker (WISA, N=40), and western bluebird (WEBL, N=18); models were not developed for other species due to low sample sizes ($\leq$5; Table 4.1).

2.2.2. Snag surveys

We conducted a complete census of snags during the summer of 2003. We recorded and mapped all snags and live trees with dead tops, strip-kill, or heart rot (hereafter, snag; Thomas et al. 1979) that were >2-m in height and >20-cm diameter (DBH). Data collected for each snag included tree species, DBH, height, presence of bird foraging activity , presence of new or old cavities, snag decay class (9 classes, sensu, Thomas et al. 1979), dead top decay length and class, evidence of strip-kill, and tree insect and disease damage(s). Snags were mapped with Wide Area

Augmentation System (WAAS)-enabled Garmin GPS map 76S GPS receivers, using point averaging. Horizontal accuracy of the snag coordinates is ±2.1-12.70-m depending on canopy density and terrain, with an average error of ± 6.20-m (Wing 2008). In total, 12,547 snags were located, measured, and georeferenced across the seven study units. These snags were considered as available habitat in both 2002 and 2003, although we acknowledge that some snags may have been added or lost between sampling periods.

2.2.3. Vegetation data

Vegetation conditions were characterized using 0.24-m resolution 3-band color digital orthoimagery (Appendix 4.1). We generated the orthoimagery by georeferencing high-resolution (15 μm) digitally-scanned aerial photographs. Ortho-rectification was conducted with ERDAS Imagine software (Leica Geosystems 2006), and the orthoimagery was horizontally georeferenced to 1998 USDA Forest Service 10-m digital ortho-quadrangles, and vertically to a USGS 1/3 arc-second (10-m) Digital Elevation Model (DEM).

Two vegetation products were derived from the orthoimagery (Appendix 4.1). First, we derived a raster image (0.24-m) of the orthoimagery to develop a high-resolution discrete vegetation classification; and second, an unsupervised polygon delineation (0.5 ha minimum mapping unit (MMU)) was developed and a photo-interpretation was used to quantify forest structure attributes. Both data sets were summarized to 60-m (1.13 ha) and 300-m (28.27 ha) radius neighborhoods around all censused snags. These two scales represented approximate lower and upper bounds on breeding season territory size for the four CNB species modeled (Thomas et al. 1979, Gutzwiller and Anderson 1987, Elchuk and Wiebe 2003).

Raster-based vegetation classification – An image classification was conducted to derive a vegetation raster (0.24-m resolution) with seven discrete classes. The color orthoimagery was prepared in ERDAS Imagine 2011, and classification was conducted in eCognition Devloper 8.4 software (Trimble 2011). Image objects were created with the multi-resolution segmentation algorithm, and the resulting image objects were classified through an iterative supervised fuzzy logic method within the eCognition environment. Resulting classes were conifer, shrub, grass, shadow, bare ground, rock, and water. Misclassification rates were not quantified, but visual assessments suggested high accuracy (Appendix 4.1) and conifer cover summarized to the delineated patches (see below) was similar to the photo-interpreted cover estimates (Pearson's r >0.9 across units).

Patch-based classification – Patch delineation (0.5 ha MMU) was completed primarily within the eCognition Developer 8.7 environment using the multi-resolution segmentation algorithm. Four datasets were used to delineate patches of like vegetation: (1) 2-m, three-band color orthoimagery (resampled from the 0.24-m orthoimagery), (2) 30-m Landsat TM (NASA Landsat Program 2011) 1998 imagery (red and near infrared bands, and derived normalized differenced vegetation index (NDVI)), (3) hillshade derived from 10-m DEM, and (4) aspect derived from 10-m DEM. Delineations were edited within ArcGIS 10.0 software, where patches were split or merged based on visual inspection of forest and non-forest vegetation attributes. The resulting polygon feature class contained a total of 998 photo-interpreted vegetation patches with a median size of 2.7 ha (1.4-5.3 ha interquartile range (IQR)).

The polygon attributes were interpreted from the 0.24-m color orthoimagery to quantify percentage total tree cover, percentage overstory tree cover, mean crown diameter of overstory trees, mean crown diameter of understory trees, number of canopy layers, degree of spatial tree

clustering, distance to patch edge, and distance to patches with tree cover <30%. Tree densities were determined by counting all stems within a 50x50-m square while viewed at a map scale of 1:300. Tree crown diameters were estimated using the measure tool in ArcGIS at the same scale for 5-15 trees, depending on the size distribution of tree crowns within each canopy layer. Percent cover was estimated as the proportion of the polygon covered by tree crowns, as viewed from above.

2.2.4. Spatial environmental data

Topographic data were calculated from a 10-m DEM, and included aspect, slope steepness (%), curvature (parallel to slope, perpendicular to slope and total), and standardized topographic position index (STPI) at 100 and 150-m radius neighborhoods (Weiss 2001). STPI is an index that quantifies topographic position of a DEM cell in relation to its surrounding. The classification ranges from -4 to 4, with positive values representing upper slope and ridge positions, negative values representing lower slope and valley positions, and values near zero represent flat areas or those with constant slope. Potential total incoming solar radiation was also calculated with the ArcGIS 10.0 Solar Analyst extension for three time periods March – May, April – June, and annual for 2003.

Following Plotkin et al. (2000), we quantified "clusters" of snags across each unit, with a clustering algorithm that uses a predefined inter-snag distance (D_{crit}) to determine cluster membership, where all snags within D_{crit} are part of a cluster. The algorithm sequentially collects all neighboring snags of each successive cluster member until no new snags are within D_{crit}. Clusters were created using D_{crit} = 5, 10, 15, and 25-m.

Distance to patch edge and distance to the nearest patch with <30% total canopy cover were calculated for each snag using the patch-based vegetation dataset.

The relative abundance of bark beetles and woodborers for each unit as a foraging resource was also estimated. For each snag in the database (12,547), the surface area of the bole that exhibited signs or symptoms of bark beetle or woodborer attack was estimated in the field with birding binoculars by recording the length of attack along the tree bole, and the diameter at the base and upper terminus of the attack. Surface area of tree boles with bark beetle or woodborer attack was calculated using the upper and lower diameters, and assuming the shape of the bole followed that of a frustum. A point density function was then used to calculate a continuous 10-m raster surface where each cell represented the density of bark beetle or woodborer infested tree bole surface area. The analysis was repeated using search radii of 75, 150, and 300-m. Because the 2003 survey recorded the year of beetle attack (present year, previous year, and prior to previous year), density surface rasters were generated that corresponded to the years of CNB survey with present year (2003), previous year (2002), and older attacks (prior to 2002), and for woodborers that attacked in the year of the survey (2003).

In total, predictor variables were collected at four spatial scales: (1) the snag scale, which described characteristics of the snag itself, (2) a "local" scale, which quantified the juxtaposition of snags to features such as patch boundaries, snag clusters, bark beetle densities, nearest CNB nests, etc., (3) a "proximate" scale, which represented the remotely-sensed vegetation data summarized to 1.1 ha (60-m radius) vicinities around each snag, and (4) a "neighborhood" scale, which summarized vegetation data to 28.3 ha (300-m radius) neighborhoods around each snag (Figure 4.2; Appendix 4.1).

2.3. Data analyses

2.3.1. Snag and CNB nest site characteristics

Mean differences in continuous snag, vegetation, and environmental variables (e.g., snag diameter, mean tree density, elevation) for nest versus available snags were tested using a non-parametric Wilcoxon-Mann-Whitney U test. Chi-squared analysis was used for categorical predictor variables. Available snags were sampled at 100-m intervals across each unit to reduce sample size imbalances between nest and available snags. Parametric analysis of variance (ANOVA) was used to compare significant differences in continuous variables among the four main CNB species. Significance was based on an α-level of 0.05 with Tukey's correction for all pairwise comparisons. See Appendix 4.2 for the set of all predictor variables, and Appendix 4.3 for all R packages used in analyses.

2.3.2. Habitat modeling

Binomial (presence-absence) regression models were used to (1) predict the probability of CNB occupancy for all georeferenced snags in the study region, (2) identify key covariates of nest selection, and (3) determine the spatial scale(s) most influential to nest site selection. The response variable for all models was 0 = available (reference) snag without a nest, and 1 = nest snag. The analysis workflow described below is summarized in Appendix 4.3.

For each of the four main species and all CNB species combined, models were designed for (1) each of the four spatial scales, where model parameters and number of predictors was selected based on specific performance measures (see below), and (2) all factorial combinations at each spatial scale were combined to determine the amount of variance explained within and

shared among spatial scales (i.e., cross-scale correlations) using variance decomposition (Whittaker 1984, Cushman and McGarigal 2002, Lawler and Edwards 2006).

In addition to the habitat predictors described above, we built models with and without predictors that quantified the distance of each snag to the nearest CNB nest snag. These models were used to directly test their importance in influencing nest selection relative to other habitat features. Distances were calculated from each snag to the nearest nest occupied by each of the four main CNB species separately, and all CNB species combined. Distances were measured separately for nests in 2002 and 2003; distances from non-nest snags (in both 2002 and 2003) were measured to 2003 nest snags only.

Multicollinearity in predictors (Appendix 4.2) was reduced prior to model development. Predictors with a Pearson's r correlation score ≥ 0.7 were submitted to simple logistic regression models, and the variable with the higher AIC score was eliminated.

Statistical model -- Logistic regression (LR) and random forest (RF; Breiman 2001, Liaw and Wiener 2002) models were used to model CNB habitat-use patterns in a side-by-side comparison. RF is an ensemble version of traditional classification and regression tree (CART) analysis (De'ath and Fabricius 2000), where hundreds to thousands of CART trees are built using only a portion of the predictors and a portion of the training data. Models are evaluated using out-of-bag error rates, similar to cross-validation. Out-of-bag samples are those data left out of the training set of individual RF trees.

Data resampling -- Because CNB nest trees represented <1% of the total snag population, we reduced data imbalances by using combinations of two resampling techniques: over-sampling

(i.e., randomly replicating) nest snags, and under-sampling (i.e., randomly eliminating) available snags.

The Synthetic Minority Oversampling Technique (SMOTE; Chawla et al. 2002) was used to oversample the number of nest snags. For each nest snag, the SMOTE algorithm (1) used k-means clustering to identify a single nearest neighbor nest site in data space, (2) subtracted the vector of predictor variables of the target nest tree from the nearest neighbor, (3) multiplied the differences by a random number between 0-1, and (4) added this to the original vector of predictors, to create a synthetic replicate nest site. We chose three levels of SMOTE sampling for each model: no resampling, low resampling (oversample until the number of nest sites equaled 15% of the number of reference snags), and high resampling (oversampling to achieve 30% of the reference snag sample size).

Under-sampling proceeded by randomly selecting available snags at intervals of 50, 100 and 200-m, so that each unit was evenly sampled (Appendix 4.4). Resultant snag sample sizes were approximately 3200 (50-m intervals), 1300 (100-m intervals), and 500 (200-m intervals).

Variable selection – Models were evaluated using the top 10, 8, 6, 4, and 2 predictors. Variable importance was based on z-scores for individual LR coefficients. RF variable importance is assessed by randomly permuting the values of individual predictor variables, recalculating out-of-bag error rates and calculating the average decrease in model performance when that variable is randomly shuffled. The goal of the modeling was to balance model parsimony with prediction.

Model validation –Performance at each spatial scale was evaluated using sensitivity and specificity rates. In general, models with high levels of oversampling and many predictor variables exhibited high sensitivity, but low specificity. To balance both measures, only models

with specificity rates >85% were included as potential final models. We determined *a priori* that this rate was acceptable given the ample supply of potential nest sites (i.e., <1% snags were occupied by CNBs). To balance model parsimony with prediction, models that included fewer predictors, had sensitivity scores within 15% of the top performing model, and had a specificity score >85% were chosen over other more complex models.

Variance decomposition – Variance decomposition was used to evaluate the variability explained by each spatial scale within the habitat models and shared among scales (cross-correlation; Whittaker 1984, Cushman and McGarigal 2002, Lawler and Edwards 2006). Spatial scales consisted of the snag, local, proximal, and neighborhood scales. A total of 15 models were developed such that all factorial combinations of single, two-way, three-way, and four-way combinations of spatial-scale predictors were included. We used Efron's R^2 (Efron 1978) as a measure of the variance explained in nest site selection:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{\pi}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

Where y_i is an observed (0 or 1) response, $\hat{\pi}_i$ is the predicted probability of the i^{th} observation, and $\bar{y}$ is the mean of the responses. Details on the calculation of pure and shared variance are found in Lawler and Edwards (2006).

Model selection - Each combination of under-sampling, over-sampling, and choice of model (three over-sampling distances * three under-sampling SMOTE levels * two models = 18 models) was replicated 10 times and model performance averaged across replications (Appendix 4.3). Models were chosen for each species and all species combined, which balanced total variance explained with variance shared among spatial scales (i.e., cross-scale correlation).

Model predictions – The above methods resulted in 10 models per species (and combined species), and predicted probabilities were averaged across all 10 model iterations. The probability cutoff used to determine whether a snag was classified as a nest snag was selected to meet the sensitivity and specific criteria outlined above.

2.3.3. Niche overlap modeling

Niche overlap estimates identify similarities in resource-use patterns among species by quantifying the proportion of niche space shared by another competing species (Geange et al. 2011, Rödder and Engler 2011). Niche overlap was calculated for the four main CNB species following the methods of Geange et al. (2011). Empirical distributions were estimated for each environmental variable and for each species, using non-parametric density estimation, and the proportion of overlap in the distributions was calculated. Overlap values were averaged across the environmental variables and values ranged between 0 (no overlap) and 1 (complete overlap). We estimated weighted averages and standard deviations with weights equal to the cumulative number of times a variable was included in the niche models (see Results).

We used a neutral model permutation to test for significant differences in resource-use patterns among species. Pseudo-data were created by assigning each individual bird in the database to a randomly drawn (without replacement) vector of habitat conditions, and NO values for each species pair were calculated as described above. A P-value for each pairwise species comparison was then calculated as the proportion of permuted NO scores $\leq$ observed NO scores. A significant finding ($P \leq 0.05$) indicated that the niche overlap between two species was lower than would be expected from random, and species were differentiated in niche space.

A measure of the distribution of resource use among all species across the total niche space was also calculated. This measure accounted for both mean and variance in niche overlap. Low variability in pairwise niche overlap scores indicated an even distribution across niche space, whereas high variability indicated clustering in niche space. The clustering index is:

$$n = \frac{s^2}{\bar{x}(1-\bar{x})}$$

Where n is the clustering index, $\bar{x}$ is the mean, and s^2 is the variance in niche overlap values among species (Geange et al. 2011). Values of the index ranged from 1 (clustered pattern) to 0 (evenly distributed pattern). Evidence of significant patterns was determined using a similar permutation procedure described above (see also, Geange et al. 2011).

3. RESULTS

3.1. Summary vegetation statistics

Dry mixed-conifer forest was the dominant vegetation cover across all seven study drainages, averaging 53 (±10.8)% (±SD across units) of the study area. Grass and shrub cover covered 23 (±5.6)% and 21 (±7.3)%, and bare ground and rock accounted for <0.5% of the remaining area. Forest cover was variable across the seven study units with >60% cover in the two northern units and 37% in the units to the south (Figure 4.1). From the photo interpretation, mean unit-level live-tree density ranged from 107-230 trees ha^{-1}. Non-forest (0% tree cover) patches represented 26-33% of the area within the units, and patches with <30% tree cover, 33-58%.

3.2. Snag survey summary

Snag (>20.3 cm DBH) densities averaged 5.1-6.5 ha^{-1} across the units, and snag density >50 cm diameter at breast height (DBH) ranged from 1.8-3.1 ha^{-1}. Median snag diameter was 45.7 cm

(35.6-55.9 inter-quartile range (IQR)), and median snag height was 18.3 m (4.6-22.9 IQR). Approximately 65% of snags were ponderosa pine (*Pinus ponderosa*; 3.9 ha^{-1}) and 30% were Douglas-fir (*Pseudotsuga menziesii*; 1.8 ha^{-1}). The remaining 5% of snags were western larch (*Larix occidentalis*), lodgepole pine (*Pinus contorta*), and *Populus spp* (quaking aspen and cottonwood). A total of 5162 snags (2.4 ha^{-1}) were in decay classes 5-8, and represented 41% of all snags.

3.3. Cavity-nesting bird densities

A total of 123 CNBs of the selected study species were identified across the 2,107 ha study area over the two years of the study (50 in 2002 (0.024 ha^{-1}), and 73 in 2003 (0.035 ha^{-1}); Table 4.1). Accounting for snag reuse and double occupancies, 101 individual snags were occupied by the CNB species monitored over two years; <1% of the population of available snags. Populations of individual CNB species were generally $\leq$5, and WISA, WEBL, HAWO, and NOFL represented >90% of the population (Table 4.1). Only these four species were included in habitat modeling.

3.4. Nearest neighbor nest distances

Nearest-nest (NN) distances for heterospecific CNBs ranged between 0-65 m, and NOFL+WEBL, HAWO+WISA, and NOFL+WISA generally nested closer to one another compared to other species pairs (Table 4.2). Conspecific NN distances were between 88-689m, with WEBL and HAWO at the lower and upper extremes, respectively. Fifteen snags were used in both 2002 and 2003 (38.5%, and 24.2% of 2002 and 2003 nest snag population, respectively). Five snags were doubly occupied in at least one year, and two snags were doubly occupied in both years. The latter were inhabited by WISA-WEBL pairs, the remaining double occupancies were Lewis's woodpecker (*Menanerpes lewis*)-WEBL, HAWO-WEBL, and NOFL-WISA pairs (Table 4.2).

3.5. Snag and nest site characteristics

Snags used for nesting were generally larger in diameter at breast height (DBH; 55.7 used vs. 48.9 cm available) and height (19.4 used vs. 14.0 m available), and nearly all nest trees were ponderosa pine (85.4% used vs. 64.8% available; P=0.0002) (Figure 4.3). Half of the CNBs nested in live trees in decay class 2 (47.2% used vs. 11.5% available; P<0.001), and three-quarters of nest snags had previous CNB excavated cavities (72.4%, nest snags, 13.8% available; P<0.001) (Figure 4.4). Bark beetle top-killed and strip-killed snags were disproportionately selected as nest trees ($P \leq 0.01$; Table 4.3). Vegetation surrounding CNB nests was generally less dense (P<0.001), had less canopy cover (P<0.001), and a higher density of snags (P=0.001, Appendix 4.5) compared to available habitat. Approximately 90% of all nest snags had signs of past insect, disease, or abiotic damages or diseases, and 64% had multiple damages. Of these, significantly more CNB nest snags displayed evidence of woodpecker insect foraging, top-kill, strip-kill, lightning strike damage, and heart rot tree fungus than available snags (Table 4.3).

WEBL, in particular, nested in areas with lower mean canopy cover (40.1 %), and tree density (172.3 ha^{-1}), and higher median shrub and grass cover compared to other CNBs (Figure 4.3; Appendix 4.5). NOFL also nested in open forests, and mean tree cover and densities were generally lower compared to HAWO and WISA (Figure 4.3, Appendix 4.5). WISA nested in tall snags (>20 m on average) with larger diameters (>60 cm), dead tops and large strip-killed sections (Figure 4.3; Appendix 4.5), and few proximate and neighborhood vegetation variables were significant (Appendix 4.5). HAWO preferred taller nest snags, with significant top-kill lengths (Appendix 4.5), but nest snag diameters were significantly smaller compared to WISA (Figure 4.3). HAWO nested in lower mean and maximum tree densities, but few variables were significant (Appendix 4.5).

3.6. CNB habitat model selection and evaluation

SMOTE resampling, snag resampling intensities, and model choice (LR vs. RF) all influenced the amount of total variance explained and shared among scales. LR models explained less of the total variance in nest-use patterns than RF, but exhibited lower levels of shared variance across spatial scales (Figure 4.5). For individual species, SMOTE increased both total variance explained and total shared variance among spatial scales, particularly for RF models (Figure 4.5).

Final models included LR+SMOTE resampling and reference snags sampled at 100m intervals. Models explained ~52-66% of the variance in CNB nest use (Table 4.4), and shared variance among scales was between 16-35%. Shared variance was highest at the snag + local scale (two finest scales), except for WEBL, which had the highest shared variance at the snag + proximate + neighborhood scales.

The snag scale (finest scale) accounted for 21-44% of the pure variance explained; more than all other scales combined (Table 4.4). Pure variance explained at the local, proximate, and neighborhood scales was <10% across all species and generally decreased with increasing spatial scale.

Including nearest neighbor distances increased total variance explained by 1-25%, with the WEBL model showing modest increases and the HAWO model improving the most (Table 4.4, Appendix 4.6). Nearest-nest distance variables increased local-scale pure variance in all models. Total pure variance explained was reduced for the WEBL and WISA models, and increased for HAWO and NOFL. Changes in pure variance were associated with a doubling of shared variance between snag and local models for all species (Table 4.4, Appendix 4.6). HAWO

was most influenced by the presence of conspecifics, nested at least 700-m from other HAWO nests, and was never its own nearest neighbor (Table 4.2).

Variable selection – Variables describing the presence of past cavity excavations and snag decay class were top predictors in all models (Table 4.5). Few damage or disease agents were selected as predictors, with the exception of Comandra rust disease (*Cronartium comandrae*) for HAWO, and bark beetle top-kill for NOFL (Table 4.5). Snag diameter was included only for the WISA model, and the probability of nesting increased with increasing diameter for this species. Snag height was not included in any of the models. Most models included variables related to nest distances to patch edges, elevation, snag cluster size, and bark beetle density (Table 4.5). Variables related to conifer density and cover, and snag density were included in both proximate (1.1 ha) and neighborhood (28.3 ha) models. For example, mean overstory canopy cover and overstory density variables were selected in all individual CNB models and all CNB combined model (Table 4.5).

Model predictions – Model predictions from the full multi-scale CNB model (without inter-nest distances), using a 0.2 probability cutoff value (i.e., the predicted probability above which a snag is classified as a nest snag; Sensitivity=88.1%, Specificity=86.2%), predicted that a total of 1808 snags (0.86 ha^{-1}) were suitable as CNB nesting habitat (Figure 4.6). When probability cutoffs between 0.1-0.5 were tested, the density of predicted nest snags ranged between 0.2-1.1 ha^{-1} for all species except WEBL, which had predicted densities <0.5 ha^{-1}. When nearest nest distances were included, predicted densities declined to roughly 0.05-0.6 ha^{-1} for individual species models.

3.7. Niche overlap analyses

When all spatial scales were combined, four out of six pairwise species comparisons were significant (NO=0.69-0.76; P≤0.006), indicating that most species nested in significantly different niche space (Table 4.6). Only HAWO-NOFL (NO=0.783, P=0.117), and WISA-HAWO (NO=0.777, P=0.06) nested in similar niche space. All NO estimates were ≥0.6, regardless of the species pair or spatial scale tested (Table 4.6).

At the snag scale, WISA occupied significantly different niche space compared to all other species (P≤0.023), except NOFL (P=0.072). All other snag-scale comparisons were insignificant (NO, 0.77-0.79).

WEBL nested in significantly different niche space from all other species (P≤0.006; Table 4.6). WEBL selected nests in significantly different niche space at the proximate and neighborhood scales compared to all other species, with the exception of NOFL at the neighborhood scale. WEBL and NOFL nested in similar habitat at all scales, except the proximate scale, but overall nest environments were different (Table 4.6). WEBL nested in smaller diameter snags and preferred lower canopy cover levels and lower live-tree densities compared to NOFL (Appendix 4.6). WISA and WEBL occupied significantly different habitat at all scales except the local scale. WISA selected nest snags that were larger in diameter and at lower elevations, surrounded by forests with higher levels of canopy cover and live-tree densities and lower levels of shrub cover at both proximate and neighborhood scales (Appendix 4.6).

HAWO and WEBL occupied different niche spaces at proximate, neighborhood, and all scales combined. HAWO and NOFL occupied similar niche spaces at all scales except the local scale, and niche space for HAWO and WISA were similar for all scales, except the snag scale.

At the community level, CNB nest sites were found to be randomly distributed across niche space at the snag, local, and all scales combined ($P \geq 0.057$). At these scales, individual species neither maximized the amount of niche differentiation among nest sites (i.e., even distribution), nor did certain species cluster in certain areas of niche space (i.e., clustered distribution). Clustering across niche space was found at the proximate ($P=0.003$), and neighborhood ($P=0.046$) scales.

4. DISCUSSION

4.1. Characteristics of cavity-nesting bird nest sites

Previous studies suggest that CNBs prefer nesting in well-decayed snags (Saab et al. 2004, Russell et al. 2007, Saab et al. 2007). CNBs in our study showed a strong preference for nesting in living trees despite having access to an abundance of large well decayed trees. Well decayed snags require less time and energy to excavate, and may allow for earlier breeding in the spring and higher fecundity (Short 1979, Jackson and Jackson 2004). These snags may also support a richer insect community used as a food source for some CNB species (Elchuk and Wiebe 2003). However, in our study, all four CNB species preferred nesting in Decay Class 2 (live trees with evidence of necrotic tissue along the bole or major branches) over other decay classes. Snags with more advanced decay (Classes 5-8) represented 41% of available snags (2.4 ha^{-1}), but were nested in only 12% of the time.

Nest site selection is a compromise among the time required for excavation, the vulnerability of a nest site to predation, competition from other conspecifics, and other influences that that may compromise a nest (Jackson and Jackson 2004). Live trees with dead stem sections may offer more protection from larger predators (Short 1979), buffer against

inclement weather, have longer lifespans, and persist through some natural disturbances, such as wildfires (Jackson and Jackson 2004). CNBs at our study sites appeared to favor the durability and protection offered by live trees despite the associated drawback of increased excavation time. These results suggest that snag management should include living trees with dead stem portions within the total population of snags that can potentially support CNB nests.

We initially expected that decays resulting from specific damage or mortality agents would influence nest site selection (Bull et al. 1997, Zack et al. 2002, Farris et al. 2004, Lonsdale et al. 2008). However, predictors related to disease complexes, fire, or other mechanical damages were rarely selected for inclusion in the habitat models. Snags with top kill, strip kill, and scaling (bark flaking during foraging) were used in a greater proportion than available (Table 4.3). CNBs used snags with these damages 45% of the time, compared to 22% of the available snags. Such damages can provide important infection courts for fungi, which can lead to decay of the damaged bole section. For example, the red belt fungus (*Fomitopsis pinicola*) is one such fungus that infects a damaged portion of a tree and slowly decays the affected sapwood and heartwood. CNBs then excavate nests within the decayed areas of otherwise healthy trees. Despite the preferences of CNBs for nesting in live trees, snags in the beginning stages of decay (Decay Class 3) were rarely selected as nest sites (Figure 4.4), particularly when no previous excavations were present. These snags were also used less often than snags with more advanced decay levels (Figure 4.4). Thirty percent of snags were in decay class 3, and more than 90% of them had evidence of recent ($\leq$2 yr) bark beetle attack. These snags were smaller in diameter on average compared to all other decay classes, and a large proportion were lodgepole pines (*Pinus contorta*). Recent beetle-killed trees tend to have intact bark and negligible decay. While these trees may not provide nesting habitat in the short-term, they are valuable foraging habitat for

species such as HAWO (Klenner and Huggard 1998), and may contribute to CNB nesting habitat over time as they continue to decay.

Snags with previous excavations were also favored as nest snags (Figure 4.4), and this variable was included in all models (LR and RF). CNBs are known to reuse nest snags or nest in close proximity to previously used nest snags, but reuse rates vary both within and among species (Aitken et al. 2002, Ahlering and Faaborg 2006, Kozma 2012). Targeting areas used by other CNBs in prior years may inform birds of past successful breeding, and reduce nest search time. Selection of a poor nest site can have negative consequences, for example, failure to find a mate, delayed breeding , reduced fecundity, increased competition and predation, and loss of nest to abiotic influences (Short 1979). Returning to a successful nest snag can lead to earlier breeding and higher fecundity (Nilsson et al. 1991). Similarly, new colonizers to an area also learn from evidence of past habitat use, which may prevent them from colonizing in resource-poor areas. However, predation risk increases with continued habitat reuse (Short 1979, Nilsson 1984, Nilsson et al. 1991), and birds must balance the time required to search for new habitats with the risk of predation in previously used areas. In our study, 38.5% of all 2002 nest trees were reused in 2003, suggesting that some birds used the same nest tree in successive years.

CNBs in our study preferred nesting in open forests, generally with live-tree densities averaging 175-275 trees ha^{-1}, and 40-60% total canopy cover (Appendix 4.6). Other studies have shown CNB preferences for open forested habitats (Elchuk and Wiebe 2003, Walter and Maguire 2005). WEBL, in particular, nested in or near patches (<30% cover), and nest sites were located in areas with the relatively low tree densities (<200 trees ha^{-1}). Open forests can foster the development of understory vegetation that in turn supports a diverse insect community favored by some species (Weikel and Hayes 1999). Conversely, dense conifer is associated with low

levels of ground vegetation cover. Trees growing in moderately dense forests will also self-prune limbs more than open grown trees, and the resultant branch stubs can become infection courts for fungi, and ultimately make prime CNB excavation sites (Jackson and Jackson 2004).

Most CNBs nested near the edges of open patches (<30% conifer cover), consistent with other CNB studies (Power 1966, Gutzwiller and Anderson 1987, Aitken et al. 2002, Lawler and Edwards 2006). This variable was included in all habitat models, and the probability of nesting decreased with increasing distance from patch edge for all species except WISA. Denser forests may provide new fledglings protection from predators and inclement weather, and open spaces may provide adequate foraging habitat for insectivores (Power 1966). WISA's preference for nesting in the interior of patches may be related to its reliance on large diameter live trees for foraging in the breeding season (Gyug et al. 2012). Results suggest that CNB abundance and diversity may increase within increasing landscape heterogeneity and amount of patch edge.

Some studies suggest that CNBs prefer nesting in snag clusters (Raphael and White 1984, Saab and Dudley 1998, Chambers and Mast 2005). Snags in clusters have higher retention times, which may increase the likelihood of securing potential nest sites over time (Chambers and Mast 2005). We defined cluster membership at several scales (5-25-m inter-snag distances (ISD)).The number of snags within each cluster ranged from 1-10 snags (5-m ISD), and 1-443 (25-m ISD), but most CNB nests were located in clusters with $\leq$3 snags, and this held for all species. These findings suggest that some CNBs may avoid snag clusters to reduce competition from other CNBs (Bull et al. 1997).

The habitat models suggested that proximate- and neighborhood-scale vegetation were not good predictors of CNB habitat. The vegetation data used to quantify available habitat were

summarized to two spatial scales to represent lower (1.1 ha) and upper (28.3 ha) bounds on the territory sizes of the CNB species (Gutzwiller and Anderson 1987). However, we could not guarantee that the scales used were ecologically meaningful for our study species (Lawler and Edwards 2006). Vegetation at finer or broader scales than those used in our study may have influenced nesting patterns, but this was not captured in our analysis. Furthermore, the use of remotely sensed data to quantify vegetation may have influenced model results. Despite using high-resolution aerial photography (0.24-m) in both a raster-based classification and a photo-interpreted patch-based classification, our vegetation data may not have captured the features integral to CNB nest selection. Fine-scaled vegetation features, such as the amount of downed woody debris, or canopy openings <0.5 ha were not captured by the remotely-sensed data, but may have contributed to CNB nest selection. Quantifying the scales most influential to nest selection is difficult and likely will require alternative geospatial analyses, or the incorporation of field-based measurements.

Our habitat modeling results suggested that social interactions may be influential in nest-use patterns, but effects varied by species. Overall, the inclusion of nest distances as predictors in the habitat models increased model R^2 by >18%. Coefficients for these variables were mostly positive, indicating an increased probability of nesting with increasing inter-nest distances. Exceptions were WISA in the WEBL model, and WISA in the HAWO model, which displayed negative coefficients. WEBL were relatively complacent to other heterospecifics and conspecifics, and nearest neighbor distances were generally lowest for this species. Conspecific distances were 2-8 times lower for WEBL than other species, and WEBL often cohabitated nest snags with another species (Table 4.2). Insensitivity to other CNBs may be related either to the reliance of this secondary cavity nester on nesting in previously excavated cavities, to its

preference for open areas for nesting and foraging, or a lack of competition with other CNB species in its foraging activity.

Alternatively, HAWO nested far from other CNBs, particularly conspecifics (>700 m),. Total variance explained increased >40% by models with nearest-nest distances, and distance to other HAWO nests was the leading model predictor. Some evidence suggests that strong excavators, such as HAWO, and other wood probers (Raphael and White 1984) tend to prefer large inter-bird distances because bark beetles and woodborers make up a substantial portion of their diet, and these birds will defend a sufficiently large area to maintain an adequate food source with little competition during the breeding season.

The exclusion of some CNB species from the bird surveys prevented us from including distance variables related to other secondary cavity nesters and weak excavators. The influence of these species on nest-use patterns of the four CNBs studied here is unknown, but competition for resources is likely low between these species and the primary cavity nesters at our sites (John Lehmkuhl, personal communication).

4.2. Resource partitioning among CNB nest site selection

Local interactions among individuals may influence resource use by limiting the amount of available habitat required to persist in an area (Özesmi and Özesmi 1999). Such interactions can lead to resource partitioning, resource sharing, or competitive exclusion (Glasser and Price 1988). Identifying key differences in resource selection among sympatric species can aid our understanding of processes related to species coexistence.

Results from the niche overlap analysis indicated that niche diversification among the main CNB species contributed to observed nest-use. Across 84 resource gradients that defined

the hypervolume of total niche space of nest tree environments, overall niche overlap was <0.78 for all species pairs (Table 4.6). Significant differences in nesting habitat for all but the HAWO-NOFL and WISA-HAWO comparisons suggest that these species may have idiosyncratic nesting requirements. Differences in nest habitats were related to the characteristics of the nest snag, the proximity of the nest site to certain resources, as well as the composition of vegetation at the neighborhood scale.

Unique nest site selection likely reflected different life-history strategies of the four main CNB species. WISA and WEBL had the lowest overall NO score, and these species, though similar in body size, have very different feeding habits. WEBL is an aerial and ground insectivore that prefers open habitats with abundant shrub and forb layers to support a diversity of insects (Wightman and Germaine 2006, Guinan et al. 2008, Hurteau et al. 2010). WISA feeds mostly on the sap of large living trees and ants during the breeding season, and prefers nesting in or around mature forests with access to desired forage trees (Smith 1982, Conway and Martin 1993, Gyug et al. 2012). Niche overlap values for these two species were lowest at the proximate and neighborhood scales, with WEBL preferring open habitats but WISA exhibiting few preferences at these spatial scales (Table 4.6, Appendix 4.5). WISA showed a clear preference for large-diameter nest trees (Figure 4.3, Appendix 4.5). These species occupied the same nest snag several times throughout the study. WISA is a primary cavity nester; WEBL may benefit from nesting in old WISA cavities due to their different foraging requirements. Nest snags shared by these two species had canopy cover greater than the upper quartile level for WEBL (>55%), but were situated close to open-forest patches.

WISA and HAWO nest-site use was generally in proportion to available habitat, and nest sites of these species occupied statistically similar niche space. These birds never shared the

same nest snag, and nearest neighbor distances between individuals were large; only 10% of interactions were within 300-m of one another. Raphael and White (1984) suggest that these two species select similar foraging habitats. HAWO will also feed on sap holes created by WISA to fulfill part of its dietary requirements (Jackson et al. 2002). Similarities in nesting and foraging requirements likely prevented these species from nesting in close proximity to one another.

HAWO and NOFL also nested in statistically similar habitats, and niche overlap scales for these species were high across all spatial scales, with the exception of the local scale. Both species appeared to be relative habitat generalists in our study (Figure 4.3, Appendix 4.5), and these species selected statistically similar habitat features (Figure 4.3). HAWO generally preferred nesting in areas with high levels of grass cover and low tree density, but few were significant in the univariate analyses (Appendix 4.5) due to high variability in nest site conditions. NOFL also had similar habitat preferences and nested in open forests. High overlap among these species was likely related to generalist habitat preferences, and availability of suitable habitat.

5. CONCLUSION

We studied habitat preferences of four CNBs within a mixed-conifer dry forest that had not experienced recent wildfire. CNBs preferred nesting in live trees with little decay, despite high availability of large, well-decayed snags. Few specific tree disease or damage complexes were associated with CNB nest selection. Preferences for open forested habitat were exhibited by most birds, particularly WEBL. Birds that occupied similar niche space often selected their nests to avoid species with similar nesting and foraging habits. HAWO in particular selected nest sites to avoid conspecific competition. Results from our study suggest management can foster the development of large ponderosa pine snags at wide spacing, particularly those with dead areas on

live trees. Examples are trees previously struck by lightning, those with old dwarf mistletoe, stem cankers, top-kill or strip-attack from bark beetles, or other bole injuries.

6. REFERENCES

Ahlering, M. A. and J. Faaborg. 2006. Avian habitat management meets conspecific attraction: If you build it, will they come? The Auk **123**:301-312.

Aitken, K. E. H., K. L. Wiebe, and K. Martin. 2002. Nest-Site Reuse Patterns for a Cavity-Nesting Bird Community in Interior British Columbia. The Auk **119**:391-402.

Battin, J. and J. J. Lawler. 2006. Cross-Scale Correlations and the Design and Analysis of Avian Habitat Selection Studies. The Condor **108**:59-70.

Bednarz, J. C., D. Ripper, and P. M. Radley. 2004. Emerging concepts and research directions in the study of cavity-nesting birds: Keystone ecological processes. The Condor **106**:1-4.

Blewett, C. M. and J. M. Marzluff. 2005. Effects of urban sprawl on snags and the abundance and productivity of cavity-nesting birds. The Condor **107**:678-693.

Brawn, J. D. and R. P. Balda. 1988. Population biology of cavity nesters in northern Arizona: do nest sites limit breeding densities? The Condor **90**:61-71.

Breiman, L. 2001. Random forests. Machine learning **45**:5-32.

Bull, E. L., A. A. Clark, and J. F. Shepherd. 2005. Short-term effects of fuel reduction on pileated woodpeckers in northeastern Oregon: a pilot study. Research Paper PNW-RP-564. USDA Forest Service Pacific Northwest Research Station, Portland, OR.

Bull, E. L., C. G. Parks, T. R. Torgersen, and U. S. F. Service. 1997. Trees and Logs Important to Wildlife in the Interior Columbia RiverBasin. Gen. Tech. Rep. PNW-GTR-391. USDA Forest Service Pacific Northwest Research Station, Portland, OR.

Bunnell, F. L., M. Boyland, and E. Wind. 2002. How should we spatially distribute dying and dead wood. Pages 739-752 *in* W. F. Laudenslayer, P. J. Shea, B. E. Valentine, C. P. Weatherspoon, and T. E. Lisle, editors. Proceedings of the symposium on the ecology and management of dead wood in western forests. USDA Forest Service Pacific Southwest Research Station, Reno, NV.

Bunnell, F. L., L. L. Kremsater, and E. Wind. 1999. Managing to sustain vertebrate richness in forests of the Pacific Northwest: relationships within stands. Environmental Reviews **7**:97-146.

Chambers, C. L. and J. N. Mast. 2005. Ponderosa pine snag dynamics and cavity excavation following wildfire in northern Arizona. Forest Ecology and Management **216**:227-240.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research **16**:321-357.

Cody, M. L. 1985. An introduction to habitat selection in birds. Pages 3-56 *in* M. L. Cody, editor. Habitat selection in birds. Academic Press, Orlando, FL.

Conner, R. N., R. G. Hooper, H. S. Crawford, and H. S. Mosby. 1975. Woodpecker nesting habitat in cut and uncut woodlands in Virginia. The Journal of Wildlife Management **39**:144-150.

Conway, C. J. and T. E. Martin. 1993. Habitat suitability for Williamson's sapsuckers in mixed-conifer forests. The Journal of Wildlife Management **57**:322-328.

Cushman, S. A. and K. McGarigal. 2002. Hierarchical, multi-scale decomposition of species-environment relationships. Landscape Ecology **17**:637-646.

Cushman, S. A. and K. McGarigal. 2004. Hierarchical analysis of forest bird species-environment relationships in the Oregon Coast Range. Ecological Applications **14**:1090-1105.

Daily, G. C., P. R. Ehrlich, and N. M. Haddad. 1993. Double keystone bird in a keystone species complex. Proceedings of the National Academy of Sciences **90**:592.

Daly, C., G. H. Taylor, W. P. Gibson, T. W. Parzybok, G. L. Johnson, and P. A. Pasteris. 2000. High-quality spatial climate data sets for the United States and beyond. Transactions of the American Society of Agricultural Engineers **43**:1957-1962.

De'ath, G. and K. E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology **81**:3178-3192.

Dobkin, D. S., A. C. Rich, J. A. Pretare, and W. H. Pyle. 1995. Nest-Site Relationships among Cavity-Nesting Birds of Riparian and Snowpocket Aspen Woodlands in the Northwestern Great Basin. The Condor **97**:694-707.

Eadie, J., P. Sherman, and B. Semel. 1998. Conspecific brood parasitism, population dynamics, and the conservation of cavity-nesting birds. Pages 306-340 *in* T. Caro, editor. Behavioral ecology and conservation biology. Oxford Univeristy Press, New York, New York.

Efron, B. 1978. Regression and ANOVA with zero-one data: Measures of residual variation. Journal of the American Statistical Association **73**:113-121.

Elchuk, C. and K. Wiebe. 2003. Home-range size of northern flickers (Colaptes auratus) in relation to habitat and parental attributes. Canadian Journal of Zoology **81**:954-961.

Farris, K. L., M. J. Huss, and S. Zack. 2004. The role of foraging woodpeckers in the decomposition of ponderosa pine snags. The Condor **106**:50-59.

Fretwell, S. D. and H. L. Lucas. 1969. On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheoretica **19**:16-36.

Ganey, J. L. and S. C. Vojta. 2004. Characteristics of snags containing excavated cavities in northern Arizona mixed-conifer and ponderosa pine forests. Forest Ecology and Management **199**:323-332.

Geange, S. W., S. Pledger, K. C. Burns, and J. S. Shima. 2011. A unified analysis of niche overlap incorporating data of different types. Methods in Ecology and Evolution **2**:175-184.

Glasser, J. W. and H. J. Price. 1988. Evaluating expectations deduced from explicit hypotheses about mechanisms of competition. Oikos **51**:57-70.

Guinan, J. A., P. A. Gowaty, and E. K. Eltzroth. 2008. Western Bluebird (Sialia mexicana). The Birds of North America Online. Ithaca: Cornell Lab of Ornithology.

Gutzwiller, K. J. and S. H. Anderson. 1987. Multiscale associations between cavity-nesting birds and features of Wyoming streamside woodlands. The Condor **89**:534-548.

Gyug, L. W., R. C. Dobbs, T. E. Martin, and C. J. Conway. 2012. Williamson's Sapsucker (Sphyrapicus thyroideus), The Birds of North America Online (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology.

Holt, R. F. and K. Martin. 1997. Landscape modification and patch selection: the demography of two secondary cavity nesters colonizing clearcuts. The Auk **114**:443-455.

Hunter, M. L. 1990. Wildlife, forests, and forestry. Principles of managing forests for biological diversity. Prentice Hall, Edglewood Cliffs, NJ.

Hurteau, S., T. Sisk, B. Dickson, and W. Block. 2010. Variability in nest density, occupancy, and home range size of western bluebirds after forest treatments. Forest Science **56**:131-138.

Jackson, J. A. and B. J. S. Jackson. 2004. Ecological relationships between fungi and woodpecker cavity sites. The Condor **106**:37-49.

Jackson, J. A., H. R. Ouellet, and B. J. S. Jackson. 2002. Hairy Woodpecker (*Picoides villosus*).*in* A. Poole, editor. The Birds of North America Online, Ithaca: Cornell Lab of Ornithology.

Johnson, M. D. 2007. Measuring habitat quality: a review. The Condor **109**:489-504.

Klenner, W. and D. Huggard. 1998. Nesting and foraging habitat requirements of woodpeckers in relation to experimental harvesting treatments at Opax Mountain. Page 299 *in* Managing the Dry Douglas-fir Forests of the Southern Interior: Workshop Proceedings, Kamloops, British Columbia, Canada.

Knapp, E. E., J. E. Keeley, E. A. Ballenger, and T. J. Brennan. 2005. Fuel reduction and coarse woody debris dynamics with early season and late season prescribed fire in a Sierra Nevada mixed conifer forest. Forest Ecology and Management **208**:383-397.

Kotliar, N. B. and J. A. Wiens. 1990. Multiple scales of patchiness and patch structure: A hierarchical framework for the study of heterogeneity. Oikos **59**:253-260.

Kozma, J. M. 2012. Nest-site characteristics of three woodpecker species in managed ponderosa pine forests of the eastern Cascade Range. Northwestern Naturalist **93**:111-119.

Lawler, J. J. and T. C. Edwards. 2006. A variance-decomposition approach to investigating multiscale habitat associations. The Condor **108**:47-58.

Lehmkuhl, J. F., M. Kennedy, E. D. Ford, P. H. Singleton, W. L. Gaines, and R. L. Lind. 2007. Seeing the forest for the fuel: Integrating ecological values and fuels management. Forest Ecology and Management **246**:73-80.

Leica Geosystems, G. I., LLC. 2006. ERDAS Imagine (R). Norcross, GA.

Li, P. and T. E. Martin. 1991. Nest-site selection and nesting success of cavity-nesting birds in high elevation forest drainages. The Auk **108**:405-418.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R News **2**:18-22.

Lonsdale, D., M. Pautasso, and O. Holdenrieder. 2008. Wood-decaying fungi in the forest: conservation needs and management options. European Journal of Forest Research **127**:1-22.

Manly, B. F. J. 2002. Resource selection by animals: statistical design and analysis for field studies. Springer, Norwell, MA.

Martin, K., K. E. H. Aitken, and K. L. Wiebe. 2004. Nest sites and nest webs for cavity-nesting communities in interior British Columbia, Canada: nest characteristics and niche partitioning. The Condor **106**:5-19.

Meyer, C. B. 2007. Does scale matter in predicting species distributions? Case study with the marbled murrelet. Ecological Applications **17**:1474-1483.

NASA Landsat Program. 2011. Landsat TM Scene LT50450261998216PAC03, Orthorectified,.*in* USGS, editor., Sioux Falls, SD.

Newton, I. 1994. The role of nest sites in limiting the numbers of hole-nesting birds: A review. Biological Conservation **70**:265-276.

Nilsson, S. G. 1984. The evolution of nest-site selection among hole-nesting birds: the importance of nest predation and competition. Ornis Scandinavica:167-175.

Nilsson, S. G., K. Johnsson, and M. Tjernberg. 1991. Is avoidance by black woodpeckers of old nest holes due to predators? Animal Behaviour **41**:439-441.

Özesmi, S. L. and U. Özesmi. 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. Ecological Modelling **116**:15-31.

Plotkin, J. B., M. D. Potts, N. Leslie, N. Manokaran, J. Lafrankie, and P. S. Ashton. 2000. Species-area curves, spatial aggregation, and habitat specialization in tropical forests. Journal of Theoretical Biology **207**:81-99.

Power, H. 1966. Biology of the Mountain Bluebird in Montana. The Condor **68**:351-371.

Raphael, M. G. and M. White. 1984. Use of snags by cavity-nesting birds in the Sierra Nevada. Wildlife Monographs **86**:3-66.

Ritchie, M. W., E. E. Knapp, and C. N. Skinner. 2013. Snag longevity and surface fuel accumulation following post-fire logging in a ponderosa pine dominated forest. Forest Ecology and Management **287**:113-122.

Rödder, D. and J. Engler. 2011. Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. Global Ecology and Biogeography **20**:915-927.

Rodewald, A. D., R. H. Yahner, and J. Brawn. 2001. Avian nesting success in forested landscapes: Influence of landscape composition, stand and nest-patch microhabitat and biotic interactions. The Auk **118**:1018-1028.

Russell, R. E., V. A. Saab, and J. G. Dudley. 2007. Habitat suitability models for cavity-nesting birds in a postfire landscape. The Journal of Wildlife Management **71**:2600-2611.

Saab, V. 1999. Importance of spatial scale to habitat use by breeding birds in riparian forests: a hierarchical analysis. Ecological Applications **9**:135-151.

Saab, V. A., J. Dudley, and W. L. Thompson. 2004. Factors influencing occupancy of nest cavities in recently burned forests. The Condor **106**:20-36.

Saab, V. A. and J. G. Dudley. 1998. Responses of cavity-nesting birds to stand-replacement fire and salvage logging in ponderosa pine/Douglas fir forests of southwestern Idaho. USDA Forest Service, Rocky Mountain Research Station, Ogden, UT.

Saab, V. A., R. E. Russell, and J. G. Dudley. 2007. Nest densities of cavity-nesting birds in relation to postfire salvage logging and time since wildfire. The Condor **109**:97-108.

Sauer, J. R., J. E. Hines, J. E. Fallon, K. L. Pardieck, D. J. Ziolkowski Jr., and W. A. Link. 2011. The North American Breeding Bird Survey, Results and Analysis 1966 - 2010. Version 12.07.2011.*in* U. P. W. R. Center, editor., Laurel, MD.

Scott, V. E. 1978. Characteristics of Ponderosa Pine Snags Used by Cavity - Nesting Birds in Arizona. Journal of Forestry **76**:26-28.

Short, L. L. 1979. Burdens of the picid hole-excavating habit. The Wilson Bulletin **91**:16-28.

Slagsvold, T. 1980. Habitat Selection in Birds: On the Presence of Other Bird Species with Special Regard to Turdus pilaris. Journal of Animal Ecology **49**:523-536.

Smith, K. G. 1982. On habitat selection of Williamson's and" red-naped" yellow-bellied sapsuckers. The Southwestern Naturalist **27**:464-466.

Stephens, S. L. and J. J. Moghaddas. 2005. Fuel treatment effects on snags and coarse woody debris in a Sierra Nevada mixed conifer forest. Forest Ecology and Management **214**:53-64.

Thomas, J., R. Anderson, C. Maser, and E. L. Bull. 1979. Snags. Pages 60-77 *in* J. W. Thomas, editor. Wildlife habitats in managed forests the Blue Mountains of Oregon and Washington, Washington, D.C.

Trimble. 2011. eCognition Developer. Munich, Germany.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. The Journal of Wildlife Management **47**:893-901.

Virkkala, R. 2006. Why study woodpeckers? The significance of woodpeckers in forest ecosystems. Annales Zoologici Fennici **43**:82-85.

Walter, S. T. and C. C. Maguire. 2005. Snags, cavity-nesting birds, and silvicultural treatments in western Oregon. Journal of Wildlife Management **69**:1578-1591.

Weikel, J. M. and J. P. Hayes. 1999. The Foraging Ecology of Cavity-Nesting Birds in Young Forests of the Northern Coast Range of Oregon. The Condor **101**:58-66.

Weiss, A. 2001. Topographic position and landforms analysis.*in* ESRI International Users Conference, San Diego, CA.

Whitham, T. G. 1980. The theory of habitat selection: Examined and extended using pemphigus aphids. The American Naturalist **115**:449-466.

Whittaker, J. 1984. Model interpretation from the additive elements of the likelihood function. Applied Statistics **33**:52-64.

Wiebe, K. L. and W. S. Moore. 2008. Nothern Flicker (Colaptes auratus).*in* A. Poole, editor. The Birds of North America. Cornell Lab of Ornithology, Ithaca, NY.

Wightman, C. S. and S. S. Germaine. 2006. Forest stand characteristics altered by restoration affect western bluebird habitat quality. Restoration Ecology **14**:653-661.

Wing, M. G. 2008. Consumer-grade Global Positioning Systems (GPS) receiver performance. Journal of Forestry **106**:185-190.

Zack, S., T. L. George, and W. F. Laudenslayer Jr. 2002. Are there snags in the system? Comparing cavity use among nesting birds in "snag-rich" and "snag-poor" eastside pine forests. Pages 2-4 *in* Proceedings of the symposium on the ecology and management of dead wood in western forests. USDA Forest Service Pacific Southwest Research Station, Reno, NV.

Zarnowitz, J. E. and D. A. Manuwal. 1985. The effects of forest management on cavity-nesting birds in northwestern Washington. The Journal of Wildlife Management **49**:255-263.

TABLES

Table 4.1. Number of cavity-nesting bird nest sites by species located during the 2002 and 2003 breeding seasons. A total of 12,547 non-nest snags were identified in 2003. Species in bold were included in individual habitat modeling and niche overlap analyses.

	----------Woodpeckers----------					---Sapsuckers---		Flickers	--------Bluebirds-------		
Year	Black-backed	Lewis'	Pileated	White-headed	**Hairy**	**William-sons'**	Red-naped	**Northern**	Mountain	**Western**	**Total**
2002	0	2	1	1	**9**	**18**	2	**11**	0	**6**	**50**
2003	1	0	0	0	**11**	**22**	3	**23**	1	**12**	**73**
Total	1	2	1	1	**20**	**40**	5	**34**	1	**18**	**123**

Table 4.2. Statistics describing the distances (m) among nest sites for four species of cavity-nesting birds in the Methow Valley region of eastern Washington State, USA. The number of interactions was calculated using the cumulative number of nest pairs, constrained by unit, at each time period. WEBL is Western Bluebird; WISA is Williamson's sapsucker; NOFL is Northern Flicker; HAWO is Hairy Woodpecker.

		Number of interactions		Distance	
		Total	<300m	Minimum	Quartile (Q_1)
Intraspecific interactions					
WEBL	WEBL	37	7	88.3	416.8
WISA	WISA	115	2	299.4	1064.7
HAWO	HAWO	28	0	689.2	1426.1
NOFL	NOFL	99	6	150.8	1236.6
Interspecific interactions					
WISA	WEBL	68	7	0.0	886.4
HAWO	WEBL	45	7	0.0	436.9
NOFL	WEBL	131	10	45.2	908.1
HAWO	WISA	126	12	53.9	828.4
NOFL	WISA	179	10	0.0	845.3
NOFL	HAWO	97	7	64.5	973.6

Table 4.3. Percentage of nest and non-nest snags with observed damages, and diseases. P-values were based on chi-square tests, and significant differences (P<0.05) are in bold.

Sign of disease or damage	Nest	Non-nest	P-value
CNB foraging holes[1]	**43.1**	**33.0**	**0.002**
Blue stain fungus	42.0	55.4	0.113
Top kill	**24.3**	**10.0**	**0.000**
Scaling[2]	**21.8**	**10.3**	**0.010**
Comandra Rust	13.8	8.4	0.153
Bark beetle strip kill	**12.2**	**2.5**	**0.000**
Saprot	10.8	7.1	0.110
Lightning strike	**4.2**	**1.7**	**0.007**
Fungal Heartrot	**1.6**	**0.4**	**0.005**
Fire scarred	0.8	1.3	0.500
Dwarf mistletoe infected	0.0	1.2	0.757
Visible conks	0.0	0.0	0.748
Root disease signs and symptoms	0.0	0.6	0.520

[1] Indicates CNB foraging for woodborer larvae (Buprestidae and Cerambicidae)
[2] Bark flaking by CNB; indicates foraging in ponderosa pine during the season of mass attack by bark beetle.

Table 4.4. Variance partitioning results depicting the percent of variance explained by each pure and shared component. Values represent the mean and variability (±SD) associated with the variance explained for models built using each of the 10 Monte Carlo simulations of non-nest snags. Results shown are for the best performing models as shown in Figure 5. Negative values indicate that models including only the individual components of the multi-level model explained as much or more variance. All metrics are based on 10-fold cross-validation. LR is logistic regression. CNB models included all 10 species observed during the study. WEBL is Western Bluebird; WISA is Williamson's sapsucker; NOFL is Northern Flicker; HAWO is Hairy Woodpecker.

	WEBL	**WISA**	**HAWO**	**NOFL**	**CNB**
Model	LR	LR	LR	LR	LR
Sample distance	100	100	100	100	100
SMOTE level	Low	High	High	High	High
Pure components					
Snag (S)	21.2 (3.3)	44.0 (6.3)	25.9 (3.3)	29.1 (3.6)	28.9 (4.5)
Local (L)	3.9 (1.2)	1.7 (1.8)	4.8 (2.2)	4.0 (1.9)	5.9 (3.0)
Proximate (P)	1.7 (1.6)	1.5 (1.5)	2.8 (0.8)	2.1 (1.2)	1.0 (0.7)
Neighborhood (N)	4.2 (1.5)	1.5 (1.2)	2.3 (1.1)	0.4 (0.4)	-0.4 (0.6)
Total Pure	31.0 (4.0)	48.8 (6.2)	35.9 (3.2)	35.6 (3.2)	35.4 (2.9)
Shared components					
S + L	4.4 (3.5)	7.1 (5.2)	12.4 (2.2)	8.0 (3.2)	11.7 (3.2)
S + P	2.7 (0.9)	0.3 (0.8)	-0.1 (0.8)	1.0 (0.8)	0.5 (0.8)
S + N	2.3 (1.5)	2.6 (1.8)	2.9 (0.7)	2.1 (1.3)	2.2 (1.2)
L + P	3.3 (1.7)	4.8 (1.8)	2.4 (1.3)	0.5 (0.5)	0.1 (0.7)
L + N	3.9 (1.3)	0.4 (0.9)	0.8 (1.1)	0.5 (0.6)	0.4 (0.8)
P + N	2.6 (1.6)	0.7 (1.1)	1.1 (1.0)	1.1 (0.8)	1.3 (0.7)
S + L + P	4.6 (2.9)	0.4 (1.1)	1.2 (1.6)	0.2 (1.4)	0.3 (0.7)
S + L + N	-0.1 (2.5)	-0.4 (1.5)	0.9 (1.3)	0.5 (1.4)	-0.6 (1.2)
S + P + N	9.9 (1.9)	0.7 (1.1)	-0.1 (0.9)	0.5 (1.4)	-0.5 (1.3)
L + P + N	3.8 (1.9)	0.4 (1.0)	0.7 (1.2)	1.3 (1.1)	2.5 (1.5)
S + L + P + N	-2.0 (2.3)	-0.5 (1.3)	-0.8 (1.8)	1.0 (1.1)	-0.7 (1.8)
Total Shared	35.4 (4.4)	16.3 (6.6)	21.6 (2.9)	16.8 (3.3)	17.1 (4.1)
Total variance explained	66.3 (2.3)	65.1 (3.2)	57.5 (3.8)	52.4 (3.4)	52.5 (2.2)

Table 4.5. Variable importance for single-level models used to model CNB nest site selection in the north Cascades, Washington State. Rankings represent the number of times a variable was included in each model out of 10 total model iterations. Variables with ≤5 total occurrences across all species were removed from the table. Abbreviations are OS, overstory; CC, canopy cover, ISD is inter-snag distance. WEBL is Western Bluebird; WISA is Williamson's Sapsucker; NOFL is Northern Flicker; HAWO is Hairy Woodpecker.

Snag level	WEBL	WISA	HAWO	NOFL	CNB
Bark beetle attack year	1	7	6	8	2
Past cavity excavations	10	10	10	10	10
Snag decay class	10	9	10	10	10
Snag diameter		10		2	4
Woodpecker foraging holes			9		
Rust			10		
Snag species			10		2
Length of topkill			1	5	1
Local level	**WEBL**	**WISA**	**HAWO**	**NOFL**	**CNB**
Solar insolation (April-June)	7	5		2	2
Woodborer density	6				
Snag clump density (15m ISD)	3	2	4	6	
Elevation	10	1	1	10	10
Distance to patch edge	10	3	9	8	8
Distance to patch (<30% CC) edge	10	3	5	5	2
Slope	2			9	2
Standardized topographic index (150m window)	9	3	1	4	1
Beetle density (50m window size)	5			1	
Previous year beetle density (75m window size)		7	9	4	1
Previous year beetle density (300m window size)		2	1	3	6
Snag clump density (25m ISD)		4	8		8
Snag clump density (5m ISD)		7			1

Table 4.5 (cont.)

Proximate level	WEBL	WISA	HAWO	NOFL	CNB
Density of snags	2	4		8	6
Percent conifer cover	9	6		1	9
Bare ground cover	5	6	1	5	10
Mean OS CC	7	3	1	9	6
Percent cover patches (<30% CC)	1	8	8	6	10
OS tree density	3	8	8	10	7
Percent grass cover		4	10		5
Mean total tree density		3	10	10	9
Neighborhood level	**WEBL**	**WISA**	**HAWO**	**NOFL**	**CNB**
Mean crown size	8				
Percent bare ground cover	5	1	10	10	7
Percent cover patches (<30% CC)	10	10		4	
Minimum tree density	2		10		4
Variance in crown size		3	9	1	6
Percent grass cover		2	5	10	6
Percent shrub cover		7	4		5
Max OS CC		5	1	3	3
Variance OS CC		8	3	2	
Variance OS density		10	3		
Max tree density		10	8		
Max CC			10	1	5
Variance CC			10	1	5
Min OS density				7	2

Table 4.6. Results from the niche overlap depicting the mean (±SD) overlap of resource use patterns across all spatial scales combined for four species of CNB. Values indicate the mean (±SD) overlap among pairwise species comparisons, and range from 0 (no overlap) to 1 (complete overlap). Bold values indicate where the two species occupy significantly different niche space ($P < 0.05$). NOFL is northern flicker (n=34), WEBL is western bluebird (n=18), WISA is Williamson's sapsucker (n=40), HAWO is hairy woodpecker (n=20).

Snag	NOFL	WEBL	WISA
HAWO	0.771(0.174)	0.786(0.175)	**0.689(0.245)**
NOFL		0.784(0.156)	0.787(0.156)
WEBL			**0.731(0.244)**
Local	NOFL	WEBL	WISA
HAWO	**0.727(0.135)**	0.710(0.159)	0.804(0.135)
NOFL		0.780(0.118)	**0.734(0.118)**
WEBL			0.726(0.144)
Proximate	NOFL	WEBL	WISA
HAWO	0.821(0.057)	**0.662(0.175)**	0.808(0.060)
NOFL		**0.690(0.143)**	**0.752(0.093)**
WEBL			**0.598(0.197)**
Neighborhood	NOFL	WEBL	WISA
HAWO	0.823(0.072)	**0.663(0.147)**	0.786(0.138)
NOFL		0.734(0.119)	0.782(0.131)
WEBL			**0.679(0.167)**
All levels	NOFL	WEBL	WISA
HAWO	0.783(0.124)	**0.702(0.170)**	0.777(0.160)
NOFL		**0.748(0.138)**	**0.762(0.128)**
WEBL			**0.686(0.193)**

FIGURES

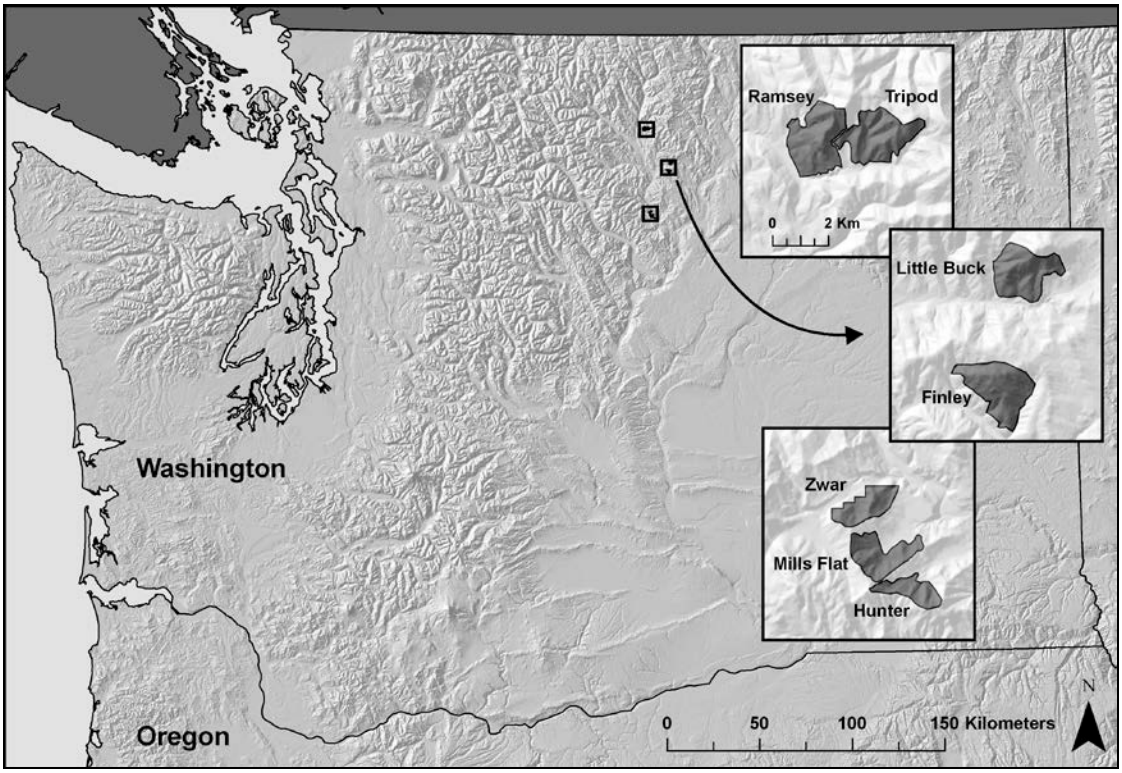


Figure 4.1. Vicinity map of the seven study units located in the eastern Cascade region of Washington State, USA.

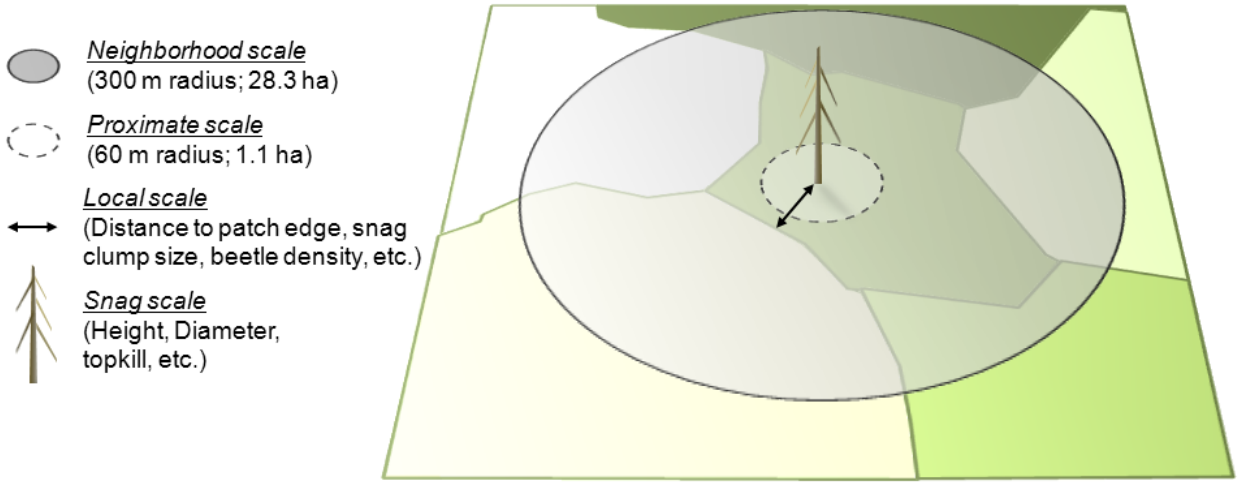


Figure 4.2. Schematic of the multi-scale design (analysis). All vegetation data were summarized to the "neighborhood" and "proximate" scales around each censused snag. The "snag" scale included attributes of the snag itself, and the "local" scale represents "distance or access to resources" from the snag. A complete description of the variables used in the analysis can be found in Appendix 4.2. Neighborhood and proximate scales are drawn approximately to scale.

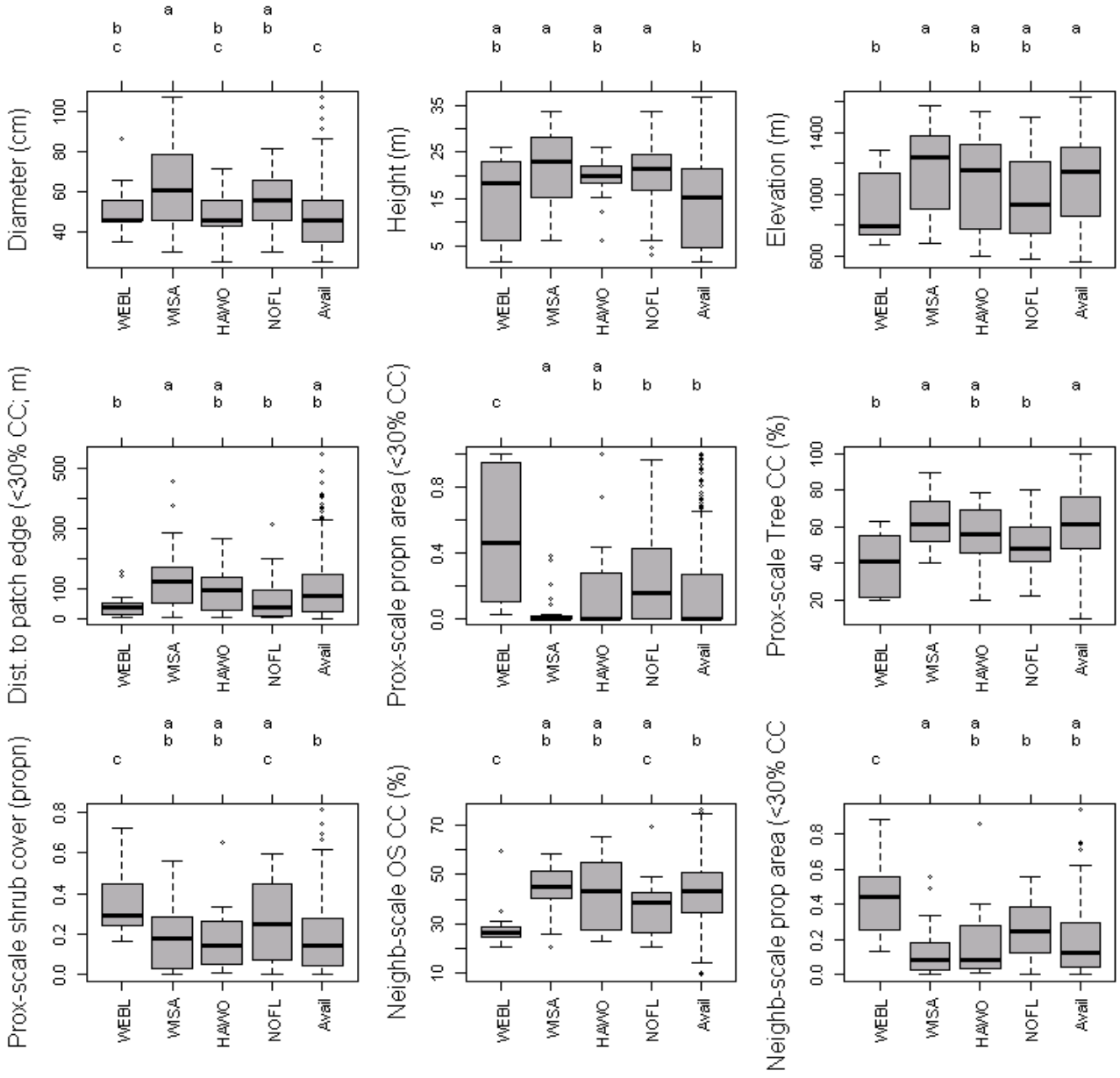


Figure 4.3. Results of selected univariate analyses depicting significant differences in habitat use of four CNB in eastern Washington State, USA. Abbreviations are: OS, overstory, CC, canopy cover; propn, proportion; prox, proximate; neighb, neighborhood. Proximate level (1.1 ha), and neighborhood level (28.3 ha) refer to the area around each snag that vegetation variables were summarized. NOFL is northern flicker (n=34), WEBL is western bluebird (n=18), WISA is Williamson's sapsucker (n=40), HAWO is hairy woodpecker (n=20).

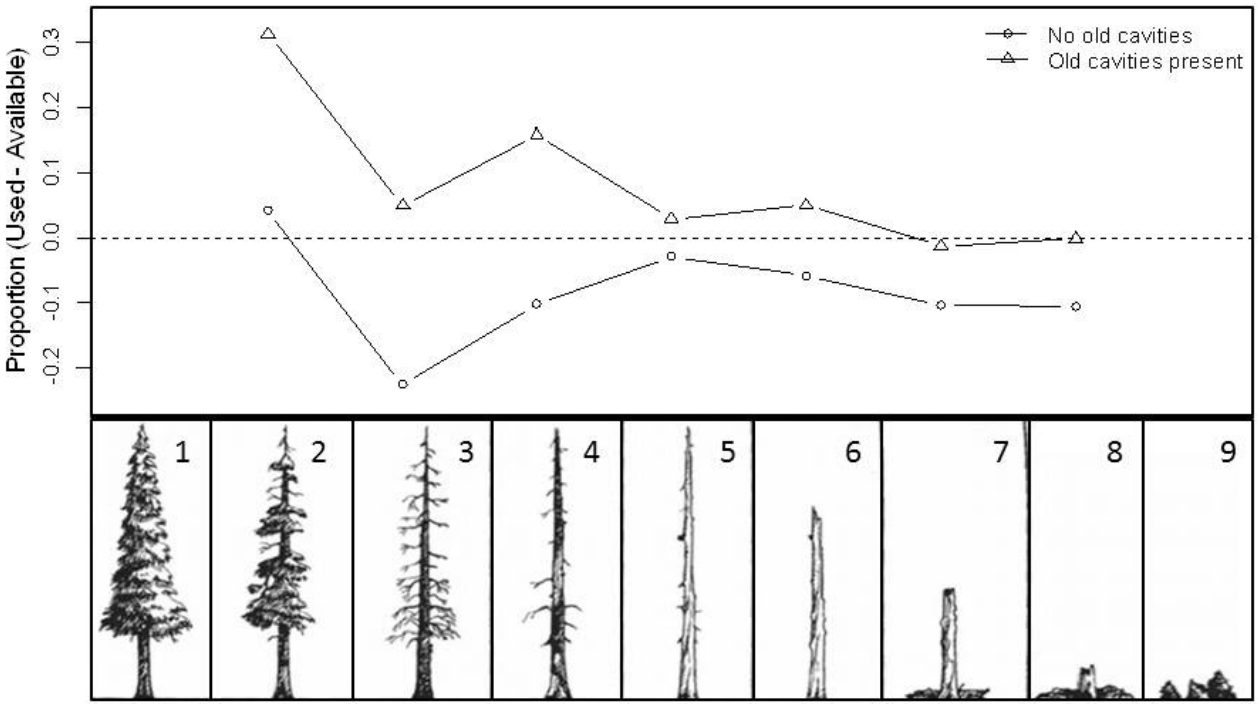


Figure 4.4. Comparison of the total proportion of used (N=123) versus available (N=12,547) snags classified by decay class, and the presence of old cavities. Points above the horizontal dashed line indicate the resource type was used at a greater proportion than what was available; points below the dashed line indicate the resource was used less than what was available. The bottom figure was taken from Thomas et al. (1979), and numbers represent decay classes.

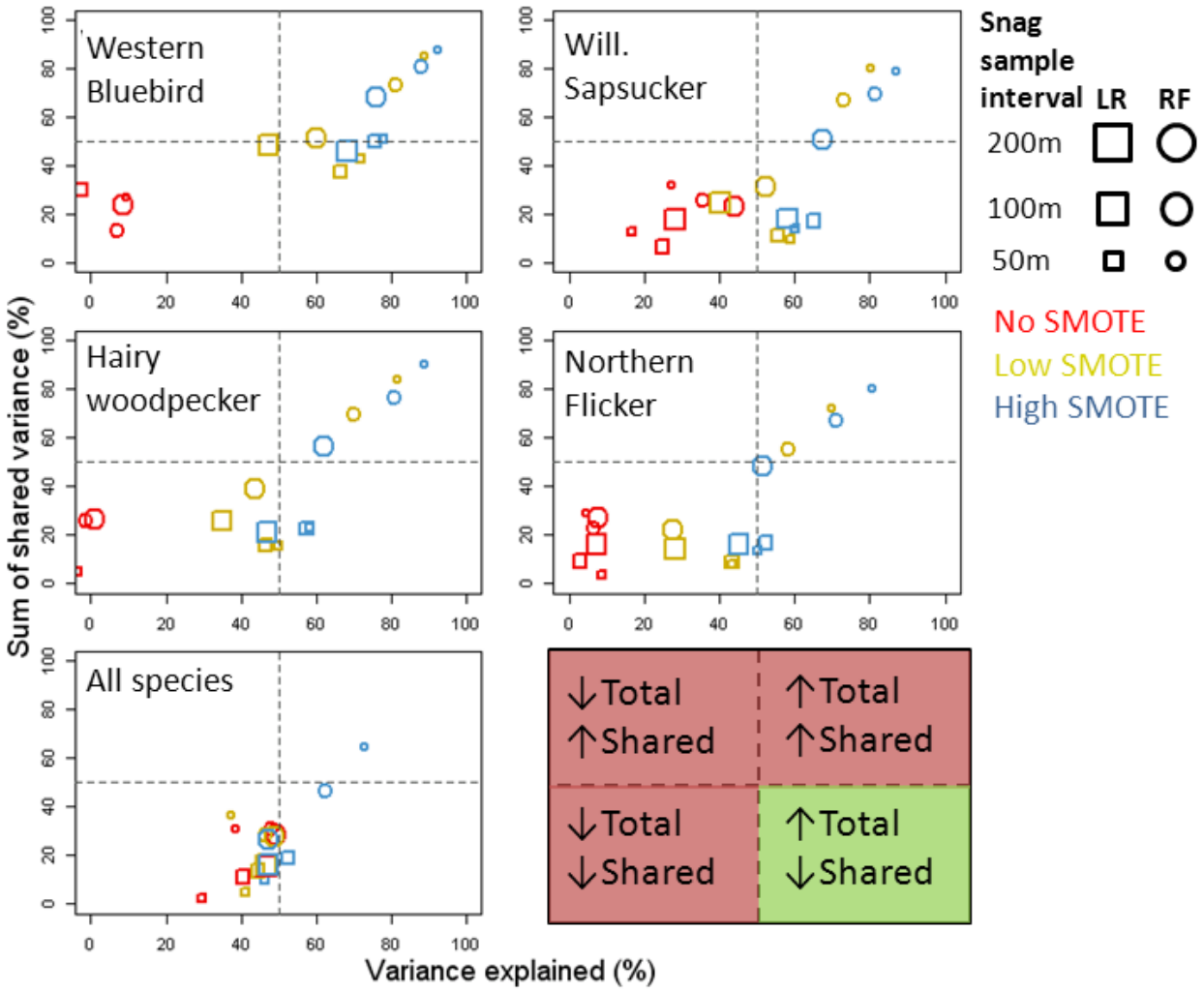


Figure 4.5. Variance partitioning results for the logistic regression and random forest analysis used to model CNB habitat in the eastern Cascades of Washington State, USA. Points that are closest to the lower right corner are preferred as they explain the highest level of variance in nest site selection and have the lowest shared variance among spatial scales. The size of the points indicates the inter-snag distance used to select non-nest (reference) snags for the niche models (50, 100, and 200-m inter-snag distances). Squares represent logistic regression models, and circles represent random forest models. Points colored red included no oversampling of nest sites, gold points are models with nest site oversampling until nest sites equaled 15% of non-nest sites, and blue points are models with nest site oversampling of 30%. NOFL is northern flicker (n=34), WEBL is western bluebird (n=18), WISA is Williamson's sapsucker (n=40), HAWO is hairy woodpecker (n=20). LR is logistic regression, and RF is random forest.

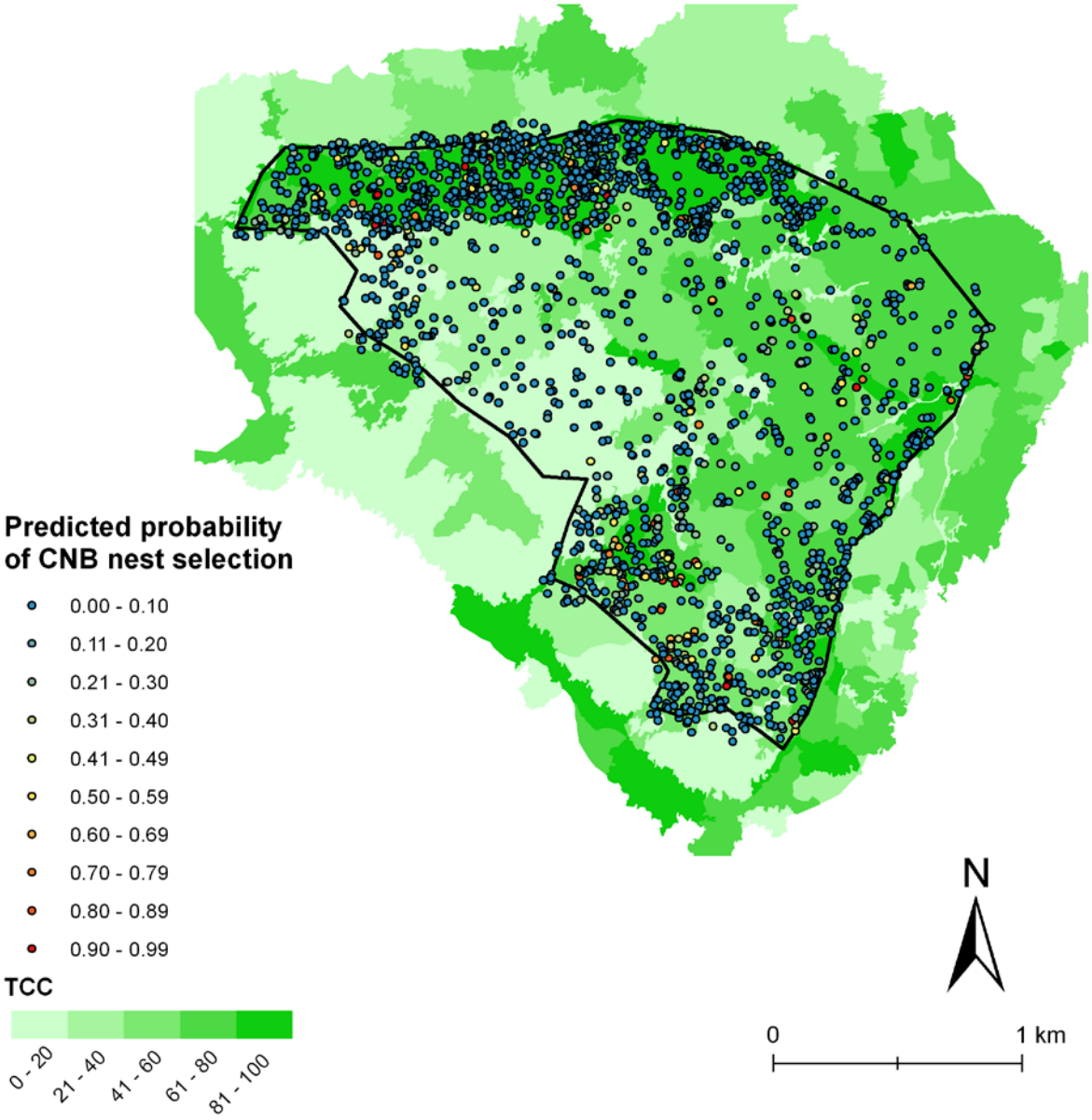
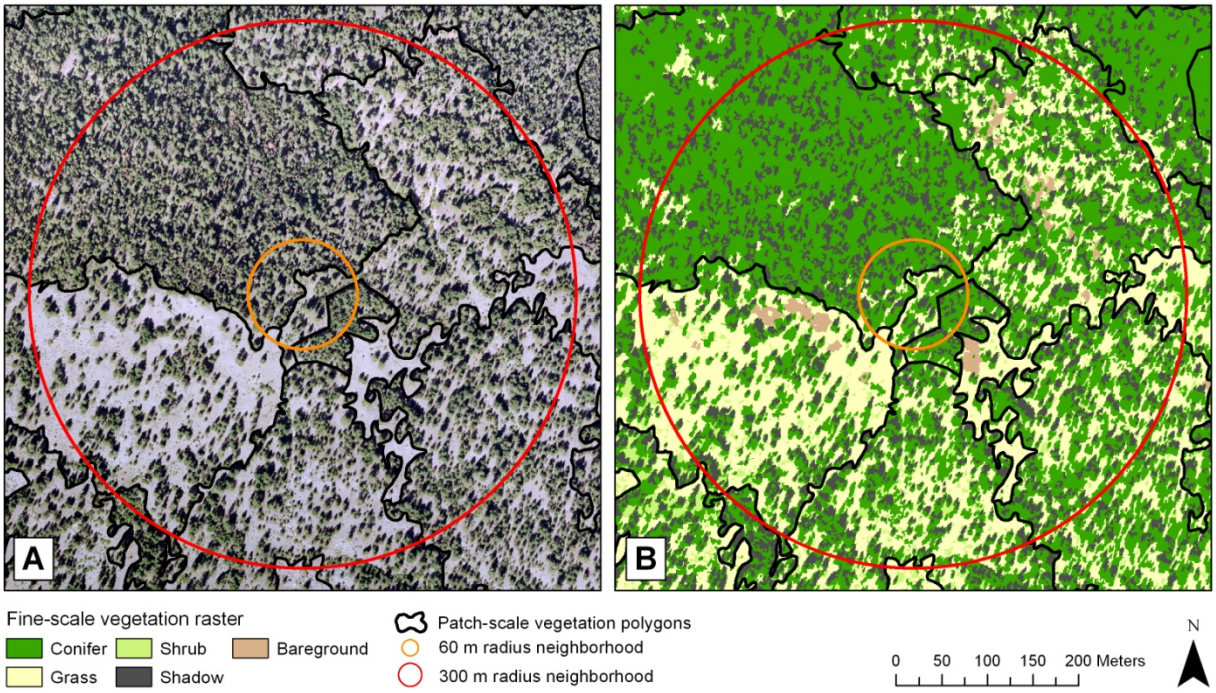


Figure 4.6. Example model predictions for the Finley unit. Probability of nest-site selection by all CNB species was modeled using the full multi-scale logistic regression model. For reference, the area unit is 390 ha. Cool-colored points represent snags with low nesting probability, and successively warmer colors indicate increased probability of occupancy. TCC is total canopy cover (%).

APPENDICES



Appendix 4.1. Example of the vegetation sampling method for the raster- and patch-based vegetation analysis, showing (A) the original aerial ortho-image, and (B) the resultant raster-based classification results. Black lines represent eCongnition unsupervised patch delineation (with manual edits), orange and red circles represent the proximate (1.1 ha) and neighborhood (28.3 ha) sampling levels used to summarize both sets of vegetation data.

Appendix 4.2. Summary of the predictor variables used in the analysis.

Variable Name	Data source	Description	Values/range (n=12,547 snags)
Snag-scale			
Species	Snag survey	Snag species	*Pinus ponderosa, P. contorta, Pseudotsuga menziesii, Populus trichocarpa, P. tremuloides, Larix occidentalis*
Status	Snag survey	Snag status	Live, dead
Diam	Snag survey	Snag diameter	(20.3-132.1cm)
Height	Snag survey	Snag height	(1.5-45.9m)
OldCavity	Snag survey	Presence of old CNB cavities	Binary
SnagDecay	Snag survey	Stage of snag decay	8 class system (Thomas et al. 1979)
TopKill	Snag survey	Evidence of mortality in upper tree bole	Binary
TopKill_Length	Snag survey	Length of upper tree bole mortality	(0-24.4m)
TopKill_Decay	Snag survey	Decay class of upper tree bole mortality	8 class system (Thomas et al. 1979)
TopKill_DiamDeath	Snag survey	Diameter of snag where topkill starts	(0-76.2cm)
StripKill	Snag survey	Evidence of vertical strip mortality along tree bole	Binary
StripKill_Length	Snag survey	Length of strip kill along bole	(0-32.0m)
HeartRot	Snag survey	Presence of heart rot fungal decay	Binary
HeartRot_DecayEx	Snag survey	Length of visible heart rot decay	(0-22.9m)
Conks	Snag survey	Presence of conks (fruiting bodies of heart rot fungi)	Binary
Beetle_Blu	Snag	Presence of blue stain	Binary

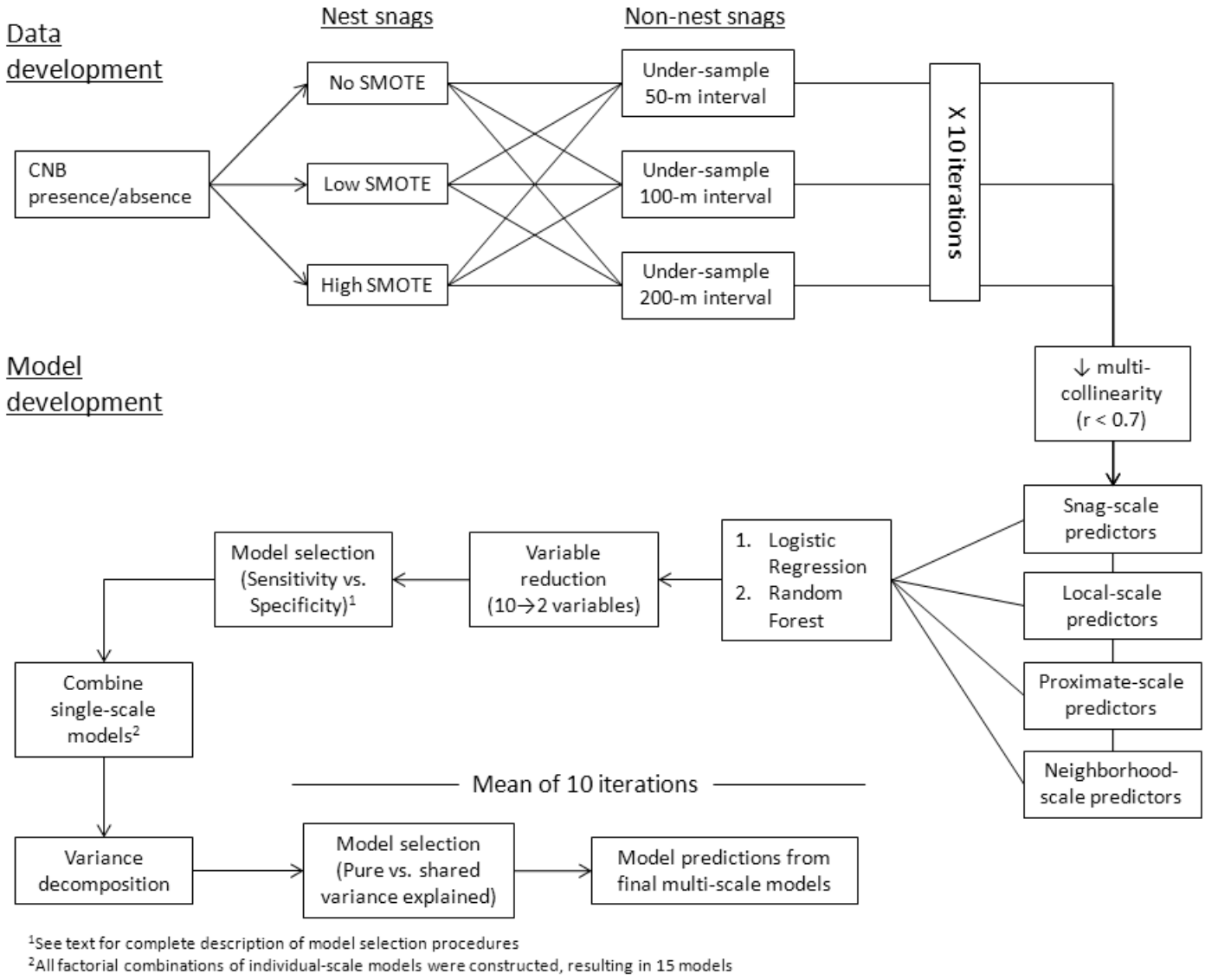
	survey	fungus	
AttackYear	Snag survey	Year of bark beetle attack, if present	0: no attack, 1: current year, 2: previous year, 3: 2+ years
Hits	Snag survey	Presence of CN bird foraging activity (holes)	Binary
Scaling	Snag survey	Presence of CN bird foraging scaling (strips)	Binary
RootDiseas	Snag survey	Presence of root disease	Binary
Rust	Snag survey	Presence of rust disease	Binary
Fire	Snag survey	Evidence of past damage/death due to fire	Binary
Mistletoe	Snag survey	Presence of mistletoe	Binary
Saprot	Snag survey	Presence of saprot (fungal infection in outer sapwood)	Binary
Lightning	Snag survey	Evidence of past lightning damage	Binary
Mechanical	Snag survey	Evidence of mechnical damage (e.g., tree fall)	Binary
Animal	Snag survey	Evidence of animal damage to tree	Binary
Other	Snag survey	Other damage agents present	Binary
Local-scale			
dem_focmea	DEM (10m)	Elevation	(504.0-1657.3m)
mvbb_aspec	DEM (10m)	Linear transformed aspect	(0-200)
mvbb_slope	DEM (10m)	Slope percent	(0-119%)
mvbb_curv	DEM (10m)	Slope curvature	(-6.9-5.6)
stpi_100	DEM (10m)	Standardize topographic index (100m window)	(-4.3-4.5)
stpi_150	DEM (10m)	Standardize topographic index (150m window)	(-3.7-3.8)
areasol_03	DEM (10m)	Incoming solar radiation (March-May)	(164206.7-463038.1 WH/m2)
areasol_04	DEM (10m)	Incoming solar radiation (April-June)	(249811.1-546464.9 WH/m2)
areasol_an	DEM (10m)	Incoming solar radiation (Annual)	(485363.9-1432637.0 WH/m2)

Vegetation Data			
clust5m	Snag survey	Number of snags within a cluster defined (5m inter-tree distance; Plotkin et al. (2002))	(1-10)
clust15m	Snag survey	Number of snags within a cluster defined (15m inter-tree distance; Plotkin et al. (2002))	(1-75)
clust25m	Snag survey	Number of snags within a cluster defined (25m inter-tree distance; Plotkin et al. (2002))	(1-443)
dist2edge	PI patch	Distance to nearest patch edge	(0-207.7m)
dist2edge30	PI patch	Distance to nearest patch with <30% total canopy cover	(0-582.1m)
dist2edge50	PI patch	Distance to nearest patch with <50% total canopy cover	(0-311.1m)
BtlCrnt_YYYw	Snag survey	Surface area of tree boles within	
BtlPrvs_YYYw	Snag survey		
Borer_YYYw	Snag survey		
Proximate-scale			
fsc60mV1	Raster vegetation	Proportion of conifer cover within 60-m of a snag	(0.02-1.0)
fsc60mV2	Raster vegetation	Proportion of shrub cover within 60-m of a snag	(0-0.7)
fsc60mV3	Raster vegetation	Proportion of grass cover within 60-m of a snag	(0-0.8)
fsc60mV5	Raster vegetation	Proportion of bare ground cover within 60-m of a snag	(0-0.3)
crnsizeMN_60m	PI patch	Weighted mean crown size of canopy trees within the 60-m radius neighborhood (weighted by percent cover of dominant canopy and subcanopy tree layers)	(0.36-7.7m)

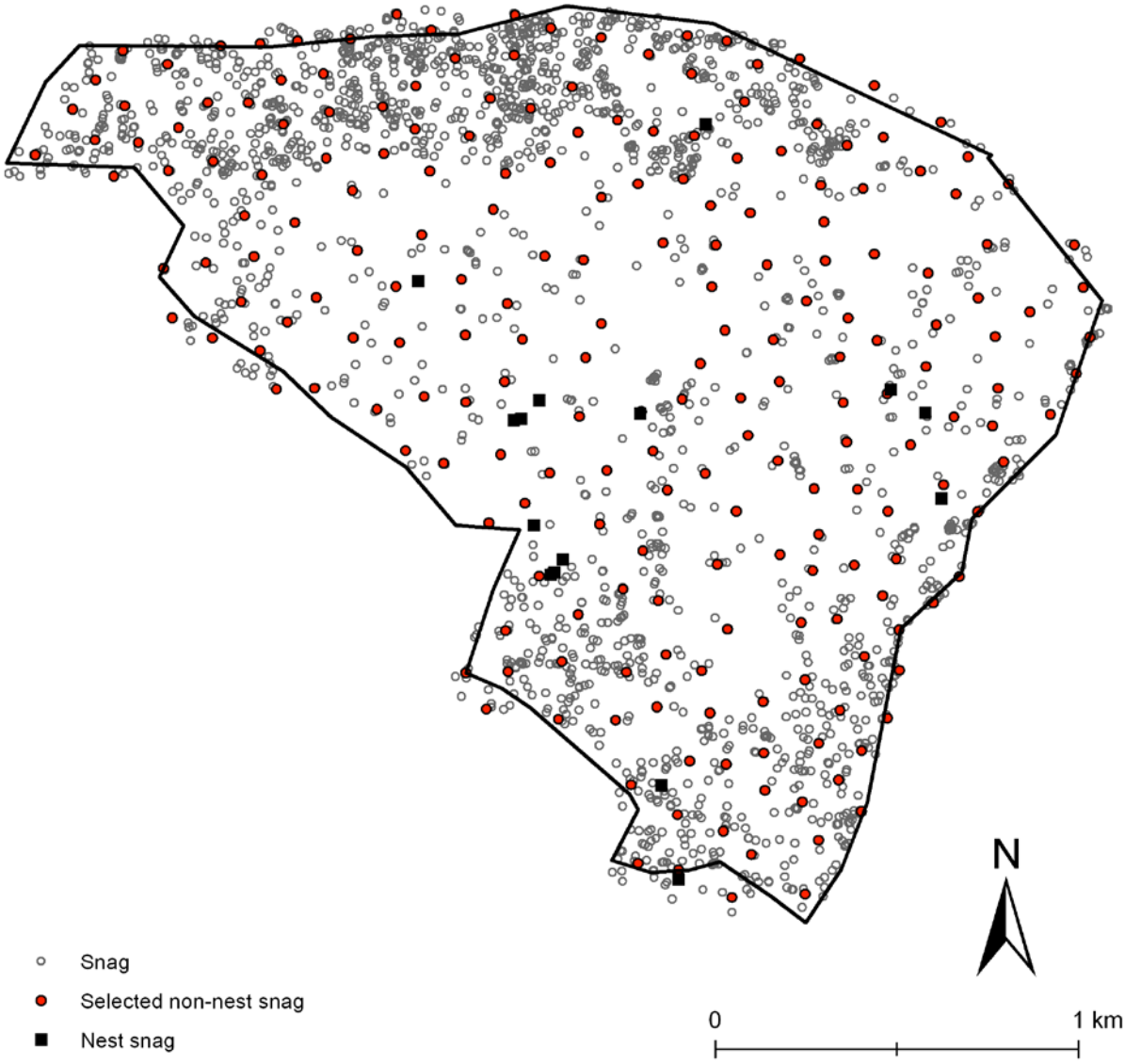
osccMN_60m	PI patch	Mean percent cover of the dominant tree layer within the 60-m radius neighborhood	(3.8-100%)
tccMN_60m	PI patch	Mean percent cover for all tree layers within the 60-m radius neighborhood	(5-100%)
tcc30MN_60m	PI patch	Proportion of area with <30% canopy cover within the 60-m radius neighborhood	(0-1)
tcc50MN_60m	PI patch	Proportion of area with <50% canopy cover within the 60-m radius neighborhood	(0-1)
tphosMN_60m	PI patch	Mean density of trees in the dominant canopy layer within the 60-m radius neighborhood	(5-1411.0 trees/ha)
tphtotMN_60m	PI patch	Mean total tree density within the 60-m radius neighborhood	(5-2500 trees/ha)
dSnags_all_60m	Snag survey	Density of all snags within 60-m neighborhood	(0-750 snags/ha)
dSnags_50_60m	Snag survey	Density of snags >50cm within 60-m neighborhood	(0-400 snags/ha)
Neighborhood-scale			
crnsizeMN_300m	PI patch	Weighted mean crown size of canopy trees within the 300-m radius neighborhood (weighted by percent cover of dominant canopy and subcanopy tree layers)	(1.5-6.4m)
crnsizeMIN_300m	PI patch	Minimum weighted mean crown size within the 300-m radius neighborhood	(0-5.1m)
crnsizeMAX_300m	PI patch	Maximum weighted mean crown size within the 300-m radius neighborhood	(4.1-9.3m)
crnsizeSTD_300m	PI patch	Standard deviation of weighted mean crown size within the 300-m radius neighborhood	(0.2-2.1m)

osccMN_300m	PI patch	Mean percent cover of the dominant tree layer within the 300-m radius neighborhood	(6.5-80.2%)
osccMIN_300m	PI patch	Minimum percent cover of the dominant tree layer within the 300-m radius neighborhood	(0-50%)
osccMAX_300m	PI patch	Maximum percent cover of the dominant tree layer within the 300-m radius neighborhood	(30-100%)
osccSTD_300m	PI patch	Standard deviation of percent cover of the dominant tree layer within the 300-m radius neighborhood	(2.8-43.4%)
tccMN_300m	PI patch	Mean total percent tree cover within the 300-m radius neighborhood	(21.9-91.5%)
tccMIN_300m	PI patch	Minimum total percent tree cover within the 300-m radius neighborhood	(0-60%)
tccMAX_300m	PI patch	Maximum total percent tree cover within the 300-m radius neighborhood	(50-100%)
tccSTD_300m	PI patch	Standard deviation of the total percent tree cover within the 300-m radius neighborhood	(6.4-38.8%)
tcc30MN_300m	PI patch	Proportion of area with <30% canopy cover within the 300-m radius neighborhood	(0-1.0)
tcc50MN_300m	PI patch	Proportion of area with <50% canopy cover within the 300-m radius neighborhood	(0-1.0)

tphosMN_300m	PI patch	Mean density of trees in the dominant canopy layer within the 300-m radius neighborhood	(19.8-923.5 trees/ha)
tphosMIN_300m	PI patch	Minimum density of trees in the dominant canopy layer within the 300-m radius neighborhood	(0-160 trees/ha)
tphosMAX_300m	PI patch	Maximum density of trees in the dominant canopy layer within the 300-m radius neighborhood	(65-1850 trees/ha)
tphosSTD_300m	PI patch	Standard deviation of the density of trees in the dominant canopy layer within the 300-m radius neighborhood	(7.3-681.9 trees/ha)
tphtotMN_300m	PI patch	Mean total tree density within the 300-m radius neighborhood	(71.4-1222.9 trees/ha)
tphtotMIN_300m	PI patch	Minimum total tree density within the 300-m radius neighborhood	(0-276 trees/ha)
tphtotMAX_300m	PI patch	Mean total tree density within the 300-m radius neighborhood	(110-2500 trees/ha)
tphtotSTD_300m	PI patch	Mean total tree density within the 300-m radius neighborhood	(7.0-1113.0 trees/ha)
dSnags_all_300m	Snag survey	Density of all snags within 300-m neighborhood	(1-350 snags/ha)
dSnags_50_300m	Snag survey	Density of snags >50cm within 300-m neighborhood	(0.2-123.6 snags/ha)



Appendix 4.3. Analysis workflow for the cavity-nesting bird habitat models. This process was used to model the four individual CNB species and all observed CNB species combined. SMOTE is synthetic minority oversampling technique. Low and high SMOTE indicated that the alogirthm was run until the number of nest snags was equal to 15% and 30% of the number of non-nest snags, respectively.



Appendix 4.4. An example of the non-nest snag undersampling routine, where all non-nest snags were randomly sampled within each unit (here, the Finley unit) such that no two selected snags were <100-m from one another.

Appendix 4.5. Wilcox-Mann-Whitney test results evaluating the null hypothesis of no difference between nest sites and available habitat. A '+' indicates CNB nest sites and a '-' indicates available habitat specific to that species. Only variables with at least one incidence of a significant nest site selection were included. Significance is indicated with bold type and grey shading, and was based on a two-sample Wilcox test (P<0.05). See Appendix 4.2 for predictor variable definitions.

	WEBL		WISA		HAWO		NOFL		CNB	
	-	+	-	+	-	+	-	+	-	+
dem_focmea	***1092.6***	***899.6***	1092.1	1138.1	1096.0	1072.5	***1096.7***	***974.4***	***1099.6***	***1026.0***
mvbb_asp_1	***86.2***	***121.3***	85.3	85.8	85.4	76.7	84.0	100.8	84.2	92.5
stpi_150	***0.0***	***0.6***	0.0	-0.1	0.0	0.2	0.0	0.4	***0.0***	***0.2***
dist2edge30	***103.7***	***43.0***	103.4	127.4	103.8	99.4	***103.3***	***61.3***	***103.6***	***85.4***
Diam	49.8	51.4	***48.9***	***63.4***	49.1	49.3	***49.0***	***56.4***	***48.9***	***55.7***
Height	13.5	15.5	***14.0***	***21.8***	***13.8***	***19.6***	***13.9***	***19.5***	***14.0***	***19.4***
TopKill_Length	3.0	3.9	***3.2***	***6.3***	***3.2***	***7.5***	***3.3***	***5.4***	***3.2***	***5.4***
TopKill_DiamDeath	1.4	2.7	***1.5***	***3.4***	***1.6***	***3.1***	***1.5***	***2.6***	***1.4***	***2.8***
StripKill_Length	0.6	2.5	***0.7***	***2.9***	0.5	0.3	***0.5***	***2.8***	***0.5***	***2.2***
clust10m	1.5	1.2	1.6	1.3	1.5	1.5	1.5	1.4	***1.5***	***1.3***
fsc60mV1	***0.6***	***0.3***	0.6	0.6	0.6	0.5	***0.6***	***0.5***	***0.6***	***0.5***
fsc60mV3	***0.2***	***0.4***	0.2	0.2	0.2	0.2	***0.2***	***0.3***	***0.2***	***0.2***
crnsizeMAX_60m	***5.1***	***4.7***	5.0	5.2	5.1	5.4	5.0	5.2	5.1	5.2
osccMN_60m	***44.3***	***24.6***	44.1	44.2	44.1	38.5	***44.3***	***34.2***	***44.8***	***36.8***
osccMIN_60m	27.0	12.2	27.4	28.9	27.2	21.8	27.5	20.3	***28.0***	***21.9***
osccMAX_60m	***59.9***	***38.9***	59.6	60.0	59.7	54.8	***59.9***	***47.6***	***60.6***	***51.8***
osccQLOW_60m	***28.9***	***13.3***	29.3	32.1	29.1	23.3	***29.3***	***21.9***	***29.8***	***24.0***
osccQUPP_60m	***58.4***	***38.9***	57.8	59.3	58.2	54.8	***58.1***	***47.2***	***58.5***	***51.4***
tccMN_60m	***62.7***	***40.1***	62.4	63.2	62.6	55.2	***62.8***	***50.4***	***62.9***	***54.6***
tccMIN_60m	***39.9***	***18.1***	40.4	41.4	40.4	33.5	***40.9***	***28.1***	***41.5***	***32.0***
tccMAX_60m	***78.0***	***57.8***	77.5	76.3	77.6	73.0	***77.7***	***70.3***	***78.4***	***71.8***
tccQLOW_60m	***42.5***	***23.6***	43.7	48.6	43.2	36.0	***43.4***	***30.7***	***44.0***	***36.6***
tccQUPP_60m	***76.7***	***57.8***	76.3	76.0	76.4	73.0	***76.6***	***68.9***	***76.8***	***71.3***
tcc30MN_60m	***0.1***	***0.5***	***0.2***	***0.0***	0.1	0.2	***0.1***	***0.2***	***0.1***	***0.2***
tcc50MN_60m	***0.4***	***0.7***	0.4	0.3	0.4	0.5	***0.4***	***0.6***	***0.4***	***0.5***
tphosMN_60m	***161.3***	***79.4***	156.1	120.6	159.3	109.0	***160.2***	***112.0***	***161.4***	***107.2***
tphosMIN_60m	***78.8***	***39.7***	79.9	77.9	78.9	59.3	***80.4***	***57.5***	***80.4***	***59.1***
tphosMAX_60m	***271.8***	***139.4***	271.0	219.2	267.6	159.5	***273.0***	***171.7***	***274.1***	***179.5***
tphosQLOW_60m	***85.5***	***44.1***	86.0	83.5	85.8	63.0	***86.5***	***61.2***	***87.5***	***63.7***
tphosQUPP_60m	***260.8***	***139.4***	251.6	210.8	254.6	159.5	***257.6***	***167.6***	***254.3***	***174.2***
tphtotMN_60m	***307.8***	***172.3***	304.2	276.8	***306.1***	***219.7***	***308.9***	***207.9***	***308.8***	***228.4***
tphtotMIN_60m	***172.0***	***64.1***	174.2	138.4	172.0	127.8	***177.5***	***111.7***	***178.8***	***114.9***
tphtotMAX_60m	***438.5***	***251.3***	433.6	455.1	***433.1***	***298.0***	***436.4***	***350.9***	***437.7***	***362.6***
tphtotQLOW_60m	***184.5***	***92.4***	188.3	163.9	185.8	137.5	***188.5***	***124.2***	***191.8***	***135.9***
tphtotQUPP_60m	***424.5***	***251.3***	418.4	455.1	422.0	297.5	***423.6***	***316.9***	***422.4***	***352.9***
dSnags_all_60m	9.3	7.7	***8.3***	***10.4***	8.5	9.3	8.7	9.6	***9.1***	***9.6***
dSnags_50_60m	***4.1***	***4.0***	3.5	4.6	3.6	3.4	***3.8***	***5.0***	***3.8***	***4.4***
fsc60mV2	0.2	0.3	0.2	0.2	0.2	0.3	0.2	0.2	***0.2***	***0.2***

Appendix 4.5 (cont).

	WEBL		WISA		HAWO		NOFL		CNB	
	-	+	-	+	-	+	-	+	-	+
fsc300mV1	***0.6***	***0.4***	0.6	0.6	0.6	0.6	***0.6***	***0.5***	***0.6***	***0.5***
fsc300mV3	***0.2***	***0.3***	0.2	0.2	0.2	0.2	***0.2***	***0.2***	***0.2***	***0.2***
fsc300mV5	***0.0***	***0.0***	0.0	0.0	0.0	0.0	***0.0***	***0.0***	***0.0***	***0.0***
crnsizeMN_300m	***4.5***	***4.0***	4.5	4.6	4.5	4.5	4.5	4.5	4.5	4.5
crnsizeMIN_300m	***2.8***	***2.0***	2.9	3.0	2.8	2.7	2.8	2.9	2.8	2.7
crnsizeQUP_300m	***5.8***	***5.4***	5.9	5.9	5.9	6.0	5.9	6.0	5.9	5.9
osccMN_300m	***43.4***	***28.2***	43.3	43.8	43.3	43.2	***43.5***	***36.2***	***43.7***	***38.7***
osccQUPP_300m	***79.6***	***68.9***	79.4	79.3	79.4	83.3	79.7	74.1	79.7	76.7
tccMN_300m	***60.6***	***44.3***	60.4	59.8	60.5	58.4	***60.6***	***54.4***	***60.7***	***55.7***
tcc30MN_300m	***0.2***	***0.5***	0.2	0.1	0.2	0.2	***0.2***	***0.3***	***0.2***	***0.2***
tcc50MN_300m	***0.4***	***0.6***	0.4	0.4	0.4	0.4	***0.4***	***0.5***	***0.4***	***0.5***
tphosMN_300m	***162.1***	***88.4***	160.0	140.7	160.7	155.1	***161.7***	***126.7***	***162.0***	***128.9***
tphosQLOW_300m	34.9	27.5	35.2	37.0	35.2	29.5	***35.1***	***24.0***	***35.6***	***28.5***
tphtotMN_300m	***300.8***	***193.6***	298.4	288.8	298.9	283.9	***300.5***	***253.4***	***300.7***	***260.5***
tphtotQLOW_300m	60.5	32.9	61.0	56.6	60.8	46.8	***60.6***	***45.0***	***61.7***	***45.4***

Appendix 4.6. Variance partitioning results with inter-nest distance variables included in the habitat models at the local scale.

	WEBL	**WISA**	**HAWO**	**NOFL**	**CNB**
Model	LR	LR	LR	LR	LR
Sample distance	100	100	100	100	100
SMOTE level	Low	High	High	High	High
	WEBL	WISA	HAWO	NOFL	CNB
Snag (S)	14.8 (3.0)	32.3 (5.4)	14.6 (1.8)	21.1 (1.9)	25.5 (2.1)
Local (L)	7.8 (1.7)	7.8 (2.2)	28.2 (2.0)	13.8 (2.6)	7.8 (1.6)
Proximate (P)	1.3 (1.3)	1.8 (0.7)	0.7 (0.6)	1.7 (1.0)	1.1 (0.9)
Neighborhood (N)	2.8 (2.0)	1.3 (1.3)	2.7 (1.2)	1.4 (1.2)	-0.1 (0.6)
Total Pure	26.7 (3.5)	43.1 (3.8)	46.2 (3.3)	38.0 (3.2)	34.3 (1.6)
Shared components					
S + L	9.8 (2.7)	18.7 (2.9)	26.1 (2.4)	16.7 (2.1)	14.2 (2.1)
S + P	2.6 (1.2)	0.3 (1.1)	1.4 (1.3)	1.4 (0.7)	0.0 (0.6)
S + N	1.1 (1.2)	3.4 (1.6)	2.2 (1.4)	1.9 (1.1)	2.0 (1.3)
L + P	2.6 (0.9)	2.3 (1.3)	3.3 (2.1)	2.1 (1.3)	0.2 (0.7)
L + N	5.0 (1.8)	0.3 (0.9)	0.3 (1.2)	-0.8 (1.0)	0.3 (1.1)
P + N	4.1 (2.1)	1.6 (0.9)	0.6 (1.3)	1.1 (0.7)	1.2 (0.8)
S + L + P	6.7 (2.2)	1.7 (1.8)	-0.0 (2.0)	-1.2 (0.9)	0.2 (1.6)
S + L + N	1.1 (1.4)	-1.5 (1.4)	1.5 (1.7)	0.9 (1.0)	-0.2 (1.2)
S + P + N	7.7 (1.9)	0.3 (1.0)	0.7 (1.5)	0.8 (1.0)	-0.1 (1.1)
L + P + N	4.8 (3.0)	2.1 (0.8)	2.5 (1.0)	0.9 (0.9)	2.4 (1.6)
S + L + P + N	-3.1 (4.1)	-1.8 (1.4)	-2.9 (1.0)	1.2 (0.9)	-0.9 (2.0)
Total Shared	42.4 (4.4)	27.4 (2.9)	35.7 (3.4)	25.1 (2.7)	19.3 (2.6)
Total variance explained	*69.1 (2.5)*	*70.5 (2.6)*	*82.0 (1.4)*	*63.1 (2.5)*	*53.7 (2.4)*