

Spring 5-12-2018

New Approaches to Mapping Forest Conditions and Landscape Change from Moderate Resolution Remote Sensing Data across the Species-Rich and Structurally Diverse Atlantic Northern Forest of Northeastern North America

Kasey R. Legaard

University of Maine, kasey.legard@maine.edu

Follow this and additional works at: <https://digitalcommons.library.umaine.edu/etd>

 Part of the [Artificial Intelligence and Robotics Commons](#), [Environmental Monitoring Commons](#), [Forest Management Commons](#), [Natural Resources Management and Policy Commons](#), and the [Other Forestry and Forest Sciences Commons](#)

Recommended Citation

Legaard, Kasey R., "New Approaches to Mapping Forest Conditions and Landscape Change from Moderate Resolution Remote Sensing Data across the Species-Rich and Structurally Diverse Atlantic Northern Forest of Northeastern North America" (2018). *Electronic Theses and Dissertations*. 2834.
<https://digitalcommons.library.umaine.edu/etd/2834>

This Open-Access Dissertation is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.

**NEW APPROACHES TO MAPPING FOREST CONDITIONS AND LANDSCAPE CHANGE FROM MODERATE
RESOLUTION REMOTE SENSING DATA ACROSS THE SPECIES-RICH AND STRUCTURALLY DIVERSE
ATLANTIC NORTHERN FOREST OF NORTHEASTERN NORTH AMERICA**

By

Kasey R. Legaard

B.S. Idaho State University, 2001

M.S. University of Maine, 2004

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Forest Resources)

The Graduate School

The University of Maine

May 2018

Advisory Committee:

Steven Sader, Professor Emeritus of Forest Resources, Co-Advisor

Aaron Weiskittel, Associate Professor of Forest Biometrics and Modeling, Co-Advisor

Jeremy Wilson, Executive Director of the Harris Center for Conservation Education

Robert Seymour, Curtis Hutchins Professor Emeritus of Silviculture

Daniel Harrison, Professor of Wildlife Ecology

William Halteman, Professor Emeritus of Biostatistics

© 2018 Kasey R. Legaard

**NEW APPROACHES TO MAPPING FOREST CONDITIONS AND LANDSCAPE CHANGE FROM MODERATE
RESOLUTION REMOTE SENSING DATA ACROSS THE SPECIES-RICH AND STRUCTURALLY DIVERSE
ATLANTIC NORTHERN FOREST OF NORTHEASTERN NORTH AMERICA**

By Kasey Legaard

Dissertation Co-Advisors: Dr. Steven Sader and Dr. Aaron Weiskittel

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Forest Resources)

May 2018

The sustainable management of forest landscapes requires an understanding of the functional relationships between management practices, changes in landscape conditions, and ecological response. This presents a substantial need of spatial information in support of both applied research and adaptive management. Satellite remote sensing has the potential to address much of this need, but forest conditions and patterns of change remain difficult to synthesize over large areas and long time periods. Compounding this problem is error in forest attribute maps and consequent uncertainty in subsequent analyses. The research described in this document is directed at these long-standing problems.

Chapter 1 demonstrates a generalizable approach to the characterization of predominant patterns of forest landscape change. Within a ~1.5 Mha northwest Maine study area, a time series of satellite-derived forest harvest maps (1973-2010) served as the basis grouping landscape units according to time series of cumulative harvest area. Different groups reflected different harvest histories, which were linked to changes in landscape composition and configuration through time series of selected landscape metrics. Time series data resolved differences in landscape change attributable to passage of the Maine Forest Practices Act, a major change in forest policy. Our approach should be of value in supporting empirical landscape research.

Perhaps the single most important source of uncertainty in the characterization of landscape conditions is over- or under-representation of class prevalence caused by prediction bias. Systematic error is similarly impactful in maps of continuous forest attributes, where regression dilution or attenuation bias causes the overestimation of low values and underestimation of high values. In both cases, patterns of error tend to produce more homogeneous characterizations of landscape conditions. Chapters 2 and 3 present a machine learning method designed to simultaneously reduce systematic and total error in continuous and categorical maps, respectively. By training support vector machines with a multi-objective genetic algorithm, attenuation bias was substantially reduced in regression models of tree species relative abundance (chapter 2), and prediction bias was effectively removed from classification models predicting tree species occurrence and forest disturbance (chapter 3). This approach is generalizable to other prediction problems, other regions, or other geospatial disciplines.

PREFACE

Spatial data play important roles in forest research and management. Forest attribute maps provide a basis for planning and executing field studies, developing and calibrating ecological models, quantifying ecosystem processes or services, and evaluating environmental change. Forest managers need as much relevant information as possible on the spatial distribution and condition of forest resources within their management areas and in the surrounding ecosystem to set management objectives, project changes, and plan management actions. With growing emphasis on the sustainable provision of non-timber ecosystem services and on the effects of rapidly changing external drivers (e.g., climate, forest pests, market conditions), research and management will increasingly require spatial data sources that can provide information on forest conditions at multiple scales, with frequent and timely updates, and at reduced cost (Franklin, 2001). Satellite remote sensing has the potential to satisfy these information needs. From programs such as Landsat or Sentinel, for example, the routine acquisition of satellite imagery supports estimation and mapping of current forest conditions as well as regular updates; new images can be readily compared to older images to detect changes in landscape conditions using well-established methodologies, some of which were initially developed at the University of Maine (e.g., Sader and Winne, 1992; Wilson and Sader, 2002). The ~40-year depth of the Landsat image archive in particular facilitates studies of forest landscape dynamics. However, the spatial and temporal heterogeneity of forests and forest change are difficult to synthesize over large areas and long time periods. New methods are needed to identify patterns of change, associate change with driving forces, and quantify the impact of map error and other sources of uncertainty on subsequent analyses.

The Atlantic Northern Forest of Maine provides a worthy setting for this sort of work. Maine lies within a transition zone between the northern boreal forest and the southern temperate deciduous-dominant forest (Likens and Franklin, 2009), and includes approximately 4 Mha of nearly contiguous,

undeveloped forestland across northern and western sections. Tree species diversity is relatively high, and the combined impacts of natural disturbance and a long history of timber harvesting includes structural diversity at stand and landscape scales. Rapid post-disturbance recovery and frequent cloud cover contribute additional technical difficulties to the prediction and monitoring of forest conditions via satellite remote sensing. However, the importance of characterizing and quantifying landscape change cannot be overstated, as major changes in disturbance impacts and management practices have elicited concerns regarding the sustainable provision of forest values.

The spruce-fir forests of the region are subject to periodic infestations of the eastern spruce budworm (*Choristoneura fumiferana* (Clem.)), a native pest that causes widespread defoliation and mortality of balsam fir (*Abies balsamea*) and spruce (*Picea spp.*) trees (Irland et al., 1988; Seymour, 1992). Maine's last outbreak occurred ca. 1972–1988 and stimulated broad-scale salvage harvesting by clearcut at rates well above recognized long-term allowable levels (Irland et al., 1988). The legacy of salvage clearcutting remains in large continuous tracts of young forest, much of which was converted from spruce-fir to deciduous and mixed types following extensive regeneration failures (Seymour, 1992). Public concern over the size and prevalence of salvage clearcuts led to the passage of the Maine Forest Practices Act (FPA; 12 MRSA §8867-A to §8888) in 1989 and its implementation in 1991. The FPA fundamentally changed management practices by placing restrictions and disincentives on clearcutting, and marked a transition between two very different disturbance regimes. State records indicate that annual harvest area roughly doubled during the 1990s (Maine Forest Service, 2000, 1994) as landowners maintained similar extraction rates via partial harvest practices that require a larger footprint to achieve the same volume removal. Post-FPA partial harvesting is dominated by the nonselective removal of merchantable timber within and adjacent to machine trails, leaving a matrix of unharvested or lightly harvested area between trails and high variability in disturbance intensity over small scales.

Chapter 1 characterizes predominant patterns of cumulative landscape change caused by both pre- and post-FPA harvest practices across a ~1.5 Mha study area in northwest Maine. Landsat imagery and forest inventory data were used to develop and validate forest composition maps and a time series of forest harvest maps spanning the period 1973-2010. Time series of cumulative harvest area were used to segment the study area into groups of landscape units with similar harvest histories. These were linked to changes in landscape composition and configuration in order to characterize the evolution of landscape conditions in response to forest management practices before and after the abrupt change induced by the FPA. In some groups (24% of landscape units), budworm salvage logging caused rapid loss and subdivision of intact mature forest. Persistent landscape change was created by large salvage clearcuts and conversion of spruce-fir to deciduous and mixed forest. In groups that were little affected by salvage (56% of landscape units), post-FPA partial harvesting caused loss and subdivision of intact mature forest at even greater rates. Patch shape complexity and edge density reached high levels even where cumulative harvest area was relatively low. Contemporary practices introduced more numerous and much smaller patches of stand-replacing disturbance (typically averaging <15 ha) and a correspondingly large amount of edge. Pre- and post-FPA management regimes impacted different areas to different degrees, producing different trajectories of landscape change that should be recognized when studying the impact of policy and management practices on forest ecology. Chapter 1 demonstrates a relatively simple yet novel means of synthesizing predominant patterns of change associated with specific landscape units, with the spatial and temporal resolution needed to attribute change to different management regimes.

The forest cover and disturbance maps used in the chapter 1 analysis were produced using well-established unsupervised classification techniques previously developed and employed to good effect in Maine (Jin and Sader, 2005; Sader and Legaard, 2008; Sader and Winne, 1992; Wilson and Sader, 2002). Although these techniques remain effective and useful, they are extremely inefficient, requiring

extensive visual interpretation and on-screen editing to achieve high accuracy. More importantly, the data production and analysis methods used in chapter 1 largely ignored the potential impact of map error on landscape pattern. Small quantities of map error can have large impacts on landscape pattern metrics used to characterize and quantify landscape change (Langford et al., 2006; Wickham et al., 1997). Uncertainty in metric values appears to increase exponentially with increasing map error (Shao and Wu, 2008), and the effects of different error characteristics likely vary depending on landscape characteristics and the specific landscape metrics applied (Li and Wu, 2004; Shao et al., 2001; Wickham et al., 1997). Maps of depicting landscape conditions at different times should generally have different error characteristics, simply because there is no established methodology to control error characteristics during map production. Similarly, when landscape pattern metrics are used to compare the effects of markedly different disturbance or management regimes in effect at different times, results will be affected to some extent by differences in map error. In the specific case of the chapter 1 analysis, change caused by clearcut harvesting was more accurately represented in maps than change caused by partial canopy removal, and differences in spatial error patterns presumably contributed to observed differences in landscape pattern.

With respect to the representation of landscape pattern in categorical maps, one of the most impactful aspects of map error is the systematic over- or under-representation of class prevalence caused by prediction bias (Shao et al., 2003; Shao and Wu, 2008). A classification algorithm is biased when omission and commission error rates are imbalanced. This can be caused by any number of factors, but for supervised classification algorithms, bias is often a product of training data imbalance (He and Garcia, 2009). When training samples from one class substantially outnumber those from another, a classification algorithm can achieve higher accuracy by preferentially assigning samples to the majority class. Bias is a similarly impactful problem with maps intended to represent continuous quantities rather than discrete classes. In this context, regression dilution or attenuation bias is caused

by predictor variable uncertainty and leads to the systematic over-estimation of low values and under-estimation of high values (Rejou-Mechain et al., 2014; Xu et al., 2009). The effects are similar to those of a classification bias in that landscapes are often represented as less variable or more homogeneous than would otherwise be the case. Strong systematic error can result in the gross misrepresentation of forest conditions.

Chapters 2 and 3 present a machine learning (ML) method that is capable of minimizing both total and systematic error in both continuous and categorical maps of forest conditions. Chapter 2 addresses attenuation bias in models of tree species relative abundance (percent of total aboveground live biomass) based on multitemporal Landsat and topoclimatic predictor data. Following extensive development, a multi-objective support vector regression (MOSVR) algorithm was used to simultaneously minimize both total prediction error and systematic error caused by attenuation bias. Applied to 13 tree species in a northwest Maine study area, MOSVR performed well compared to other prediction methods including single-objective SVR minimizing total error (SOSVR), Random Forest (RF), gradient nearest neighbor (GNN), and Random Forest nearest neighbor (RFNN) algorithms. MOSVR produced the least systematic error for all species with total error that was markedly less or comparable to that of other methods. Predicted patterns of dominance/codominance matched observations well. Although others have presented means of reducing attenuation bias in parametric regression models, MOSVR provides an effective nonparametric approach, and should be fully generalizable to other remote sensing applications and prediction problems.

Chapter 3 presents a multi-objective support vector classification algorithm (MOSVC) that simultaneously minimizes classification bias and either omission or commission error. Multi-objective optimization is used to produce alternative solutions that express tradeoffs between class accuracy and bias under the expectation that different tradeoffs may be more or less beneficial for specific applications. Applied to the prediction of tree species occurrence in northwest Maine, MOSVC produced

diverse sets of alternative models and maps including solutions with zero bias. In contrast, the traditional single-objective approach to model training produced inconsistent and biased outcomes because individual training criteria could not adequately control the balance of omission/commission error. MOSVC solutions with different levels of bias produced different representations of class distributions depending on spatial patterns of omission and commission error. This was made most apparent when used to map forest disturbance because omission/commission errors were visible through comparison of pre- and post-disturbance imagery. MOSVC produced disturbance maps with uniformly high overall accuracy but with different error characteristics demonstrating different impacts of omission and commission error on disturbance class configuration. By approaching error reduction as a multi-objective optimization problem, MOSVC produces alternative solutions that can be used to meet specific application needs or compared to evaluate the sensitivity of application outcomes to map error characteristics.

The algorithms described in chapters 2 and 3 are highly accurate and highly adaptive to different data characteristics. They require very little oversight or intervention, and are therefore amenable to large-scale application. They are, however, also very computationally demanding. Producing data at high volume and low cost will require a computationally efficient software implementation that capitalizes on high-performance computing resources. Current work has therefore focused on the development of an efficient code base, including new approaches to accelerate the execution of genetic algorithms. Efficient front-end and back-end data handling will facilitate high-volume data production from advanced land-imaging satellite systems, including Landsat 8 and Sentinel 2. Work is proceeding on user interfaces to simplify algorithm execution and map production across large areas, with user control of key output map characteristics. The goal is to enable low-cost delivery of products statewide using software implementations that will work well in other regions. Future efforts will bring this work full-

circle, such that new maps are used to better characterize forest conditions and landscape change throughout the state.

References

- Franklin SE. Remote Sensing for Sustainable Forest Management. New York, NY: Lewis Publishers; 2001.
- He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 2009;21:1263–84. doi:10.1109/TKDE.2008.239.
- Irland LC, Dimond JB, Stone JL, Falk J, Baum E. The spruce budworm outbreak in Maine in the 1970's - assessment and directions for the future. Orono, Maine, USA: Maine Agricultural and Forest Experiment Station, University of Maine; 1988.
- Jin S, Sader SA. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sens Environ* 2005;94:364–72. doi:10.1016/j.rse.2004.10.012.
- Langford WT, Gergel SE, Dietterich TG, Cohen W. Map misclassification can cause large errors in landscape pattern indices: examples from habitat fragmentation. *Ecosystems* 2006;9:474–88. doi:10.1007/s10021-005-0119-1.
- Li H, Wu J. Use and misuse of landscape indices. *Landsc Ecol* 2004;19:389–399. doi:10.1023/B:LAND.0000030441.15628.d6.
- Likens GE, Franklin JF. Ecosystem thinking in the Northern Forest - and beyond. *Bioscience* 2009;59:511–3.
- Maine Forest Service. Silvicultural activities report. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 2000.
- Maine Forest Service. Silvicultural activities report. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 1994.
- Rejou-Mechain M, Muller-Landau HC, Detto M, Thomas SC, Le Toan T, Saatchi SS, et al. Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. *Biogeosciences* 2014;11:6827–40. doi:10.5194/bg-11-6827-2014.
- Sader SA, Legaard KR. Inclusion of forest harvest legacies, forest type, and regeneration spatial patterns in updated forest maps: A comparison of mapping results. *For Ecol Manage* 2008;255:3846–56. doi:10.1016/j.foreco.2008.03.047.
- Sader SA, Winne JC. RGB-NDVI colour composites for visualizing forest change. *Int J Remote Sens* 1992;13:3055–67.
- Seymour RS. The red spruce-balsam fir forest of Maine: evolution of silvicultural practice in response to stand development patterns and disturbances. In: Kelty MJ, Larson BC, Oliver CD, editors. *Ecol. Silv. Mix. For.*, Norwell, Massachusetts, USA: Kluwer Academic Publishers; 1992, p. 217–44.

Shao G, Liu D, Zhao G. Relationships of image classification accuracy and variation in landscape statistics. *Can J Remote Sens* 2001;27:35–45.

Shao G, We W, Wu G, Zhou X, Wu J. An explicit index for assessing the accuracy of cover-class areas. *Photogramm Eng Remote Sens* 2003;69:907–13. doi:10.14358/PERS.69.8.907.

Shao G, Wu J. On the accuracy of landscape pattern analysis using remote sensing data. *Landsc Ecol* 2008;23:505–11. doi:10.1007/s10980-008-9215-x.

Wickham JD, Neill RVO, Riitters KH, Wade TG, Jones KB. Sensitivity of selected landscape pattern metrics to land-cover misclassification and differences in land-cover composition. *Photogramm Eng Remote Sens* 1997;63:397–402.

Wilson EH, Sader SA. Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sens Environ* 2002;80:385–96.

Xu Y, Dickson BG, Hampton HM, Sisk TD, Palumbo JA, Prather JW. Effects of mismatches of scale and location between predictor and response variables on forest structure mapping. *Photogramm Eng Remote Sensing* 2009;75:313–22.

ACKNOWLEDGEMENTS

Funding for the work reported in chapter 1 was provided by the USDA Forest Service Agenda 20/20 Program (project entitled "Predicting responses of forest landscape change on wildlife umbrella species: Modeling future effects of alternative forest harvesting scenarios on vertebrate diversity across multiple spatial scales on commercial forestlands in Maine") and the USDA Forest Service Northeastern States Research Cooperative (project entitled "Developing and testing a third party landscape forest sustainability and biodiversity monitoring system"). The work reported in chapters 2 and 3 was supported by the U.S. Carbon Cycle Science Program funded jointly by NASA and USDA National Institute of Food and Agriculture (2011-67003-30351), by the National Science Foundation Dynamics of Coupled Natural and Human Systems Program (DEB-1313688), and by the Northeastern States Research Cooperative (projects entitled "Merging Landsat time-series and FIA data to develop vulnerability maps for spruce budworm decision support", "Evaluating the interacting effects of forest management practices and periodic spruce budworm infestation on broad-scale, long-term forest productivity", and "Long-term outcomes and tradeoffs of forest policy and management practices on the broad-scale sustainability of forest resources: wood supply, carbon, and wildlife habitat"). Confidential coordinates of Forest Inventory and Analysis (FIA) field plots were made available through a collaborative agreement with the USDA Forest Service Northern Research Station FIA Program (FS Agreement No. 2014-MU-11242305-055). I thank Elizabeth Burrill for her patient and good-humored assistance in obtaining and working with FIA spatial data.

I would like to extend my deep gratitude to my committee members Steve Sader, Dan Harrison, Jeremy Wilson, Bill Halteman, Bob Seymour, and Aaron Weiskittel. I am indebted to Steve for accepting me as a doctoral student, despite a dubious background in mathematics and oceanography, and very little directly relevant experience. I thank Dan, Steve, and Jeremy for their leadership in the research that set me on my way in forestry. All committee members have shown extraordinary patience and

commitment as I have slowly worked toward the completion of my degree. My attention has necessarily been divided by many different projects over the years, and my research has strayed from what was originally envisioned. I find it remarkable that everyone has stuck with this endeavor nonetheless. Lastly, I thank Aaron for joining as co-chair of my committee following Steve's retirement. This was only fitting, giving his close involvement in my work dating back to 2010. I can't thank Aaron enough for his collaboration and support over that time.

I should also express my gratitude to the faculty, staff, and students of the School of Forest Resources. It's been a good home. I've been around for quite a long time, and have shared too many relationships to single out. But that said, I would like to extend a special thanks to Matt Noone and Wil Mercier, whose friendships helped to anchor me during a tricky period. Beyond SFR, I should acknowledge the Advanced Computing Group, and Bruce Segee, Larry Whitsel and Chris Wilson in particular, for their varied contributions to the work presented in chapters 2 and 3. Most of the content of those chapters predates recent work with the ACG, but I have no doubt that hours of conversation with Bruce, Larry, and Chris over the past few months have left their mark.

Lastly, I have to find a way to thank my colleague, collaborator, and wife, Erin Simons-Legaard. Erin has been with me in all ways throughout all of these years. We've been working together since 2005, sharing everything. I can no longer envision any other way. Since the completion of her Ph.D. in 2009, we have had two remarkable kids – Finn, born in 2009, and Avery, born in 2017. Regarding my appreciation of her, personal and professional, there's really nothing that can be said in a few moments or a few lines of text - I need a lifetime.

TABLE OF CONTENTS

PREFACE	iii
ACKNOWLEDGEMENTS.....	xi
LIST OF TABLES.....	xvii
LIST OF FIGURES.....	xviii
CHAPTER 1. EVALUATING THE IMPACT OF ABRUPT CHANGES IN FOREST POLICY AND	
MANAGEMENT PRACTICES ON LANDSCAPE DYNAMICS: ANALYSIS OF A LANDSAT IMAGE	
TIME SERIES IN THE ATLANTIC NORTHERN FOREST.....	1
1.1 Abstract	1
1.2. Introduction	2
1.3. Methods	6
1.3.1. Study Area	6
1.3.2. Data Production.....	8
1.3.2.1. Forest Harvest Mapping, 1973-2010.....	9
1.3.2.2. Forest Type Mapping, 1975 and 2004.....	11
1.3.3. Data Validation	12
1.3.3.1. Harvest Time Series Validation.....	12
1.3.3.2. Forest Type Validation.....	14
1.3.4. Data Analysis	14
1.3.4.1. Empirical Orthogonal Functions (EOF) Analysis of Cumulative Harvest	
Time Series.....	15
1.3.4.2. Predominant Patterns of Harvesting and Landscape Change	17
1.4. Results	18
1.4.1. Data Validation	18

1.4.2. EOF Analysis of Cumulative Harvest Time Series	21
1.4.3. Harvesting Trends.....	23
1.4.4. Patterns of Landscape Change	25
1.5. Discussion.....	29
CHAPTER 2. MULTI-OBJECTIVE SUPPORT VECTOR REGRESSION REDUCES SYSTEMATIC ERROR IN	
MODERATE RESOLUTION MAPS OF TREE SPECIES ABUNDANCE.....	35
2.1 Abstract	35
2.2. Introduction	36
2.3. Methods	41
2.3.1. Study Area	41
2.3.2. Reference and Predictor Data	42
2.3.3. Background.....	47
2.3.3.1. Support Vector Regression (SVR)	47
2.3.3.2. Genetic Algorithms	49
2.3.4. Multi-objective SVR (MOSVR) Algorithm Description	51
2.3.5. MOSVR Algorithm Execution.....	55
2.3.6. Model Comparisons	57
2.4. Results	59
2.5. Discussion.....	70
2.6. Conclusions	78
CHAPTER 3. CONTROLLING MAPPED CLASS PREVALENCE AND THE BALANCE OF ERROR BY	
MULTI-OBJECTIVE OPTIMIZATION OF SUPPORT VECTOR CLASSIFICATION MODELS.....	79
3.1 Abstract	79
3.2. Introduction	80

3.3. Background	85
3.3.1. Support Vector Machines.....	85
3.3.2. Genetic Algorithms.....	87
3.3.3. Model Performance Metrics	89
3.4. Algorithm Description and Implementation	91
3.4.1. Algorithm Overview.....	91
3.4.2. Algorithm Implementation.....	92
3.5. Example Applications: Study Area and Methods	97
3.5.1. Study Area	97
3.5.2. Tree Species Occurrence	98
3.5.2.1. Reference Data and Spatial Covariates	99
3.5.2.2. Algorithm Execution	101
3.5.3. Forest Canopy Disturbance	102
3.5.3.1. Reference Data and Spatial Covariates	102
3.5.3.2. Algorithm Execution	104
3.6. Example Applications: Results.....	105
3.6.1. Tree Species Occurrence	105
3.6.2. Forest Canopy Disturbance	113
3.7. Discussion.....	117
BIBLIOGRAPHY	123
APPENDICES	136
Appendix A. Chapter 1 Landsat Image Processing.....	136
Appendix B. Validation of Chapter 1 Maps	139
Appendix C. Preparation of Spatial Predictor Data Used in Chapter 3	147

BIOGRAPHY OF THE AUTHOR.....	153
------------------------------	-----

LIST OF TABLES

Table 1.1.	Landsat images used to map forest harvesting (1973-2010) and forest composition (1975 and 2004)	8
Table 1.2.	Error matrix and accuracy estimates for validation classes aggregated from the 1973-2010 forest harvest time series	19
Table 1.3.	Error matrix and accuracy estimates for the 2010 cumulative harvest map	20
Table 1.4.	Error matrix and accuracy estimates for the 1975 forest type map	21
Table 1.5.	Error matrix and accuracy estimates for the 2004 forest type map	21
Table 2.1.	Landsat images used for predictive modeling of tree species relative abundance	43
Table 2.2.	Terrain and climate variables used to model and map tree species relative abundance	45
Table 2.3.	Reference data characteristics of the 13 modeled tree species	56
Table 2.4.	Predictive performance by species and model type	62
Table 2.5.	Multi-objective support vector regression variable and training sample selection	73
Table 3.1.	Confusion matrix for a binary classification or presence-absence model	90
Table 3.2.	Model evaluation or map accuracy metrics for a binary classification or presence-absence model	90
Table C.1.	Landsat images used to model and map tree species occurrence (ca. 2004) and canopy disturbance (2004-2007)	148
Table C.2.	Terrain and climate variables used to model and map tree species occurrence (ca. 2004)	150

LIST OF FIGURES

Figure 1.1.	Northern Maine, U.S.A. study area with 5 km square sample landscape units superimposed	7
Figure 1.2.	Forest harvest trends and landscape change for two sample grid cells	22
Figure 1.3.	Predominant patterns of harvesting.....	24
Figure 1.4.	Time series of regenerating and partially harvested forest area.....	25
Figure 1.5.	Time series of intact mature forest configuration metrics	26
Figure 1.6.	Time series of regenerating forest configuration metrics	27
Figure 1.7.	Initial landscape composition and changes in composition, 1975-2004	28
Figure 1.8.	Changes in average harvest patch size through time	30
Figure 2.1.	Study area	41
Figure 2.2.	Genetic algorithm chromosome design.....	52
Figure 2.3.	Multi-objective support vector regression algorithm implementation.....	53
Figure 2.4.	Pareto front and sample of Pareto-optimal models for species ABBA (balsam fir)	60
Figure 2.5.	Trends in residual values for selected model types fit to species ABBA (balsam fir)	66
Figure 2.6.	Trends in residual values for selected model types fit to individual tree species.....	67
Figure 2.7.	Observed and predicted patterns of species codominance	69
Figure 2.8.	Spatial predictions of relative abundance for species ABBA (balsam fir)	77
Figure 3.1.	Multi-objective support vector classification algorithm overview	91
Figure 3.2.	Genetic algorithm chromosome design.....	93
Figure 3.3.	Multi-objective support vector classification algorithm implementation.....	94

Figure 3.4.	Genetic recombination and mutation operations	97
Figure 3.5.	Study area	98
Figure 3.6.	Black ash occurrence - model performance and estimation uncertainty.....	106
Figure 3.7.	Black ash occurrence maps	108
Figure 3.8.	Black ash occurrence - replicate runs compared to single-objective optimization outcomes	109
Figure 3.9.	Eastern white pine occurrence - replicate runs compared to single-objective optimization outcomes	111
Figure 3.10.	Red spruce occurrence - replicate runs compared to single-objective optimization outcomes	112
Figure 3.11.	Forest canopy disturbance - replicate runs.....	113
Figure 3.12.	Forest disturbance maps.....	114
Figure 3.13.	Cloud- and shadow-induced error in forest disturbance maps	116

CHAPTER 1

EVALUATING THE IMPACT OF ABRUPT CHANGES IN FOREST POLICY AND MANAGEMENT PRACTICES

ON LANDSCAPE DYNAMICS: ANALYSIS OF A LANDSAT IMAGE TIME SERIES IN THE ATLANTIC

NORTHERN FOREST

1.1. Abstract

Sustainable forest management is based on functional relationships between management actions, landscape conditions, and forest values. Changes in management practices make it fundamentally more difficult to study these relationships because the impacts of current practices are difficult to disentangle from the persistent influences of past practices. Within the Atlantic Northern Forest of Maine, U.S.A., forest policy and management practices changed abruptly in the early 1990s. During the 1970s-1980s, a severe insect outbreak stimulated salvage clearcutting of large contiguous tracts of spruce-fir forest. Following clearcut regulation in 1991, management practices shifted abruptly to near complete dependence on partial harvesting. Using a time series of Landsat satellite imagery (1973-2010) we assessed cumulative landscape change caused by these very different management regimes. We modeled predominant temporal patterns of harvesting and segmented a large study area into groups of landscape units with similar harvest histories. Time series of landscape composition and configuration metrics averaged within groups revealed differences in landscape dynamics caused by differences in management history. In some groups (24% of landscape units), salvage caused rapid loss and subdivision of intact mature forest. Persistent landscape change was created by large salvage clearcuts (often averaging > 100 ha) and conversion of spruce-fir to deciduous and mixed forest. In groups that were little affected by salvage (56% of landscape units), contemporary partial harvesting caused loss and subdivision of intact mature forest at even greater rates. Patch shape complexity and edge density reached high levels even where cumulative harvest area was relatively low. Contemporary practices introduced more numerous and much smaller patches of stand-replacing disturbance (typically

averaging <15 ha) and a correspondingly large amount of edge. Management regimes impacted different areas to different degrees, producing different trajectories of landscape change that should be recognized when studying the impact of policy and management practices on forest ecology.

1.2. Introduction

Forest policy and management practices within the U.S. have changed substantially following widespread dissatisfaction with management overly focused on the production of wood fiber and game species habitat. Over the past several decades, managers of public and private lands have to varying degrees incorporated a much wider set of objectives including the protection or provision of amenities, biodiversity, and ecosystem services (Kohm and Franklin, 1997; Seymour and Hunter, 1999). Much of this change followed from recognition that management practices had undermined the landscape conditions needed to support certain forest values. Advances in scientific knowledge, stakeholder engagement, and government oversight of public interests have led to changes in public policy and private forest practices intended to improve the function of managed forest landscapes (Cubbage and Newman, 2006; Kohm and Franklin, 1997; Seymour and Hunter, 1999). There are many, varied mechanisms of change. Management has evolved in response to public perception and market incentives. More abrupt changes have resulted from legislation and implementation of forest policy by government at all levels, from municipal to federal. State governments have been particularly active in legislating and enforcing regulatory programs (Cubbage and Newman, 2006; Ellefson et al., 2007). Due to the complexity of ecological, economic, and social issues intertwined in the problem of forest management, regulatory programs are put into place with incomplete knowledge of future effects.

The sustainable management of forest landscapes and development of effective forest policy requires an understanding of the functional relationships between management practices, changes in landscape conditions, and ecological response. Abrupt changes in forest policy or other drivers of landscape dynamics make it fundamentally more difficult to evaluate these relationships. Because

ecological processes operate over a wide range of temporal scales, responses to landscape change are time-dependent. Changes in species presence or abundance are frequently delayed following periods of rapid landscape change, and ecological communities take time to equilibrate to new landscape dynamics imposed by new management practices (Ewers and Didham, 2006; Jackson and Sax, 2010; Schrott et al., 2005; With, 2007). Delayed responses may effectively decouple ecological processes from recent patterns of landscape change (With, 2007). The degree to which this occurs will vary depending on species life histories and the spatiotemporal dynamics of forest disturbance and recovery (Schrott et al., 2005; With, 2007), but in general the ecological effects of forest policy change may emerge over long timeframes. This may be particularly true where past management practices imposed landscape conditions that persist for long periods. Legacies of past management practices (e.g., forest composition, spatial configuration of stand types) persist because they limit management options or alter patterns of natural disturbance or succession (James et al., 2007; Sturtevant et al., 2014). Unrecognized legacies and lagged responses may confound the attribution of observed ecological impacts to specific management practices.

Empirical studies of forest loss or fragmentation effects commonly rely on a space-for-time substitution (Pickett, 1989), where replicate landscapes or patches are selected based on the current amount or configuration of forest (e.g., McGarigal and McComb, 1995; Radford et al., 2005). Although the intent is to study a fundamentally dynamic process, replication occurs in space rather than time, and landscape disturbance history is treated as an extraneous variable that is not controlled by experimental design. Inferences require the assumption that disturbance history acts as a random error term (Eberhardt and Thomas, 1991) when in fact it may be confounded with the experimental variables of current forest amount or configuration (Schrott et al., 2005). Studies that are intended to reveal impacts of landscape change should integrate disturbance history or temporal variability of landscape condition into study design (e.g., Price et al., 2013). Similarly, where different management practices have been

imposed at different times, knowledge of management history is needed to differentiate the consequences of contemporary practices from persistent impacts of past practices. Empirical evidence will otherwise be difficult to establish following abrupt changes in management regimes, when empirical study is perhaps most needed.

Satellite images provide the synoptic views needed to characterize forest conditions and landscape change. The ~40-year depth of the Landsat image archive in particular facilitates studies of forest landscape dynamics. However, there are relatively few retrospective analyses of landscape dynamics following abrupt changes in forest management practices. In the Pacific Northwest region of the U.S., Landsat image time series have been used to address the consequences of federal forest policy change (Healey et al., 2008; e.g., Kennedy et al., 2012). Landsat-derived forest cover maps and disturbance time series have been used to evaluate changes in forest conditions following the collapse of socialism in Eastern Europe and the former Soviet Union (e.g., Baumann et al., 2012; Bergen et al., 2008). In these cases, disturbance rates or measures of landscape change were summarized over time periods of interest (i.e., periods before and after policy change or sociopolitical reform) and over study areas defined by political boundaries, ecoregions, or image extents. Results provide summaries of change in useable forms, but the spatiotemporal dynamics of landscape change are resolved only in so far as they are partitioned by predetermined time periods or study areas. Empirical study of ecological processes affected by management requires knowledge of how management practices have influenced landscape dynamics across a range of ecologically relevant scales, but the spatial and temporal heterogeneity of management effects are difficult to synthesize over large areas and long time periods.

The Atlantic Northern Forest of the northeastern U.S. encompasses roughly 11 million hectares within a transition zone between the northern boreal forest and the southern temperate deciduous-dominant forest. A substantial portion of this area lies within northern Maine, the largest contiguous block of undeveloped forestland in the nation (~4 Mha). Despite a long history of logging and

commercial management for fiber production, major changes in management practices within recent decades have led to contemporary landscape conditions with little historical precedent. The spruce-fir forests of the region are subject to periodic infestations of the eastern spruce budworm (*Choristoneura fumiferana* (Clem.)), a native pest that causes widespread defoliation and mortality of balsam fir (*Abies balsamea*) and spruce (*Picea spp.*) trees (Irland et al., 1988; Seymour, 1992). Maine's last outbreak occurred ca. 1972-1988 and stimulated broad-scale salvage harvesting by clearcut (Irland et al., 1988). Public concern over the size of salvage clearcuts led to the passage of the Maine Forest Practices Act (FPA; 12 MRSA §8867-A to §8888) in 1989 and its implementation in 1991. The FPA fundamentally changed management practices by placing restrictions and disincentives on clearcutting. As a proportion of annual harvest area, clearcuts fell from 44% in 1989 to 10% in 1994 (Maine Forest Service, 1994) and less than 5% by 2000 (Maine Forest Service, 2000).

Management practices in Maine have elicited concerns regarding the sustainable provision of forest values. During the budworm outbreak, salvage logging rates were well above recognized long-term allowable levels (Irland et al., 1988). Regeneration failures within salvage clearcuts resulted in the conversion of large areas of spruce-fir forest to deciduous and mixed types (Seymour, 1992). Following implementation of the FPA, state records indicate that annual harvest area roughly doubled during the 1990s (Maine Forest Service, 2000, 1994) as landowners maintained similar extraction rates via partial harvest practices that require a larger footprint to achieve the same volume removal. The spatial dynamics associated with implementation of the FPA have been partially assessed. Analysis of a Landsat-derived disturbance time series (1988-1999) found that implementation of the FPA coincided with a change toward fewer and smaller clearcut patches, and fewer but larger partial harvest patches (Sader et al., 2003). A subsequent analysis of Landsat-derived forest cover and disturbance data found that harvest patches of the 1980s were larger and more compact than patches of the 1990s, but this study did not differentiate clearcuts from partial harvests (Sader et al., 2006). Management practices and

harvest rates differ between private forestland owners (Hagan et al., 2005; Jin and Sader, 2006), suggesting important differences in post-FPA landscape change. However, rates and patterns of landscape change attributable to pre- and post-FPA management regimes have not been sufficiently resolved to support a more complete assessment of policy impact on landscape dynamics.

The objective of our research was to characterize predominant patterns of cumulative landscape change in the Atlantic Forest of northern Maine, and to evaluate how pre- and post-FPA management regimes have influenced landscape conditions across space and time. We used Landsat imagery and forest inventory data to develop and validate forest composition maps and a time series of forest harvest maps (1973-2010). We modeled predominant temporal patterns of harvesting and segmented a large study area into groups of landscape units with similar harvest histories. We then linked harvest history with changes in landscape composition and configuration in order to characterize the evolution of landscape conditions in response to forest management practices before and after abrupt change induced by the FPA. Our approach provided an objective synthesis of predominant patterns of change associated with specific landscape units, with the spatial and temporal resolution needed to attribute change to different management regimes.

1.3. Methods

1.3.1. Study Area

Our northern Maine, U.S.A. study region (Fig. 1.1) was defined by the overlap of Landsat images and includes ~1.5 Mha of forestland. Rural development and agriculture are concentrated in a few small areas. Topography is generally flat or rolling with occasional low mountains and an extensive network of rivers, lakes, and wetlands. Forest types are typical of the Atlantic Northern Forest and generally occur in predictable patterns associated with climatic gradients and soil conditions determined by glacial deposition (Seymour, 1995). Northern hardwood species (*Acer rubrum*, *Acer saccharum*, *Betula alleghaniensis*, *Betula papyrifera*, *Fagus grandifolia*) predominate across lower hilltops and at mid-slope.

Spruce-fir species (*Abies balsamea*, *Picea glauca*, *Picea mariana*, *Picea rubens*) predominate where soil or microclimatic conditions exclude the more demanding hardwoods. Mixedwood stands commonly occur along ecotones or as a result of successional dynamics following disturbance. Shade-intolerant hardwood species (e.g., *Populus tremuloides*, *Betula papyrifera*) are commonly found following intense disturbance. Periodic defoliation by spruce budworm is the most prominent form of natural disturbance. Windthrow is common but generally results in small canopy gaps (Lorimer and White, 2003). Virtually all forestland is considered commercially productive (Seymour, 1995) and roughly 90% is private. Public lands are interspersed and primarily state-owned.

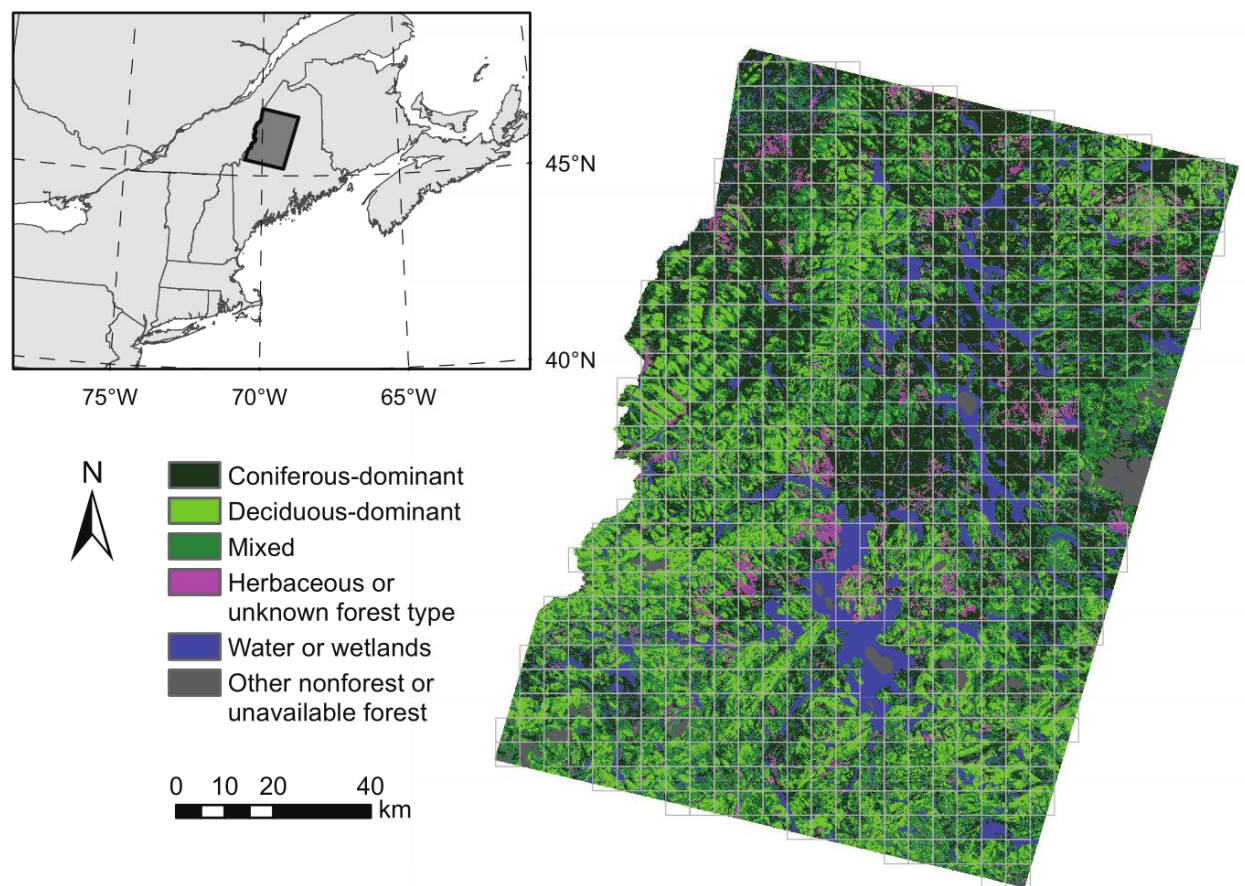


Figure 1.1. Northern Maine, U.S.A. study area with 5 km square sample landscape units superimposed. Harvesting trends and patterns of landscape change were calculated for forestland assumed available for harvest. Mapped forest composition classes demonstrate the spatial distribution of general forest types at the onset of our study period (1975). State and provincial boundaries displayed in the inset map were obtained from the National Atlas of the U.S. (Political Boundaries) and the Atlas of Canada (National Frameworks Data, Census Subdivisions and Population Ecumene).

Table 1.1. Landsat images used to map forest harvesting (1973-2010) and forest composition (1975 and 2004). Images were acquired over Landsat Worldwide Reference System (WRS)-2 path 12, row 28 (1985-2010) and WRS-1 path 13, row 28 (1973-1982). Unless otherwise indicated, images were obtained from the U.S. Geological Survey Earth Resources Observation and Science Center.

Acquisition date	Landsat sensor	Landsat satellite	% forestland under cloud/shadow
2010, August 30	TM	5	<0.1
2007, June 17 ^a	TM	5	0.8
2004, June 10	TM	5	0.8
2001, May 25	ETM+	7	< 0.1
2000, August 26	ETM+	7	1.6
1999, June 13	TM	5	< 0.1
1997, June 23	TM	5	16.0
1995, July 4	TM	5	5.0
1993, September 16 ^b	TM	5	15.9
1991, June 7 ^b	TM	5	< 0.1
1988, September 2	TM	5	0.9
1988, September 2	MSS	5	0.9
1985, June 22 ^c	MSS	5	4.3
1982, July 30	MSS	3	0.9
1978, August 11	MSS	2	< 0.1
1975, August 9	MSS	2	0
1973, July 23 ^c	MSS	1	0.1

^aAreas of cloud cover filled with TM image data acquired on 22 August 2007.

^bAvailable through the Maine GAP Analysis Project.

^cAvailable through the North American Landscape Characterization project.

1.3.2. Data Production

Forest harvest and composition maps were assembled from a time series of Landsat Multispectral Scanner (MSS), Thematic Mapper (TM), and Enhanced Thematic Mapper Plus (ETM+) images acquired during summer leaf-on conditions (Table 1.1). Consecutive images were spaced 1-4 years apart, as determined by the availability of high quality, predominantly cloud-free imagery. Mapping procedures were applied to forested pixels as identified by the 1993 Maine Gap Analysis Program (GAP) land cover map. The GAP map represents conditions near the midpoint of our time series, and discriminated forest from non-forest with an estimated 100% accuracy within our study area (Hepinstall et al., 1999). Not all forestland within our study area is operable or available for harvesting. To normalize harvest rates and metrics of landscape change according to the amount of available

forestland in different landscape units, we masked forest pixels over 823 m (2700 ft) in elevation or >40% slope, as determined from the 1 arc-second National Elevation Dataset. Harvesting under these conditions has historically been allowed by special permit only, and we consider these areas inoperable or otherwise unavailable for harvest. Forested islands were masked as well, with the exception of one large island with a history of harvesting. Less than 3% of forestland was masked as unavailable. Refer to Appendix A for a detailed description of image processing performed prior to forest harvest and composition mapping.

1.3.2.1. Forest Harvest Mapping, 1973-2010

Forest harvest maps were produced using a change detection procedure based on vegetation index values calculated from sequential Landsat images. As initially described by Sader and Winne (1992), forest canopy disturbance and recovery can be visualized using a three-band color composite image incorporating values of the normalized difference vegetation index ($NDVI = [near-infrared - red] / [near-infrared + red]$) acquired on three separate dates. Classification of the three-date NDVI data produces a thematic map depicting forest canopy changes (Sader et al., 2003). Other vegetation indices may be substituted for the NDVI and the normalized difference moisture index ($NDMI = [near-infrared - mid-infrared] / [near-infrared + mid-infrared]$) has been found particularly effective in discriminating partial canopy disturbance using TM/ETM+ data (Jin and Sader, 2005; Wilson and Sader, 2002). Whereas the NDVI represents a normalized contrast between near-infrared and red reflectance, the NDMI contrasts near-infrared and mid-infrared reflectance. The improved sensitivity of the NDMI to partial canopy disturbance is generally attributable to the heightened sensitivity of mid-infrared wavelengths to differences in forest canopy structure, leaf area, and biomass (Cohen and Goward, 2004; Jin and Sader, 2005).

We classified three-date NDMI and NDVI composites to produce forest change maps from TM/ETM+ and MSS image sequences, respectively. MSS imagery lacks a mid-infrared band required for

calculation of the NDMI. This difference, coupled with reduced spatial and radiometric resolution, limits the efficacy of MSS imagery for detection of partial canopy disturbance. Disturbances mapped using MSS imagery (1973-1988) represent stand-replacing events, predominantly spruce budworm salvage clearcuts. Disturbances mapped using TM/ETM+ imagery (1988-2010) represent a wide range of intensities, and we differentiated two intensity classes interpreted as stand-replacing and partial canopy disturbance. The stand-replacing class was intended to represent harvests in which a new cohort was established following removal of a large proportion of the canopy, whether by clearcut as defined by the FPA (12 MRSA §8868, Maine Forest Service Rules Chapter 20) or by other harvest types. Mapped disturbance events were almost exclusively the result of harvest operations and we therefore refer to our data as a time series of forest harvest maps.

Harvest maps were produced by unsupervised classification of overlapping three-date NDVI or NDMI image sequences (e.g., 1973-1975-1978, 1975-1978-1982, ...). Classification of a three-date sequence mitigates the impact of cloud cover in the second image provided the first and third give clear views. An ISODATA algorithm applied to each three-date composite produced 50 statistical classes that were interpreted into forest disturbance, regrowth, and no-change information classes. Stand-replacing and partial harvest classes derived from TM/ETM+ imagery were differentiated based on the relative magnitude of NDMI change, guided by visual interpretation of Landsat imagery and available aerial photography. Confusion between light partial harvests and changes induced by factors such as atmospheric effects or interannual variability in forest phenology were resolved through on-screen editing (Sader and Legaard, 2008). Individual harvest maps were compiled for each time interval (e.g. 1973-1975, 1975-1978, ...) by combining equivalent harvest classes from overlapping three-date change maps. Harvest patches less than 0.81 ha in size were removed, and a 3x3 pixel majority filter was applied to consolidate patch boundaries and simplify the patch structure of maps produced from TM/ETM+ imagery to more closely match maps produced from the lower resolution MSS imagery.

We produced a time series of maps depicting cumulative harvest impact (1975-2010) by overlaying successive harvest maps. For each time series date, a pixel was labeled as regenerating forest if preceding intervals included a harvest 1973-1988 or a stand-replacing harvest 1988-2010. A pixel was labeled as partially harvested if preceding intervals included only a single partial harvest. When preceding intervals included multiple partial harvests, pixels were labeled as regenerating forest, reflecting the anticipated ecological and silvicultural effects of multiple entries within the ~20-year period over which partial harvests were mapped (1988-2010). For each date of our time series, the result depicts the cumulative footprint of harvest operations since 1973.

1.3.2.2. Forest Type Mapping, 1975 and 2004

We mapped forest composition using equivalent unsupervised classification methods applied to each of the 1975 MSS and 2004 TM images. Dates were selected on the basis of cloud cover and image quality. For the purpose of forest type mapping, small areas of cloud cover in the 2004 image were replaced with data from the 2001 ETM+ image. Statistical classes produced from an ISODATA algorithm were aggregated to coniferous-dominant (>75% coniferous), deciduous-dominant (>75% deciduous), and mixed type classes through visual interpretation of Landsat imagery, with reference to available aerial photography and existing land cover maps. In some previously disturbed areas, exposed soils, woody debris, or herbaceous vegetation precluded the assignment of forest type and pixels were instead assigned to an indeterminate class. Patches less than 0.81 ha in size were removed and a 3x3 majority filter was applied to each map to consolidate patch boundaries and simplify the 2004 patch structure to more closely match the 1975 data.

Assignment of ISODATA classes to forest types was subjective and sometimes difficult. A mistaken assignment could lead to bias in the representation of forest type extent. If for example pixels representing forest with a deciduous component of 70-75% were mistakenly committed to the deciduous-dominant class rather than the mixed class, the extent of the deciduous class would be

overestimated according to the class definition of >75% deciduous. We used validation data obtained from field plots (described below and in Appendix B) to iteratively refine the aggregation and labeling of ISODATA classes to ensure that the 1975 and 2004 maps provide unbiased representations of forest type classes at the same thresholds of forest composition. For each map, we identified coniferous- and deciduous-dominant class thresholds for which omission and commission errors were balanced. To do so, we varied coniferous and deciduous threshold values from 50-95% in increments of 5%, assigned reference class labels based on threshold values, and calculated omission and commission error rates. We iteratively refined the maps and reevaluated error rates until a reasonable balance was achieved at the same threshold for both maps, facilitating meaningful comparisons of class extent between maps.

1.3.3. Data Validation

The U.S. Forest Service Forest Inventory and Analysis (FIA) Program provides quality-assured measurements of forest attributes from a national network of field plots adhering to a systematic sampling design (McRoberts et al., 2005). We made extensive use of FIA data as a statistically rigorous basis for map validation. However, use of field plot data for map validation is subject to uncertainty arising primarily from mismatches in location and scale between field plots and map pixels. Validation using FIA data should be considered an assessment of agreement with an accepted and widely utilized source of information on forest conditions, rather than an assessment of accuracy against ground truth. Here we provide an overview of our validation procedures; details are provided in Appendix B.

1.3.3.1. Harvest Time Series Validation

FIA estimates of forest age have been used to validate Landsat-derived disturbance time series under the assumption that trees sampled for age estimation germinated at the time of disturbance (Thomas et al., 2011). However, age is an imprecise estimate of the timing of past disturbance due to estimation uncertainty and variation in the timing of germination with respect to canopy removal. A new cohort may have been established from a seed source several years following disturbance or as

advance regeneration prior to disturbance. Alternatively, visual interpretation of Landsat imagery is a credible means of dating disturbance events (Cohen et al., 2010; Sader et al., 2003; Sader and Legaard, 2008). Unfortunately, visual discrimination of harvest intensity at the pixel scale is highly subjective. We developed a validation procedure based on visual interpretation of Landsat imagery over FIA plot locations. Image interpretation was used to date harvest events; FIA plot data were used to discriminate stand-replacing and partial harvests.

In Maine, the contemporary FIA inventory design was established in 1999, with 20% of plots surveyed annually during sequential 5-year cycles. Although data are available from plots measured during earlier inventories, coordinate locations are known for only a fraction of those plots. We therefore used data collected during contemporary inventory cycles to discriminate past harvest intensity. A harvest event identified by image interpretation was labeled stand-replacing provided FIA age dated stand origin to 1970 or later (allowing for advance regeneration prior to 1973) and field crews labeled the stand as either sapling or poletimber. However, for plots harvested after 1999, recorded stand age was an unreliable indicator of stand-replacing disturbance because age estimates frequently corresponded to a few residual stems rather than a newly established cohort. In these cases, intensity classes were discriminated using plot measurements made during consecutive 5-year inventory cycles; a harvest was labeled stand-replacing if plot basal area (cross-sectional area of stems measured at 1.37 m) was reduced by >70%.

Our validation sample of 509 plots was insufficient to produce reasonably precise accuracy estimates for individual time series intervals. We therefore aggregated intervals into six validation classes: 1973-1988 stand-replacing harvest, 1988-1999 stand-replacing harvest, 1988-1999 partial harvest, 1999-2010 stand-replacing harvest, 1999-2010 partial harvest, and intact mature forest (no history of harvest, 1973-2010). Map and reference validation class labels were assigned in a manner consistent with the construction of cumulative harvest maps. Where multiple entries occurred, labels

were assigned based on the date of first stand-replacing disturbance. Where multiple partial harvests occurred, labels of either 1988-1999 or 1999-2010 stand-replacing were assigned based on the date of second entry. Map and reference labels were compiled into an error matrix. Overall accuracy, user accuracy (the complement of class commission error), producer accuracy (the complement of class omission error), and corresponding standard error estimates were calculated by poststratification (Card, 1982). Additionally, we evaluated the accuracy of our 2010 cumulative harvest map by further aggregating validation classes into regenerating, partially harvested, and intact mature forest.

1.3.3.2. Forest Type Validation

The 1975 and 2004 forest type maps were validated using FIA plot measurements of coniferous and deciduous live tree basal area collected during 1980-1982 and 1999-2003 inventories, respectively. Differences in dates between maps and field inventories were resolved by excluding samples where intervening harvests occurred. For 2004 map validation, we excluded plots mapped as harvested from 1999-2004; for 1975, we excluded plots mapped as harvested from 1975-1982. A sample of 445 plots remained for validation of the 2004 map; only 70 plots were available for validation of the 1975 map. As previously described, we identified coniferous-dominant and deciduous-dominant class thresholds for which errors were best balanced and mapped class extents least biased. Following refinements made to improve consistency between maps, an error matrix was compiled for each map based on selected threshold values. Estimates of overall, user, and producer accuracy were calculated by poststratification (Card, 1982).

1.3.4. Data Analysis

To quantify harvest rates through time, identify trends, and associate trends with changes in landscape conditions, we tessellated our study area into landscape units using a 5 km square grid (Fig. 1.1). A 5 km grid cell size was a somewhat arbitrary compromise: small enough to resolve spatial variations in harvest history and consequent landscape change, but large enough to calculate

meaningful trends in harvest rates and landscape pattern metrics. We excluded grid cells with <50% available forest or <5% of available forest harvested from 1975-2010 (17 cells). A sample of 608 grid cells remained.

1.3.4.1. Empirical Orthogonal Function (EOF) Analysis of Cumulative Harvest Time Series

An EOF analysis identifies a sequence of uncorrelated patterns or modes of variability that characterize variation within a two-dimensional data set (Preisendorfer, 1988). EOF analysis is commonly employed in meteorology and oceanography, where conventional applications decompose time series of geospatial data into characteristic spatial patterns whose contributions to observed variation change through time. EOF outcomes can just as readily be interpreted as characteristic temporal patterns whose contributions to observed variation differ between locations (e.g., Small and Elvidge, 2013). We performed an EOF analysis to identify characteristic temporal patterns of variation in cumulative harvest area sampled across our 5 km grid. Cumulative harvest time series were arranged as rows within a matrix X ($M = 608$ rows; $N = 15$ columns). EOF analysis decomposed X into matrices A and B such that $X = A \cdot B$ (A is $M \times N$; B is $N \times N$). The rows of B represented a sequence of mutually uncorrelated patterns of temporal variability referred to as empirical orthogonal functions (EOFs). The columns of A represented a complementary set of spatial patterns referred to as amplitude functions. The observed cumulative harvest time series were thereby represented as linear combinations of temporal EOFs, whose contributions were given by the spatial amplitude functions. EOF analysis is mathematically equivalent to principal component analysis (PCA). The temporal EOFs are computed as the eigenvectors of the dispersion matrix $X^T X$ and are equivalent to the loading vectors or principal components of a PCA. The spatial amplitudes correspond to the PCA scores obtained by projecting the time series of X onto the subspace spanned by the EOFs.

A traditional EOF analysis or PCA is sensitive to extreme observations and outliers, which can distort the outcome such that dominant modes of variability represent contrasts between anomalous

and regular observations rather than patterns of variability within regular observations (Hubert and Engelen, 2004). We performed our EOF analysis using a variant of the robust algorithm ROBPCA that is also suitable for skewed distributions (Hubert et al., 2009; Verboven and Hubert, 2010). Cumulative harvest area distributions were significantly skewed for 9 of the 15 time series dates (medcouple, $p < 0.05$; (Brys et al., 2004)). The ROBPCA algorithm is based on robust estimation of the covariance matrix from a specified proportion of samples with minimal outlyingness. The proportion of samples used may range from 0.5 to 1, and the value selected represents a compromise between the robustness and efficiency of the estimate. We used a sample proportion of 0.9 following exploratory analysis which suggested that relatively few outliers were present. Outcomes were not sensitive to the exact value used. Prior to analysis, we centered and scaled the cumulative harvest time series by removing the median and dividing by the median absolute deviation (computed across all cells, for each observation date). Scaling improved the fit of the EOF model for intervals near the beginning and end of the study period.

Paired EOFs and amplitude functions comprise orthogonal modes of variability, ordered by the amount of total variance explained. We modeled cumulative harvest time series as linear combinations of dominant EOFs, selected based on the proportion of overall variance explained by successive modes. By including only dominant modes, modeled time series represent statistically coherent variability in harvesting patterns that occurred over large portions of the study area. The ROBPCA algorithm provides a measure of orthogonal distance between samples and the subspace defined by dominant EOF modes. Unusually large orthogonal distances indicate outlying samples that do not conform to characteristic patterns defined by dominant modes. We identified 12 orthogonal outliers using the ROBPCA nominal cutoff value (Hubert et al., 2009) and excluded them from subsequent analyses.

1.3.4.2. Predominant Patterns of Harvesting and Landscape Change

To classify predominant temporal patterns of harvesting from the EOF analysis and to associate those patterns with groups of grid cells, we performed an agglomerative hierarchical clustering (Everitt et al., 2011) of modeled time series. Using Ward's minimum variance method (Ward, 1963), we produced a dendrogram and identified clusters of grid cells with similar harvest history. The mean of the modeled time series from each cluster demonstrated a predominant pattern of harvesting through time, representative of a group of landscape units.

Landscape composition metrics were calculated for grid cells and averaged within groups to evaluate changes associated with predominant harvesting trends. Within our time series of cumulative harvest maps, available forestland was classified as either regenerating, partially harvested, or intact mature forest (no harvesting, 1973-2010). The EOF and cluster analyses produced time series of cumulative harvest area, the reciprocal of intact mature forest area. We also produced time series of cumulative partial harvest and regenerating forest area to evaluate changes in harvest intensity associated with predominant harvesting trends. Available forestland was summarized by 1975 forest type to associate harvest history and landscape change with initial landscape composition. To evaluate composition change as a legacy of harvest practices, we quantified forest type change between 1975 and 2004 for areas harvested before 2004. Composition change in unharvested forestland was not evaluated as part of this research.

Early successional and intact late successional forest patches are landscape elements of particular interest, given the expansion of partial harvest practices in the 1990s. Cumulative changes in the patch configuration of regenerating and intact mature forest were evaluated by calculating time series of landscape metrics. We selected a small number of metrics of general relevance to forest ecology that quantify primary aspects of class configuration thought to have been affected by changes in management practices. Metrics were calculated using Fragstats version 4.2 (McGarigal et al., 2012).

Area-weighted mean patch size (Area_AM; ha) and area-weighted mean fractal dimension (Frac_AM; unitless) were calculated as measures of average patch area and shape complexity. Patch density (PD; patches/100 ha) was calculated as a simple measure of patch subdivision. The prevalence of edge conditions was quantified by edge density (ED; m/ha). An eight-neighbor rule was used for patch delineation. Non-forest and unavailable forest classes were treated as external to the landscape in order to normalize metric values across grid cells containing different amounts of managed forestland. Grid cell borders and non-forest edges were not included in the calculation of ED. Unavailable forest edges were included in the calculation of regeneration ED but not intact mature ED (i.e., unavailable forest was treated as intact mature forest for the purpose of calculating ED).

1.4. Results

1.4.1. Data Validation

Harvest validation classes were mapped with an overall accuracy of 88% (Table 1.2). User and producer accuracies for the intact mature class were high (>95%) and well balanced, indicating an accurate depiction of overall harvest footprint. Stand-replacing harvests of 1973-1988 were mapped with high accuracy (89-91%) compared to subsequent periods in which confusion between stand-replacing and partial harvests reduced accuracies for both classes (75-91%). Individual class accuracies were reasonably well balanced save for 1988-1999 stand-replacing harvests, which may have been systematically under-represented. However, the criteria used to establish reference class labels differed between periods, and differences in class accuracy estimates may partly reflect inconsistency in discriminating harvest intensity from available validation data. The 2010 cumulative harvest map depicted regenerating and partially harvested forest with >86% and >75% accuracy, respectively (Table 1.3). Overall accuracy associated with regenerating, partially harvested, and intact mature forest classes approached 90%.

Table 1.2. Error matrix and accuracy estimates for validation classes aggregated from the 1973-2010 forest harvest time series. Standard error estimates are provided in parentheses.

Mapped validation class	Reference validation class						Total	User accuracy
	1973-1988 stand-replacing	1988-1999 stand-replacing	1988-1999 partial harvest	1999-2010 stand-replacing	1999-2010 partial harvest	Intact mature ¹		
1973-1988 stand-replacing	73	1	2	0	0	4	80	91.3% (3.2%)
1988-1999 stand-replacing	1	49	3	1	0	0	54	90.7% (4.0%)
1988-1999 partial harvest	3	7	36	0	0	2	48	75.0% (6.3%)
1999-2010 stand-replacing	1	1	1	30	7	0	40	75.0% (6.9%)
1999-2010 partial harvest	1	0	0	7	33	3	44	75.0% (6.6%)
Intact mature ¹	4	2	2	0	3	232	243	95.5% (1.3%)
Total	83	60	44	38	43	241	509	
Producer Accuracy ²	88.6% (2.7%)	77.2% (4.4%)	83.9% (5.0%)	81.6% (5.0%)	76.4% (5.1%)	95.3% (1.5%)		
Overall Accuracy ²	87.9% (1.4%)							

^aNo harvesting, 1973-2010.

^bEstimated by poststratification using known pixel counts.

Table 1.3. Error matrix and accuracy estimates for the 2010 cumulative harvest map. Standard error estimates are provided in parentheses.

Mapped forest class	Reference forest class			Total	User accuracy
	Regenerating	Partially harvested	Intact mature ¹		
Regenerating	157	13	4	174	90.2% (2.3%)
Partially harvested	18	69	5	92	75.0% (4.5%)
Intact mature ¹	6	5	232	243	95.5% (1.3%)
Total	181	87	241	509	
Producer accuracy ²	86.6% (2.0%)	81.3% (3.6%)	95.4% (1.5%)		
Overall accuracy ²	89.3% (1.3%)				

^aNo harvesting, 1973-2010.

^bEstimated by poststratification using known pixel count

Forest type classes for 1975 and 2004 were mapped with overall accuracies of 76% and 68%, respectively (Tables 1.4 and 1.5). Individual class accuracy estimates were similarly lower for 2004 than for 1975, presumably due to more heterogeneous landscape conditions. Off-diagonal entries in error matrices indicated confusion between the mixed class and both coniferous- and deciduous-dominant classes. There was little confusion between coniferous and deciduous classes. Error matrices and accuracy estimates were derived using class definitions for which omission and commission errors were best balanced and class accuracies acceptably high for both maps. Using coniferous-dominant and deciduous-dominant class thresholds of >80% and >70% basal area, respectively, errors were very well balanced for 2004 forest type classes (Table 1.5). User and producer accuracies for the 1975 map (Table 1.4) suggested under-representation of coniferous forest area and over-representation of mixed forest under these same class definitions, but the relatively small validation sample and correspondingly large standard error estimates made this inconclusive. Available validation data suggested that user and producer accuracies were best balanced under these class definitions.

Table 1.4. Error matrix and accuracy estimates for the 1975 forest type map. Standard error estimates are provided in parentheses.

1975 mapped forest type	Reference forest type			Total	User accuracy
	Coniferous ^a	Mixed	Deciduous ^b		
Coniferous ^a	15	3	0	18	83.3% (9.0%)
Mixed	6	22	4	32	68.8% (8.3%)
Deciduous ^b	1	3	16	20	80.0% (9.2%)
Total	22	28	20	70	
Producer accuracy ^c	74.0% (6.6%)	76.5% (6.4%)	78.9% (8.5%)		
Overall accuracy ^c	76.2% (5.0%)				

^a>80% coniferous basal area.

^b>70% deciduous basal area.

^cEstimated by poststratification using known pixel counts.

Table 1.5. Error matrix and accuracy estimates for the 2004 forest type map. Standard error estimates are provided in parentheses.

2004 mapped forest type	Reference forest type			Total	User accuracy
	Coniferous ^a	Mixed	Deciduous ^b		
Coniferous ^a	107	45	1	153	69.9% (3.7%)
Mixed	43	122	36	201	60.7% (3.5%)
Deciduous ^b	6	31	118	155	76.1% (3.4%)
Total	156	198	155	509	
Producer accuracy ^c	70.7% (2.9%)	59.9% (2.7%)	76.5% (2.9%)		
Overall accuracy ^c	68.3% (2.0%)				

^a>80% coniferous basal area.

^b>70% deciduous basal area.

^cEstimated by poststratification using known pixel counts.

1.4.2. EOF Analysis of Cumulative Harvest Time Series

By 2010, 61% of available forestland was mapped as harvested, and 40% regenerated by stand-replacing or multiple partial harvests. Averaged across all grid cells, harvest rates increased ca. 1985 and then remained quite constant at about 2% per year (median cumulative harvest time series; Fig. 1.2). The EOF analysis decomposed cumulative harvest time series into characteristic patterns of deviation from the median series. We retained 3 dominant EOF modes, which collectively explained 92% of total variance of the centered and standardized time series (62%, 23%, and 7% of total variance). This 3-mode EOF model provided an adequate representation of temporal trends for the great majority of individual time series (>90% of variance captured at 78% of grid cells; <70% of variance captured at 2% of cells).

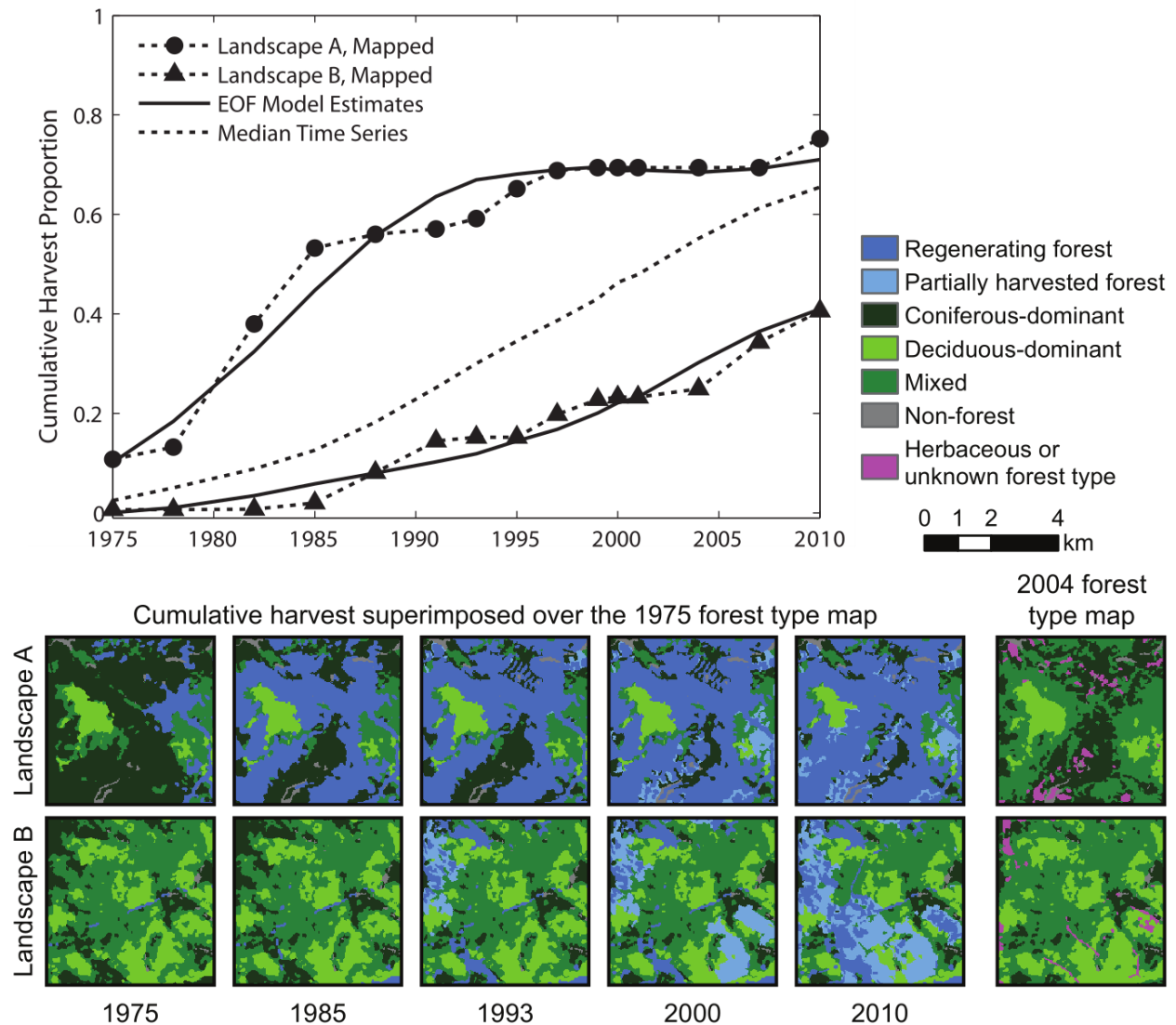


Figure 1.2. Forest harvest trends and landscape change for two sample grid cells. Mapped and modeled cumulative harvest time series for two arbitrary sample grid cells, expressed as a proportion of available forestland. The median cumulative harvest time series ($n = 608$) is shown for reference. Images of landscape conditions include cumulative harvest impact superimposed over the 1975 forest type map for a subset of time series dates, and the 2004 forest type map. Comparison of the 1975 and 2004 forest type data indicates areas where intervening harvests induced changes in forest type.

Time series of mapped and modeled cumulative harvest area at two sample locations (Fig. 1.2) illustrate the suitability of the EOF model for representing trends and smoothing the irregularities resulting from more erratic year-to-year changes in harvest rates. For the first of these sample landscapes (Landscape A; Fig. 1.2), harvest area increased rapidly through the first half of the study period (compared to the median time series), and then changed very little during the second half. The

extensive harvesting of the 1970s and 1980s was predominantly stand-replacing and directed at coniferous forest. By 1985, more than half of available forestland was regenerated. As indicated by the 2004 forest type map, much of that area was converted from coniferous-dominant to mixed. Within the second sample landscape (Landscape B; Fig. 1.2), harvesting consisted of both stand-replacing and partial canopy removals primarily during the second half of the study period within deciduous and mixed forest. Little harvesting occurred prior to 1985. Harvest rates were relatively modest between 1985 and 2004 and were somewhat elevated thereafter.

1.4.3. Harvesting Trends

From hierarchical clustering of modeled cumulative harvest time series, we identified six well-defined groups (Fig. 1.3a) ranging in size from 10% to 22% of grid cells. The mean time series from each group represented a predominant pattern of harvesting through time (Fig. 1.3b). For groups 1-2 (24% of grid cells), harvest rates exceeded median rates during the first half of the study period, particularly for group 1, and then dropped during the second half. The group 3 time series closely resembled the median time series. Groups 4-6 (56% of grid cells) shared the characteristics of relatively little harvesting early on followed by elevated rates during later years. Group 5 was notable in that harvest rates were exceptionally low through the 1980s but very high through the 1990s and 2000s.

Time series of cumulative regenerating and partially harvested forest area (Fig. 1.4) differ substantially between groups. Note that partial harvests were not mapped prior to 1988. Groups 1-2 were notable for rapid, heavy harvesting during the first half of the study period, followed by moderated rates of both stand-replacing and partial harvesting through the second half. Groups 4 and 5 were most strongly differentiated from other groups by high rates of partial harvesting, although high harvest rates during the 1990s and 2000s were sustained by both stand-replacing and partial harvests. Group 6 stand-replacing and partial harvest rates were low or moderate throughout the study period.

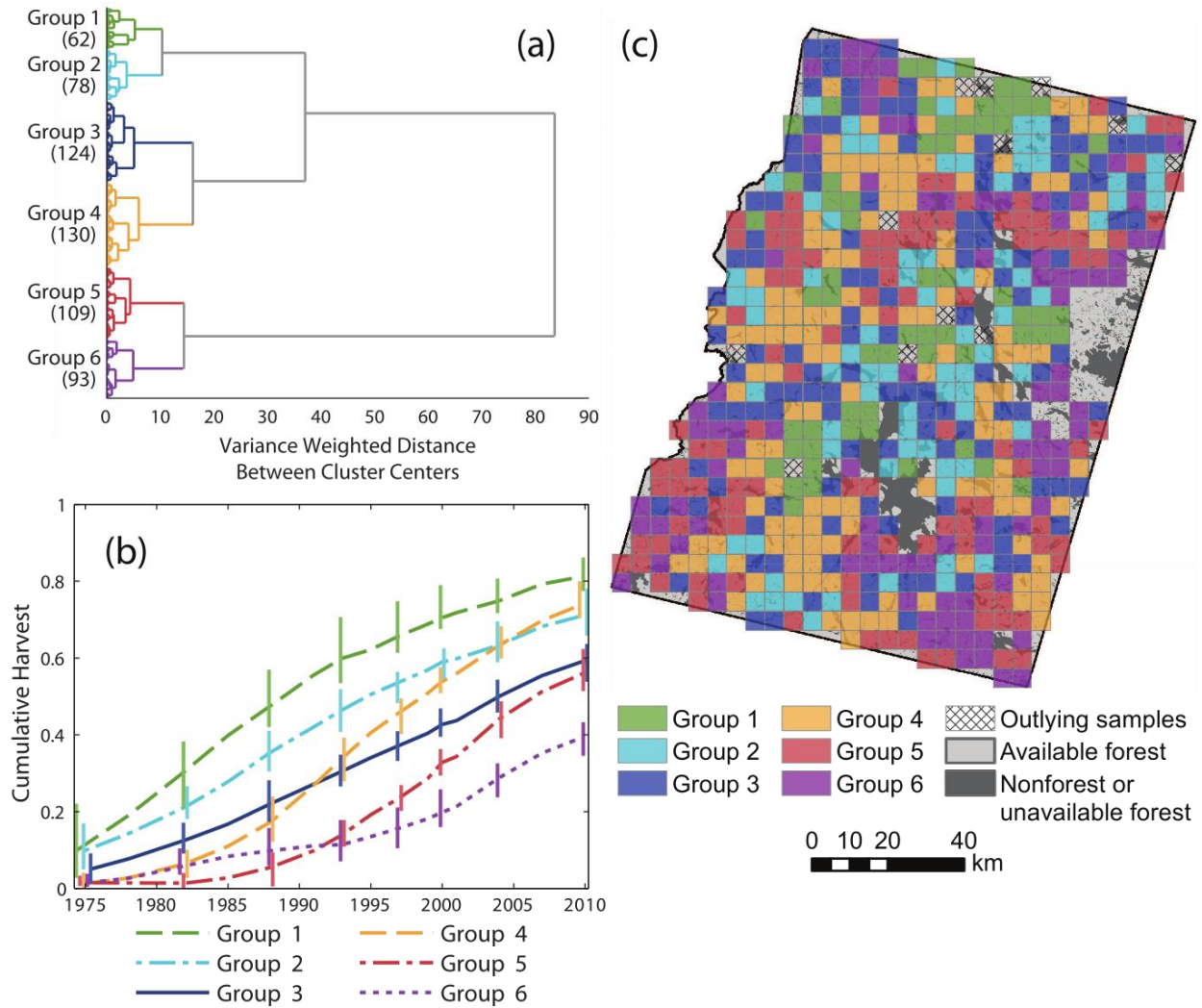


Figure 1.3. Predominant patterns of harvesting. (a) Dendrogram produced by agglomerative hierarchical clustering of modeled cumulative harvest time series. Six groups of landscape units were identified for subsequent analysis (sample sizes provided in parentheses). (b) Mean cumulative harvest area time series for each of the groups identified in (a), expressed as a proportion of available forestland. Vertical bars represent the interquartile range. Bars are provided at a subset of dates and are offset horizontally to improve visual clarity. (c) Spatial distribution of groups identified in (a). Hatching indicates outlying samples excluded from further analysis.

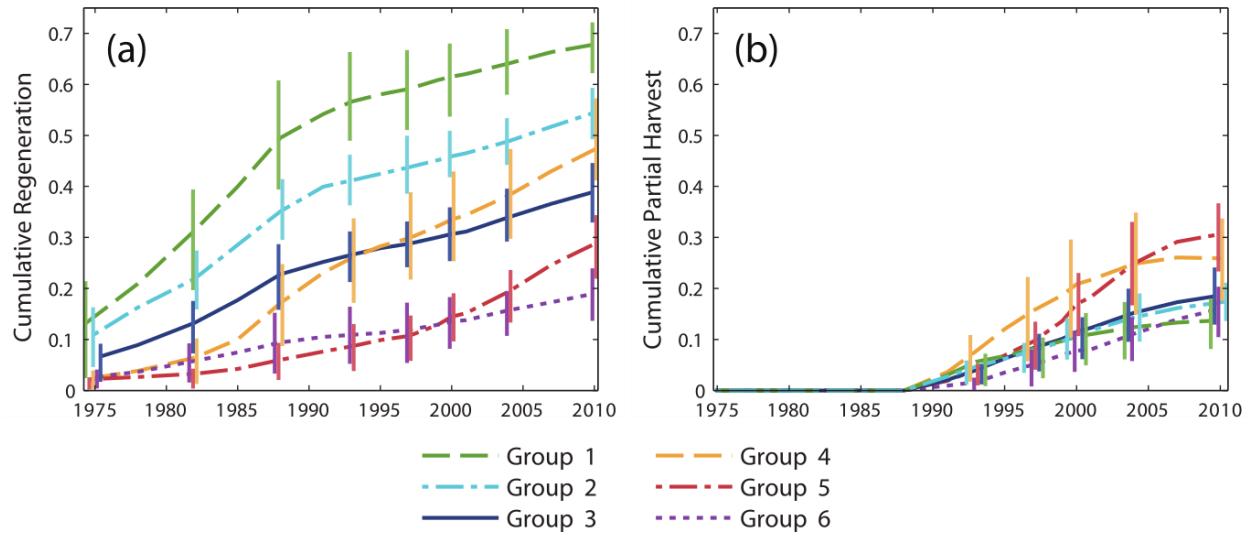


Figure 1.4. Time series of regenerating and partially harvested forest area. Cumulative time series of (a) regenerating forest area and (b) partially harvested forest area, expressed as a proportion of available forestland and averaged within groups identified by cluster analysis of modeled cumulative harvest time series. Vertical bars represent the interquartile range. Bars are provided at a subset of dates and are offset horizontally to improve visual clarity.

1.4.4. Patterns of Landscape Change

Time series of landscape metrics quantified the cumulative effects of harvesting on forest configuration. Time series of average patch size for intact mature forest (Fig. 1.5a) reflected trends in cumulative harvest area (Fig. 1.3b); periods of rapid patch size reduction coincided with periods of rapid harvesting. Patch density (Fig. 1.5b) increased through time, most rapidly in groups 4 and 5 during the 1990s and 2000s. For groups 1-5, the amount of edge between intact mature forest and harvested forest (Fig. 1.5c) increased and then peaked as harvest area approached and then surpassed 50% of available forestland. The increase in edge density was most rapid in groups 4 and 5 during the 1990s. Trajectories of average patch shape complexity (Fig. 1.5d) were similar in general character to those of edge density, but with peak values occurring somewhat earlier and with little change in groups 1-2 over the course of our study period.

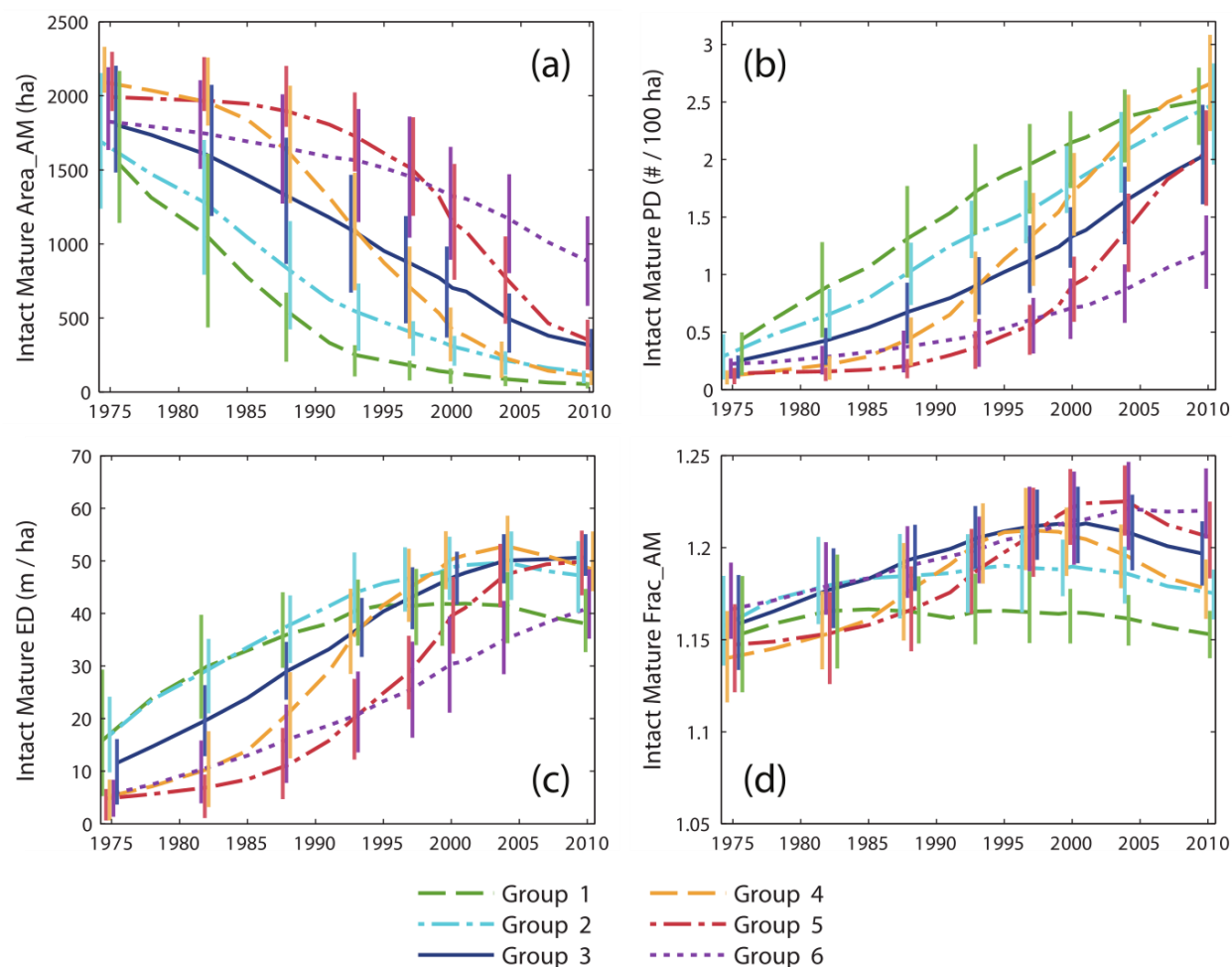


Figure 1.5. Time series of intact mature forest configuration metrics. Time series of cumulative change in (a) area-weighted mean patch size, (b) patch density, (c) edge density, and (d) area-weighted mean fractal dimension for intact mature forest, averaged within groups identified by cluster analysis of modeled cumulative harvest time series. Vertical bars represent the interquartile range. Bars are provided at a subset of dates and are offset horizontally to improve visual clarity.

Large differences between groups in time series of regenerating forest area (Fig. 1.4a) were only partly reflected in configuration metrics. Changes in average regenerating forest patch size (Fig. 1.6a) were greatest for group 1, increasing from less than 200 ha in 1975 to more than 800 ha by 1988. In contrast, average patch size remained low for groups 3-6. Group 5 values remained well below 200 ha despite relatively rapid increases in regenerating forest area during the 2000s (Fig. 1.4a). Regenerating patch density (Fig. 1.6b) generally increased throughout the study period, but this trend was less pronounced for group 2 and not apparent for group 1. The largest values of patch density were attained

by group 5 following rapid increase during the 1990s and 2000s. Despite low rates of stand-replacing disturbance in group 6 (Fig. 1.4a), patch density increased steadily and was quite high by 2010. The amount of regenerating forest edge (Fig. 1.6c) was greatest for groups 1-2 until the 2000s when group 4 attained comparable levels following rapid gains beginning in the late 1980s. Group 5 edge density increased rapidly during the 2000s. Average patch shape complexity of regenerating forest (Fig. 1.6d) increased during the first half of the study period, generally leveled somewhat during the 1990s, and then increased once more during the 2000s, most markedly for groups 4 and 5. Throughout the study period, regenerating patch shape complexity was greatest for groups 1-2.

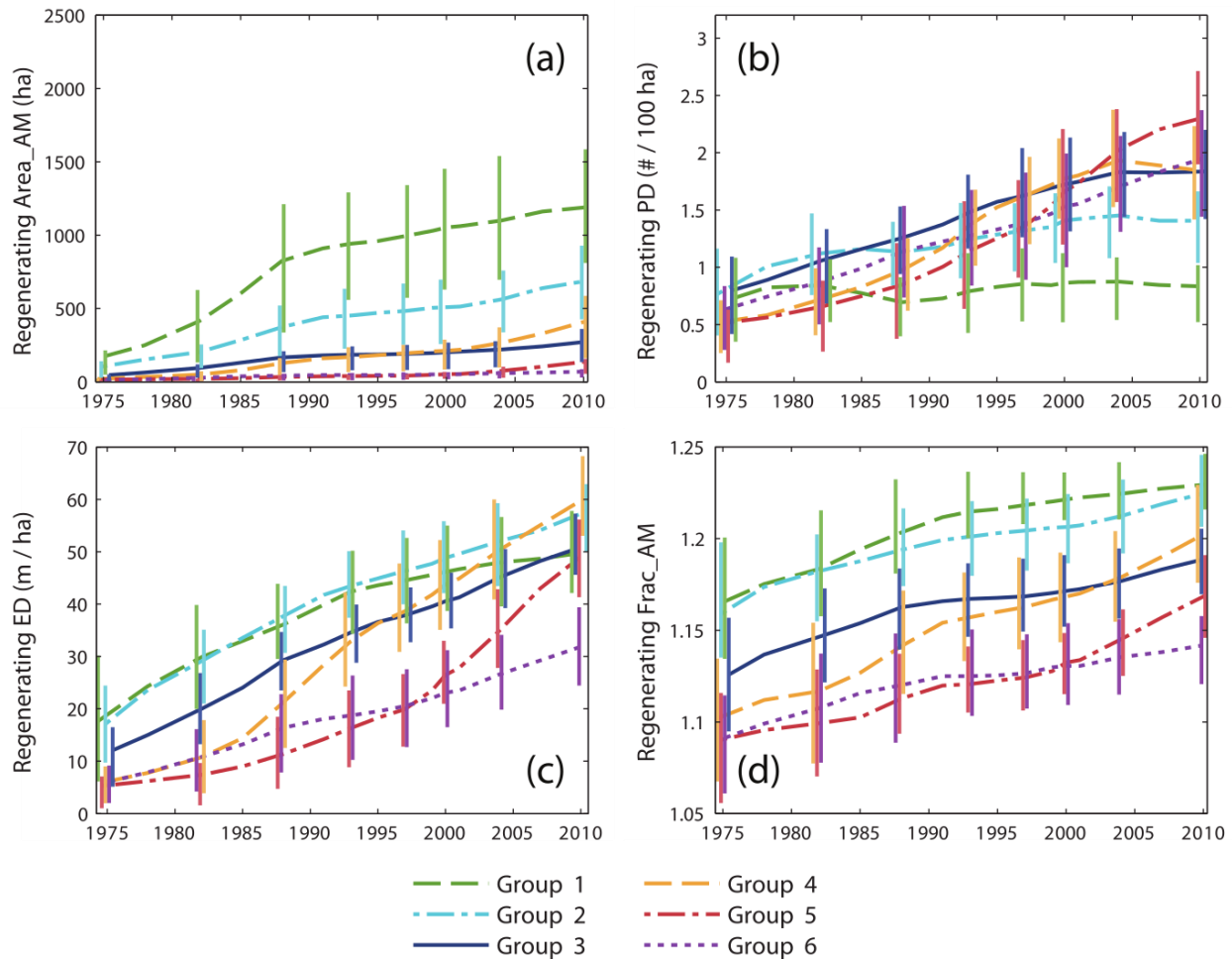


Figure 1.6. Time series of regenerating forest configuration metrics. Time series of cumulative change in (a) area-weighted mean patch size, (b) patch density, (c) edge density, and (d) area-weighted mean fractal dimension for regenerating forest, averaged within groups identified by cluster analysis of modeled cumulative harvest time series. Vertical bars represent the interquartile range. Bars are provided at a subset of dates and are offset horizontally to improve visual clarity.

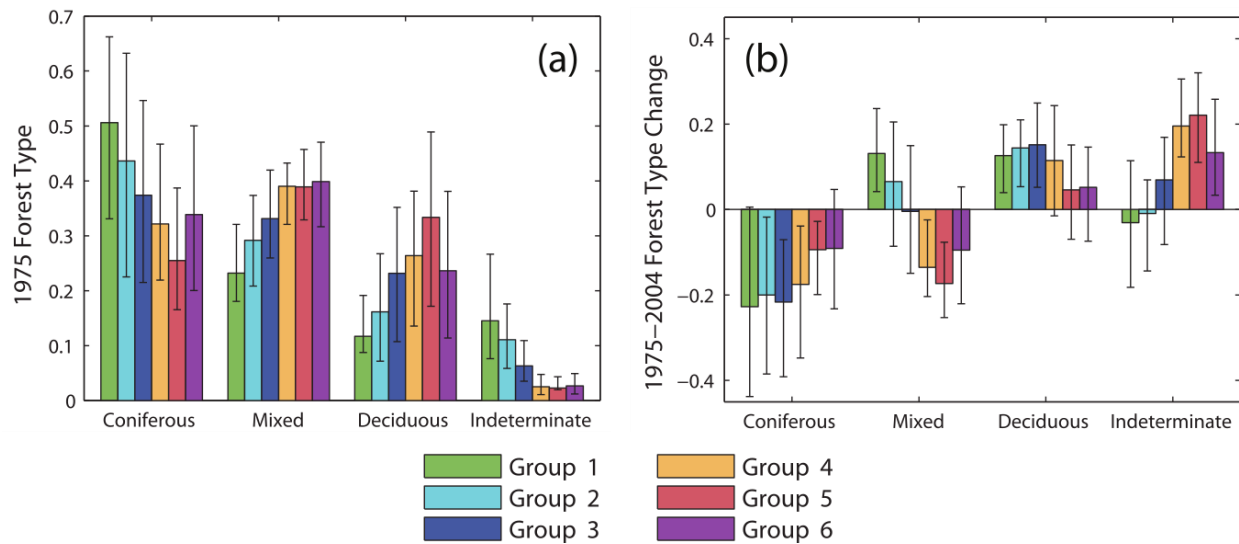


Figure 1.7. Initial landscape composition and changes in composition, 1975-2004. (a) Proportion of available forestland classified by 1975 forest type. (b) Change in forest type between 1975 and 2004, expressed as a proportion of forestland harvested prior to 2004 (negative values indicate loss; positive values indicate gain). Values were calculated for individual sample grid cells and then averaged within groups identified by cluster analysis of modeled cumulative harvest time series. Error bars represent the interquartile range.

The average initial composition of sample landscapes differed between groups, although there was a large amount of variability between landscapes in any single group (Fig. 1.7a). In 1975, groups 1-2 contained more coniferous-dominant forest and less deciduous-dominant and mixed forest than other groups. Forest type classes were least balanced for group 1, with coniferous forest comprising 51% and deciduous forest 12% of available forestland. Conversely, group 5 contained more deciduous (33%) and less coniferous forest (25%) than other groups. The composition of groups 4 and 6 were very similar. The amount of forest of indeterminate type in 1975 was greatest for groups 1-2, a result of harvesting during the early 1970s.

Between 1975 and 2004, harvesting and subsequent forest recovery resulted in substantial changes in landscape composition (Fig. 1.7b). On average all groups lost coniferous-dominant forestland. For groups 1-4, the coniferous forest lost amounted to about 20% of harvested forestland. For groups 1-3, much of this area transitioned to mixed or deciduous forest types, and the amount of forest classified as indeterminate remained little changed (recovery from early disturbance was

balanced by disturbance in later years). Groups 4 and 5 lost both coniferous and mixed forest. This was partially balanced by an increase in deciduous forest for group 4, but a substantial proportion of total harvest area (~20%) was mapped as indeterminate due to high harvest rates during the 1990s and early 2000s.

1.5. Discussion

During the spruce budworm outbreak of ca. 1972-1988, there were no legislative definitions or standards in place to regulate the practice of clearcutting. As the outbreak progressed, landowners engaged in extensive pre-salvage and salvage logging operations that typically took the form of large commercial clearcuts, much larger than would have been planned in the absence of the outbreak (Irland et al., 1988). The FPA was designed to regulate the execution of clearcuts larger than 14 ha (revised to 8 ha in 1999; 12 MRSA §8869), and its implementation in 1991 marked a fundamental and abrupt change in forest policy and management.

Averaged across all grid cells, cumulative harvest area increased more or less linearly (Fig. 1.2). From the cluster analysis, group 3 reflected this trend but contained only 20% of grid cells (Fig. 1.3). The cumulative harvest time series of the other five groups differed substantially from the area-wide average. These groups of grid cells comprised segments of the study area with different management histories. Groups 1-2 (24% of grid cells) were differentiated from other groups by elevated rates of stand-replacing harvests during the budworm outbreak (Fig. 1.4). Harvesting continued at moderated rates throughout the post-FPA period, then set against landscape conditions created by salvage logging. In contrast, harvesting within groups 4-6 (56% of grid cells) predominantly occurred during the post-FPA period. A large increase in group 5 harvest rates coincided with the end of the budworm outbreak and enactment of the FPA (Fig. 1.3b). Group 4 harvest rates increased during the late 1980s, but most harvest area accrued post-FPA with particularly high partial harvest rates during the 1990s. Similarly,

although group 6 harvest rates remained low to moderate post-FPA, most harvest area accrued during that time.

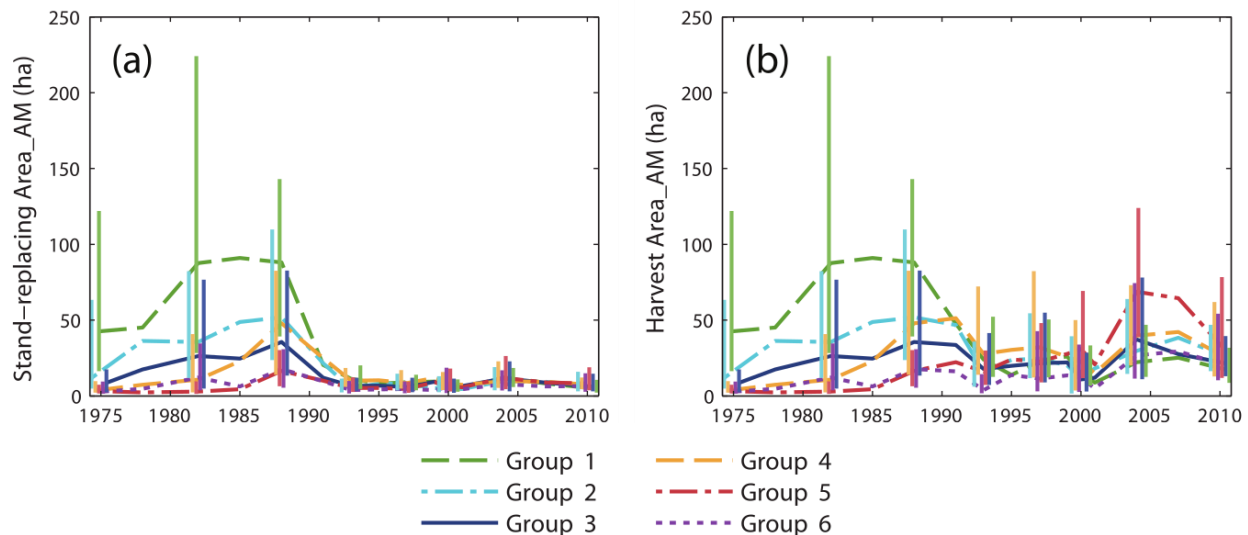


Figure 1.8. Changes in average harvest patch size through time. Area-weighted mean patch size for (a) the stand-replacing harvest class and (b) the combined stand-replacing and partial harvest class, calculated at each time series interval. Values were averaged within groups identified by cluster analysis of modeled cumulative harvest time series. Vertical bars represent the interquartile range. Bars are provided at a subset of dates and are offset horizontally to improve visual clarity.

Implementation of the FPA had the intended effect of reducing the size of clearcuts (Sader et al., 2003), and more generally the size of stand-replacing patches. Although our stand-replacing harvest class did not adhere to the FPA clearcut definition (12 MRSA §8868), average stand-replacing patch size calculated from individual harvest maps (rather than cumulative harvest maps) dropped dramatically between 1988 and 1991 for groups of grid cells affected by pre-FPA logging (Fig. 1.8a). Average patch size of stand-replacing harvests varied dramatically between cells prior to 1991, often exceeding 100 ha for groups 1 and 2. By comparison, post-FPA stand-replacing patch sizes were uniformly low for all grid cells, with group averages below 15 ha. In contrast, overall harvest patch sizes (stand-replacing and partial harvest classes combined) remained relatively high post-FPA, with group 5 averages approaching the pre-FPA values of groups 1-2 (Fig. 1.8b). The FPA placed a strong disincentive on clearcutting. State records indicate that clearcutting fell from 44% of annual harvest area in 1989 to <5% by 2000, and

annual harvest area roughly doubled during that time (Maine Forest Service, 2000, 1994). Within typical post-FPA partial harvest blocks, timber is removed within and adjacent to machine trails, leaving a matrix of unharvested or lightly harvested area between trails. Partial harvests are composed of typically small areas of complete or nearly complete canopy removal intermixed with areas of light or negligible canopy disturbance.

Predominant patterns of cumulative landscape change created by pre- and post-FPA management regimes were revealed by time series of intact mature and regenerating forest metrics. Groups 1-2 most clearly represented salvage logging impact. During the 1970s and 1980s, salvage caused a rapid decrease in the average patch size of intact mature forest (Fig. 1.5a) and a rapid increase in the average patch size of regenerating forest (particularly for group 1; Fig. 1.6a). Intact mature forest patch density increased during this period (Fig. 1.5b), but regenerating patch density changed relatively little (Fig. 1.6b). Edge density between intact mature forest and regenerating forest increased (Figs. 1.5c and 1.6c), but the average patch shape complexity of intact mature forest changed little, and the patch shape complexity of regenerating forest was comparable (Figs. 1.5d and 1.6d). These trends were consistent with the subdivision of mature forest by salvage clearcut of large contiguous tracts of spruce-fir (e.g., Landscape A, Fig. 1.2). Subsequent harvesting during the post-FPA period resulted in continued subdivision of intact mature forest at rates similar to the salvage period (Fig. 1.5b). Otherwise, changes in configuration metrics of both intact mature forest and regenerating forest were considerably reduced. The primary effect of the post-FPA regime in grid cells with a prominent salvage logging legacy appears to have been the production of more small patches of intact mature forest, with little influence on other metrics.

Groups 4-6 represented segments of our study area that were little affected by salvage logging but heavily impacted by post-FPA harvesting. Similar to salvage in groups 1-2, over time there was substantial loss and subdivision of intact mature forest. In groups 4 and 5, average patch size decreased

and patch density increased at rates that actually exceeded those of groups 1 and 2 pre-FPA (Figs. 1.5a and 1.5b). Edge density and patch shape complexity increased sharply during the 1990s as well (Figs. 1.5c and 1.5d). In group 6, cumulative harvest area was lower (Fig. 1.3b) and the loss and subdivision of intact mature forest correspondingly reduced (Figs. 1.5a and 1.5b), yet edge density and patch shape complexity increased to levels approaching or exceeding all other groups (Figs. 1.5c and 1.5d). Post-FPA partial harvest practices resulted in complex patches of intact mature forest and high edge densities presumably due to residual inclusions of mature forest within harvest blocks. In sharp contrast to salvage logging in groups 1-2, average regenerating forest patch sizes in groups 5 and 6 remained very low (Fig. 1.6a). Group 5 regenerating patch density increased rapidly in the 1990s and 2000s, surpassing all other groups by 2004 (Fig. 1.6b). Group 5 edge density increased rapidly during the 2000s, ultimately exceeding the values of groups 1-2 at the end of the salvage period (Fig. 1.6c) despite considerably less regenerating forest area (Fig. 1.4a). Group 6 regenerating forest patch size, edge density, and shape complexity remained relatively low, but patch density steadily increased throughout the study period (Fig. 1.6). Patterns within these groups indicate that post-FPA stand-replacing harvest patches were more numerous, much smaller, and simpler in shape compared to the pre-FPA salvage logging period (e.g., Landscape B, Fig. 1.2).

Not surprisingly, groups that were most heavily impacted by budworm salvage logging were also those with the greatest amount of coniferous-dominant forestland in 1975 (groups 1-2, Fig. 1.7a). Groups 4-6 contained less coniferous forestland and we attribute the contrast in management history and landscape change between groups 1-2 and groups 4-6 in large part to differences in initial landscape composition and vulnerability to the budworm outbreak. Initial composition set different segments of the study area on fundamentally different trajectories of landscape change. However, groups 4 and 6 differed very little in the relative abundance of 1975 forest types (Fig. 1.7a), and comparisons between them suggest the influence of some other factor that affected post-FPA harvest patterns. Private

ownership diversified greatly during the 1990s as industrial forest products companies restructured and sold their lands to investment entities, nonprofit conservation groups, high net-worth individuals, and other owner types (Hagan et al., 2005; Jin and Sader, 2006). We hypothesize that post-FPA differences in harvest rates, intensities, and trajectories of landscape change may have been influenced by differences in management incentives between different landowners (e.g., fiber production vs. resource conservation). Previous research documented differences in harvest rates between categories of ownership and ownership change (2006), but the influence of owner-to-owner variability on patterns of landscape change remains unclear.

The relative importance of individual landowner behavior, public forest policy, and management or disturbance legacies on contemporary trajectories of landscape change is an important question with implications extending to regional forest planning, management, and conservation. Because multiple forest values are often maintained only when actions are integrated over large areas with diverse forest conditions, it is important to understand the relative influence of factors that act to either enhance or reduce landscape-scale heterogeneity. Within northern Maine, salvage logging introduced persistent heterogeneity at the scale of 5 km landscape units due to large clearcut operations. However, another important aspect of salvage legacy is loss of coniferous-dominant forest area (Fig. 1.7b) and consequent homogenization of forest composition due to clearcut operations that failed to adequately regenerate spruce and fir (Seymour, 1992). Management under the FPA further homogenized landscape structure by effectively eliminating large clearcuts and incentivizing the expansion of partial harvesting. Under the post-FPA management regime, differences between landowners in management incentives, objectives, or strategies may provide an important source of landscape heterogeneity. Given the small amount of publicly owned forestland within the state of Maine (approximately 7% (McCaskill et al., 2011)), the sustainable management of Maine's forest resources will require a clearer understanding of landscape

dynamics and management outcomes under various forms of private ownership, as well as closer consideration of the ways in which public policy may constrain outcomes.

The changes in landscape composition and configuration we have quantified imply potentially important impacts on forest ecology and wildlife. Salvage clearcuts created large blocks of early successional forest habitat, benefitting the federally threatened Canada lynx (*Lynx canadensis*) (Simons, 2009). In Maine, the primary prey of lynx, snowshoe hare (*Lepus americanus*), are found at highest density within coniferous and mixed regenerating forest ~15-35 years post-harvest (Fuller and Harrison, 2005; Homyack et al., 2007; Robinson, 2006). The current amount and configuration of this high-quality lynx foraging habitat is largely a product of pre-FPA salvage logging. Post-FPA harvest practices produce smaller and more numerous regenerating forest patches, promoting the subdivision of high-quality lynx foraging habitat (Simons, 2009). Additionally, the large annual footprint of post-FPA partial harvesting and accelerated loss and subdivision of intact mature forest suggest rapid loss of suitable habitat or reduction of habitat quality for species that require features associated with mid- and late-successional forest, such as the American marten (*Martes americana*) (Fuller and Harrison, 2005). For species that are either dependent upon landscape legacies or potentially impacted by rapid habitat modification, responses to contemporary management may be difficult to establish without knowledge of landscape history and disturbance trends. Our analysis demonstrated one approach by which landscape disturbance history can be defined and evaluated using a time series of Landsat-derived forest disturbance maps.

CHAPTER 2

MULTI-OBJECTIVE SUPPORT VECTOR REGRESSION REDUCES SYSTEMATIC ERROR IN MODERATE RESOLUTION MAPS OF TREE SPECIES ABUNDANCE

2.1. Abstract

When forest conditions are mapped from empirical models, uncertainty in remotely sensed predictor variables can cause the systematic overestimation of low values, underestimation of high values, and suppression of variability. This regression dilution or attenuation bias is a well-recognized problem in remote sensing applications, with few practical solutions. Attenuation is of particular concern for applications that are responsive to prediction patterns at the high end of observed data ranges, where systematic error is typically greatest. We addressed attenuation bias in models of tree species relative abundance (percent of total aboveground live biomass) based on multitemporal Landsat and topoclimatic predictor data. We developed a multi-objective support vector regression (MOSVR) algorithm that simultaneously minimizes total prediction error and systematic error caused by attenuation bias. Applied to 13 tree species in the Atlantic Northern Forest of the northeastern U.S., MOSVR performed well compared to other prediction methods including single-objective SVR minimizing total error (SOSVR), Random Forest (RF), gradient nearest neighbor (GNN), and Random Forest nearest neighbor (RFNN) algorithms. SOSVR and RF yielded the lowest total prediction error but produced the greatest systematic error, consistent with strong attenuation bias. Underestimation at high relative abundance caused strong deviations between predicted patterns of species dominance/codominance and those observed at field plots. In contrast, GNN and RFNN produced dominance/codominance patterns that deviated little from observed patterns, but predicted species relative abundance with lower accuracy and substantial systematic error. MOSVR produced the least systematic error for all species with total error often comparable to SOSVR or RF. Predicted patterns of dominance/codominance matched observations well, though not quite as well as GNN or RFNN. MOSVR

provides an effective nonparametric approach to the reduction of systematic prediction error, and should be fully generalizable to other remote sensing applications and prediction problems.

2.2. Introduction

As forest ecosystems are pushed beyond historic conditions by anthropogenic disturbance and environmental change, there is increasing need to forecast future conditions as a basis for policymaking and management planning. Ecological forecasting requires quantitative understanding of ecological processes, and how existing conditions are likely to affect processes moving forward. Detailed measurements and observations are needed to develop this understanding. Field data and specifically forest inventory measurements provide great detail at high accuracy, but are collected from a sample of small plots. Forecasting at broader scales relevant to forest policy and management requires extrapolation of plot measurements across large contiguous areas.

Forest conditions are typically predicted across landscapes and regions from empirical relationships between field measurements and remote sensing data. Uncertainty in remotely sensed predictor variables can cause severely detrimental patterns of prediction error. At moderate spatial resolutions, a prominent and impactful source of uncertainty in predictor variables may be physical differences in measurements between field plots and image pixels (Xu et al., 2009). Whereas the ideal remotely sensed predictor data would represent the same ground area as reference plot data, scale and location mismatches introduce uncertainty in predictor values. Forest inventory measurements are typically obtained over plots that are a fraction of the size of moderate resolution image pixels. The USDA Forest Service Forest Inventory and Analysis (FIA) program, for example, provides measurements from a national network of field plots composed of a cluster of four subplots, each <0.02 ha (McRoberts et al., 2010). Subplots span an area roughly equivalent to a 3x3 neighborhood of 30 m Landsat pixels. Subplot area however, equates to only 8% of that pixel neighborhood. In the presence of sub-pixel variation of forest conditions, average conditions across a pixel neighborhood likely will not correspond

to those measured at FIA subplots (Rejou-Mechain et al., 2014; Xu et al., 2009). Image georeferencing or registration error coupled with GPS error in plot coordinates further interferes with the physical correspondence of pixels and plots (Xu et al. 2009). Potentially compounding these problems are differences in timing between image acquisitions and plot measurements and additional sources of predictor uncertainty associated with remote sensing platforms, instrumentation, viewing conditions, and data handling.

Without correcting for uncertainty in predictor variables, regression algorithms generally assign variation in the predictors to variation in the response given the predictors, causing what is known as regression dilution or attenuation bias (Bartlett et al., 2009; Frost and Thompson, 2000). The strength of the relationship between predictors and response is underestimated, resulting in a characteristic pattern of prediction error where low values tend to be overestimated, high values tend to be underestimated, and the variability of predictions relative to reference data is suppressed. Although attenuation bias is a long-recognized problem in remote sensing (Curran and Hay, 1986), recent studies have emphasized a general lack of appreciation of its impacts, and a lack of suitable options for their correction. Rejou-Mechain et al. (2014) emphasized the importance of attenuation bias in the estimation of aboveground biomass by regression against field plot measurements. They asserted that attenuation bias was largely ignored in remote sensing applications, and demonstrated that established statistical approaches to reducing bias may be inadequate. Xu et al. (2009) asserted that attenuation bias is a pervasive problem in remote sensing of forest attributes, and suggested that no analytical method was capable of eliminating this bias. After analyzing causes using simple error models, they suggested that field data be collected over a spatial support similar to the size of pixels or pixel neighborhoods used for prediction. Yet they also demonstrated that location mismatches can cause severe attenuation regardless. Robinson et al. (2013) recognized strong attenuation bias when estimating aboveground biomass from FIA plot data and airborne L-band radar, and suggested that FIA

plots may not provide suitable reference observations. However, FIA or similar inventory data are commonly used for model training, often resulting in patterns of error consistent with attenuation bias (e.g., Blackard et al., 2008; Frescino et al., 2001; Ohmann et al., 2014; Ohmann and Gregory, 2002; Powell et al., 2010; Riemann et al., 2010). Attenuation is a potential problem for applications that are specifically dependent upon or influenced by patterns of prediction at high or low values, and outcomes may be affected in ways that are difficult to identify or correct.

Our interest in attenuation bias stemmed from a specific need to map tree species distributions for initialization of the LANDIS-II forest landscape model (FLM) (Gustafson et al., 2000; Mladenoff, 2004; Scheller et al., 2007). FLMs are complex process models that simulate the spatiotemporal dynamics of forest ecosystems (Xi et al., 2009). FLMs operate over a raster representation of forest conditions, typically at moderate resolutions (30-250 m). They are therefore well suited to forecasting ecosystem dynamics based on conditions mapped by remote sensing, and are increasingly useful in a predictive capacity for strategic decision support (Gustafson, 2013). Within LANDIS-II, grid cells are populated by cohorts of trees defined by species and age. Processes including establishment, growth, competition, and mortality are tracked for each cohort in each cell, and cells are linked by the spatial processes of seed dispersal and disturbance (Scheller et al., 2007). To simulate dynamics stemming from existing forest conditions, LANDIS-II applications require detailed map data for model initialization. However, application across large areas (millions of hectares) generally requires simplification of cohort structure to reduce computational requirements. We adopted a LANDIS-II initialization strategy that defined cohort structure based on the relative abundance of the three most abundant species in a given cell. This strategy required accurate predictions of species relative abundance and dominance/codominance from moderate resolution imagery and geospatial data. FIA measurements were our sole source of reference data.

A potentially favorable approach to producing the species distribution data needed for LANDIS-II initialization is nearest neighbor imputation, where response data is assigned to a pixel by averaging reference observations from a set of k similar pixel locations (Dijak, 2013; Duveneck et al., 2015; Scheller et al., 2008). Similarity is determined by spectral or environmental covariates. Nearest neighbor methods are advantageous in that they can be used to populate pixels with an entire suite of observations obtained at reference plot locations. In case of $k = 1$, observations from individual plots are imputed to pixels, retaining the plot-level covariance structure between variables and ensuring ecological realism. However, nearest neighbor methods cannot extrapolate beyond the range of observed data, and for $k = 1$ imputed values (and combinations of values) are strictly limited to those observed. Regression methods that model species distributions individually may achieve greater accuracy for individual species, particularly if reference data are limited. On the other hand, combining multiple species models subject to strong attenuation bias may compound error in predictions of codominance.

In the context of species distribution modeling, little attention appears to have been paid to the specific problem of estimating species codominance, though impacts of prediction bias have been considered for co-occurrence. In this case, stacked predictions from a set of individual species models tend to overestimate species co-occurrence or richness, and this has been linked to attenuation bias in models of species probability of occurrence (Calabrese et al., 2014). More generally, a number of approaches have been advanced to reduce attenuation bias in parametric species distribution models, using error-in-variables (Foster et al., 2012) or Bayesian methods (Denham et al., 2011; McNerny and Purves, 2011). Here we present an approach based on nonparametric machine learning models, specifically support vector machines (SVMs). SVMs induce relationships between remote sensing data and field measurements without pre-specification of a form for the modeled relationship. Predictions based on these relationships are often more accurate than those based on parametric statistical models

because relationships between variables are often too complex or too little understood to pre-specify an appropriate parametric model form. SVMs were originally developed for binary classification (Vapnik, 1995) but have been widely applied to regression problems (Mountrakis et al., 2011; Salcedo-Sanz et al., 2014).

Use of SVMs requires the specification of several free parameters that determine model fit, and optionally the selection of a subset of predictor variables. SVMs are sensitive to parameter settings, which adds complexity to the model selection process because adequate performance cannot be assured under any pre-specified or default values (Brereton and Lloyd, 2010). Our approach exploits the complexity of SVM model selection to obtain a set of solutions with different levels of prediction error and attenuation bias. Because attenuation arises from the minimization of error in the presence of predictor uncertainty, regression model training presents a trade-off between total prediction error and the systematic component of error caused by attenuation bias. Total and systematic error can be simultaneously minimized as partially conflicting objectives using multi-objective optimization methods (Jin, 2006; Konak et al., 2006). Multi-objective model training produces a set of solutions with different trade-offs between objectives - for example, an increase in systematic error associated with a reduction of total error. Our goal was to obtain solutions with reduced systematic error at acceptable levels of total error.

We present a support vector regression method based on model training by a multi-objective genetic algorithm (GA), and evaluate its use in predicting tree species relative abundance and codominance from multitemporal Landsat imagery and topoclimatic surfaces. We include a detailed algorithm description, and compare predictions of species relative abundance and codominance with those obtained from other modeling approaches including Random Forests and nearest neighbor methods. Finally, we discuss the merits of multi-objective model training for this specific application and for the reduction of attenuation bias more generally.

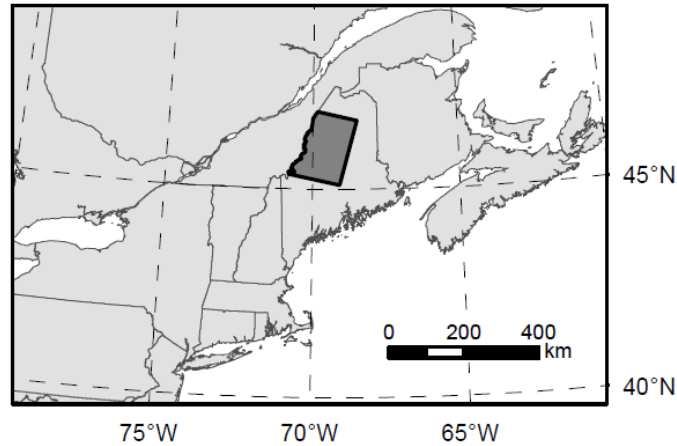


Figure 2.1. Study area. Northern Maine, U.S.A. study area encompassing 1.9 Mha of forestland. State and provincial boundaries obtained from the National Atlas of the U.S. (Political Boundaries) and the Atlas of Canada (National Frameworks Data, Census Subdivisions and Population Ecumene).

2.3. Methods

2.3.1. Study Area

The Atlantic Northern Forest of the northeastern U.S. occupies a transition zone between the northern boreal forest and the southern temperate deciduous-dominant forest (Likens and Franklin, 2009), and includes approximately 4 Mha of nearly contiguous, undeveloped forestland across northern and western Maine. Our 1.9 Mha study region (Fig. 2.1) was defined by the overlap of Landsat Worldwide Reference System path 12, row 28 and the political boundary between northwestern Maine and Quebec, Canada. Topography is generally flat or rolling with occasional low mountains and an extensive network of rivers, lakes, and wetlands. Tree species diversity is relatively high as the northern limit of southern species overlaps with the southern limit of northern species (Nightingale et al., 2008). Forest type distributions are associated with climatic gradients, topo-edaphic conditions, and disturbance history (Seymour, 1995). Northern hardwood species (*A. saccharum*, *Betula alleghaniensis*, *Fagus grandifolia*) predominate across lower hilltops and at mid-slope. Spruce-fir species (*Abies balsamea*, *Picea rubens*, *P. mariana*) predominate where soil or microclimatic conditions exclude the more demanding hardwoods. Mixedwood stands commonly occur along ecotones or as a result of

succession following disturbance. Shade-intolerant hardwood species (*Betula papyrifera*, *Populus spp.*) are commonly found following intense disturbance.

2.3.2. Reference and Predictor Data

Predictive models of species relative abundance are based on reference data provided by the USDA Forest Service FIA Program. The contemporary network of field plots adheres to an equal probability sampling design, with plots randomly located within 2428 ha hexagonal tiles (McRoberts et al., 2005). The FIA program maintains the confidentiality of true plot locations to protect the privacy of landowners and to preserve plot integrity (Smith, 2002). True locations were made available for use through a collaborative agreement with the USFS Northern Research Station FIA Program. Tree measurement data were used to calculate species relative abundance as a proportion of estimated live aboveground biomass (stems >2.54 cm diameter, measured at 1.37 m; DRYBIOT variable, FIADB v3.0). Since 1999, 20% of plots within Maine have been surveyed annually during 5-year inventory cycles (McRoberts et al., 2005).

Our primary source of spatial predictor data was Landsat TM and ETM+ imagery. We selected images from the early 2000s to maximize the availability of cloud-free imagery prior to the failure of the ETM+ scan line corrector in May 2003 and to support a LANDIS-II simulation start date of 2010. We targeted images acquired at different times throughout the growing season (late April through early October) to exploit species-specific phenological patterns. Frequent and extensive cloud cover dictated the use of images acquired over multiple years, and we ultimately selected eight relatively cloud- and snow-free images spanning a roughly 5-year observation period (2001-2006) matching a FIA inventory cycle (Table 2.1). Images were obtained from the U.S. Geological Survey Earth Resources Observation and Science Center and the Multi-Resolution Land Characteristics Consortium at 30 m resolution with standard terrain correction applied. Clouds and cloud shadows were masked using a semi-automated procedure developed in-house. Errors were corrected by on-screen digitization. Bands 1-5 and 7 (visible

and reflective infrared) were extracted for further processing as spatial covariates. Visible snow cover in early-season imagery was masked by unsupervised classification using an ISODATA algorithm and visual interpretation of snow-covered classes. Images were converted to top-of-atmosphere reflectance and then corrected for topographic illumination effects using the SCS+C algorithm (Soenen et al., 2005), with slope and aspect calculated from the 1 arc-second (30 m) National Elevation Dataset (NED).

Table 2.1. Landsat images used for predictive modeling of tree species relative abundance. Images were acquired over Landsat Worldwide Reference System-2 path 12, row 28. Images were obtained from the U.S. Geological Survey Earth Resources Observation and Science Center unless indicated otherwise.

Acquisition date	Landsat sensor	Landsat satellite	% forest under cloud/shadow	% forest under snow
Species abundance:				
April 29, 2006	TM	5	-	1.0
May 12, 2005	TM	5	1.7	0.8
May 25, 2001	ETM+	7	1.2	<0.1
June 10, 2004	TM	5	0.4	-
July 20, 2001 ¹	TM	5	0.9	-
Sept. 14, 2004	TM	5	0.3	-
Sept. 30, 2001	ETM+	7	-	-
Oct. 6, 2006	TM	5	3.0	-
Canopy change: ²				
July 20, 2001 ¹	TM	5	0.9	-
June 10, 2004	TM	5	0.4	-
June 19, 2007	TM	5	9.0	-

¹Available from the Multi-Resolution Land Characteristics consortium.

²Images used to mask spectral change resulting from disturbance and regrowth over the observation period used for prediction of species relative abundance. Cloud-contaminated data in the June 19, 2007 image were replaced with data from a Landsat 5 image acquired on Aug. 22, 2007.

Forest canopy change during the 5-year observation period dissociated image characteristics from field measurements at affected plot locations. We therefore masked locations of apparent canopy cover change using available leaf-on images acquired in 2001, 2004, and 2007 (Table 2.1). The iteratively-reweighted multivariate alteration detection transformation (Canty and Nielsen, 2008) was applied to 2001-2004 and 2004-2007 image pairs to estimate a probability of spectral change during each interval. Intervals were combined by selecting the maximum probability of change, and a threshold

was selected to identify 20% of forest pixels as change pixels. Threshold selection was arbitrary, but visual inspection of the resulting 2001-2007 change mask indicated close correspondence with canopy disturbance and visible regrowth in previously disturbed stands.

Additional spatial covariates included climate and terrain attributes thought to be relevant to tree establishment or growth. Terrain data included 10 morphometry, 8 lighting/visibility, and 11 hydrology variables (Table 2.2) calculated from the 1 arc-second (30 m) NED and the National Hydrography Dataset (NHD) using the System for Automated Geoscientific Analyses software, version 2.1.4 (Conrad et al., 2015). The NED was lightly smoothed with a Gaussian filter to reduce the effects of random error and systematic artifacts (circular filter element, radius = 90 m, $\sigma = 1.5$). Terrain slope, aspect, and curvature were calculated from a second-order polynomial fit (Zevenbergen and Thorne, 1987). Direct insolation was calculated at mid-month, April-September, by assuming a uniform 65% atmospheric transmittance, a value that produced insolation estimates in good agreement with a previously published regional climate model (Ollinger et al., 1995). Hydrology variables including catchment area, flow path length, and distance to stream channel were calculated using a bidimensional flow routing algorithm (Quinn et al., 1991) after filling sinks in the NED (Wang and Liu, 2006). Synthetic stream channel networks were derived from the catchment area raster after masking and dilating NHD water bodies using a 5x5 filter element. The dilated water body mask reduced the tendency for channels to initiate near the edges of water bodies, where the flow routing algorithm produced large estimates of flow accumulation. Climate data were obtained from the USDA Forest Service Rocky Mountain Research Station, Moscow Forestry Sciences Laboratory, and included 17 variables (Table 2.2) derived from monthly temperature and precipitation surfaces interpolated from weather station data for the climate normal period of 1961-1990 (Rehfeldt, 2006). Climate data were available at approximately 1 km spatial resolution.

Table 2.2. Terrain and climate variables used to model and map tree species relative abundance.

Terrain variables were calculated using the System for Automated Geoscientific Analyses (SAGA) software (Conrad et al., 2015) with default settings unless otherwise specified. Climate variables were obtained directly from the USDA Forest Service Rocky Mountain Research Station, Moscow Forestry Sciences Laboratory.

Terrain morphometry (10 variables)		
Elevation		
Slope	Local terrain slope, from fit of second-order polynomial	
Aspect	Local terrain aspect, from fit of second-order polynomial	$\cos(\text{aspect} - 45^\circ) + 1$ (Beers et al., 1966)
Curvature	Local terrain curvature, from fit of second-order polynomial	Tangential, profile, and plan curvature
Topographic position index	Difference between elevation and mean elevation of circular neighborhood (Guisan et al., 1999)	150 m, 300 m, 1000 m, 2000 m neighborhood radii
Lighting/visibility (8 variables)		
Visible sky	Proportion of hemisphere unobstructed by terrain (Häntzschel et al., 2005)	10,000 m search radius
Sky view factor	Ratio of diffuse irradiance to that of an unobstructed horizontal surface (Häntzschel et al., 2005)	10,000 m search radius
Direct insolation	Potential incoming solar radiation	single day estimate at mid-month, April-September; 65% atmospheric transmittance
Hydrology (11 variables)		
Catchment area	Upslope area or flow accumulation	log10 transformed
Catchment height	Difference between elevation and mean elevation of upslope pixels	
Catchment slope	Mean slope of upslope pixels	
Catchment aspect	Mean aspect of upslope pixels	$\cos(\text{aspect} - 45^\circ) + 1$ (Beers et al., 1966)
Flow path length	Mean distance of flow from upslope pixels	
Distance to stream channel	Shortest distance (or distance component) to synthetic stream channel network calculated by flow routing algorithm	overland distance and horizontal, vertical distance components; stream networks from 10 ha and 50 ha flow initiation thresholds

Table 2.2 continued

Climate (17 variables)	
d100	Julian date on which the sum of degree-days >5°C reaches 100
dd0	Degree-days <0°C (from mean monthly temperatures)
dd5	Degree-days >5°C (from mean monthly temperatures)
fday	Julian date of the first autumn freeze
ffp	Length of the frost-free period (days)
gsdd5	Degree-days >5°C accumulated over the frost-free period
gsp	Growing season precipitation (April-September)
map	Mean annual precipitation
mat_tenths	Mean annual temperature
mmax_tenths	Mean maximum temperature of warmest month
mmindd0	Degree-days <0°C (from mean minimum monthly temperatures)
mmin_tenths	Mean minimum temperature of coldest month
mtcm_tenths	Mean temperature of coldest month
mtwm_tenths	Mean temperature of warmest month
sday	Julian date of last spring freeze
smrpb	Summer precipitation balance (July+Aug.+Sept. / April+May+June)
smrsprpb	Summer/spring precipitation balance (July+Aug. / April+May)

Covariate values were extracted at forested FIA plots. Landsat and terrain data were compiled by averaging values from forest pixels within 3x3 neighborhoods surrounding plot centers; climate predictor data were extracted as 1 km pixel values. Forest pixels were differentiated from non-forest using the 1993 Maine Gap Analysis Program (GAP) land cover map, augmented with the agricultural classes of the 2001 National Land Cover Database (NLCD). The 1993 GAP map differentiated forest from non-forest with an estimated 100% accuracy in our study area (Hepinstall et al., 1999). Incorporation of the 2001 NLCD agricultural classes accounted for a small amount of land cover change predating our 2001-2006 observation period. We excluded reference locations with missing data due to forest cover change, cloud/shadow cover, or snow cover. SVMs are generally incapable of working with incomplete predictor data and for the purposes of algorithm development and evaluation, we elected to exclude samples with missing data rather than incorporate an additional algorithm for imputing missing data. Remaining plot locations yielded a training/validation data set consisting of 349 samples.

2.3.3. Background

2.3.3.1. Support Vector Regression (SVR)

Following Vapnik (1998, 1995), development of the SVR algorithm is based on estimation of a linear regression function

$$(1) \quad f(x) = \langle w, x \rangle + b$$

where $w, x \in \mathbb{R}^d$. The regression estimate is obtained by minimizing

$$(2) \quad \frac{1}{2} \|w\|^2 + \frac{1}{n} \sum_{i=1}^n L_\epsilon(f(x_i), y_i)$$

where the Euclidean norm of the weight vector w gives the flatness of the regression function (a geometric representation of model complexity) and L_ϵ is the so-called ϵ -insensitive loss function:

$$(3) \quad L_\epsilon(f(x_i), y_i) = \begin{cases} 0, & \text{if } |y_i - f(x_i)| \leq \epsilon \\ |y_i - f(x_i)| - \epsilon, & \text{otherwise} \end{cases}$$

L_ϵ defines a margin of width ϵ bounding the regression function, with nonzero loss applied only to training samples lying outside that margin. Minimization of Eq. (2) is equivalent to the minimization of

$$(4) \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*)$$
$$\text{subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \epsilon + \xi_i^* \\ \epsilon, \xi_i, \xi_i^* \geq 0 \end{cases}$$

where ξ_i and ξ_i^* are nonzero slack variables that quantify the deviation of the observed values above or below the margin and the constant C specifies a trade-off between the minimization of flatness (model complexity) and minimization of prediction error for a given value of ϵ . The training samples lying outside the margin are referred to as support vectors (SVs). They alone determine the regression function estimate $\hat{f}(x)$. A smaller margin width ϵ generally corresponds to a larger number of SVs and a more complex solution that may fit the training data well but may not generalize to new data. A unique solution to the constrained minimization problem of Eq. (4) is found through the introduction of

Lagrange multipliers α_i to derive the dual formulation, followed by the use of standard quadratic programming techniques to obtain the optimal weight vector

$$(5) \quad w_0 = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i$$

and solution

$$(6) \quad \hat{f}(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \langle x, x_i \rangle + b$$

The SVs are the training samples with nonzero α_i .

A nonlinear representation is obtained by projecting the training data into a high dimensional feature space via a mapping function $\Phi(x)$:

$$(7) \quad x \in \mathbb{R}^n \rightarrow \Phi(x) = [\phi_1(x), \phi_2(x), \dots, \phi_n(x)]^T \in \mathbb{R}^f$$

A linear function is approximated in this new feature space, resulting in a solution that is potentially highly nonlinear when expressed in the original variable space. Rather than define the mapping directly, SVMs rely on an implicit definition of $\Phi(x)$ provided by a kernel function defined as

$$(8) \quad K(x, x_i) = \langle \Phi(x), \Phi(x_i) \rangle$$

Using the kernel function K , the similarities between samples given by inner products in the high dimensional feature space are computed directly in the original variable space, providing the solution

$$(9) \quad \hat{f}(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x, x_i) + b$$

K must satisfy particular criteria and in practice users typically specify a function from one of a few families. Each requires the specification of one or more free parameters. A popular choice is the Gaussian radial basis function (RBF) of width γ :

$$(10) \quad K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\gamma^2}\right)$$

Narrower kernels correspond to more complex solutions when expressed in the original variable space.

The RBF kernel typically performs well due to several computational and practical advantages, including the need to specify only a single free parameter (Brereton and Lloyd, 2010; Salcedo-Sanz et al., 2014).

The parameters ϵ , C , and γ collectively determine the complexity of the regression function and its ability to generalize to new data. Optimal values are problem-specific, varying with the available training data and set of predictor variables. There is often little basis for their selection apart from testing a very large number of possible combinations against validation data. SVR parameterization is therefore equivalent to a search for an optimal combination of values from a multidimensional search space. The complexity of the problem is further increased if the search is expanded to include the selection of a subset of predictor variables. As with other statistical or machine learning methods, variable selection reduces computational complexity and can facilitate interpretation of modeled relationships. Of greater significance to our application, variable selection can alter predictive performance, perhaps reducing systematic error. Similar benefits may follow from the selection of a subset of available training data (Blum and Langley, 1997). All aspects of model specification are ideally performed simultaneously, and several classes of heuristic optimization or search algorithms are suitable, including ant colony optimization (e.g., Huang, 2009; Samadzadegan et al., 2012), particle swarm optimization (e.g., Li and Tan, 2010; Lin et al., 2008), and genetic algorithms (e.g., Bazi and Melgani, 2006; Friedrichs and Igel, 2005; Huang and Wang, 2006).

2.3.3.2. Genetic Algorithms

GAs are population-based optimization algorithms founded on the analogy of evolution by natural selection. A population of possible solutions is subjected to a selection pressure, leading to the evolution of traits associated with improved outcomes (Holland, 1975; Zäpfel et al., 2010). A GA optimization of SVR parameters ϵ , C , and γ treats each parameter as a gene and a combination of parameter values as a genotype of an individual solution. Observed characteristics of a trained SVR model constitute the phenotype of an individual solution, and these could include various measures of regression error. A GA designed to minimize a specific error metric (e.g., RMSE) applies a selection pressure to the population of solutions by favoring the perpetuation of genes that are associated with

lower metric values. Over successive generations the population evolves to include models with progressively lower error.

A GA creates a new generation of solutions by recombining properties of existing solutions (Zäpfel et al., 2010). A certain proportion of individuals are selected as parents, with probability of selection determined by a specific fitness metric. A recombination operation is applied to pairs of parent solutions to generate members of the new generation. Recombination ensures inheritance of genetic information from fit individuals, while introducing novelty to the next generation. The iterative recombination and replacement of solutions will generally lead to a loss of population diversity and a less comprehensive search for solutions (Zäpfel et al., 2010). Random mutations of genes are interjected at each generation to promote diversity and encourage a more expansive search. Evolution proceeds until a specified criterion is met, for example, convergence of population traits such that additional iterations result in little further improvement, or execution of a specified number of generations. Upon termination the typical GA returns the individual with maximum fitness as the optimal solution. By combining a guided search with a certain level of randomization, GAs are capable of obtaining near-optimal solutions from a large and complex search space (Goldberg, 1989).

GAs have been applied to a variety of SVM optimization problems, including parameter selection for both classification and regression (e.g., Friedrichs and Igel, 2005; Lorena and De Carvalho, 2008; Üstün et al., 2005; Wu et al., 2009), variable selection (e.g., Li et al., 2011), and simultaneous parameter and variable selection (e.g., Bazi and Melgani, 2006; Huang and Wang, 2006). The predominant approach is to identify a single best solution according to a single model performance objective. However, GAs are very well suited to multi-objective optimization. As a population-based algorithm, they explore different portions of the search space simultaneously and from a single run they can provide a large and diverse set of solutions expressing trade-offs between objectives (Konak et al., 2006). Ghoggali et al. (2009) applied a multi-objective GA to a pair of semi-supervised image

classification problems, and demonstrated that the simultaneous minimization of SVM classification error and number of SVs (as a measure of model complexity) produced better results than the minimization of either criterion alone. Suttorp and Igel (2006) described the use of a multi-objective GA to train SVMs for pedestrian detection from infrared images obtained by driver assistance systems. Control of classifier performance was attained through the simultaneous minimization of omission error, commission error, and number of SVs. Pasolli et al. (2011) selected parameter values for SVR models to predict soil moisture from synthetic aperture radar imagery using a multi-objective GA designed to optimize RMSE, R^2 , and the slope of the relationship between predicted and observed values. Here we apply a similar approach to the prediction of tree species abundance.

2.3.4. Multi-objective SVR (MOSVR) Algorithm Description

We implemented a multi-objective SVR (MOSVR) algorithm that includes parameter selection, variable selection, and a form of training sample selection. Variable and sample selection are primarily employed as a basis for diversifying SVR models, supporting the evolution of models with a wide range of error characteristics. Our approach to sample selection is to specify a subset of reference samples as eligible for exclusion from model training. All reference samples are used for model validation within a k-fold cross-validation (CV) procedure. Use of a GA requires the expression of individual models in the form of a genotype subject to selection, genetic recombination, and mutation. Each SVR model is represented by a bit string chromosome, composed of segments encoding parameter values, variable selection, and sample exclusion (Fig. 2.2). The lengths of segments representing parameter values determine the levels of precision with which real values are represented by binary encoding. Variable selection is encoded as a bit string segment with length equal to the number of available covariates, interpreted as a binary mask specifying selection of specific covariates. Sample exclusion is similarly encoded as a segment with length equal to the number of samples made eligible for exclusion,

indicating specific samples to be excluded from model training. The GA is initiated with a uniform random population of a user-specified size.

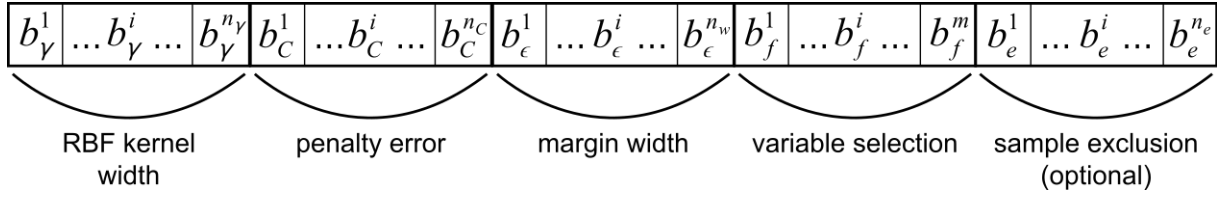


Figure 2.2. Genetic algorithm chromosome design. Bit string chromosomes are composed of segments encoding model parameter values, predictor variable selection, and training sample exclusion.

Numerous multi-objective GAs have been published and reviewed (Konak et al., 2006). Our approach is based on the popular NSGA-II algorithm (Deb et al., 2002) implemented in the MATLAB Global Optimization Toolbox, Release 2014a (The MathWorks, Inc., Natick, Massachusetts, USA). The LIBSVM open source software (Chang and Lin, 2011) is used for SVM training and prediction. The MATLAB implementation of NSGA-II acts as a wrapper for LIBSVM. A diagrammatic representation of algorithm details is provided in Fig. 2.3.

For each generation of solutions, NSGA-II differentiates groups of parents (P) and offspring (Q) of equal size. Initially all individuals are random and specification of P_0 and Q_0 is arbitrary. The chromosome representing each member of the current population ($P_t \cup Q_t$) is decoded into real-valued SVM parameters and variable selection and sample exclusion masks, used to extract variables and training samples from reference data. Individual models are trained and validated by a k-fold CV. Continuous variables are scaled to unit range ([0,1]) at each CV iteration to prevent the disproportionate influence of those with larger numeric ranges. The CV procedure is repeated a user-specified number of times and results averaged to reduce the uncertainty of objective function estimates (Kim, 2009). CV estimates of objective function values are assigned to each member of the current population. Objective functions quantify total error ($RMSE_T$) and systematic error ($RMSE_{sys}$):

$$g_1 = RMSE_T = \left[\frac{1}{n} \sum_{i=1}^n (\hat{f}(x_i) - y_i)^2 \right]^{1/2}$$

$$g_2 = RMSE_{SYS} = \left[\frac{1}{n} \sum_{i=1}^n (a + by_i - y_i)^2 \right]^{1/2}$$

where a and b are the intercept and slope of the least squares regression between predicted values $\hat{f}(x_i)$ and observed values y_i (Willmott, 1981).

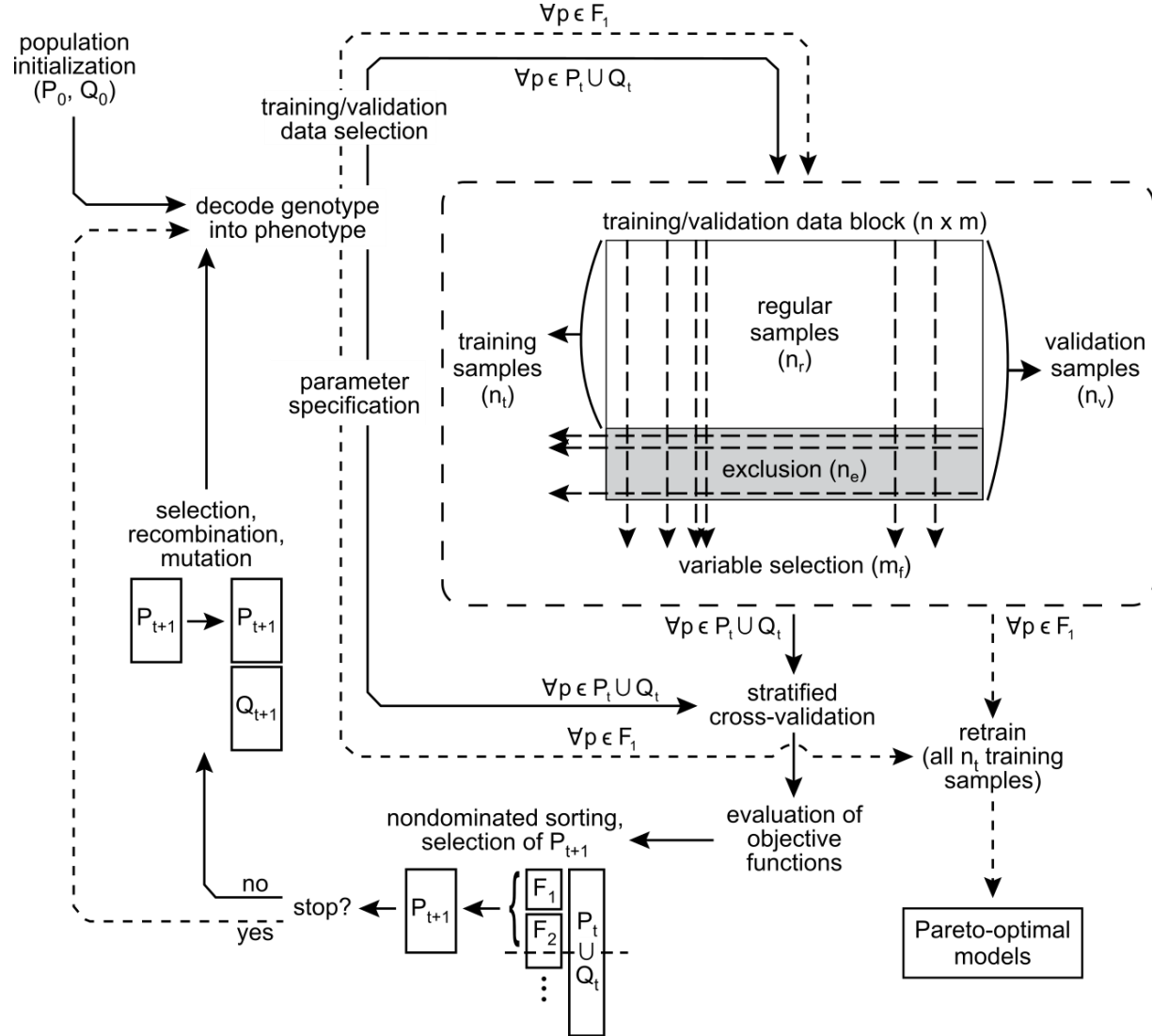


Figure 2.3. Multi-objective support vector regression algorithm implementation. Following selection of training and validation data, SVR models are fit and predictions made using LIBSVM. Objective function values are estimated by repeated cross-validation, and serve as the basis for population sorting, parent selection, and genetic operations embedded within the nondominated sorting genetic algorithm (NSGA-II).

The objective functions map solutions into a two-dimensional objective space $\Phi = \{g_1(p) \ g_2(p) \mid p \in \Omega\}$, where Ω is the set of all possible solutions. Solution p_i is said to dominate solution p_j provided

$g_1(p_i) \leq g_1(p_j)$ and $g_2(p_i) \leq g_2(p_j)$ with at least one strict inequality. In other words, one solution dominates another if it is better in one objective and at least as good in the other. A solution is nondominated if neither objective can be improved further without a worsening of the other. The set of nondominated solutions in Ω is referred to as the Pareto set, and the image of the Pareto set in the objective space Φ is the Pareto front. The goal of NSGA-II is to closely approximate the true Pareto set by driving evolution toward the Pareto front.

At each generation, NSGA-II sorts the current population of solutions ($P_t \cup Q_t$) into a sequence of nondominated fronts (F_1, F_2, \dots). The first front F_1 includes all nondominated solutions from the total population and is the current best approximation of the Pareto front. Once F_1 is obtained, these solutions are removed from the population, and the next front F_2 is obtained as nondominated solutions from the reduced population. The process is iterated until all population members have been assigned to a front. NSGA-II subsequently identifies one half of the population as the next generation of parents (P_{t+1}), selecting solutions from successive fronts. The maximum number of parent solutions selected from F_1 is constrained to a user-specified proportion of the total population in order to promote population diversity throughout algorithm execution. Additional fronts are added to P_{t+1} in succession until one cannot be accommodated in its entirety. At that point, solutions are selected from sparse or less crowded portions of the front to further promote population diversity.

The next generation of offspring (Q_{t+1} , equal in size to P_{t+1}) are obtained through genetic recombination and mutation of parent solutions. A user-specified proportion of offspring are produced through genetic recombination of a pair of parent solutions, and the remainder through mutation of a single parent. Individual parents are identified by tournament selection (Zäpfel et al., 2010), where a user-specified number of solutions are randomly selected from P_{t+1} and the best is selected as a parent. Better solutions lie on lower ranked fronts and in less crowded regions along their front. Genetic recombination may occur through one of several standard crossover operations in which an offspring is

constructed from bit string segments copied from its parents. Different recombination operations determine the manner in which information is exchanged and the potential degree of novelty introduced through exchange (Zäpfel et al. 2010). An offspring produced by mutation is a copy of its parent subjected to a mutation operation that switches individual bit values with a user-specified probability. Once offspring have been produced, parent and offspring chromosomes ($P_{t+1} \cup Q_{t+1}$) are decoded and the process repeats.

Stopping criteria are evaluated at each generation after solutions are sorted into nondominated fronts. The algorithm is assumed to have converged to a close approximation of the Pareto front when the change in spread of solutions along F_1 averaged over a user-specified number of generations is less than a user-specified threshold. Alternatively, the algorithm stops when the generation count exceeds a user-specified maximum. Once stopped, members of F_1 are retrained using all available training samples and returned as a set of alternative solutions expressing tradeoffs between $RMSE_T$ and $RMSE_{SYS}$.

2.3.5. MOSVR Algorithm Execution

From the FIA data compiled for our study area, we modeled and mapped the relative abundance of 13 tree species for eventual inclusion in our LANDIS-II applications (Table 2.3). SVR parameter values were constrained within reasonable ranges ($\log(\gamma) \in [-4,0]$; $\log(C) \in [-1,3]$; $\log(\epsilon) \in [-4,0]$). A set of 78 reference samples was made eligible for exclusion from model training, including samples with high spectral variability within pixel neighborhoods, averaged across all images and bands, as well as samples for which FIA records indicated presence of non-forest cover types. Note that we retained samples for which FIA records indicated multiple forest types. All reference samples were used for model validation in a 10-fold, 10 times repeated CV.

We specified GA parameters that balanced population diversity against execution time. The GA operated on a population of 500 chromosomes, with a maximum of 20% maintained on the approximate Pareto front. Parent chromosomes were selected by tournament with 10 participants. 70%

Table 2.3. Reference data characteristics of the 13 modeled tree species.

Species	Common Name	Species Code ¹	Prevalence ²	Mean Relative Abundance ³	Median Relative Abundance ³	Maximum Relative Abundance ³
<i>Abies balsamea</i>	Balsam fir	ABBA	0.84	0.15	0.089	0.92
<i>Acer rubrum</i>	Red maple	ACRU	0.69	0.11	0.045	0.91
<i>Acer saccharum</i>	Sugar maple	ACSA3	0.38	0.12	0	0.99
<i>Betula alleghaniensis</i>	Yellow birch	BEAL2	0.62	0.11	0.045	0.82
<i>Betula papyrifera</i>	Paper birch	BEPA	0.65	0.11	0.036	0.74
<i>Fagus grandifolia</i>	American beech	FAGR	0.26	0.036	0	0.64
<i>Fraxinus americana</i>	White ash	FRAXI	0.14	0.013	0	0.43
<i>Picea glauca</i>	White spruce	PIGL	0.24	0.017	0	0.59
<i>Picea mariana</i>	Black spruce	PIMA	0.099	0.037	0	0.99
<i>Picea rubens</i>	Red spruce	PIRU	0.81	0.16	0.079	0.90
<i>Pinus strobus</i>	White pine	PIST	0.19	0.031	0	0.57
<i>Thuja occidentalis</i>	Northern white cedar	THOC2	0.41	0.072	0	0.81
<i>Tsuga canadensis</i>	Eastern hemlock	TSCA	0.084	0.012	0	0.45

¹Species codes used by the USFS FIA Program.

²Prevalence was calculated as the proportion of FIA plots at which the species was present.

³Relative abundance was calculated as a proportion of estimated live aboveground biomass (of stems >2.54 cm diameter, measured at 1.37 m).

of offspring were generated by crossover of parent chromosomes, using the ‘scattered’ crossover operation in which bits were selected from each parent at random. 30% of offspring were generated by mutation of parents, with a mutation rate of 2.5%. Scattered crossover and a relatively high mutation rate promoted population diversity and exploration of the solution space. Approximate Pareto fronts typically stabilized by 80-100 generations, and algorithm execution was limited to 120 generations.

The estimation of RMSE_{sys} by linear least squares regression of CV predictions onto observed values was in some cases sensitive to outlying samples whose CV predictions deviated strongly from those of other samples with similar observed values. In these cases, removal of outliers was required to ensure that a small number of influential samples did not drive the GA toward less desirable solutions by distorting RMSE_{sys} estimates. We implemented an automated outlier removal strategy at 30, 60, 90 generations based on the identification of influential outliers for each member of the F₁ front. Outlying samples were identified by applying a threshold to absolute studentized residuals. Influential outliers

were identified as those whose removal resulted in a change in $RMSE_{SYS}$ exceeding a threshold level, when expressed as a proportion of $RMSE_T$. Samples identified as influential outliers in the majority of F_1 solutions were removed from both training and validation data. For most species, we applied a residual threshold of 3 and a $RMSE_{SYS}$ threshold of 1%. For FRAXI, PIMA, and TSCA we used more conservative threshold values of 4 and 2% to reduce the number of outliers removed. The number of outliers removed for each species ranged from zero to seven, and averaged four.

Lastly, at the end of MOSVR execution, an individual solution was selected from the midsection of the Pareto front where solutions represented a compromise between $RMSE_{SYS}$ and $RMSE_T$. We selected the model positioned nearest to the origin after unit-scaling $RMSE_{SYS}$ and $RMSE_T$ values to normalize for differences in magnitude between the two. Other selection methods were informally evaluated and appeared to have little practical influence on outcomes as long as they targeted the midsection of the Pareto front.

2.3.6. Model Comparisons

We compared MOSVR results to those obtained from Random Forest (RF; Breiman 2001), gradient nearest neighbor (GNN; Ohmann and Gregory, 2002), Random Forest nearest neighbor (RFNN; Crookston & Finley 2008), and single-objective SVR (SOSVR) algorithms. RF is an ensemble algorithm based on regression trees, and has been widely applied in species distribution modeling and remote sensing applications due to its typically high predictive accuracy and ease of use. RF requires specification of several parameters, but results are not overly sensitive to parameter selections and default values are often used. GNN has been used extensively for regional tree species distribution modeling based on moderate resolution remote sensing and geospatial data. GNN was originally developed and has been commonly applied as a $k = 1$ nearest neighbor algorithm, with proximity calculated within a feature space defined by a canonical correspondence analysis (CCA) of plot measurements and image or environmental predictor data. RFNN is another $k = 1$ nearest neighbor

variant, with proximity obtained from the nodes of one or more RF models. RFNN shares the advantages of GNN, but is based on a novel, non-Euclidean distance metric that may lead to improved outcomes (Hudak et al., 2008). We also implemented a single-objective approach to SVR model training (SOSVR) using a traditional GA (MATLAB Global Optimization Toolbox, Release 2014a) minimizing $RMSE_T$, because SVR model selection is typically based on minimization of overall prediction error. Finally, to evaluate the relative benefits of variable and sample selection strategies employed by MOSVR, we compared results to two alternative MOSVR execution strategies that included parameter selection only, and parameter plus variable selection but no sample selection.

All MOSVR execution strategies used the same GA settings and the same outlier removal strategy. SOSVR runs were executed using the same values for applicable GA settings, and included parameter, variable, and sample selection. For the remaining algorithms we adopted typical parameter settings and execution strategies using R v 3.0.3 (R Core Team, 2017). RF models were fit with the R package randomForest, v 4.6-12 (Liaw and Wiener, 2002), with an ensemble size of 2000 and default parameter settings ($mtry$ = one third of the number of predictor variables; $nodesize$ = 5). For GNN, CCA models were first fit with the R package vegan, v 2.4-3 (Oksanen et al., 2017) using the relative abundance of all species as the multivariate response. Following Ohmann and Gregory (2002), we performed a forward stepwise variable selection procedure based on AIC, permutation testing, and a check of variance inflation factors. Variables were considered for addition in the order of their contribution to constrained inertia (equivalent to AIC when all variables are continuous). Variables were added provided they were deemed significant by a permutation test ($p = 0.01$, 99 permutations) and provided all variance inflation factors remained below 20. Nearest neighbor imputation was based on Euclidian distance calculated from the first seven CCA axes (accounting for >95% of total variation explained), scaled by their constrained eigenvalues. GNN imputation, and execution of the RFNN algorithm, was performed using the R package yalImpute, v 1.0-26 (Crookston and Finley, 2008). The

RFNN imputation was based on a combined nodes matrix obtained by three separate RF models, fit to total live aboveground biomass, the species with maximum relative abundance based on aboveground live biomass, and the relative abundance of that species.

Fitted models of all types were run through a 10-fold CV 100 times with different random partitions. To ensure fair comparisons amongst model types, we removed CV predictions associated with influential outliers in MOSVR, on a species by species basis. We compared mean model performance metrics from CV predictions ($RMSE_T$, $RMSE_{SYS}$, linear slope, and R^2), and calculated 95% confidence intervals under the assumption that metrics obtained by repeated CV were approximately normal. We also compared outcomes by visual evaluation of the relationship between CV predictions and observed values, focusing on systematic deviations from a 1:1 relationship. We calculated dominance and codominance as the frequency with which any species or pair of species occurred or co-occurred as one of the three most abundant species, based on our anticipated use of model outcomes for LANDIS-II initialization. Mean predicted dominance/codominance frequencies and corresponding confidence intervals were calculated from CV repetitions, and compared against observed values.

2.4. Results

The approximate Pareto fronts obtained by MOSVR generally shared a common geometry. Solutions were distributed more or less evenly across a curvilinear front suggesting that the true Pareto front was ostensibly continuous, with incremental change in one objective balanced by incremental change in the other (Fig. 2.4a). At one end, models had comparatively low total error, but high systematic error, apparent as a deviation from the 1:1 relationship between predicted and observed values (Fig. 2.4b). At the other end, models had comparatively low systematic error but higher total error (Fig. 2.4d). As expected, SVRs accomplished a reduction of total error only with an associated increase in systematic error, consistent with greater attenuation bias. Fronts were convex toward the origin (Fig. 2.4a) such that nearer to either end the value of one objective function changed much more

quickly than the other. Rather than select models with minimal systematic error from one end of the front, where small decreases in $RMSE_{SYS}$ were associated with large increases in $RMSE_T$, we selected models from the midsection (Fig. 2.4a), where total prediction error represented more of a compromise between systematic and total error (Fig. 2.4c).

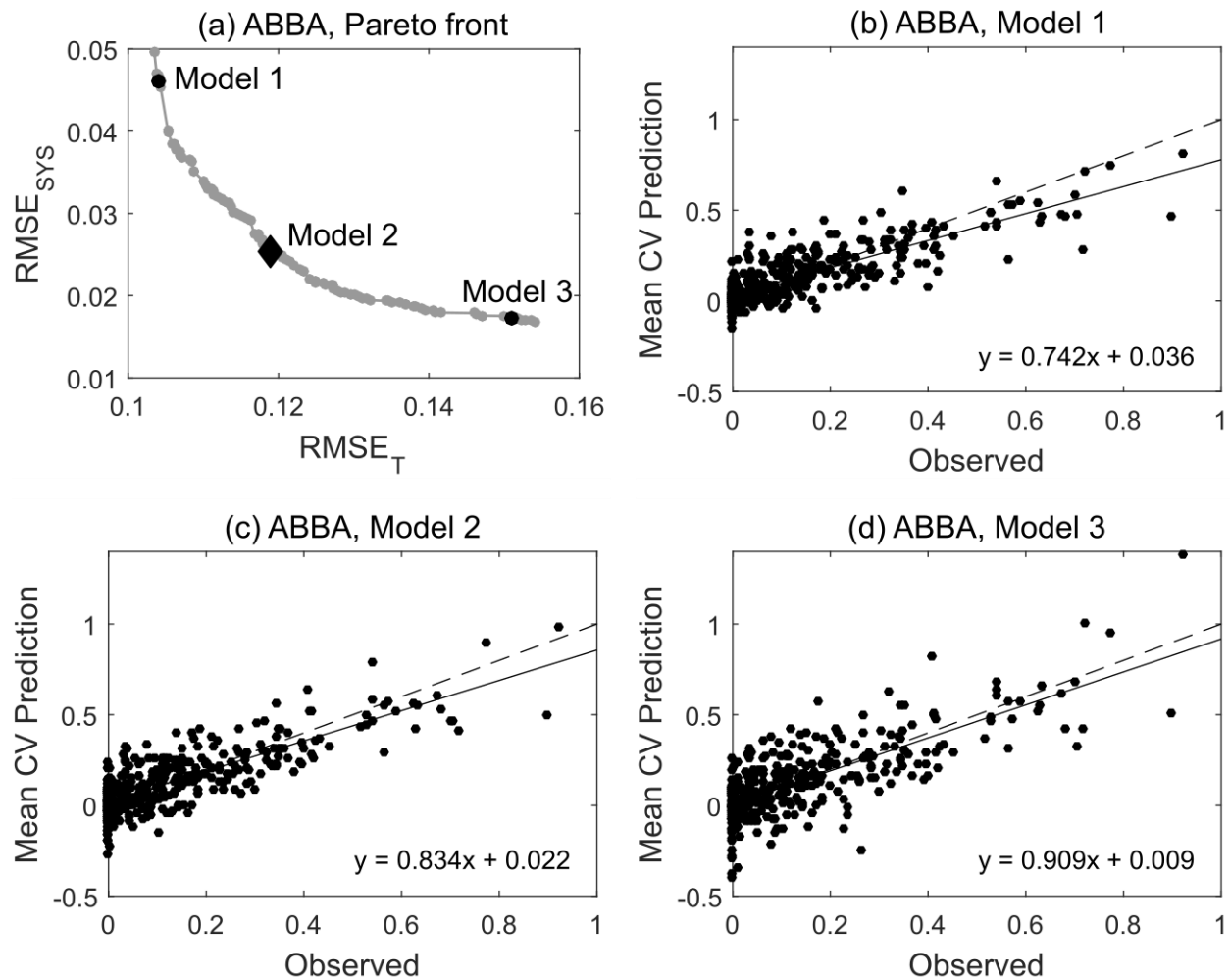


Figure 2.4. Pareto front and sample of Pareto-optimal models for species ABBA (balsam fir). (a) The approximate Pareto front obtained by MOSVR (including parameter, variable, and sample selection). (a-c) Mean predicted values obtained from 100 repetitions of a 10-fold CV plotted against observed values for models lying at different positions along the Pareto front. For comparison to other prediction methods, and for use in forest mapping, a model was selected from the midsection of the (Model 2).

Several patterns appeared when comparing model performance metrics across all model types and all species (Table 2.4). SOSVR attained the least total error for all but a single species. The least systematic error, when expressed as a percentage of total error, was always attained by MOSVR. The

slope between predicted and observed values was also greatest (closest to one) for MOSVR models. R^2 values were generally greatest for MOSVR models as a result of reduced levels of systematic error, although the low total error obtained by SOSVR occasionally resulted in R^2 values as high or higher. Of the three MOSVR execution strategies evaluated, the least systematic error was typically attained through the simultaneous selection of parameter values, variables, and samples (MOSVR A in Table 2.4). In a number of cases, parameter and variable selection (MOSVR B) achieved similar levels of systematic error, and in one case significantly lower systematic error. However, in all of these cases total error exceeded that achieved with sample selection. Parameter selection alone (MOSVR C) failed to reduce systematic error to the levels achieved with the introduction of variable selection. For these reasons, all MOSVR results presented hereafter were obtained with parameter, variable, and sample selection. In nearly all cases, nearest neighbor methods (GNN and RFNN) resulted in the greatest total error, and RF models the greatest systematic error. Compared to the best MOSVR outcomes (MOSVR A) in which systematic error ranged from 10-42% of total error across species, systematic error in RF models accounted for 62-93% of total error. Nearest neighbor methods attained levels of systematic error between those of MOSVR and RF models, provided MOSVR models included variable selection.

Table 2.4. Predictive performance by species and model type. Performance metrics were obtained from linear regression of CV predictions against observed values, averaged across 100 repetitions of a 10-fold CV. Bracketed values provide 95% confidence intervals. Bolded entries highlight model types that produced either the best mean metric value for a given species, or a value whose confidence interval overlapped that of the best. Results are presented for three MOSVR execution strategies: parameter, variable, and sample selection (MOSVR A); parameter and variable selection (MOSVR B); and parameter selection only (MOSVR C).

		RMSE _T	RMSE _{SYS} (% RMSE _T)	OLS slope	R ²
ABBA	MOSVR A	0.1201 [0.1193, 0.1208]	23.95 [23.29, 24.61]	0.834 [0.830, 0.839]	0.604 [0.600, 0.608]
	MOSVR B	0.1325 [0.1316, 0.1334]	22.78 [22.22, 23.33]	0.826 [0.822, 0.831]	0.550 [0.545, 0.554]
	MOSVR C	0.1313 [0.1309, 0.1316]	57.89 [57.69, 58.09]	0.560 [0.558, 0.562]	0.449 [0.446, 0.452]
	SOSVR	0.1085 [0.1083, 0.1088]	58.72 [58.57, 58.88]	0.630 [0.629, 0.632]	0.605 [0.603, 0.607]
	RF	0.1272 [0.1272, 0.1273]	76.67 [76.65, 76.68]	0.435 [0.435, 0.436]	0.458 [0.457, 0.458]
	GNN	0.1858 [0.1848, 0.1869]	60.71 [60.15, 61.28]	0.348 [0.341, 0.356]	0.143 [0.138, 0.148]
	RFNN	0.1823 [0.1816, 0.1829]	45.70 [45.43, 45.97]	0.527 [0.523, 0.532]	0.240 [0.236, 0.243]
ACRU	MOSVR A	0.1712 [0.1702, 0.1723]	23.97 [23.45, 24.49]	0.733 [0.727, 0.738]	0.312 [0.307, 0.316]
	MOSVR B	0.1459 [0.1450, 0.1468]	42.84 [42.43, 43.25]	0.591 [0.585, 0.596]	0.319 [0.313, 0.324]
	MOSVR C	0.1476 [0.1472, 0.1480]	77.91 [77.76, 78.06]	0.249 [0.247, 0.252]	0.145 [0.142, 0.147]
	SOSVR	0.1262 [0.1260, 0.1263]	80.07 [80.00, 80.14]	0.338 [0.337, 0.339]	0.317 [0.316, 0.319]
	RF	0.1427 [0.1426, 0.1427]	89.30 [89.28, 89.31]	0.167 [0.167, 0.168]	0.136 [0.136, 0.137]
	GNN	0.1885 [0.1875, 0.1895]	60.29 [59.86, 60.72]	0.258 [0.251, 0.265]	0.065 [0.062, 0.068]
	RFNN	0.1888 [0.1881, 0.1895]	69.44 [69.21, 69.66]	0.142 [0.139, 0.145]	0.025 [0.024, 0.026]
ACSA3	MOSVR A	0.1189 [0.1178, 0.1200]	11.36 [10.65, 12.08]	0.937 [0.932, 0.941]	0.730 [0.726, 0.734]
	MOSVR B	0.1084 [0.1074, 0.1093]	12.19 [11.47, 12.91]	0.938 [0.934, 0.942]	0.766 [0.762, 0.769]
	MOSVR C	0.1159 [0.1155, 0.1164]	39.29 [38.97, 39.60]	0.780 [0.778, 0.782]	0.697 [0.695, 0.699]
	SOSVR	0.0869 [0.0867, 0.0871]	51.46 [51.28, 51.63]	0.785 [0.784, 0.786]	0.827 [0.826, 0.827]
	RF	0.1252 [0.1252, 0.1253]	61.83 [61.80, 61.87]	0.633 [0.632, 0.633]	0.639 [0.639, 0.640]
	GNN	0.1664 [0.1648, 0.1680]	30.77 [29.96, 31.57]	0.763 [0.756, 0.771]	0.500 [0.493, 0.506]
	RFNN	0.1698 [0.1691, 0.1706]	30.22 [29.81, 30.63]	0.760 [0.756, 0.764]	0.486 [0.482, 0.490]
BEAL2	MOSVR A	0.1433 [0.1425, 0.1441]	33.65 [33.27, 34.03]	0.654 [0.650, 0.658]	0.310 [0.306, 0.314]
	MOSVR B	0.1408 [0.1397, 0.1419]	34.59 [34.16, 35.02]	0.649 [0.645, 0.654]	0.316 [0.312, 0.321]
	MOSVR C	0.1305 [0.1300, 0.1309]	74.32 [74.14, 74.49]	0.305 [0.302, 0.307]	0.189 [0.186, 0.192]
	SOSVR	0.1101 [0.1099, 0.1102]	75.54 [75.43, 75.64]	0.403 [0.402, 0.404]	0.373 [0.371, 0.375]
	RF	0.1265 [0.1265, 0.1265]	85.54 [85.53, 85.56]	0.220 [0.220, 0.220]	0.177 [0.176, 0.177]
	GNN	0.1708 [0.1699, 0.1716]	57.63 [57.18, 58.08]	0.289 [0.282, 0.297]	0.077 [0.073, 0.080]
	RFNN	0.1723 [0.1716, 0.1729]	56.03 [55.68, 56.38]	0.302 [0.298, 0.306]	0.079 [0.077, 0.081]

Table 2.4 continued

		RMSE _T	RMSE _{SYS} (% RMSE _T)	OLS slope	R ²
BEPA	MOSVR A	0.1233 [0.1226, 0.1240]	30.53 [29.98, 31.08]	0.726 [0.721, 0.731]	0.417 [0.413, 0.421]
	MOSVR B	0.1115 [0.1110, 0.1120]	48.09 [47.62, 48.55]	0.611 [0.607, 0.615]	0.422 [0.418, 0.426]
	MOSVR C	0.1257 [0.1252, 0.1262]	64.03 [63.75, 64.31]	0.412 [0.409, 0.415]	0.254 [0.250, 0.258]
	SOSVR	0.0962 [0.0959, 0.0965]	62.54 [62.30, 62.79]	0.561 [0.558, 0.563]	0.510 [0.507, 0.513]
	RF	0.1279 [0.1279, 0.1279]	92.60 [92.58, 92.61]	0.137 [0.137, 0.137]	0.131 [0.130, 0.131]
	GNN	0.1830 [0.1820, 0.1839]	63.73 [63.29, 64.17]	0.148 [0.140, 0.155]	0.021 [0.019, 0.023]
	RFNN	0.1672 [0.1667, 0.1678]	67.33 [67.04, 67.62]	0.178 [0.174, 0.182]	0.038 [0.036, 0.039]
FAGR	MOSVR A	0.0577 [0.0570, 0.0583]	20.92 [20.12, 21.72]	0.851 [0.845, 0.856]	0.592 [0.586, 0.598]
	MOSVR B	0.0638 [0.0631, 0.0644]	9.73 [9.000, 10.47]	0.926 [0.920, 0.932]	0.575 [0.569, 0.581]
	MOSVR C	0.0725 [0.0720, 0.0730]	51.61 [51.07, 52.14]	0.531 [0.526, 0.536]	0.318 [0.312, 0.323]
	SOSVR	0.0486 [0.0483, 0.0490]	52.82 [52.17, 53.47]	0.678 [0.673, 0.684]	0.632 [0.626, 0.637]
	RF	0.0671 [0.0671, 0.0672]	73.25 [73.23, 73.28]	0.391 [0.390, 0.391]	0.317 [0.316, 0.317]
	GNN	0.1037 [0.1030, 0.1045]	48.92 [48.13, 49.71]	0.368 [0.356, 0.380]	0.097 [0.091, 0.103]
	RFNN	0.0882 [0.0875, 0.0890]	52.56 [51.96, 53.16]	0.421 [0.412, 0.430]	0.168 [0.161, 0.175]
FRAXI	MOSVR A	0.0457 [0.0455, 0.0460]	35.99 [35.47, 36.51]	0.633 [0.628, 0.639]	0.307 [0.302, 0.311]
	MOSVR B	0.0481 [0.0478, 0.0485]	40.10 [39.51, 40.68]	0.570 [0.563, 0.577]	0.252 [0.246, 0.257]
	MOSVR C	0.0489 [0.0487, 0.0490]	73.05 [72.82, 73.28]	0.206 [0.203, 0.210]	0.072 [0.070, 0.074]
	SOSVR	0.0416 [0.0416, 0.0417]	91.28 [91.19, 91.37]	0.152 [0.151, 0.153]	0.138 [0.136, 0.140]
	RF	0.0468 [0.0468, 0.0468]	90.50 [90.48, 90.51]	0.059 [0.058, 0.059]	0.017 [0.017, 0.017]
	GNN	0.0548 [0.0543, 0.0552]	71.15 [70.60, 71.71]	0.131 [0.123, 0.138]	0.025 [0.022, 0.028]
	RFNN	0.0625 [0.0623, 0.0628]	62.28 [62.03, 62.53]	0.132 [0.129, 0.136]	0.015 [0.014, 0.015]
PIGL	MOSVR A	0.0470 [0.0467, 0.0473]	41.54 [40.83, 42.25]	0.547 [0.537, 0.556]	0.233 [0.226, 0.240]
	MOSVR B	0.0446 [0.0442, 0.0450]	49.89 [49.12, 50.67]	0.482 [0.471, 0.494]	0.226 [0.217, 0.235]
	MOSVR C	0.0590 [0.0587, 0.0593]	61.64 [61.33, 61.95]	0.158 [0.152, 0.164]	0.022 [0.020, 0.024]
	SOSVR	0.0394 [0.0393, 0.0394]	88.74 [88.67, 88.82]	0.189 [0.187, 0.190]	0.167 [0.165, 0.170]
	RF	0.0443 [0.0443, 0.0443]	90.35 [90.32, 90.38]	0.080 [0.079, 0.080]	0.031 [0.031, 0.032]
	GNN	0.0558 [0.0549, 0.0568]	70.43 [68.79, 72.08]	0.093 [0.076, 0.110]	0.017 [0.012, 0.022]
	RFNN	0.0679 [0.0672, 0.0687]	65.77 [65.16, 66.39]	- [-0.038, -0.031] 0.034	0.001 [0.001, 0.001]
PIMA	MOSVR A	0.0543 [0.0538, 0.0548]	15.15 [14.21, 16.08]	0.928 [0.922, 0.933]	0.777 [0.773, 0.780]
	MOSVR B	0.0626 [0.0621, 0.0631]	17.65 [16.53, 18.78]	0.899 [0.892, 0.906]	0.713 [0.708, 0.717]
	MOSVR C	0.0899 [0.0895, 0.0904]	43.52 [43.06, 43.97]	0.639 [0.634, 0.643]	0.420 [0.416, 0.425]
	SOSVR	0.0560 [0.0554, 0.0565]	36.51 [35.72, 37.30]	0.813 [0.808, 0.818]	0.740 [0.735, 0.744]
	RF	0.0947 [0.0946, 0.0948]	75.35 [75.30, 75.40]	0.346 [0.345, 0.347]	0.265 [0.264, 0.266]
	GNN	0.0948 [0.0933, 0.0962]	39.24 [37.90, 40.58]	0.661 [0.650, 0.671]	0.404 [0.394, 0.413]
	RFNN	0.0970 [0.0962, 0.0978]	40.25 [39.69, 40.82]	0.639 [0.634, 0.644]	0.377 [0.371, 0.383]

Table 2.4 continued

		RMSE _T	RMSE _{SYS} (% RMSE _T)	OLS slope	R ²
PIRU	MOSVR A	0.1571 [0.1563, 0.1579]	36.11 [35.64, 36.58]	0.685 [0.680, 0.690]	0.414 [0.410, 0.418]
	MOSVR B	0.1666 [0.1656, 0.1675]	35.24 [34.82, 35.65]	0.674 [0.670, 0.678]	0.377 [0.373, 0.381]
	MOSVR C	0.1774 [0.1768, 0.1781]	60.99 [60.72, 61.27]	0.399 [0.395, 0.402]	0.206 [0.203, 0.210]
	SOSVR	0.1385 [0.1382, 0.1387]	75.52 [75.40, 75.64]	0.418 [0.417, 0.420]	0.407 [0.405, 0.409]
	RF	0.1568 [0.1567, 0.1568]	82.34 [82.32, 82.35]	0.283 [0.283, 0.284]	0.247 [0.246, 0.247]
	GNN	0.2114 [0.2102, 0.2125]	60.02 [59.65, 60.39]	0.297 [0.292, 0.302]	0.091 [0.088, 0.094]
	RFNN	0.1973 [0.1966, 0.1979]	57.20 [56.85, 57.56]	0.372 [0.368, 0.377]	0.146 [0.143, 0.149]
PIST	MOSVR A	0.0896 [0.0891, 0.0901]	33.73 [33.21, 34.24]	0.675 [0.670, 0.681]	0.357 [0.352, 0.362]
	MOSVR B	0.1038 [0.1032, 0.1043]	34.44 [33.94, 34.95]	0.616 [0.610, 0.622]	0.257 [0.253, 0.261]
	MOSVR C	0.0979 [0.0975, 0.0983]	67.42 [67.13, 67.72]	0.292 [0.289, 0.295]	0.124 [0.121, 0.127]
	SOSVR	0.0801 [0.0800, 0.0803]	85.22 [85.14, 85.29]	0.265 [0.264, 0.267]	0.257 [0.255, 0.259]
	RF	0.0952 [0.0952, 0.0952]	88.08 [88.06, 88.10]	0.104 [0.104, 0.104]	0.044 [0.044, 0.044]
	GNN	0.1239 [0.1229, 0.1250]	61.68 [60.99, 62.36]	0.180 [0.168, 0.192]	0.031 [0.027, 0.035]
	RFNN	0.1204 [0.1196, 0.1211]	63.82 [63.49, 64.15]	0.174 [0.170, 0.178]	0.030 [0.029, 0.032]
THOC2	MOSVR A	0.1252 [0.1244, 0.1260]	25.05 [24.58, 25.52]	0.783 [0.779, 0.787]	0.461 [0.457, 0.465]
	MOSVR B	0.1388 [0.1383, 0.1394]	32.72 [32.40, 33.04]	0.684 [0.681, 0.688]	0.358 [0.355, 0.361]
	MOSVR C	0.1376 [0.1371, 0.1381]	65.31 [65.05, 65.57]	0.373 [0.370, 0.376]	0.208 [0.205, 0.211]
	SOSVR	0.1112 [0.1109, 0.1114]	75.20 [75.05, 75.34]	0.419 [0.417, 0.420]	0.401 [0.398, 0.403]
	RF	0.1374 [0.1374, 0.1375]	87.25 [87.23, 87.26]	0.166 [0.166, 0.166]	0.111 [0.111, 0.112]
	GNN	0.1715 [0.1704, 0.1725]	61.74 [61.01, 62.46]	0.261 [0.251, 0.271]	0.073 [0.068, 0.077]
	RFNN	0.1864 [0.1857, 0.1871]	60.19 [59.80, 60.59]	0.217 [0.211, 0.222]	0.042 [0.040, 0.044]
TSCA	MOSVR A	0.0450 [0.0446, 0.0454]	20.90 [20.09, 21.71]	0.824 [0.817, 0.831]	0.495 [0.488, 0.502]
	MOSVR B	0.0506 [0.0502, 0.0510]	20.87 [19.89, 21.85]	0.802 [0.792, 0.812]	0.424 [0.416, 0.432]
	MOSVR C	0.0609 [0.0606, 0.0612]	61.89 [61.60, 62.19]	0.292 [0.288, 0.297]	0.095 [0.092, 0.098]
	SOSVR	0.0404 [0.0402, 0.0405]	82.56 [82.37, 82.76]	0.372 [0.369, 0.375]	0.427 [0.422, 0.432]
	RF	0.0532 [0.0532, 0.0532]	81.28 [81.24, 81.32]	0.184 [0.184, 0.185]	0.090 [0.089, 0.090]
	GNN	0.0640 [0.0630, 0.0651]	68.25 [66.92, 69.57]	0.178 [0.166, 0.190]	0.044 [0.038, 0.049]
	RFNN	0.0708 [0.0704, 0.0712]	65.14 [64.61, 65.67]	0.129 [0.122, 0.136]	0.017 [0.015, 0.019]

Patterns of prediction error summarized by model performance metrics were readily apparent in residuals obtained by subtracting the linear fit between predicted and observed values. Species ABBA provides a representative example (Fig. 2.5). All model types produced some degree of systematic overestimation of relative abundance for observed values <0.2 , although in this case overestimation by MOSVR was negligible for relative abundance >0.05 . All model types produced a systematic underestimation of relative abundance for observed values >0.2 . The magnitude of underestimation increased as observed relative abundance increased, producing large apparent error at high relative abundance. The magnitude of systematic error at high abundance varied between models, with GNN producing the most and MOSVR the least (Fig. 2.5f). For species ABBA, the SOSVR and RF models clearly produced the least residual scatter or unsystematic error, while RFNN and GNN produced the most. This latter result did not hold in general, as the amount of unsystematic error in MOSVR predictions exceeded that in nearest neighbor predictions for most species. This follows from the fact that, although MOSVR did typically produce lower total error than nearest neighbor methods (Table 2.4), large reductions in systematic error were achieved in part by allowing for greater levels of unsystematic error. Because MOSVR models were selected to balance systematic and total error, further reduction in either systematic or unsystematic error could have been achieved, but only at the expense of the other. The degree to which MOSVR mitigated systematic over- and under-prediction varied from species to species (Fig. 2.6), partly because we adopted a strategy for selecting a specific model from the Pareto front that was based on a balance between objectives, rather than a specified magnitude of systematic error. Regardless, MOSVR always reduced systematic underestimation at the high levels of relative abundance that would presumably most influence predicted patterns of dominance/codominance.

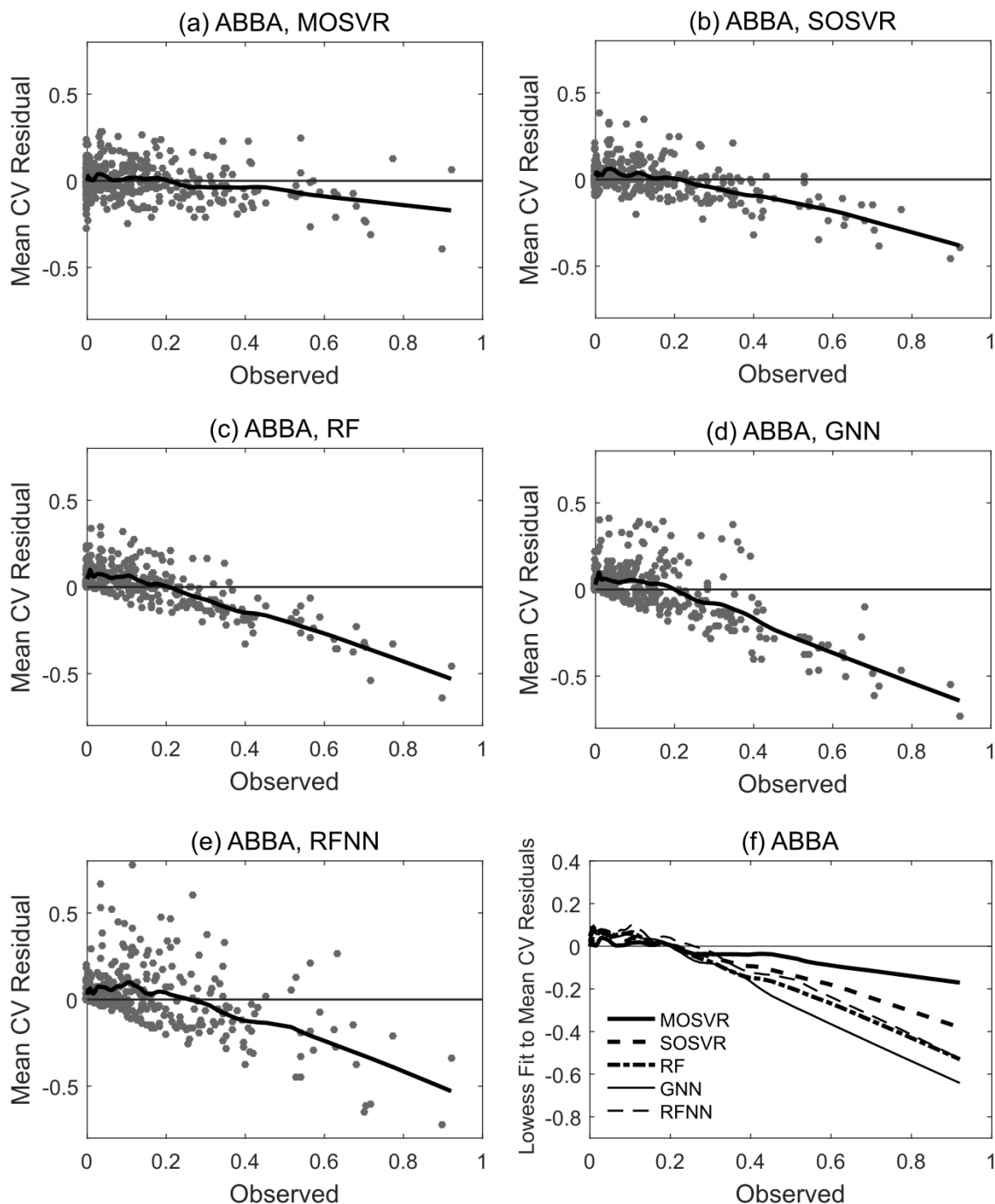


Figure 2.5. Trends in residual values for selected model types fit to species ABBA (balsam fir). Residuals plotted against observed values for (a) MOSVR, (b) SOSVR, (c) RF, (d) GNN, and (e) RFNN model types. For visualization of trends, we plotted the mean of residuals obtained from 100 repetitions of a 10-fold CV, and overlay a lowess curve (local weighted least squares regression of a first degree polynomial, spanning 20% of samples). (f) Lowess curves fit to residual plots for direct comparisons between model types.

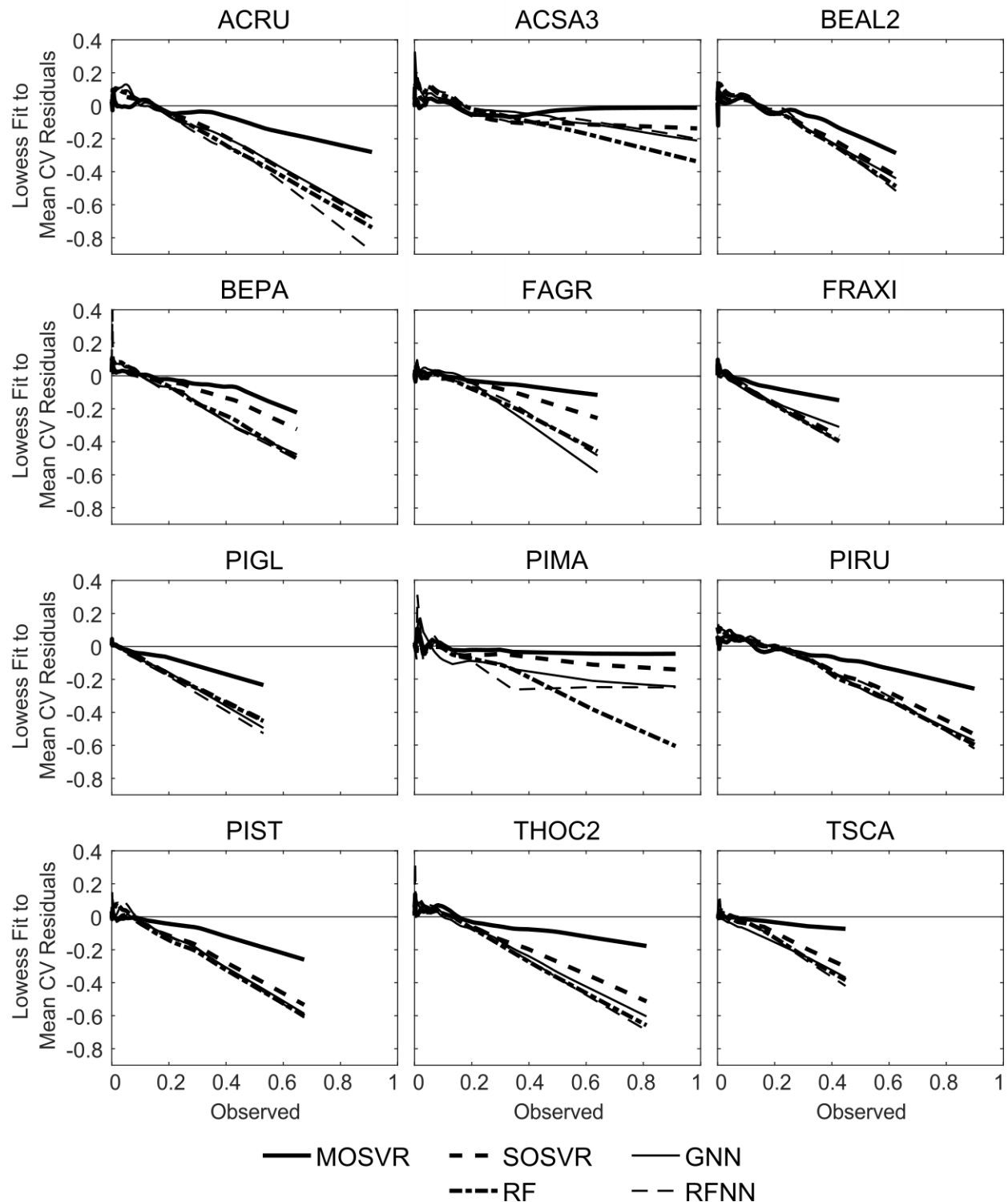


Figure 2.6. Trends in residual values for selected model types fit to individual tree species. Lowess curves fit to residual plots (demonstrated in Fig. 2.5) for each of five model types and 12 individual species. Lowess curves for species ABBA are provided in Fig. 2.5.

Species dominance and codominance was calculated as the frequency with which species occurred or co-occurred as one of the three most abundant, as observed in FIA data or as predicted. Patterns of observed dominance and codominance largely reflected species associations found in the most prevalent forest types of the region (Fig 2.7a). Elevated dominance/codominance of ABBA and PIRU were consistent with a high prevalence of upland spruce-fir. Similarly, dominance/codominance of ACSA3 and BEAL2 reflected a high prevalence of northern hardwood. Although typically occurring as a northern hardwood associate, FAGR occurred at somewhat lower relative abundance across our study area (Table 2.3), and this was reflected in dominance/codominance. A high prevalence of mixedwood associations was reflected in dominance/codominance patterns involving the hardwoods ACRU, BEAL2, and BEPA, and the softwoods ABBA and PIRU. BEPA commonly occurs at relatively high abundance following intense disturbance, which is common throughout much of our study area. A number of species are either not prevalent within our study area (FRAXI, PIMA, TSCA) or not generally found at high relative abundance (PIGL, PIST) (Table 2.3), and this was reflected in observed dominance/codominance.

Of the model types evaluated, nearest neighbor methods and RFNN in particular produced patterns of dominance and codominance that most closely conformed to those observed (Figs. 2.7e and 2.7f). The maximum absolute difference between observed dominance/codominance frequencies and those predicted by GNN and RFNN was about 4% and 2%, respectively, and absolute differences averaged less than 1% for both. In contrast, both SOSVR and RF models resulted in predicted patterns that deviated from observations much more strongly (Figs. 2.7c and 2.7d), with absolute differences averaging about 2% for each but exceeding 10% in a number of instances. The largest differences were over-estimates of codominance, and for RF several of these amounted to a near doubling of observed frequencies. MOSVR produced patterns much closer to those observed and to those predicted by the nearest neighbor methods, with a maximum absolute difference of about 6%, and an average absolute difference of about 1%. There were few consistencies in patterns of over- or under-estimation between

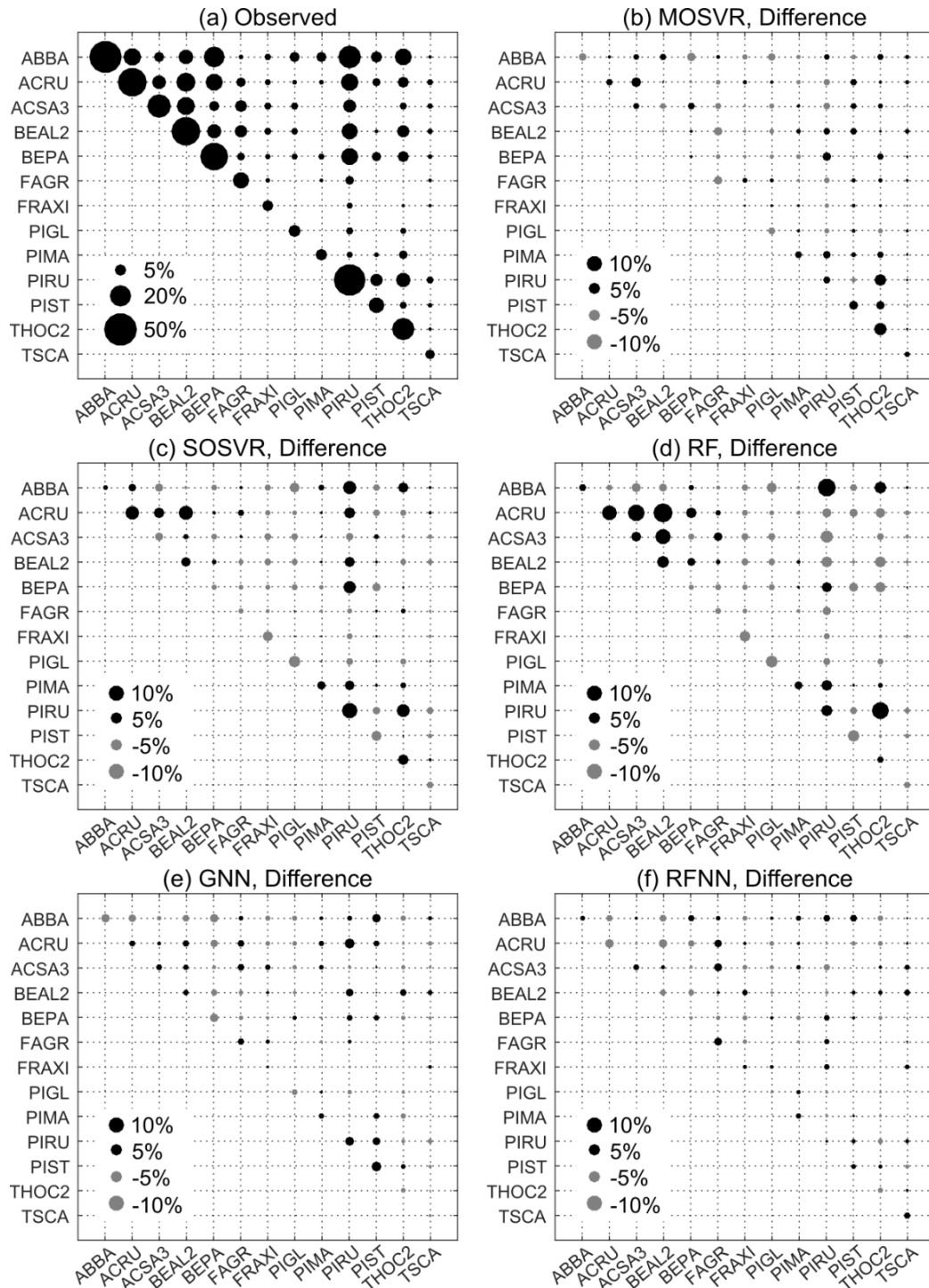


Figure 2.7. Observed and predicted patterns of species codominance. (a) Observed codominance frequency, or the proportion of FIA plots at which a pair of species co-occurs as one of the three most abundant species. Difference between predicted codominance frequency and observed codominance frequency for (b) MOSVR, (c) SOSVR, (d) RF, (e) GNN, and (f) RFNN model types. Predicted codominance values were calculated as mean values obtained from 100 repetitions of a 10-fold CV. The maximum width of corresponding 95% confidence intervals ranged from 0.4% for RF models to 0.7% for GNN.

MOSVR, GNN, and RFNN. SOSVR and RF were somewhat more consistent in their tendencies to over- or under-estimate certain dominance/codominance values.

2.5. Discussion

Motivated by our need of initialization data for the LANDIS-II forest landscape model, our goal was to develop a method of predicting individual tree species relative abundance from moderate resolution data at high accuracy and with minimal systematic error. Comparisons of multiple prediction algorithms across 13 tree species indicated that our MOSVR algorithm accomplished that goal (Table 2.4 and Figs. 2.5 and 2.6). As expected, algorithms that yielded the lowest total prediction error (RF and SOSVR) also produced the greatest systematic error, consistent with a strong attenuation bias arising from predictor variable uncertainty. Although these methods effectively minimized mean prediction error, they did so at the cost of systematic over- and underestimation at low and high ends of observed data distributions. Underestimation at high relative abundance in particular appears to have affected predicted patterns of species dominance and codominance, causing strong deviations from those observed at FIA plots (Figs. 2.7c and 2.7d). In contrast, two variants of $k = 1$ nearest neighbor methods (GNN and RFNN) reproduced observed dominance/codominance patterns with comparatively little error (Figs. 2.7e and 2.7f). By simultaneously imputing reference measurements of all species, these methods retained plot-level relationships between species and reproduced dominance/codominance patterns most closely. However, total prediction error was comparatively high for individual species, and typically included a large component of systematic error (Table 2.4). Others have emphasized the strength of nearest neighbor methods in producing reliable community-level outcomes (Henderson et al., 2014; Ohmann and Gregory, 2002). In this case, despite their reproduction of observed dominance/codominance frequencies, nearest neighbor methods yielded predictions of individual species relative abundance with comparatively low accuracy, subject to strong attenuation bias. MOSVR produced the least systematic error for all species, at levels of total error that were always less than

nearest neighbor methods and often comparable to either SOSVR or RF (Table 2.4). Predicted dominance/codominance frequencies agreed with observations much more closely than SOSVR and RF, though not quite as well as GNN or RFNN (Fig. 2.7). Nonetheless, by reducing systematic error in individual species models, MOSVR balanced the benefits of GNN and RFNN against those of SOSVR and RF. Others have developed methods of reducing systematic error in parametric statistical models of species distributions (Denham et al., 2011; Foster et al., 2012; McInerny and Purves, 2011). Here we have demonstrated an approach based on a nonparametric, machine learning algorithm.

MOSVR was able to achieve our primary objective of reducing systematic error by treating the minimization of both total and systematic error as training objectives within a multi-objective framework. Multi-objective model training requires a statistical learning process capable of generating diverse solutions through the controlled manipulation of model structure. SVMs are well-suited in the sense that manipulation of a few free parameters can dramatically alter the geometry of decision boundaries (Brereton and Lloyd, 2010). Pasolli et al. (2011) previously implemented a multi-objective method for SVR parameter selection. For our species relative abundance problems, parameter selection alone failed to achieve desired reductions in systematic error (Table 2.4). Meaningful reductions required additional complexity in model specification, achieved through the integration of variable and sample selection. Different variable combinations were expected to have different levels of impact on attenuation bias due to differences in spatial or temporal observation characteristics. Integration of variable selection into GA chromosome design enabled population diversification across a much larger search space, ultimately leading to the evolution of models with substantially reduced bias (Table 2.4). Our sample selection mechanism led to further improvements in model performance in some cases, presumably for similar reasons. SVR models are directly determined by individual samples (SVs) lying on or outside margin boundaries. The removal or addition of a SV necessarily changes model fit, whereas removal of a sample lying within the SVR margin does not. We made certain samples eligible for

exclusion based on an assumption that they were more likely to be SVs under a variety of model specifications due to observed variability in land cover or image characteristics. Enabling their exclusion further reduced bias or total error in certain cases.

The addition of variable and sample selection resulted in a very large search space. With a population of 500 individuals and a maximum generation count of only 120 (determined by available computational resources), the GA sampled a vanishingly small proportion of the total number of potential solutions. Intuitively, we might have expected parameter selection alone to yield the best results because the GA would have been much more likely to obtain a good approximation to a true global optimum. That was clearly not the case. Greater complexity in model specification and a larger search space was clearly needed to produce the desired solutions. Many predictor variables were correlated, and some samples eligible for exclusion likely had no effect on model fit because they were consistently placed within SVR margins. These factors would have reduced the effective size of the search space. Additionally, trade-offs between objectives may have followed a characteristic pattern that was determined more by the physical characteristics of the data and less by the specifics of SVM models; there may have been many, many ways to achieve the same or similar objective function values. Lastly, GAs provide a guided search mechanism, and are capable of evaluating a highly diverse set of solutions while narrowing in on aspects of model specification that are most associated with desirable performance characteristics (Goldberg, 1989). For our application, the NSGA-II algorithm was highly effective when presented with a sufficiently diverse population of potential solutions.

Use of a GA for model training bears certain implications for model interpretation, particularly in a multi-objective framework. Similar performance characteristics may be achieved with different model specifications, and models lying near to one another on the Pareto front may show substantive differences in SVR parameters or variable/sample selections. MOSVR relative abundance models included on average 21 of 94 variables, and excluded on average 24 samples from model training (Table

2.5). For nearly all species, models included spectral, terrain, and climatological variables. This was expected given known influences of climatic gradients, topo-edaphic conditions, and disturbance history on current forest type and tree species distributions. We caution against over-interpretation of variable selections and sample exclusions at this point. GAs provide a group selection mechanism, and we cannot be sure that any individual variable (or sample) provided a substantive contribution to model performance. Additional steps could be taken to evaluate the relative importance of variables in particular. Post-hoc analyses of response and covariate values could be used to illuminate important relationships between variables (e.g., Goldstein et al., 2014), and more comprehensible decision rules can be obtained by modeling the SVR predictions themselves (e.g., Martens et al., 2007). Inspection of variable selection and variable importance patterns across the Pareto front may provide insight into how certain variables may influence attenuation. As is typical of ML approaches, further work on methods of model interpretation is warranted.

Table 2.5. Multi-objective support vector regression variable and training sample selection.

Species Code¹	Number of variables selected					Training samples excluded	Training samples used
	Spring image bands ² (of 18)	Summer image bands ³ (of 12)	Fall image bands ⁴ (of 18)	Terrain (of 29)	Climate (of 17)		
ABBA	7	4	4	6	6	21	326
ACRU	4	2	0	10	3	29	320
ACSA3	2	2	3	3	4	29	314
BEAL2	2	4	5	8	2	38	308
BEPA	3	5	5	7	3	26	318
FAGR	6	1	4	5	1	18	325
FRAXI	1	5	7	2	7	25	323
PIGL	3	4	5	5	0	17	326
PIMA	4	4	7	5	2	19	323
PIRU	5	3	7	7	7	16	327
PIST	3	4	5	8	4	21	324
THOC2	6	2	8	4	4	19	328
TSCA	1	3	2	3	6	28	317

¹Species codes used by the USFS FIA Program.

²April 29, 2006; May 12, 2005; May 25, 2001.

³June 10, 2004; July 20, 2001.

⁴Sept. 14, 2004; Sept. 30, 2001; Oct. 6, 2006.

Although MOSVR effectively reduced systematic error, there is room to question when this is necessary or desirable. Riemann et al. (2010) compared biomass predictions from a GNN variant to FIA observations and found over-estimation at low biomass and under-estimation at high biomass. They recognized that the scale mismatch between moderate resolution pixels and FIA plots was at least partially responsible for this pattern, but treated it as a product of the map validation process and reference data uncertainty rather than a pattern of systematic error caused by attenuation bias in the predictive model. They reasoned that if the validation data had been collected at the same scale as the predictions (250 m pixels), this pattern of disagreement would have been less pronounced or absent. Spatial averaging of plot and pixel values across progressively larger grid cells led to a progressive reduction in the magnitude of systematic disagreement. This was cited as evidence that systematic disagreement was not the product of prediction bias, but rather an artifact of validation by direct comparison of plots and pixels. They suggested that this plot-pixel comparison may be inappropriate, even when the predictive algorithm was trained on the same plot data set.

Whether systematic disagreement between FIA plots and map pixels reveals model bias or a validation artifact depends in large part on the nature of predictor variable uncertainty. Xu et al. (2009) examined this issue in the context of ordinary linear regression. Using a field measurement protocol specifically designed to investigate the effects of mismatches in scale and location between plots and pixels, they evaluated prediction error patterns against those expected from two types of predictor variable uncertainty. When the observed predictor W is a noisy realization of the true or ideal predictor X ($W = X + U$, where the error term U has zero mean and is independent of X such that $E(W|X) = X$), the Classical error model applies. This corresponds to the situation in which plots are larger than pixels, or a species responds to a long-term average but the corresponding predictor variable reflects a shorter timeframe (as would be the case for our mid-month insolation predictors, for example). When the observed predictor is considered a smooth representation of the true or ideal predictor ($X = W + U$ and

$E(X|W) = W$), the Berkson error model applies. This corresponds to the situation in which plots are smaller than pixels, or a species responds to environmental conditions over a shorter timeframe than predictors represent (as may be the case when species are affected by extreme conditions that are not resolved by climatological predictors). Xu et al. (2009) demonstrated that although Berkson error does cause apparent systematic error in cross-validation outcomes, that pattern is no longer present when predictions are compared to new reference observations made at the same scale. Linear models are not biased by Berkson error. In contrast, Classical error does cause strong attenuation bias of the model itself, affecting coefficients and introducing systematic error that does not go away when reference data are scaled to match pixels.

The Berkson model fits the situation in which moderate resolution predictors are paired with FIA plots, and the work of Xu et al. (2009) would appear to validate the assertions of Riemann et al. (2010) on those grounds. However, several factors virtually ensure that actual predictor error deviates from the Berkson model. Other sources of predictor uncertainty undoubtedly compound error associated with scale mismatches. Location mismatches caused by georeferencing or GPS error, for example, are best represented by a mixture of Classical and Berkson error and can cause attenuation bias more severe than Classical error associated with a scale mismatch (Xu et al., 2009). Additionally, many applications build models using predictors with different patterns of uncertainty, some of which may be best represented by Berkson error and some by Classical error. For species distribution models that utilize environmental variables, the nature of predictor uncertainty may differ by species due to different responses to environmental conditions (e.g., differing sensitivity to extreme vs. average conditions). Finally, the analysis provided by Xu et al. (2009) was based on ordinary linear regression. Both Berkson and Classical error can cause attenuation bias and systematic prediction error when models are nonlinear or nonparametric (Carroll et al., 1995). For these reasons, the systematic error apparent in numerous studies, including Riemann et al. (2010), is at least partially attributable to true attenuation

bias. It is also worth noting that a comparison of spatially averaged predictions and plot values neither contradicts nor corroborates the presence of attenuation bias in the predictive model. Assuming minimal mean bias and relatively homogeneous spatial error patterns, averaging across larger and larger areas will reduce systematic disagreement even if the underlying predictive model is severely biased. This simply means that under-estimates are balanced by over-estimates, and is entirely consistent with the origin of attenuation bias in the minimization of average error. In the absence of a more thorough accounting of predictor uncertainty and its effects within a specific modeling framework, it seems safe to conclude that systematic deviations in plot-pixel comparisons are at least partially indicative of true bias.

The ultimate impact of attenuation bias on map use will presumably depend on map- and application-specific factors. Attenuation does not degrade mean predictive accuracy, and area averages should be minimally affected. However, attenuation can dramatically affect spatial prediction patterns, particularly at the high end of observed values. Species ABBA provides a convenient illustration. MOSVR and SOSVR models explained nearly identical amounts of variation in observed values, but MOSVR predictions had less systematic error and SOSVR predictions had less scatter and lower total error (Table 2.4 and Fig. 2.5). Spatial patterns of prediction were notably different at landscape scales (Fig. 2.8). Whereas MOSVR predicted values up to 100%, SOSVR predictions only infrequently exceeded 75%. The stronger attenuation bias of SOSVR generally suppressed local variability and produced a more diffuse pattern of species relative abundance than expected. MOSVR reduced attenuation bias, producing what we consider to be more realistic spatial patterns including patches of high ABBA relative abundance. These differences are sufficient to affect map use. ABBA is the primary host of the eastern spruce budworm (*Choristoneura fumiferana* Clem.), a native defoliator that causes widespread mortality during cyclic outbreaks (Morin et al., 2007). Vulnerability to spruce budworm defoliation is in large part determined by primary host relative abundance, with the greatest impact anticipated to occur in mature

stands exceeding 75% primary host abundance (Hennigar et al., 2011). Attenuation bias can cause systematic under-estimation of vulnerability, potentially affecting projected patterns of budworm impact. This is only one example of the general problem posed by attenuation bias in remote sensing applications. MOSVR provides one way to reduce bias while maintaining high overall predictive accuracy and the benefits of a nonparametric approach.

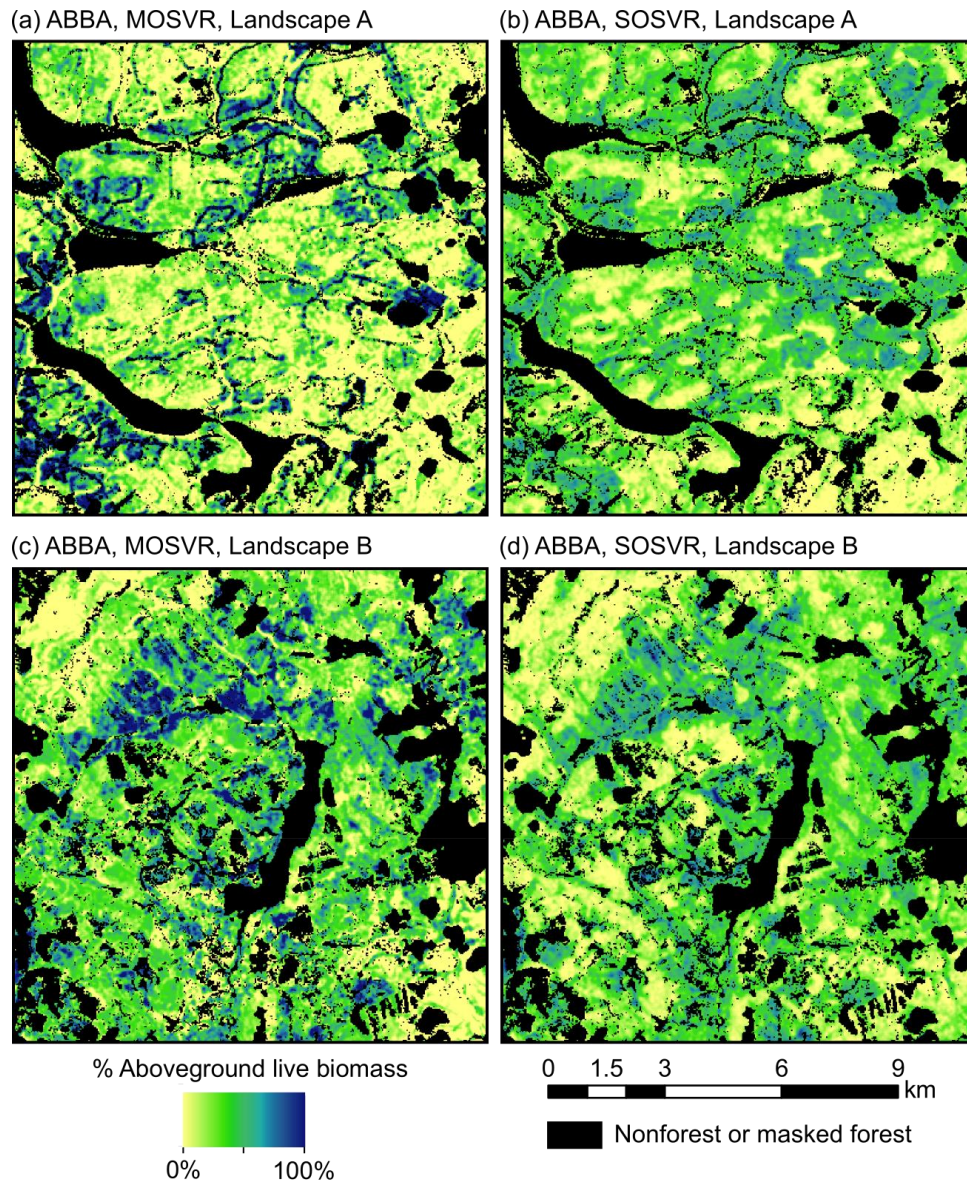


Figure 2.8. Spatial predictions of relative abundance for species ABBA (balsam fir). Relative abundance predicted from (a, c) MOSVR and (b, d) SOSVR models, across two randomly positioned sample landscapes 12 km x 12 km in size. Masked areas include nonforest pixels and forest pixels affected by canopy change during the study period or missing data due to cloud or snow cover. Predictions were truncated at 0 and 100%.

2.6. Conclusions

Patterns of error observed in predictions of tree species relative abundance were consistent with strong attenuation bias caused by uncertainty in remote sensing and geospatial predictor data. Comparing results across different predictive models, systematic error as a fraction of total error was typically greatest in regression models that achieved the lowest overall error. Pronounced underestimation at high relative abundance caused large deviations between predicted and observed patterns of species dominance and codominance. As expected, nearest neighbor methods produced better agreement with observed dominance/codominance by preserving observed species associations. Yet predictive accuracy was low and attenuation bias was high for individual species. MOSVR effectively reduced systematic error for all species while maintaining comparatively low total error, and improved predicted patterns of dominance/codominance to a level approaching that of the nearest neighbor methods.

Others have made compelling arguments that physical differences in scale and location between pixels and field plots are primary contributors to attenuation bias (Rejou-Mechain et al., 2014; Xu et al., 2009), and some have suggested that the use of FIA or similar forest inventory data for model training may be ill-advised (Robinson et al., 2013; Xu et al., 2009). Yet FIA data is used to train predictive models, and although error patterns at the scale of predictions are not always reported (e.g., Duveneck et al., 2015; Wilson et al., 2012), results are probably subject to some level of attenuation bias. Not all systematic error is indicative of model bias, but without a thorough accounting of predictor uncertainty and its impact on predictions in a specific modeling framework, it may be best to assume that some level of correction is warranted. In that case, MOSVR may provide an effective means of reducing systematic error in nonparametric regression models.

CHAPTER 3

CONTROLLING MAPPED CLASS PREVALENCE AND THE BALANCE OF ERROR BY MULTI-OBJECTIVE OPTIMIZATION OF SUPPORT VECTOR CLASSIFICATION MODELS

3.1. Abstract

The mitigation of uncertainty in remote sensing applications requires not just the reduction of prediction error but consideration of error patterns and their impacts on data use. In a thematic mapping context, commission and omission errors typically have different consequences, and an imbalance between the two results in a biased estimation of class prevalence and a biased representation of class distributions. An unbiased classification may be desirable, but given the tremendous diversity of map uses, the balance of error and level of bias achieved should be informed by intended use. We present a multi-objective support vector classification algorithm (MOSVC) that simultaneously minimizes classification bias and either omission or commission error through optimal parameterization of support vector machines (SVMs), selection of covariate subsets, and flexible use of available training data. Multi-objective optimization produces alternative solutions that express tradeoffs between class accuracy and bias under the expectation that different tradeoffs may be more or less beneficial for specific applications. We demonstrate MOSVC within a multispectral remote sensing context by mapping tree species occurrence and canopy disturbance in the temperate Atlantic Northern Forest of Maine, U.S.A. Applied to three different species occurrence problems, MOSVC produced diverse sets of alternative models and maps including solutions with zero bias. In contrast, single-objective optimization of SVMs produced inconsistent and biased outcomes because individual training criteria could not adequately control the balance of omission/commission error. MOSVC solutions with different levels of bias produced different representations of class distributions depending on spatial patterns of omission and commission error. This was most apparent for the canopy disturbance problem because omission/commission errors were visible through comparison of pre- and

post-disturbance imagery. MOSVC produced disturbance maps with uniformly high overall accuracy (>97%) despite extensive cloud contamination and the absence of image pre-processing. Yet maps with different error characteristics demonstrated different impacts of omission and commission error on disturbance class configuration. By approaching error reduction as a multi-objective optimization problem, MOSVC produces alternative solutions that collectively control predicted class prevalence and the balance of omission/commission error. Individual solutions can be selected to meet specific application needs, or multiple solutions can be compared to evaluate the sensitivity of application outcomes to map error characteristics.

3.2. Introduction

Spatial data and satellite-derived maps in particular play diverse roles in the ecological and environmental sciences. Remote sensing and map data provide a basis for planning and executing field studies, developing and calibrating models, quantifying ecosystem processes or services, and evaluating environmental change. Natural resource managers use maps to characterize resource conditions, project changes, and direct management actions. Maps are, however, abstract and imperfect representations of environmental variation. Map error heightens uncertainty in application outcomes, and for this reason new analytical methods are continually sought to improve map accuracy. Many remote sensing applications have adopted and elaborated pattern recognition or machine learning (ML) algorithms, most notably supervised algorithms based on neural networks (Mas and Flores, 2008), kernel methods including support vector machines (Mountrakis et al., 2011), decision trees (Gislason et al., 2006; Pal and Mather, 2003), or k-nearest neighbors (McRoberts et al., 2010). In the context of supervised learning, ML algorithms induce relationships between predictor and response variables without pre-specification of a form for the modeled relationship, for example by fitting a model made up of many simple components or primitive functions (e.g., decision rules in a classification tree). Predictions based on induced relationships are often more accurate than those based on more

traditional parametric statistical models because relationships between variables are often too complex or too little understood to pre-specify an appropriate parametric model form. Nonetheless, a certain level of error remains due to ecological and environmental complexity, data availability, measurement limitations, or other factors.

Inferences and decisions must be made within the context of map error, and mitigation of uncertainty in map applications requires not only reduction of error but consideration of error patterns and their impacts on map use. Here we specifically consider binary classification or presence-absence models, where observations are categorized into either positive or negative cases. Prediction errors are one of two types, either false positives or false negatives, or equivalently commission or omission errors for the positive class. The balance between commission and omission error dictates the degree to which class prevalence is either over- or underestimated by the model, and hence biased high or low on the map.

Many applications would benefit from classification methods that not only reduce overall error but control patterns of error and consequent bias in mapped class prevalence. Although there are established methods of deriving unbiased estimates of class area from biased maps (Olofsson et al., 2013), applications that require the spatial representation provided by a map remain subject to bias. Elimination of bias from mapped class distributions has received comparatively little attention, although Puertas et al. (2013) present an approach based on numeric optimization of thresholds applied to predicted probabilities of class membership. More generally, costs associated with over- and under-representation of positive cases are likely to differ, such that different directions or levels of bias are likely to have different impacts, either positive or negative (e.g., Loiselle et al., 2003; Václavík and Meentemeyer, 2009; Wilson et al., 2005). For example, the efficiency and effectiveness of conservation planning and reserve design are differentially impacted by both omission and commission error, but in ways that depend on how conservation goals are formulated (Rondinini et al., 2006). When the costs of

over- or under-representation differ, it may be desirable to not just reduce bias but to control bias. Atkinson et al. (2007) reduced commission error at the expense of elevated omission error when mapping the occurrence of individual ash and sycamore trees to obtain a large and accurate sample of tree locations for point pattern analysis. Maps often serve as a basis for sampling design, and control of omission/commission error and class bias could assist sampling in a number of ways, for example by improving the efficiency of stratified sampling through manipulation of the uniformity of map-based strata (Cochran, 1977). Another common map analysis objective is the quantification of class configuration, and configuration metrics are known to be sensitive to map error (Langford et al., 2006; Shao and Wu, 2008). Shao and Wu (2008) suggested that configuration metrics should be more reliable when omission and commission error are balanced, but landscape ecologists rarely quantify the impact of classification error (Lechner et al., 2012) and we are not aware of any specific assessment of the impact of biased class prevalence on calculated metric values. A certain level of bias could provide a more reliable representation of certain aspects of spatial configuration, depending on the spatial patterning of omission and commission error. This suggests not just control of map bias, but a systematic evaluation of sensitivity to different levels of bias.

Biased estimation of class prevalence has received considerable attention in ML research due to the well-recognized impact of imbalanced training data on classifier performance. When class proportions are imbalanced, ML methods (and statistical learning techniques in general) commonly produce undesirable levels of bias by favoring the majority class when attempting to fit a parsimonious model. Strategies exist to compensate for this effect, either by resampling the training data or restructuring the learning problem (e.g., cost-sensitive learning) (He and Garcia, 2009). However, ML model training typically involves tuning model structure to minimize overall error or optimize some other measure of model performance. If the training criterion cannot adequately distinguish different degrees or directions of imbalance between omission and commission error, the model training process

cannot adequately control the degree or direction of classification bias (Mouton et al., 2010). Alternatively, the balance of error and the level of bias in predicted class prevalence can be controlled by modeling continuous probabilities of class membership and then applying different thresholds to obtain presence/absence maps (Freeman and Moisen, 2008; Puertas et al., 2013). Ecological applications have focused considerable attention on this approach, using for example logistic regression or Random Forests (Breiman, 2001) to model probability of occurrence. Thresholds are often selected based on measures of presence/absence predictive performance (Freeman and Moisen, 2008). This approach to controlling error patterns is therefore similar to the selection of a model training criterion in that it requires prior specification of an appropriate performance metric or an appropriate balance between omission and commission error. It also assumes that a single probability model will produce sufficiently accurate binary classifications from various plausible thresholds, when in fact it may not (Calabrese et al., 2014; Jiménez-Valverde et al., 2013).

A different approach to controlling the balance between omission and commission error is to simultaneously minimize both as conflicting objectives. Multi-objective optimization of conflicting performance metrics results in a set of models that collectively express tradeoffs between different modeling objectives, rather than a single model identified as best according to a single objective (Jin, 2006; Konak et al., 2006). Multi-objective optimization therefore recognizes the ambiguity in evaluating models for different aspects of performance that cannot be optimized independently, such as omission and commission error rates. Consider the set of all possible model solutions Ω and let commission and omission error be represented by the objective functions f_{CE} and f_{OE} . The objective functions map solutions from Ω into a two-dimensional objective space $\Phi = \{f_{CE}(p) \ f_{OE}(p) \mid p \in \Omega\}$. Solution p_i is said to dominate solution p_j provided $f_{CE}(p_i) \leq f_{CE}(p_j)$ and $f_{OE}(p_i) \leq f_{OE}(p_j)$ with at least one of these inequalities being a strict inequality. In other words, one solution dominates another if it is better in one objective and at least as good in the other objective. A solution is nondominated if neither objective can be

improved further without a worsening of the other. The set of nondominated solutions in Ω is referred to as the Pareto set, and the image of the Pareto set in the objective space Φ is referred to as the Pareto front. The Pareto front describes tradeoffs between objectives - for example, the increase in commission error f_{CE} associated with a given reduction of omission error f_{OE} , or vice versa. The goals of simultaneously minimizing omission and commission error by multi-objective optimization would be to obtain a diverse set of alternative models and maps, each expressing a near-optimal solution for a particular balance of omission/commission error or equivalently a particular level of bias, including zero bias.

We present a multi-objective optimization algorithm for presence-absence or binary classification problems. Our approach utilizes support vector machines (SVMs) optimized by a multi-objective genetic algorithm (GA). SVMs were originally developed for binary classification (Vapnik, 1995), although they have been widely applied to multiclass, single-class, and regression problems (Mountrakis et al., 2011; Salcedo-Sanz et al., 2014). As binary classifiers SVMs are capable of discriminating classes with complex, overlapping distributions within a high-dimensional feature space, producing decision boundaries that generalize well to new data (Brereton and Lloyd, 2010). Use of SVMs requires the specification of several free parameters that determine model fit, and optionally the identification of an optimal subset of predictor variables or training samples. SVMs are sensitive to parameter settings, which adds complexity to the model selection process because adequate performance cannot be assured under any pre-specified or default parameter values. Our approach exploits the complexity of SVM model selection to obtain a diverse set of solutions that tradeoff commission and omission error. From that set, specific models may be used to control map bias or to explore the impact of different levels of bias. In the sections that follow, we provide a description of our multi-objective support vector classification algorithm (MOSVC) and its implementation (section 3). We then demonstrate MOSVC within a forest remote sensing context by mapping tree species occurrence

and canopy disturbance in the temperate Atlantic Northern Forest of Maine, U.S.A. (sections 4 and 5). We first review relevant aspects of SVMs, GAs, and presence-absence model performance (section 2; readers familiar with these topics may wish to skip ahead).

3.3. Background

3.3.1. Support Vector Machines

Binary classification by SVMs is based on the concept of fitting a decision boundary between classes based on training samples, with no assumptions regarding their statistical distribution. Here we follow the conceptual framework provided by Brereton and Lloyd (2010) in their comprehensive review of SVMs. In the simplest case, training data are sampled from two classes whose distributions are separable by a linear margin within a multi-dimensional feature space, with some small subset of observations lying along margin boundaries. Different subsets of observations define different orientations of the linear margin between the two classes. SVMs identify the optimal decision boundary as the one that corresponds to a margin of maximum width. The samples that define the maximum margin, and hence the decision boundary, are called support vectors (SVs). The SVs are often few, and no other training samples have any bearing on the location of the boundary.

When classes cannot be separated using linear boundaries defined in the original feature space, SVMs fit decision boundaries by projecting the data into a new space of higher dimension in which linear separation is possible. Training data are mapped across higher dimensions using a kernel function defined within the original feature space, and a boundary is fit by linear margin maximization within the new space. When expressed in the original feature space, the boundary is nonlinear and potentially highly complex, but still defined by a set of SVs. The kernel function must satisfy particular criteria and in practice users typically specify a function from one of a few families. Each requires the specification of one or more free parameters. The Gaussian radial basis function (RBF) of a specified width (γ) is popular because it is typically very effective and requires specification of only one parameter (Brereton and

Lloyd, 2010). Narrow RBF kernels essentially allow for the projection of training data into extremely high dimensions with a large number of SVs. Narrow kernels therefore correspond to more complex boundaries.

To control complexity a regularization term is introduced into the boundary optimization problem that allows for the definition of a margin that includes certain training samples within its width. For classes with overlapping distributions, the margin expands to accommodate overlap, allowing for a simpler boundary between training samples and presumably less error when the model is generalized to predict new observations. Another free parameter, the penalty error term (C), determines the degree of regularization by stipulating a certain level of tolerance of training samples within the margin and potentially on the wrong side of the decision boundary (i.e., misclassified training samples). A high penalty error implies low tolerance, resulting in a narrow margin and complex boundary that fits the training data closely. A single penalty error term implies an equal misclassification cost amongst all training samples. When training data are heavily imbalanced, equal treatment of misclassified samples causes the decision boundary to migrate toward the minority class, biasing predictions in favor of the majority (He and Garcia, 2009; Tang et al., 2009). In the so-called cost-sensitive SVM (Tang et al., 2009), different penalty error values are applied to samples from different classes. A larger penalty error applied to the minority class reduces the tolerance of minority samples within the margin, pushing the decision boundary toward the majority class and offsetting the effect of class imbalance. The cost-sensitive SVM requires the specification of two penalty error terms, or equivalently a single penalty error term (C) and a weight ($C_w \geq 1$) applied to the penalty error for the minority class.

Optimal kernel functions and penalty error values are problem-specific, varying with the available feature set and training data. There is little basis for their selection apart from testing a very large number of possible combinations against validation data. Some form of data partitioning (e.g., cross-validation) is used to estimate the expected prediction error under different combinations of

parameter values (Brereton and Lloyd, 2010). Model parameterization is therefore equivalent to a search for an optimal combination of values from a multidimensional search space. The complexity of the problem is further increased when the search includes selection of an optimal subset of predictor variables, or an optimal subset of training samples. SVMs perform well given a large number of predictors since classification does not require estimation of class distributions (Bazi and Melgani, 2006), but selection of an optimal variable subset can improve results (Huang and Wang, 2006). Similar benefits may follow from selection of an optimal training sample (Blum and Langley, 1997). Parameter, variable, and sample selection should ideally be performed within a single search, and several classes of heuristic optimization or search algorithms are suitable, including ant colony optimization (e.g., Huang, 2009; Samadzadegan et al., 2012), particle swarm optimization (Li and Tan, 2010; e.g., Lin et al., 2008), and genetic algorithms (Bazi and Melgani, 2006; e.g., Friedrichs and Igel, 2005; Huang and Wang, 2006).

3.3.2. Genetic Algorithms

GAs are population-based optimization algorithms founded directly on the analogy of evolution by natural selection. A population of possible solutions is subjected to a selection pressure, leading to the evolution of traits associated with improved outcomes (Goldberg, 1989; Holland, 1975; Zäpfel et al., 2010). Considering optimization of a cost-sensitive SVM with a RBF kernel and no variable or training sample selection, individual solutions correspond to specific combinations of γ , C , and C_w . These three parameters are treated as analogous to genes, and a specific combination of parameter values equivalent to the genotype of an individual solution. The observable characteristics of a trained model or corresponding map constitute the phenotype of the individual solution, and these could include various measures of classification accuracy, model complexity, or map attributes. A GA designed to minimize classification error as a phenotypic trait applies a selection pressure to the population of solutions by favoring the perpetuation of genes that are associated with lower error. Over successive

generations or iterations of the GA, the population evolves to include models with progressively lower classification error.

At each iteration of the search process, a GA creates a new generation of solutions by recombining properties of existing solutions (Zäpfel et al., 2010). A certain proportion of individuals are selected as parents, with probability of selection determined by a specific fitness metric. A recombination operation is applied to pairs of parent solutions to generate members of the new generation. Recombination ensures inheritance of genes or genetic information from fit individuals, while introducing novelty to the next generation. The iterative recombination and replacement of solutions will generally lead to a loss of population diversity and a less comprehensive search for an optimal solution (Zäpfel et al., 2010). Random mutations of genes are interjected at each generation to promote diversity and encourage a more expansive search. Evolution proceeds until a specified criterion is met, for example, convergence of population traits such that additional iterations result in little further improvement, or execution of a specified number of generations. Upon termination the typical GA returns the individual with maximum fitness as the optimal solution.

By combining a guided search with a certain level of randomization, GAs are capable of obtaining globally optimal solutions from a large and complex search space (Goldberg, 1989). GAs have been successfully applied to a variety of SVM optimization problems, including parameter selection for both classification and regression problems (Friedrichs and Igel, 2005; Lorena and De Carvalho, 2008; Üstün et al., 2005; Wu et al., 2009), feature selection (Li et al., 2011), and simultaneous parameter and feature selection (Bazi and Melgani, 2006; Huang and Wang, 2006). The predominant approach is to identify an optimal or near-optimal solution according to a single model performance objective. However, GAs are very well suited to multi-objective optimization. As a population-based algorithm, they explore different portions of the search space simultaneously and from a single run they can provide a large and diverse set of solutions expressing tradeoffs between objectives (Konak et al., 2006).

Ghoggali et al. (2009) applied a multi-objective GA to a pair of semi-supervised image classification problems, and demonstrated that the simultaneous minimization of SVM classification error and number of SVs (as a measure of model complexity) produced better results than the minimization of either criterion alone. Suttorp and Igel (2006) described the use of a multi-objective GA to train SVMs for pedestrian detection from infrared images obtained by driver assistance systems. Control of classifier performance was attained through the simultaneous minimization of omission error, commission error, and number of SVs. We are unaware of any similar attempt to control classifier performance for mapping applications.

3.3.3 Model Performance Metrics

By comparing predicted class membership to observed class membership, omission and commission errors can be cross-tabulated in a confusion matrix (Table 3.1) from which various evaluation metrics can be calculated (Fielding and Bell, 1997; Stehman and Czaplewski, 1998). Different metrics reflect different aspects of model performance or map accuracy (Table 3.2). One of the simplest is overall accuracy or the proportion of samples predicted correctly (PPC). Sensitivity and specificity are the proportions of true positive and true negative cases that were correctly predicted by the model and hence correctly labeled by the map. Positive and negative predictive values (PPV and NPV) are the proportions of predicted positive and predicted negative cases that were truly positive and negative. Sensitivity and PPV are often referred to as the producer's and user's accuracy (PA and UA) for the positive class (Story and Congalton, 1986), reflecting the generic interests of a map producer to capture known locations of occurrence on the map, and of a map user to identify new locations of occurrence from the map. PA and UA are directly determined by class omission and commission error, and the balance between PA and UA corresponds to the degree to which class prevalence is either over- or underestimated. For example, high UA (low commission error) and low PA (high omission error) indicate that class prevalence is underestimated by the model and hence positive cases are systematically under-

represented on the map. Balanced PA and UA implies that the mapped class prevalence is unbiased relative to the reference data. Alternative measures of model performance have been devised to quantify tradeoffs between omission and commission error, for example by contrasting sensitivity and specificity (e.g., the G metric, Kubat and Matwin, 1997) or UA and PA (e.g., the F metric, van Rijsbergen, 1979) (Table 3.2).

Table 3.1. Confusion matrix for a binary classification or presence-absence model. We assume a probability sample design and express matrix entries as proportions of the sampled population (e.g., map pixels).

		Reference Class		Predicted Proportions
		Positive (+)	Negative (-)	
Predicted or Mapped Class	Positive (+)	p_{++}	p_{+-}	p_{+}
	Negative (-)	p_{-+}	p_{--}	p_{-}
	Reference Proportions	p_{+}	p_{-}	

Table 3.2. Model evaluation or map accuracy metrics for a binary classification or presence-absence model. See Table 3.1 for notation.

Metric	Calculation
Proportion Predicted Correctly (PPC)	$p_{++} + p_{--}$
Sensitivity	p_{++} / p_{+}
Specificity	p_{--} / p_{-}
Positive Predictive Value (PPV)	p_{++} / p_{+}
Negative Predictive Value (NPV)	p_{--} / p_{-}
Commission error ¹ or false positive proportion	p_{+-} / p_{+}
Omission error ¹ or false negative proportion	p_{-+} / p_{+}
Producer's accuracy (PA) ¹	$p_{++} / p_{+} = 1 - \text{omission error}$
User's accuracy (UA) ¹	$p_{--} / p_{-} = 1 - \text{commission error}$
G ²	$(\text{sensitivity} \cdot \text{specificity})^{1/2}$
F ^{1,3}	$(2 \cdot \text{UA} \cdot \text{PA}) / (\text{UA} + \text{PA})$

¹Calculated for the positive class

²Geometric mean of sensitivity and specificity (Kubat and Matwin, 1997)

³Harmonic mean of UA and PA (van Rijsbergen, 1979)

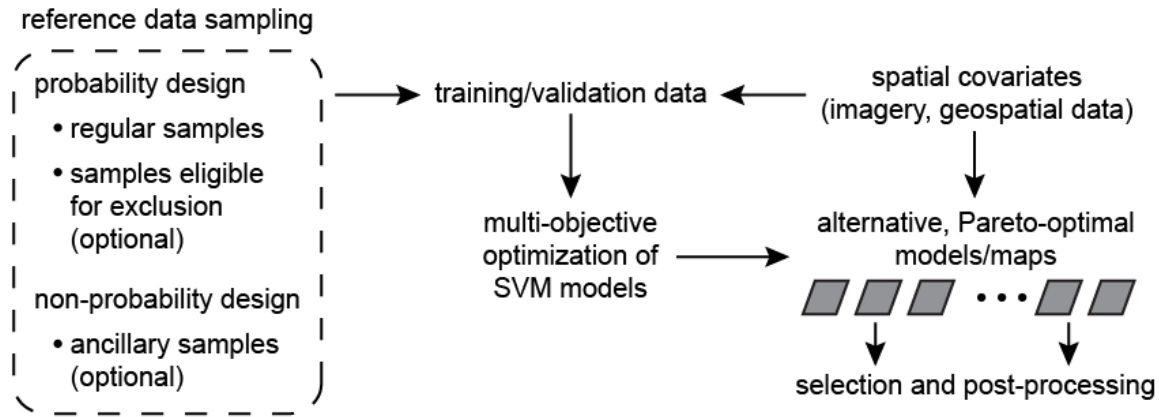


Figure 3.1. Multi-objective support vector classification algorithm overview. Classification is based on SVMs optimized using a multi-objective GA. Simultaneous optimization of multiple model performance metrics results in a set of maps that express different degrees of balance between performance objectives.

3.4. Algorithm Description and Implementation

3.4.1. Algorithm Overview

The MOSVC algorithm is based on optimization of SVMs using a multi-objective GA, and allows for flexible use of available reference data and spatial covariates (Fig. 3.1). Because control of classification bias requires consistent estimation of map error or accuracy metrics, we assume the use of reference data collected by a probability sampling design, where the probability of including any mapped location is known and nonzero. Complex probability sampling designs are accommodated through specification of sample inclusion probabilities (Cochran, 1977; Stehman and Czaplewski, 1998). Probability samples are partitioned for model training and validation by k-fold cross-validation (CV). All probability samples are used for model validation, but we allow a specified subset to be made eligible for exclusion from model training, either to provide additional flexibility for model fit or to reduce the potential influence of certain samples on model fit. We further allow for the use of ancillary training/validation data collected by a non-probability, haphazard, or unknown sampling design. Inclusion of ancillary samples may be beneficial, either by eliminating specific sources of error, improving results for a rare class that may be inadequately represented in the probability sample, or by

leveraging additional data collected for other purposes or by other means. Ancillary samples are included in model training but not model validation, ensuring consistent estimation of model performance metrics and map accuracy estimates. Spatial covariates may include both continuous and categorical variables, reformatted as numeric indicator or dummy variables. All continuous variables are scaled to unit range $([0,1])$ to prevent the disproportionate influence of those with larger numeric ranges. Variable or feature selection is employed to eliminate noisy or uninformative variables and to reduce computational complexity.

3.4.2. Algorithm Implementation

Optimization by GA requires the expression of individual SVM models in the form of a genotype subject to selection, genetic recombination, and mutation. We express individual models in the form of a bit string chromosome, composed of segments encoding parameter values, variable selection, and optional sample exclusion (Fig. 3.2). RBF kernel width (γ), penalty error (C), and penalty error weight (C_w) values are encoded as bit string segments whose lengths (coupled with user-specified minimum and maximum values) determine the precision with which each parameter is represented by binary encoding. Variable selection is encoded as a bit string segment with length equal to the number of available covariates, interpreted as a binary mask specifying selection of specific covariates. Optional sample exclusion is similarly encoded as a segment with length equal to the number of samples eligible for exclusion, indicating specific samples to be excluded from model training. The maximum number of samples permitted to be excluded in any individual chromosome is capped at a user specified percentage of those eligible for exclusion. The GA is initiated with a uniform random population of a user-specified size.

Numerous multi-objective GAs have been published and reviewed (Konak et al., 2006). Our approach is based on the popular NSGA-II algorithm (Deb et al., 2002) as implemented in the MATLAB Global Optimization Toolbox, Release 2014a (The MathWorks, Inc., Natick, Massachusetts, USA). We use

the LIBSVM open source software (Chang and Lin, 2011) for SVM training and prediction. The MATLAB implementation of NSGA-II acts as a wrapper for model induction using LIBSVM, and includes a customized fitness evaluation function. A diagrammatic representation of algorithmic details is provided in Fig. 3.3.

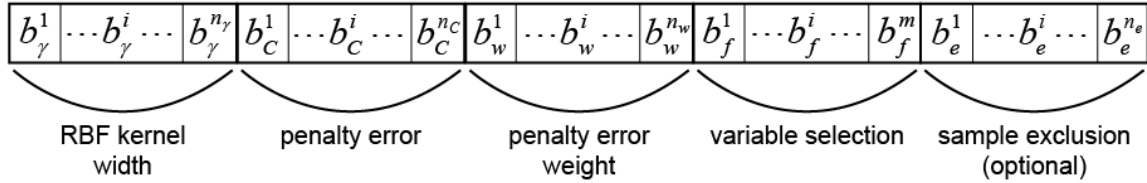


Figure 3.2. Genetic algorithm chromosome design. Bit string chromosomes are composed of segments encoding model parameter values (RBF kernel width, penalty error, and penalty error weight), predictor variable selection, and optional training sample exclusion.

At each iteration of the search, or for each generation of solutions, NSGA-II differentiates groups of parents (P) and offspring (Q) of equal size. Initially all individuals are random and specification of P_0 and Q_0 is arbitrary. The chromosome representing each member of the current population ($P_t \cup Q_t$) is decoded into real-valued SVM parameters, a variable selection mask, and a sample exclusion mask. The masks are used to extract variables from the original reference data block and to identify training and validation samples. Individual models are trained and validated by stratified k-fold CV, with strata defined by reference class label. Stratification ensures the uniform distribution of samples belonging to each class across folds for consistent evaluation of cost-sensitive SVMs. When available, ancillary samples are included in the training data for each of the k iterations of the CV procedure. Data scaling is applied at each CV iteration. The entire CV procedure is repeated a user-specified number of times and results averaged to reduce estimation uncertainty (Kim, 2009). CV estimates of model performance metrics are used to assign objective function values to each member of the current population.

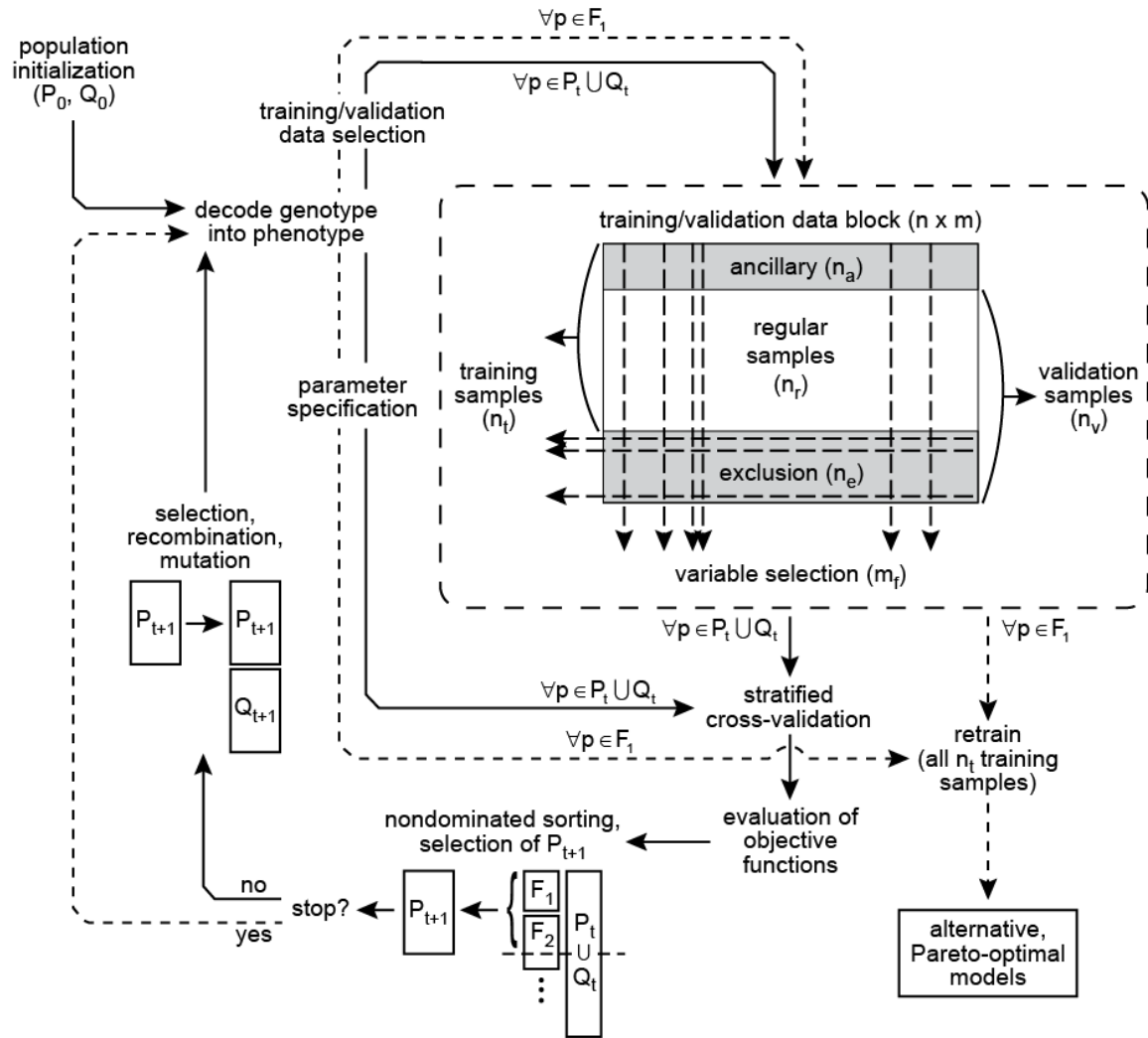


Figure 3.3. Multi-objective support vector classification algorithm implementation. Following selection of training and validation data, SVMs are fit and predictions made using the LIBSVM open-source software. Objective function values are estimated by cross-validation, and serve as the basis for population sorting, parent selection, and genetic operations embedded within the nondominated sorting genetic algorithm (NSGA-II).

In principle, simultaneous minimization of commission and omission error should produce a set of solutions expressing a full range of classification bias. In practice, users may wish to find solutions that are biased in only one direction, and for any given problem it may be easier to find near-optimal alternative solutions if bias is restricted to one direction. We therefore minimize either commission or omission error, and a measure of the extent to which class prevalence is biased either high or low. When minimizing commission error, models will tend to under-predict class prevalence, and we therefore pair

commission error with an objective function based on the ratio of reference to predicted class prevalence:

$$1) \quad f_1 = \text{commission error} = p_{+-} / p_{+}$$

$$2) \quad f_2 = |p_{+} / p_{+-} - 1| + 1$$

(following the notation used in Tables 3.1 and 3.2). For solutions that underestimate prevalence, the ratio p_{+} / p_{+-} is minimized by obtaining predicted prevalence approaching the reference prevalence, resulting in a ratio approaching 1. However, ratios <1 will occur for solutions within the GA population that overestimate class prevalence. We assume that solutions biased in the wrong direction may be valuable for the evolution of favorable population traits if their bias is relatively small. For solutions biased toward overestimation, Eq. 2 essentially reassigns an objective function value >1 , and we permit these solutions to persist in the population. By simultaneously minimizing Eqs. 1 and 2, we obtain a set of nondominated models with minimum commission error or maximum UA for different degrees of class under-representation. Alternatively, to obtain models expressing minimum omission error or maximum PA for different degrees of class over-representation, we simultaneously minimize omission error and an objective function based on the ratio of predicted to reference class prevalence:

$$3) \quad f_1 = \text{omission error} = p_{-+} / p_{+}$$

$$4) \quad f_2 = |p_{+-} / p_{+} - 1| + 1$$

We additionally constrain proportional over- or under-representation to a reasonable range to prevent an accumulation of undesirable solutions. A solution is considered undesirable if the predicted prevalence of either class differs from the reference prevalence by more than a factor of six (an arbitrary value found suitable for the problems we consider here).

Once objective functions have been evaluated for all members of the current population ($\forall p \in P_t \cup Q_t$), NSGA-II sorts solutions into a sequence of nondominated fronts (F_1, F_2, \dots). The first front F_1 includes all nondominated solutions from the total population. Once F_1 is obtained, these solutions

are removed from the population, and the next front F_2 is obtained as nondominated solutions from the reduced population. The process is iterated until all population members have been assigned to a front. NSGA-II subsequently identifies one half of the population as the next generation of parents (P_{t+1}), selecting solutions from successive fronts. The maximum number of parent solutions selected from F_1 is constrained to a user-specified proportion of the total population in order to promote population diversity throughout algorithm execution. Additional fronts are added to P_{t+1} in succession until one cannot be accommodated in its entirety. For the front that is only partially accommodated, solutions are selected from sparse or less crowded portions of the front to further promote population diversity.

The next generation of offspring (Q_{t+1} , equal in size to P_{t+1}) are obtained through genetic recombination and mutation of parent solutions (Fig. 3.4). A user-specified proportion of offspring are produced through genetic recombination of a pair of parent solutions, and the remainder through mutation of a single parent. Individual parents are identified by tournament selection (Zäpfel et al., 2010), where a user-specified number of solutions are randomly selected from P_{t+1} and the best is selected as a parent. Better solutions lie on lower ranked fronts and in less crowded regions along their front. Genetic recombination may occur through one of several crossover operations (Fig. 3.4a) in which an offspring is constructed from one or more bit string segments copied from each parent. Different recombination operations determine the manner in which information is exchanged and the potential degree of novelty introduced through exchange (Zäpfel et al., 2010). An offspring produced by mutation is a copy of its parent subjected to a mutation operation that switches individual bit values with a user-specified probability (Fig. 3.4b). Once offspring have been produced, parent and offspring chromosomes ($P_{t+1} \cup Q_{t+1}$) are decoded and the process repeats.

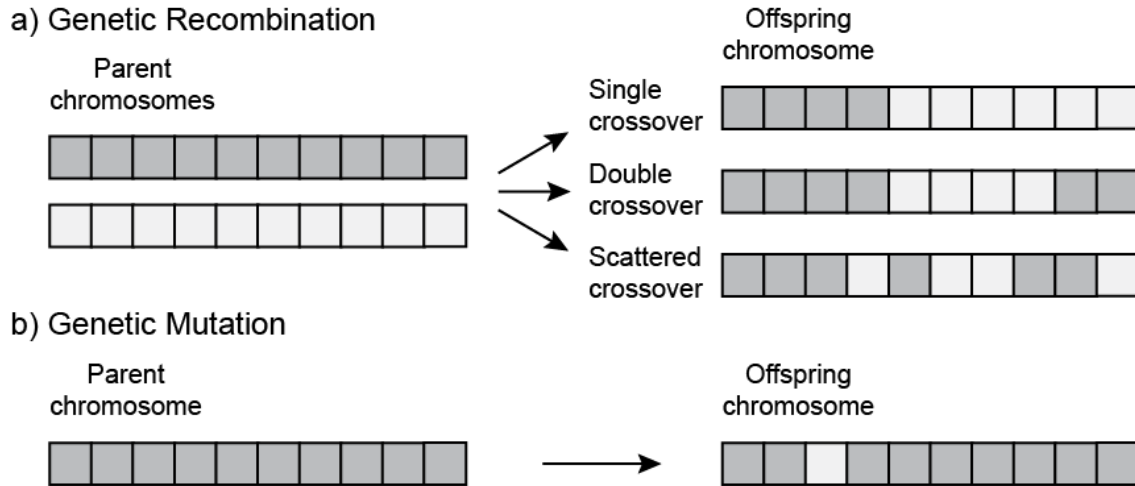


Figure 3.4. Genetic recombination and mutation operations. Graphical depiction of a) genetic recombination and b) genetic mutation operations within MOSVC.

Stopping criteria are evaluated at each generation after solutions are sorted into nondominated fronts. The algorithm is assumed to have converged to a close approximation of the Pareto front when the change in spread of solutions along F_1 averaged over a user-specified number of generations is less than a user-specified threshold. Alternatively, the algorithm stops when the generation count exceeds a user-specified maximum. Once stopped, members of F_1 are retrained using all available training samples and returned as a set of alternative solutions expressing tradeoffs between accuracy and bias objectives.

3.5. Example Applications: Study Area and Methods

3.5.1. Study Area

We demonstrate MOSVC by mapping tree species occurrence and canopy disturbance in the temperate Atlantic Northern Forest of Maine, U.S.A. The Northern Forest of the northeastern U.S. encompasses roughly 11 Mha within a transition zone between the northern boreal forest and the southern temperate deciduous-dominant forest (Likens and Franklin, 2009), including ~4 Mha of nearly contiguous, undeveloped forestland across northern and western Maine. Tree species diversity is relatively high as the northern limit of southern species overlaps with the southern limit of northern species (Nightingale et al., 2008). Our ~1.9 Mha study region (Fig. 3.5) was defined by the overlap of

Landsat images used to map species occurrence and canopy disturbance in northwestern Maine.

Topography is generally flat or rolling with occasional low mountains and an extensive network of rivers, lakes, and wetlands. Roughly 90% of forestland is privately owned, and private lands are predominantly managed for commercial production.

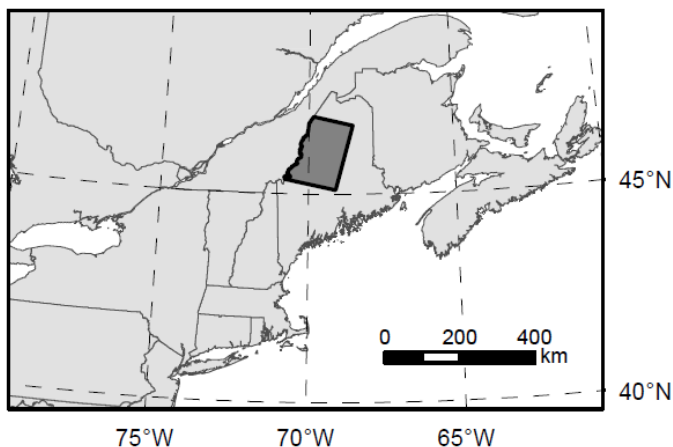


Figure 3.5. Study area. Northern Maine, U.S.A. study area encompassing ~1.9 Mha of forestland. State and provincial boundaries obtained from the National Atlas of the U.S. (Political Boundaries) and the Atlas of Canada (National Frameworks Data, Census Subdivisions and Population Ecumene).

3.5.2. Tree Species Occurrence

For the purposes of demonstrating and evaluating MOSVC, we mapped the occurrence of black ash (*Fraxinus nigra*), eastern white pine (*Pinus strobus*), and red spruce (*Picea rubens*). Black ash is a key cultural and economic resource of Maine's Native American communities because its wood is uniquely suited to basket weaving. Black ash is threatened by the invasive emerald ash borer (*Agrilus planipennis*) (Herms and McCullough, 2014; Ranco et al., 2012) and maps of its distribution are needed to better evaluate the existing resource and to plan response. Eastern white pine and red spruce are economically important timber species and the primary and alternate hosts of the pine leaf adelgid (*Pineus pinifoliae*), a native insect whose life cycle depends on intergenerational migration between primary and alternate host trees. Pine adelgid causes only minor damage to spruce but can kill up to 100% of new pine shoots during heavy infestations, leading to significant growth reduction and mortality (Balch and Underwood,

1950; Dimond and Bishop, 1968). For the first time since the early 1960s, a significant outbreak is spreading within northern Maine, with observable damage centered in our study area (Currier et al., 2015). Maps of pine and spruce distributions are needed to direct field assessments of damage, evaluate landscape risk factors, and forecast outbreak development.

3.5.2.1. Reference Data and Spatial Covariates

Predictions of species occurrence are based on reference data provided by the USDA Forest Service Forest Inventory and Analysis (FIA) Program. The FIA Program provides quality-assured measurements of forest attributes from a national network of field plots adhering to an equal-probability sampling design. The contemporary design is based on a hexagonal tessellation, with one plot randomly located within each 2428 ha tile (McRoberts et al., 2005). The FIA program maintains the confidentiality of true plot locations to protect the privacy of landowners and to preserve plot integrity (Smith, 2002). True locations were made available for use through a collaborative agreement with the USFS Northern Research Station FIA Program.

Spatial covariates included multispectral imagery, terrain attributes, and climate surfaces (details provided in Appendix C). We sought to obtain Landsat Thematic Mapper (TM) or Enhanced Thematic Mapper Plus (ETM+) images acquired at different times throughout the growing season in order to exploit species-specific foliar phenology. Frequent and extensive cloud cover necessitated the collection of imagery across multiple years. We selected imagery from the early 2000s when Landsat 5 and Landsat 7 were both fully operational. We obtained eight nearly cloud- and snow-free images acquired between late April and early October, 2001-2006. Since 1999, Maine FIA data have been collected in rolling 5-year inventory cycles, with 20% of plots surveyed annually. Due to low prevalence of ash and pine within our study area, we used FIA observations collected over a full inventory cycle (i.e., 2002-2006) coincident with the acquisition of selected Landsat images. Additional spatial covariates included climate and terrain attributes thought to be relevant to tree establishment or growth. Terrain

data included 10 morphometry, 8 lighting/visibility, and 11 hydrology variables calculated from the 1 arc-second (30 m) NED and the National Hydrography Dataset (NHD). Climate data included 17 variables mapped at approximately 1 km spatial resolution and representing the 1961-1990 climate normal period, obtained from the USDA Forest Service Rocky Mountain Research Station, Moscow Forestry Sciences Laboratory.

Covariate values were extracted at 712 forested FIA plots lying within our study area. Under the modern FIA inventory design, field plots consist of a center subplot with three satellite subplots (McRoberts et al., 2005), sampling an area loosely equivalent to a 3x3 neighborhood of 30 m pixels (Cooke, 2000). Landsat and terrain predictor data were compiled by averaging pixels within 3x3 neighborhoods surrounding plot centers; climate predictor data were extracted as 1 km pixel values. For the purposes of demonstrating and evaluating our classification method, we excluded locations where pixel neighborhoods were impacted by apparent forest cover change during the 5-year observation period (299 samples). We also masked cloud and snow cover from affected images, causing loss of data for certain acquisition dates at some plot locations. SVMs are generally incapable of working with incomplete predictor data, and rather than incorporate an additional data imputation algorithm, we elected to exclude reference locations with missing data (64 samples). Remaining plot locations yielded a training/validation data set consisting of 349 samples.

For black ash, FIA tree measurements were used to produce a binary response variable indicating presence/absence. For both pine and spruce, we used tree measurement data to calculate relative abundance as a proportion of estimated live aboveground biomass (of stems >2.54 cm diameter measured at 1.37 m) and produced binary response variables indicating presence at 5% or greater relative abundance (an arbitrary threshold imposed to exclude low levels of abundance for the study of pine adelgid dynamics).

3.5.2.2. Algorithm Execution

For each species occurrence problem, we present two sets of multi-objective optimization outcomes: 1) simultaneous minimization of Eqs. 1 and 2 (maximization of UA with minimal class under-representation), and 2) simultaneous minimization of Eqs. 3 and 4 (maximization of PA with minimal class over-representation). Each optimization problem was repeated 5 times to demonstrate stochasticity in MOSVC outcomes. We compare multi-objective outcomes with the results of single-objective optimization using either PPC, F, or G (Table 3.2) as the objective function, each repeated 5 times using a single-objective GA (MATLAB Global Optimization Toolbox, Release 2014a).

All optimization problems included parameter and feature selection. Parameter values were constrained within reasonable ranges ($\log(\gamma) \in [-4,1]$; $\log(C) \in [-1,3]$; $\log(C_w) \in [1,2]$). For pine and red spruce we executed additional MOSVC runs with certain reference samples eligible for exclusion during model training. For the maximization of UA with minimal class under-representation (Eqs. 1 and 2), we allowed any positive case (relative abundance $\geq 5\%$) to be excluded from model training under the assumption that this would provide greater flexibility in the search for models that under-predict prevalence. For the maximization of PA with minimal class over-representation (Eqs. 3 and 4), we assumed that exclusion of certain negative cases would similarly provide greater flexibility in the search for models that over-predict prevalence. We therefore allowed cases with relative abundance $< 5\%$ but > 0 to be eligible for exclusion under the reasoning that these specific cases may be most restrictive of predicted distributions. All reference samples were used for model validation in a 10-fold, 10 times repeated CV. The GA operated on a population of 500 chromosomes, with a maximum of 20% maintained on the approximate Pareto front. Parent chromosomes were selected by tournament with 10 participants. 70% of offspring were generated by scattered crossover of parent chromosomes; 30% were generated by mutation, with a mutation rate of 2.5%. Scattered crossover and a relatively high mutation rate promoted population diversity and prevented early convergence. Additionally, we

specified convergence criteria that ensured execution of >100 generations, up to a maximum of 300.

Single-objective optimization problems were executed using the same values for applicable settings, but returned a single solution rather than a set of Pareto-optimal solutions.

3.5.3. Forest Canopy Disturbance

Within the commercial forests of northern Maine, logging represents the most prevalent and widespread form of disturbance, with timber predominantly removed by mechanized partial harvesting. Clearcutting accounts for less than 5% of annual harvest area (Maine Forest Service, 2014, 2005). Within a typical partial harvest, trees are removed within and adjacent to machine trails, with light or negligible removal between trails. Although partial harvests of low to moderate intensity maintain quasi-continuous canopy cover, contemporary rates of partial harvesting cause rapid loss and fragmentation of intact mature forest, accumulation of small regenerating forest patches, and a corresponding increase in edge-affected forest area (Legaard et al., 2015). For wildlife species or forest values that are potentially sensitive to the extent or configuration of either intact mature or early successional forest conditions, management may require knowledge of past disturbance trends or regular monitoring of ongoing harvest activity. Here we demonstrate MOSVC within a multispectral change detection context by mapping forest canopy disturbance.

3.5.3.1. Reference Data and Spatial Covariates

Landsat TM images acquired during the summers of 2004 and 2007 were selected for disturbance mapping (details provided in Appendix C). A 3-year interval between relatively cloud-free, leaf-on Landsat images is not uncommon for this region (Legaard et al., 2015), contributing additional difficulty to disturbance detection due to post-disturbance vegetation growth within the observation interval (Jin and Sader, 2005). Patterns of cloud cover within these specific images afforded an opportunity to demonstrate the response of MOSVC to a common source of error in forest change detection using multispectral imagery. Cloud cover in the post-disturbance image and cloud shadow in

the pre-disturbance image introduced patterns of spectral change similar to those caused by canopy disturbance. Because rates of canopy disturbance are typically low, even relatively small amounts of cloud- or shadow-induced error can introduce an intolerably large amount of uncertainty and bias in a forest disturbance map (Huang et al., 2010). Automated masking is difficult due to the highly variable spectral characteristics of cloud- and shadow-affected pixels (Huang et al., 2010; Zhu and Woodcock, 2012), and cloud cover remains a problem for change detection applications. Small cumulus clouds and cloud shadows were sparsely scattered within the 2004 image. Stratocumulus clouds were present over a much larger portion of the 2007 image, with cloud conditions ranging from entirely opaque to largely transparent. To evaluate the performance of MOSVC for change detection under cloudy conditions, we implemented a reference sampling procedure with the intent of adequately representing 2007 cloud cover. The much less prevalent 2004 cloud/shadow conditions were not adequately represented, and we instead incorporated an ancillary, non-probability sample to reduce shadow-induced error.

Reference data were compiled as a stratified random sample with image strata defined by multispectral change patterns revealed by the iteratively-reweighted multivariate alteration detection transformation (IR-MAD; Canty and Nielsen, 2008). A threshold was applied to the IR-MAD component that best represented forest disturbance patches to define a disturbance stratum containing 19.7% of forest pixels. Most were not impacted by any discrete disturbance event, but nonetheless displayed broadly similar patterns of spectral change (including 2007 cloud cover). A random sample of 500 training/validation locations were selected from the disturbance stratum. All forest pixels not included in the disturbance stratum (including 2004 cloud shadow) were allocated to a second stratum from which 250 locations were randomly selected. Inclusion probabilities were calculated as the ratio of stratum sample size to total stratum size. To quantify the influence of cloud and cloud shadow on predictive accuracy, we drew an additional random sample of 100 reference locations from cloud/shadow masks generated for each image.

Reference labels were assigned by visual interpretation of individual sample pixels in the Landsat image pair (Cohen et al., 2010, 1998; Legaard et al., 2015; Sader and Legaard, 2008). A trained image interpreter recorded the occurrence of a canopy disturbance at reference locations provided spectral changes, image texture, and other contextual cues were consistent with either harvesting or natural canopy disturbance. Natural disturbance at the pixel scale was extremely rare, and although harvesting was generally readily apparent, pixel-level identification of disturbance could be difficult due to prior disturbance, regrowth during the 3-year observation interval, or heterogeneous forest conditions. The interpreter indicated reference locations for which confidence was low, and these were made eligible for exclusion from model training. Finally, reference locations affected by cloud or shadow were labeled as disturbed if disturbance was visually discernible. Locations were labeled as undisturbed if opaque cloud or dark shadow obscured the state of the forest canopy. We expected MOSVC to resolve disturbance if discernible to the interpreter, regardless of cloud/shadow conditions.

To model and map pixel-level disturbance we used both individual pixel values and 3x3 neighborhood statistics as predictor variables. Use of neighborhood information reduced fine-scale spatial variability in disturbance maps and lessened the need for subsequent spatial filtering, but at the cost of a larger number of predictor variables. From both 2004 and 2007, raw digital numbers for TM bands 1-7 were obtained for individual sample pixels and summarized within 3x3 neighborhoods. Summary statistics included the mean, median, standard deviation, and range of forest pixels, yielding a total of 70 predictor variables.

3.5.3.2. Algorithm Execution

For the forest disturbance problem, we present a single set of multi-objective optimization outcomes resulting from the simultaneous minimization of Eqs. 1 and 2, equivalent the maximization of UA with minimal class under-representation. For the purposes of evaluating patterns of forest disturbance (e.g., harvest patch characteristics), under-representation may be preferable to over-

representation because interannual variability in forest phenology or canopy condition tends to introduce undesirable patterns of commission error as predicted class prevalence increases. Although the characterization of error patterns is itself of interest, for the purpose of demonstrating use of MOSVC we present only the single set of results, repeated 5 times.

Optimization included parameter selection, feature selection, and exclusion of low-confidence samples. We used a 10 times repeated 5-fold CV and the same GA settings as those used for species occurrence modeling. 2004 cloud shadow was in fact consistently mapped as forest disturbance, due to its lack of representation in the training/validation sample. From a rapid visual inspection, we identified a set of 30 ancillary reference locations situated in false change patches caused by 2004 shadow. Ancillary data comprised a non-probability, purely haphazard sample obtained to reduce shadow-induced error. Ancillary data were included in another set of 5 MOSVC runs, but with GA populations initialized using the final populations from the prior set of runs. The MOSVC algorithm was iterated over an additional 100 generations and outcomes were compared with prior runs.

3.6. Example Applications: Results

3.6.1. Tree Species Occurrence

The results of a single black ash run demonstrate relationships between model performance metrics for each type of optimization problem (Fig. 3.6). Model performance estimates described a negative relationship between UA and class prevalence (the UA Pareto front, Fig. 3.6a). Higher UA estimates were attained by reducing the predicted prevalence of the positive class to include only those areas where true positive cases were predictable with higher accuracy. The UA front included solutions whose estimated performance varied from ~60% UA at 100% of reference prevalence to 100% UA at ~40% of reference prevalence. The former was an unbiased outcome, with UA equal to PA and predicted prevalence equal to reference prevalence; the latter was a heavily biased outcome, with high UA and low PA indicating that the corresponding map would identify a subset of true ash locations with high

accuracy. Performance estimates described a positive relationship between PA and class prevalence (the PA Pareto front, Fig. 3.6b), with higher PA attained by increasing predicted prevalence to include more true positive cases. The PA front included solutions that varied from ~65% PA at 100% reference prevalence to 100% PA at ~250% reference prevalence. The latter was a biased outcome where all true positive cases were predicted correctly, but at the expense of greater commission error and a more than

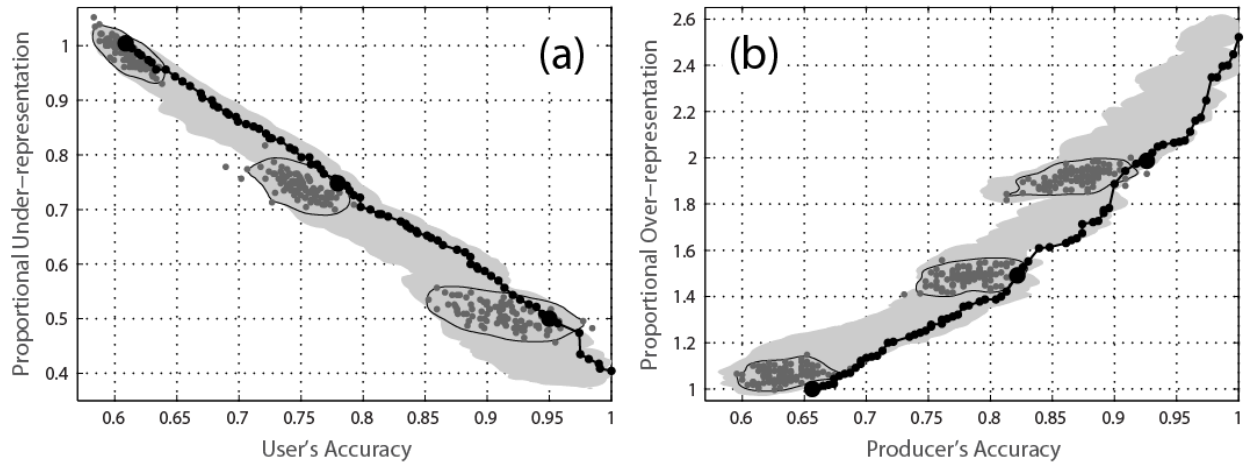


Figure 3.6. Black ash occurrence - model performance and estimation uncertainty. MOSVC outcomes and associated uncertainty (shaded) expressing tradeoffs between a) UA and under-representation of ash occurrence as a proportion of reference prevalence, and b) PA and over-representation of ash occurrence. Also shown for selected models are replicate CV estimates used to derive uncertainty envelopes by kernel density estimation. doubling of class area.

Model optimization by MOSVC is based on CV estimates of objective function values that are subject to uncertainty, particularly when training/validation samples are few or classes are heavily imbalanced (e.g., 6.6% of reference plots contained black ash). To evaluate CV estimation uncertainty, we replicated the stratified, repeated k-fold CV procedure 100 times using different random partitions. We used kernel density estimation to summarize CV variability for each MOSVC model, applying a Gaussian kernel to replicated CV estimates with optimal bandwidth provided by Bowman and Azzalini (1997, page 37). We iteratively evaluated threshold density values to obtain a 90% volume contour from

the density estimates, representing approximately 90% of the estimation uncertainty introduced by random partitioning of training/validation data.

For black ash, the 90% contour for any individual model typically spanned about 10% UA or PA, and about 10-20% of class prevalence. The superposition of 90% contours (Fig. 3.6, shaded) demonstrated aggregate uncertainty around MOSVC outcomes, primarily caused by the distribution of a small number of positive cases across CV folds. UA and PA fronts were situated near the more favorable edge of the uncertainty envelope defined by 90% contours because UA and PA estimates returned by MOSVC were biased high. Use of CV estimates for optimization presumably caused over-estimation of UA and PA because the algorithm favored models with high estimated accuracy under the specific CV partitioning used during model training/validation. Prevalence estimates were on average nearly equal to mean CV estimates obtained by repartitioning. The magnitude of CV uncertainty reflects reference data availability, and for black ash uncertainty was relatively high, as was the tendency to over-estimate class accuracy. For two solutions lying near one another on either front, CV uncertainty implies that performance estimates may not reflect true differences in mapped distributions. However, solutions that differ in estimated performance by an amount that exceeds apparent levels of CV uncertainty will produce maps with substantively different error characteristics.

In our study area black ash is thought to be most commonly distributed within deciduous and mixed forested wetland communities occupying basins, drainage bottoms, or groundwater seepage sites along gentle slopes (Maine Forest Service, 2008; Maine Natural Areas Program, n.d.). A map of black ash with predicted prevalence equal to reference prevalence generally conformed to these patterns, even though estimated class accuracy was only 61% (Fig. 3.7a). In comparison, a map with higher UA could be more desirable for certain applications, including the identification of areas with trees suitable for basket making, because their occurrence could be verified with greater efficiency. A model with predicted prevalence equal to one half of reference prevalence had an estimated UA of 95% (Fig. 3.6a).

In the corresponding map (Fig. 3.7b), ash occurrence was generally more consolidated within a smaller number of patches.

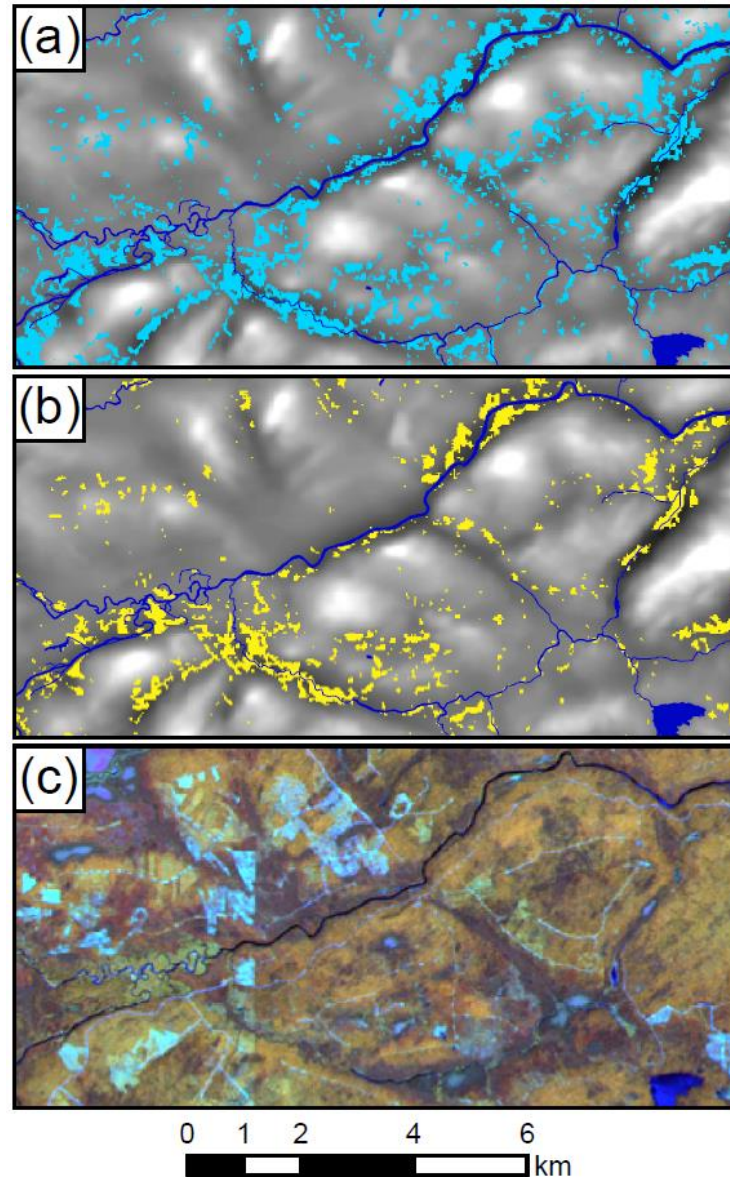


Figure 3.7. Black ash occurrence maps. Predicted occurrence of black ash from MOSVC models with performance estimates of a) 61% UA and predicted prevalence equal to reference prevalence, and b) 95% UA and predicted prevalence equal to one half reference prevalence. Predicted occurrence is superimposed over the topographic position index (1 km neighborhood radius; Guisan et al., 1999), with lighter areas indicating upper slopes and hilltops and darker areas indicating lower slopes and drainage bottoms. Hydrographic features were obtained from the National Hydrography Dataset. c) Landsat TM image acquired on June 10, 2004 (bands 4, 5, and 3 shown in red, green, and blue).

For black ash, five MOSVC runs produced parallel UA fronts (Fig. 3.8a). Differences in estimated performance were modest across all levels of performance, although aggregate uncertainty was elevated. PA fronts displayed greater variability (Fig. 3.8b). MOSVC produced unbiased models in all runs, but PA fronts were fully extended in only two runs, and 100% PA was achieved at very different levels of prevalence. Run-to-run variability was greatest at high PA but pronounced at all levels of performance, with the estimated accuracy of unbiased models ranging from 64-76%. Greater variability between fronts suggests that the simultaneous optimization of PA and prevalence was a more difficult problem, leading to more variable solutions.

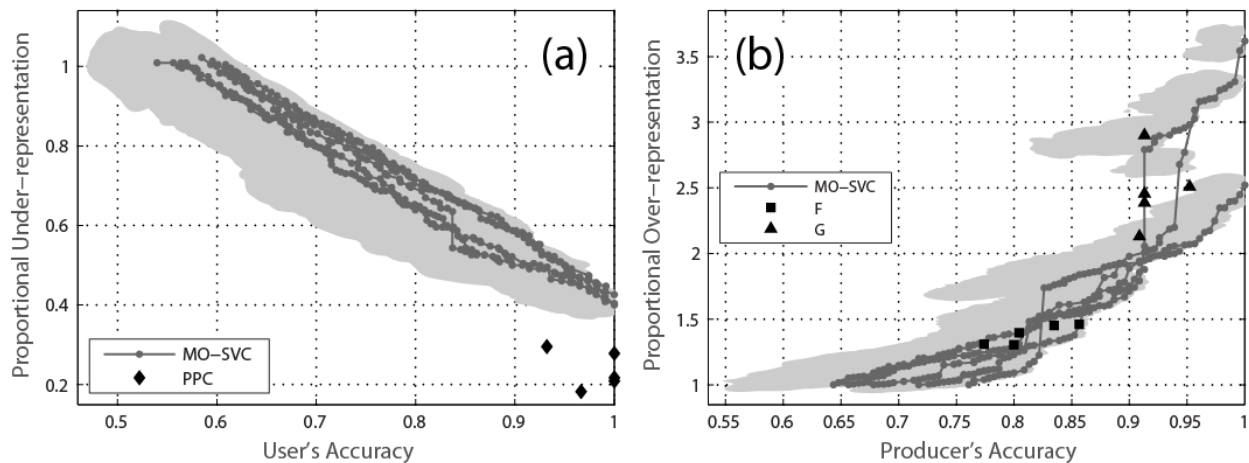


Figure 3.8. Black ash occurrence - replicate runs compared to single-objective optimization outcomes. MOSVC outcomes from five replicate runs, and associated uncertainty (shaded), expressing tradeoffs between a) UA and under-representation of ash occurrence as a proportion of reference prevalence, and b) PA and over-representation of ash occurrence. Also shown are performance estimates of models obtained by single-objective optimization.

We compared MOSVC outcomes to performance estimates of models derived from single-objective optimization of different model training criteria. Single-objective outcomes were plotted alongside either UA or PA fronts depending on whether they corresponded to a proportional under- or over-representation of prevalence (single-objective uncertainty envelopes not shown for visual clarity). Use of PPC as a training criterion resulted in gross under-representation of black ash prevalence (18-30% of reference prevalence) (Fig. 3.8a). The G metric produced solutions with very high predicted class

prevalence (213-290% of reference prevalence) (Fig. 3.8b). The F metric produced less biased outcomes although with a substantial over-representation of class prevalence nonetheless (130-146% of reference prevalence). Optimization by either F or G (but not PPC) produced solutions that were comparable to a subset of MOSVC solutions.

MOSVC outcomes followed different patterns for eastern white pine. UA fronts from runs employing parameter and feature selection, but no sample exclusion, included unbiased solutions but were truncated at <85% UA (Fig. 3.9a). UA fronts therefore offered limited representations of tradeoffs between accuracy and bias. CV uncertainty around UA fronts was comparable to that of black ash, presumably because positive cases were similarly rare (13.8% of reference plots). PA fronts however, were less variable with reduced uncertainty around CV estimates of class accuracy (Fig. 3.9b). PA fronts followed similar trajectories from about 70% PA at 100% reference prevalence to 100% PA at about 220% reference prevalence. The results of single-objective optimization were qualitatively similar to those for black ash, with PPC and G resulting in under- and over-estimation of prevalence. Use of F generally resulted in over-estimation (Fig. 3.9b), but also produced an unbiased model in one run (Fig. 3.9a). With positive cases eligible for exclusion from model training, MOSVC produced UA fronts ranging from ~70-75% UA at 100% of reference prevalence to 100% UA at ~40-55% of reference prevalence (Fig. 3.9c). UA fronts were no longer truncated, estimated performance was improved (i.e., greater UA at any given prevalence), and CV variability was generally reduced. The effects of sample exclusion on PA fronts were less pronounced, presumably because only 19 of 301 negative cases were made eligible for exclusion (those with nonzero abundance).

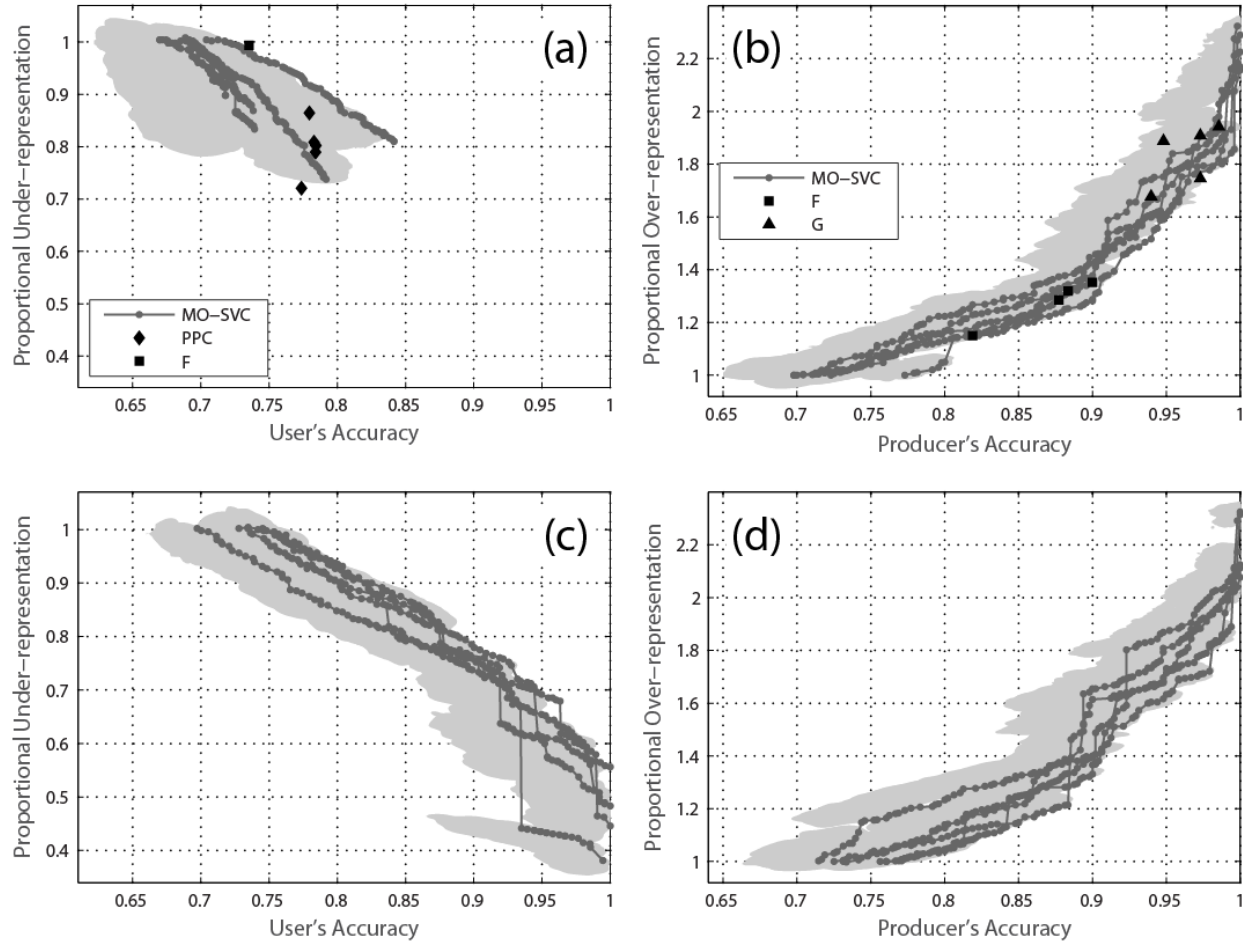


Figure 3.9. Eastern white pine occurrence - replicate runs compared to single-objective optimization outcomes. MOSVC outcomes from five replicate runs, and associated uncertainty (shaded), expressing tradeoffs between a) UA and under-representation of pine occurrence as a proportion of reference prevalence, and b) PA and over-representation of pine occurrence, obtained by parameter and feature selection but no sample exclusion. Also shown are performance estimates of models obtained by single-objective optimization. The five replicate runs shown in c) and d) were obtained as before but with a subset of samples made eligible for exclusion from model training.

Red spruce presented a different problem than either pine or black ash, with a reference prevalence of 59% and a reasonable balance between positive and negative cases. Run-to-run variability was lower and CV estimation uncertainty was greatly reduced (i.e., 90% contours for individual solutions typically spanned <2% predicted accuracy or prevalence) (Fig. 3.10). In the absence of sample exclusion, UA and PA fronts included unbiased models with estimated class accuracies of ~79-84%. UA fronts extended to 100% accuracy at roughly 30-50% of reference prevalence (Fig. 3.10a). Two of the five PA

fronts were truncated at ~85% accuracy, but three extended to nearly 100% accuracy at slightly less than 160% of reference prevalence (the limit imposed on feasible solutions based on the predicted prevalence of negative cases) (Fig. 3.10b). Patterns of single-objective outcomes differed from those of pine or ash. Use of PPC resulted in solutions with little bias (103-105% of reference prevalence) (Fig. 3.10b). Use of G tended to as well (Fig. 3.10b), although one run produced a substantial underestimation of prevalence (84% of reference prevalence) (Fig. 3.10a). F produced results with a consistent and comparatively strong negative bias (75-80% of reference prevalence) (Fig. 3.10a). Use of sample

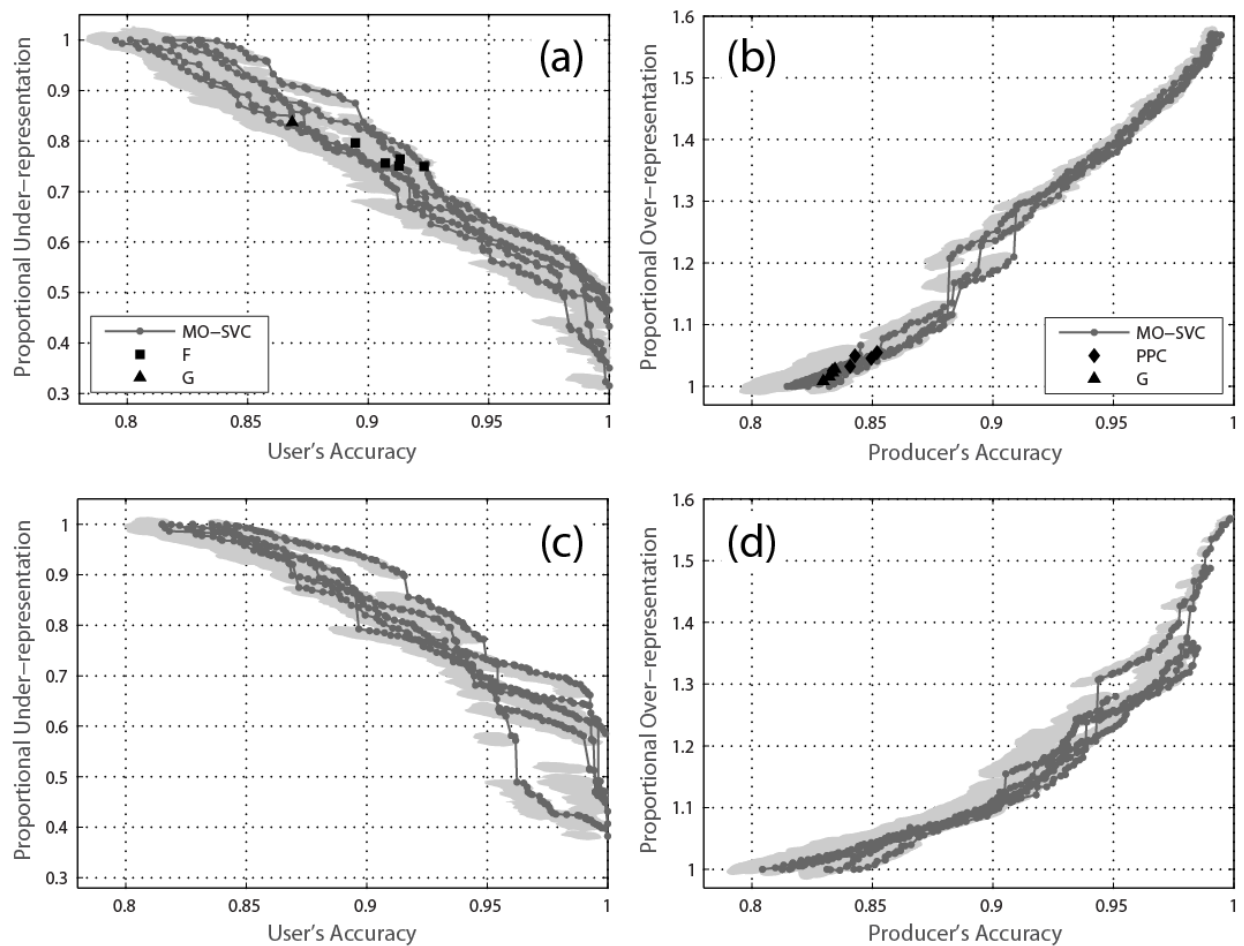


Figure 3.10. Red spruce occurrence - replicate runs compared to single-objective optimization outcomes. MOSVC outcomes from five replicate runs, and associated uncertainty (shaded), expressing tradeoffs between a) UA and under-representation of red spruce occurrence as a proportion of reference prevalence, and b) PA and over-representation of red spruce occurrence, obtained by parameter and feature selection but no sample exclusion. Also shown are performance estimates of models obtained by single-objective optimization. The five replicate runs shown in c) and d) were obtained as before but with a subset of samples made eligible for exclusion from model training.

exclusion improved MOSVC outcomes, with both UA and PA fronts consistently shifted toward more favorable solutions, although the added complexity of the optimization problem did result in somewhat greater run-to-run variability (Figs. 3.10c and 3.10d).

3.6.2. Forest Canopy Disturbance

Reference data for disturbance models were obtained as a stratified random sample. After calculating inclusion probabilities by stratum and weighting observations accordingly, we obtained a reference prevalence of 7.6%. MOSVC produced unbiased solutions with estimated disturbance class accuracies of 90-92% (Fig.11). UA fronts displayed a very consistent relationship between UA and predicted prevalence to about 99% UA and 75-80% of reference prevalence. CV uncertainty was relatively low, with 90% contours typically spanning 1-2% UA and 2-4% prevalence. CV estimates of UA were biased high, but typically by less than 1%. Execution of MOSVC included sample exclusion, with 70 low-confidence or ambiguous interpretations made eligible for exclusion. Sample exclusion had a minor (though consistently positive) effect on model performance estimates.

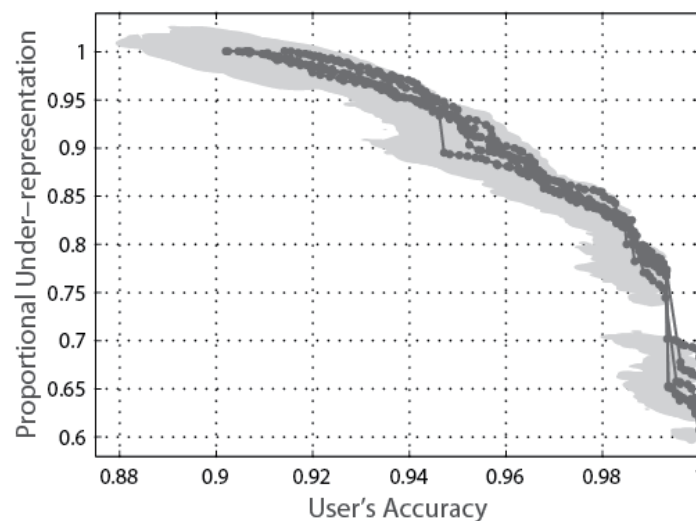


Figure 3.11. Forest canopy disturbance - replicate runs. MOSVC outcomes from five replicate runs, and associated uncertainty (shaded), expressing tradeoffs between UA and under-representation of canopy disturbance as a proportion of reference prevalence.

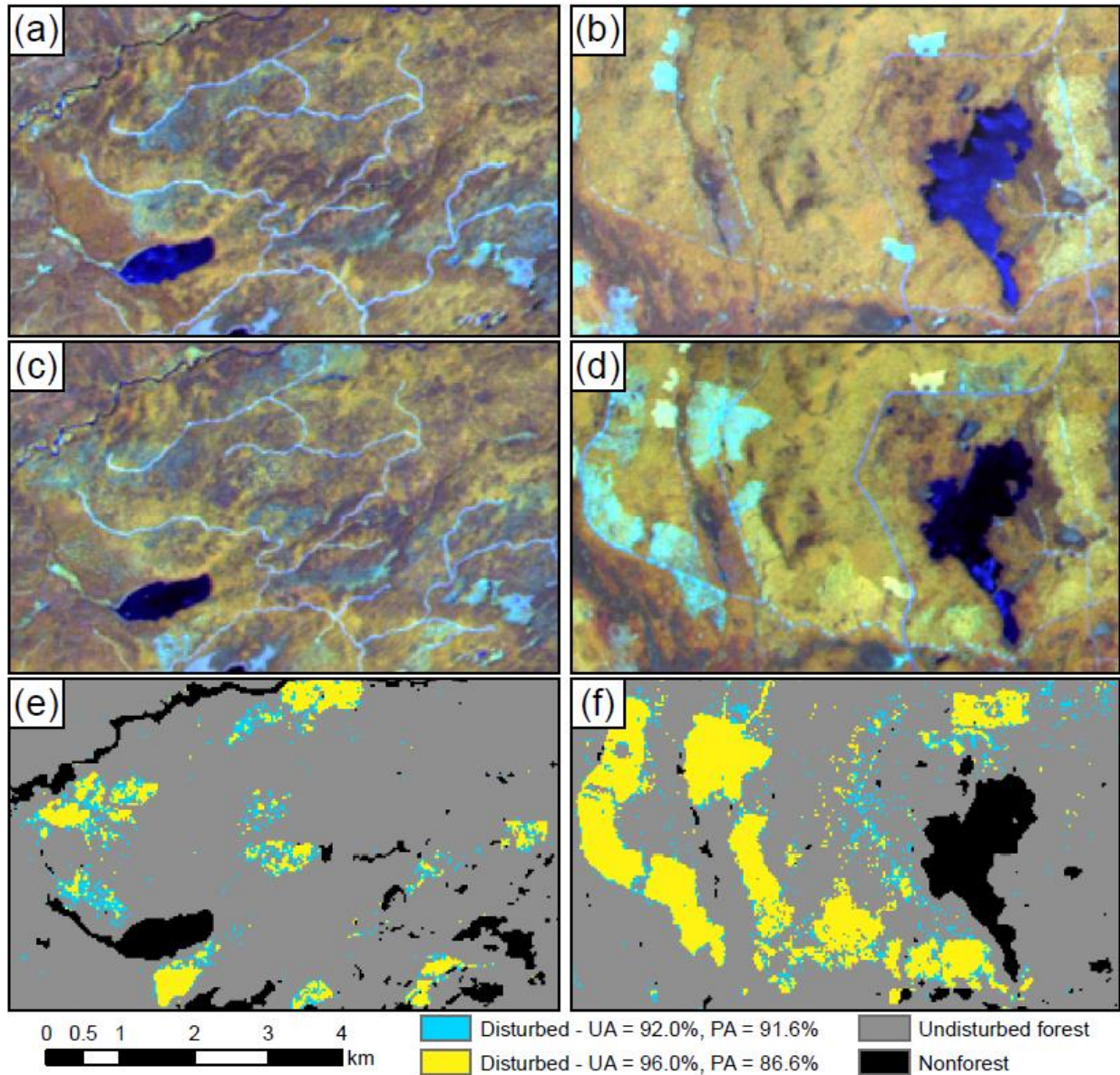


Figure 3.12. Forest disturbance maps. Predicted distributions of forest canopy disturbance for two example areas subjected to different patterns and intensities of disturbance. a-b) 2004 Landsat TM image (bands 4, 5, and 3 shown in red, green, and blue). c-d) 2007 TM image. e-f) Disturbance predicted from two MOSVC models with different patterns of error and different levels of predicted class prevalence. At an estimated 92% UA, predicted prevalence was nearly equal to reference prevalence (cyan); at an estimated 96% UA, predicted prevalence was reduced to about 90% of reference prevalence (yellow, superimposed over cyan).

Disturbance maps with different levels of predicted prevalence demonstrated differences in spatial patterns that were clearly attributable to omission and commission error, visible by inspection of TM imagery (Fig. 3.12). This was most apparent when comparing areas that were subject to different

levels of disturbance intensity. For visual comparison, we superimposed a biased, low-prevalence class distribution (UA = 96.0%, PA = 86.6%, predicted prevalence = 90.2% of reference prevalence) over a nearly unbiased one (UA = 92.0%, PA = 91.6%, predicted prevalence = 99.6% of reference prevalence). The biased predictions under-represented disturbance due to elevated levels of omission error. Omission errors were concentrated in areas affected by low-intensity disturbance (Fig. 3.12c), where the biased predictions produced a more disconnected class distribution (Fig. 3.12e). In an area impacted by harvesting at greater intensity (Fig. 3.12d), the biased predictions produced a much more adequate representation of disturbance patches, whereas the unbiased predictions marginally over-represented the extent and connectivity of disturbance due to greater commission error (Fig. 3.12f). This was also an area in which subtle differences in canopy conditions introduced spectral change where no discrete disturbance occurred. The unbiased predictions included relatively well-connected patches of commission error, whereas commission error within the biased predictions generally occurred as small groups of pixels that could be removed by spatial filtering.

Cloud cover in the 2007 image and cloud shadows in the 2004 image were expected to introduce commission error. Despite variable viewing conditions caused by 2007 cloud cover, disturbance maps included only small and scattered patches of cloud-induced commission error (Fig. 3.13, dashed lines). A random sample of 100 locations drawn from the 2007 cloud/shadow mask were predicted with 100% accuracy in the MOSVC maps shown in Fig. 3.13. Disturbance was accurately predicted beneath cloud cover provided it was partially transparent in the infrared bands. However, cloud shadows in the 2004 image introduced distinct patches of commission error (Figs. 3.13a and 3.13c, solid lines). From a random sample of 100 reference locations drawn from the 2004 cloud/shadow mask, 13% of predictions were incorrect, all commission errors contributed by shadow. Shadow impacted a small fraction of mask pixels and a very small fraction of forested pixels, but visibly shadowed pixels were almost always predicted in error. We subsequently collected an ancillary, haphazard sample of 30

observations from shadow-induced false-positive pixel locations to supplement MOSVC training data. Comparing outcomes with equal predicted class prevalence (Figs. 3.13c and 3.13d), the ancillary training data reduced shadow-induced commission error, with only 3% of the 2004 cloud/shadow reference locations predicted in error. The ancillary data had a similar positive impact on MOSVC outcomes across the full range of estimated UA or class prevalence, based on similar reductions in cloud/shadow reference data error.

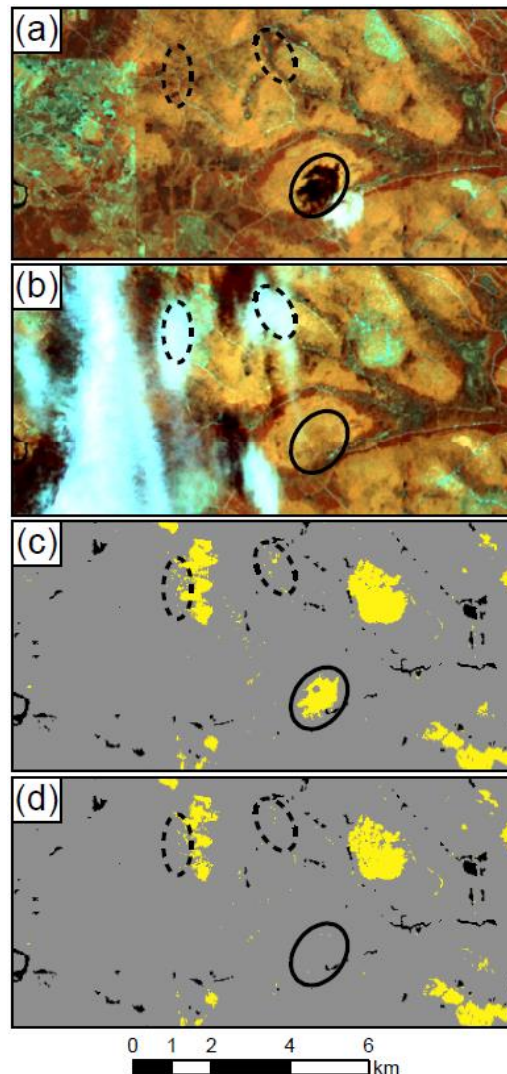


Figure 3.13. Cloud- and shadow-induced error in forest disturbance maps. Example area demonstrating the effects of cloud and cloud shadow in a) the 2004 Landsat TM image (bands 4, 5, and 3 shown in red, green, and blue) and b) the 2007 TM image on commission error patterns in c) a nearly unbiased map with estimated 92% UA. To reduce commission error caused by 2004 cloud shadow, an ancillary sample of shadowed locations was introduced into model training, with the outcome shown in d) a nearly unbiased map with estimated 91.6% UA.

3.7. Discussion

MO-SVM performs heuristic optimization of SVMs within a multi-objective framework, providing simultaneous control of both classification error and bias. Applied to three different species occurrence problems, MOSVC produced diverse sets of alternative models and maps including solutions with zero bias. In contrast, single-objective optimization using a similar GA produced inconsistent and biased outcomes because individual training metrics could not control both omission and commission error. Of the three training metrics considered for single-objective optimization (F, G, and PPC; Table 3.2), we might have expected F to consistently produce the least biased outcomes across problems with different levels of reference prevalence, since F balances UA against PA. PPC and G can clearly produce heavily biased outcomes when positive and negative reference cases are imbalanced, as was the case for black ash and pine (Figs. 8 and 9). Yet F tended to produce a substantial positive bias in these problems as well. For red spruce, where reference cases were more nearly balanced, F produced a comparatively strong negative bias while G and PPC tended to produce a small positive bias (Fig. 3.10). MOSVC controlled both class accuracy and bias by design, and always produced an unbiased solution. The multi-objective optimization framework also produced alternative solutions under the expectation that different tradeoffs between class accuracy and bias may be more or less beneficial for any specific map use.

To obtain an accurate and complete representation of tradeoffs between accuracy and bias, multi-objective optimization requires a statistical learning process capable of generating diverse solutions through the controlled manipulation of model structure. SVMs are well-suited in the sense that manipulation of a few free parameters can dramatically alter the geometry of decision boundaries (Brereton and Lloyd, 2010). Moreover, the effect of different SVM parameter combinations is partially dependent on covariate selection. Although SVMs have performed well in a wide variety of remote sensing applications (Mountrakis et al., 2011) including tree species distribution and forest cover or

disturbance mapping (Gavier-Pizarro et al., 2012; e.g., Guo et al., 2005; Huang et al., 2008; Kuemmerle et al., 2009), the complexity of SVM model specification could be considered a weakness relative to other ML approaches whose outcomes may be less affected by user-determined settings. In a multi-objective context, however, sensitivity of decision boundaries to parameter and variable selection is a strength. For species occurrence problems in particular, a wide range of model performance characteristics were obtained by through the selection of different combinations of parameter values and variable subsets. On the other hand, parameter and variable selection by stochastic search could produce different model specifications with very similar performance characteristics. Model interpretation is therefore made difficult both through the use of a "black box" ML algorithm and stochastic optimization of algorithm performance. Although post-hoc analyses of response and covariate values could be used to illuminate some important relationships between variables (e.g., Goldstein et al., 2014), the strength of MOSVC is clearly not in model interpretation, but rather the reduction of prediction error and control of its characteristics.

Within a single run the MOSVC algorithm evaluated a very large number of SVM models, but did not always yield a set of solutions expressing a full range of tradeoffs between performance objectives. Most notably, MOSVC initially failed to return pine occurrence models with high UA (Fig. 3.9a). Further elaboration of the GA chromosome (Fig. 3.2) and differentiation of training and validation data (Fig. 3.3) allowed for certain reference samples to be identified for possible exclusion from model training. Sample exclusion offered another means of affecting decision boundaries by changing which samples were available to define the SVM margin. In principle all samples could be made eligible for exclusion, in which case the algorithm could select an optimal subset. However, sample exclusion lengthens the GA chromosome and adds complexity to the search space, potentially impeding convergence to near-optimal solutions. For pine the consequences of sample exclusion were mixed. UA fronts were extended to ~100% accuracy and performance levels were generally improved, but PA fronts were little affected

(Fig. 3.9). Sample exclusion improved both UA and PA fronts for red spruce, although benefits were most pronounced for PA fronts (Fig. 3.10). Depending on the problem, selection of a subset of training samples can clearly improve SVM performance, but optimal sample selection has received little attention compared to optimal feature selection. Within a remote sensing context, however, it has long been recognized that SVMs can be well-trained with the careful selection of training samples (Foody and Mathur, 2006, 2004).

Statistical learning relies on representative training data, and for many applications the available data will be insufficient to resolve certain influential relationships between variables, leading to specific patterns of prediction error. In the forest disturbance problem, reference data were initially incapable of discriminating patterns of spectral change caused by canopy disturbance from those caused by cloud shadow in the pre-disturbance image. Visibly shadowed pixels were almost always predicted in error, producing undesirable patches of false change (Fig. 3.13). Augmentation of training/validation data to resolve specific sources of error like this generally requires careful consideration of sample inclusion probabilities for consistent estimation of predictive performance. After differentiating training and validation data within the MOSVC implementation (Fig. 3.3), ancillary training data could be introduced without affecting the consistency of CV estimates. To reduce the impact of cloud shadows on disturbance maps, we introduced a small ancillary sample obtained from a haphazard selection of shadowed locations. Because ancillary samples were not used for model validation, they did not influence MOSVC outcomes by providing new information on model performance. Instead, ancillary samples changed decision boundaries because they were formerly misclassified by SVM models and were therefore incorporated into the training process as SVs. The altered decision boundaries affected model performance characteristics, and we therefore iterated MOSVC over additional generations to achieve a new set of solutions. The new MOSVC fronts were very similar to the old fronts because shadowed pixels remained poorly represented by validation data. Yet the end effect was the near

elimination of commission error due to cloud shadow (Fig. 3.13), accomplished without adding excessive complexity to reference data collection, model performance estimation, or GA optimization.

GAs provide a solution to the difficult problem of identifying optimal SVM model structures, but themselves require the specification of a large number of settings. User-specified options generally affect the breadth of the search or the likelihood of convergence to local optima. For most applications of MOSVC, rapid convergence will probably be less desirable than a slower but fuller exploration of alternative solutions. We specified GA options that should generally promote population diversity and delay convergence (e.g., scattered crossover and a relatively high mutation rate). However, the benefits of exploring a diverse population of solutions will be partially undermined by uncertainty in estimates of model performance. CV uncertainty causes inconsistent approximation of the Pareto front, contributes to the over-estimation of model performance metrics, and probably increases the number of generations needed for convergence. We found that repeated cross-validation was necessary to reduce estimation uncertainty, despite the obvious impact to execution time. Fortunately, GA runtime can be reduced by parallel calculation of fitness metrics. The difficulties of training and validating models with limited reference data are not unique to MOSVC, but validation uncertainty does have specific implications for the degree to which model performance characteristics can be controlled. Ideally, MOSVC outcomes closely approximate the true underlying relationships between accuracy and class prevalence such that any small change in one objective is attained with minimal change in the other. In reality, optimality is gauged by uncertain estimates of model performance. However, solutions that differ in estimated performance by an amount that exceeds approximate levels of CV uncertainty will produce maps with substantively different error characteristics and levels of prediction bias.

The effects of classification bias on spatial representation or class configuration will depend on the spatial distribution of the response variable and on spatial patterns of omission and commission error. Effects will be problem specific, potentially complex, and difficult to infer from a single map when

spatial patterns of error not well known. For the canopy disturbance problem, omission and commission errors within any single disturbance map were visible through the comparison of pre- and post-disturbance imagery (Fig. 3.12). Omission errors were most prevalent in areas where disturbance intensity was light. The patchy nature of disturbance meant that omission errors were clumped and typically in close proximity to predicted disturbance. Commission errors were commonly caused by year-to-year changes in canopy condition. Although often more prevalent within certain forest stands, commission errors tended to be more widely scattered. A comparison of MOSVC maps with different degrees of bias and different levels of omission and commission error highlighted these patterns and their contribution to mapped class configuration. A biased map with comparatively high omission error tended to provide a more disconnected spatial representation of disturbance patches in areas affected by low-intensity disturbance (Fig. 3.12e). An unbiased map with comparatively high commission error had fairly large and somewhat consolidated patches of false change where fewer, smaller, and more scattered patches were present in the biased map (Fig. 3.12f). All models including the most heavily biased had very high overall accuracy (estimated PPC >97%), and in most areas models produced very similar predictions. But a comparison of maps with different error characteristics highlighted areas where a heightened prevalence of either omission or commission error impacted class configuration.

Map error may be approached from different viewpoints (see Edwards and Fortin, 2001). Error can be viewed as a potential impediment to map use that should be reduced as much as possible. To that end, a classification model is fit to minimize a specified measure of error, returning a best solution conditioned on available data and the selected training criterion. Alternatively, map error can be viewed as an inherent element of spatial representation that cannot be dissociated from a 'true' class distribution, but whose characteristics and consequences need to be better understood and more effectively controlled. These are not opposing viewpoints, but they are usually set apart by virtue of the fact that classification approaches do not provide adequate or predictable control of error

characteristics, nor do they provide alternative solutions needed to evaluate tradeoffs and their consequences. Despite important developments in the characterization of map error and spatial uncertainty (e.g., Kyriakidis, 2001; McGwire and Fisher, 2001), most map users have limited options and a limited ability to infer whether or how possible alternatives could improve application outcomes. Multi-objective optimization by MOSVC provides alternative solutions that systematically differ in terms of error characteristics and classification bias. An individual solution can be selected if application needs are known. Multiple solutions can be compared to evaluate the sensitivity of application outcomes to map error characteristics. In either case, multi-objective optimization of classification models can simultaneously reduce and control error in a way that may be of practical benefit whenever a ML modeling approach is appropriate.

BIBLIOGRAPHY

- Atkinson PM, Foody GM, Gething PW, Mathur A, Kelly CK. Investigating spatial structure in specific tree species in ancient semi-natural woodland using remote sensing and marked point pattern analysis. *Ecography (Cop)* 2007;30:88–104. doi:10.1111/j.2006.0906-7590.04726.x.
- Balch RE, Underwood GR. The life history of *Pineus pinifoliae* (Fitch) (Homoptera: Adelgidae) and its effects on white pine. *Can Entomol* 1950;82:117–23.
- Bartlett JW, De Stavola BL, Frost C. Linear mixed models for replication data to efficiently allow for covariate measurement error. *Stat Med* 2009;28:3158–78. doi:10.1002/sim.
- Baumann M, Ozdogan M, Kuemmerle T, Wendland KJ, Esipova E, Radeloff VC. Using the Landsat record to detect forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia. *Remote Sens Environ* 2012;124:174–84. doi:10.1016/j.rse.2012.05.001.
- Bazi Y, Melgani F. Toward an Optimal SVM Classification System for Hyperspectral Remote Sensing Images. *Geosci Remote Sensing, IEEE Trans* 2006;44:3374–85. doi:10.1109/TGRS.2006.880628.
- Beers TW, Dress PE, Wensel LC. Aspect transformation in site productivity research. *J For* 1966;64:691–2.
- Bergen KM, Zhao T, Kharuk V, Blam Y, Brown DG, Peterson LK, et al. Changing regimes: Forested land cover dynamics in Central Siberia 1974 to 2001. *Photogramm Eng Remote Sens* 2008;74:787–98.
- Blackard JA, Finco MV, Helmer EH, Holden GR, Hoppus ML, Jacobs DM, et al. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sens Environ* 2008;112:1658–77. doi:10.1016/j.rse.2007.08.021.
- Blum AL, Langley P. Selection of relevant features and examples in machine learning. *Artif Intell* 1997;97:245–71. doi:10.1016/S0004-3702(97)00063-5.
- Bowman AW, Azzalini A. *Applied Smoothing Techniques for Data Analysis*. New York: Oxford University Press; 1997.
- Breiman L. Random forests. *Mach Learn* 2001;45:5–32. doi:10.1023/A:1010933404324.
- Brereton RG, Lloyd GR. Support vector machines for classification and regression. *Analyst* 2010;135:230–67. doi:10.1039/b918972f.
- Brys G, Hubert M, Struyf A. A robust measure of skewness. *J Comput Graph Stat* 2004;13:996–1017. doi:10.1198/106186004X12632.
- Calabrese JM, Certain G, Kraan C, Dormann CF. Stacking species distribution models and adjusting bias by linking them to macroecological models. *Glob Ecol Biogeogr* 2014;23:99–112. doi:10.1111/geb.12102.

Canty MJ, Nielsen AA. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens Environ* 2008;112:1025–36. doi:10.1016/j.rse.2007.07.013.

Card DH. Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogramm Eng Remote Sensing* 1982;48:431–9.

Carroll RJ, Ruppert D, Stefanski LA. *Measurement Error in Nonlinear Models, Monographs on Statistics and Applied Probability* 63. London: Chapman & Hall; 1995.

Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* 2011;2:1–27. doi:10.1145/1961189.1961199.

Cochran WG. *Sampling Techniques*. 3rd ed. New York: John Wiley & Sons; 1977.

Cohen WB, Fiorella M, Gray J, Helmer E, Anderson K. An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogramm Eng Remote Sensing* 1998;64:293–300.

Cohen WB, Goward SN. Landsat's role in ecological applications of remote sensing. *Bioscience* 2004;54:535–45. doi:10.1641/0006-3568(2004)054[0535:LRIEAO]2.0.CO;2.

Cohen WB, Yang Z, Kennedy R. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. *Remote Sens Environ* 2010;114:2911–24. doi:10.1016/j.rse.2010.07.010.

Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L, et al. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci Model Dev* 2015;8:1991–2007. doi:10.5194/gmd-8-1991-2015.

Cooke WH. Forest/non-forest stratification in Georgia with Landsat Thematic Mapper data. In: McRoberts RE, Reams GA, Van Deusen PC, editors. *Proc. First Annu. For. Invent. Anal. Symp.*, St. Paul: US Forest Service, North Central Research Station; 2000, p. 28–30.

Crookston NL, Finley AO. yalmpute: An R package for kNN imputation. *J Stat Softw* 2008;23:1–16. doi:doi:

Cubbage FW, Newman DH. Forest policy reformed: A United States perspective. *For Policy Econ* 2006;9:261–73. doi:10.1016/j.forpol.2005.07.008.

Curran PJ, Hay A. The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogramm Eng Remote Sensing* 1986;52:229–41.

Currier J, Kanoti A, Ostrofsky W, Teerling C. Forest and shade tree insect and disease conditions for Maine: July 21, 2015. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 2015.

Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 2002;6:182–97. doi:10.1109/4235.996017.

Denham RJ, Falk MG, Mengersen KL. The Bayesian conditional independence model for measurement error: Applications in ecology. *Environ Ecol Stat* 2011;18:239–55. doi:10.1007/s10651-009-0130-3.

Dijak W. Landscape Builder: Software for the creation of initial landscapes for LANDIS from FIA data. *Comput Ecol Softw* 2013;3:17–25. doi:10.0000/issn-2220-721x-compuecol-2013-v3-0003.

Dimond JB, Bishop RH. Susceptibility and vulnerability of forests to the pine leaf aphid, *Pineus pinifoliae* (Fitch) (Adelgidae). Orono, Maine, USA: Maine Agricultural and Forest Experiment Station, University of Maine; 1968.

Duveneck MJ, Thompson JR, Wilson BT. An imputed forest composition map for New England screened by species range boundaries. *For Ecol Manage* 2015;347:107–15. doi:10.1016/j.foreco.2015.03.016.

Eberhardt LL, Thomas JM. Designing Environmental Field Studies. *Ecol Monogr* 1991;61:53–73. doi:10.2307/1942999.

Edwards G, Fortin M. A cognitive view of spatial uncertainty. In: Hunsaker CT, Goodchild MF, Friedl MA, Case TJ, editors. *Spat. Uncertain. Ecol. Implic. Remote Sens. GIS Appl.*, New York, NY: Springer-Verlag; 2001, p. 133–57.

Ellefson P V., Kilgore MA, Granskog JE. Government regulation of forestry practices on private forest land in the United States: An assessment of state government responsibilities and program performance. *For Policy Econ* 2007;9:620–32. doi:10.1016/j.forpol.2006.05.001.

Everitt BS, Landau S, Leese M, Stahl D. *Cluster Analysis*. 5th ed. Chichester: Wiley; 2011.

Ewers RM, Didham RK. Confounding factors in the detection of species responses to habitat fragmentation. *Biol Rev Camb Philos Soc* 2006;81:117–42. doi:10.1017/S1464793105006949.

Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence / absence models. *Environ Conserv* 1997;24:38–49. doi:10.1017/S0376892997000088.

Foody GM, Mathur A. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sens Environ* 2006;103:179–89. doi:10.1016/j.rse.2006.04.001.

Foody GM, Mathur A. Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. *Remote Sens Environ* 2004;93:107–17. doi:10.1016/j.rse.2004.06.017.

Foster SD, Shimadzu H, Darnell R. Uncertainty in spatially predicted covariates: Is it ignorable? *J R Stat Soc Ser C Appl Stat* 2012;61:637–52. doi:10.1111/j.1467-9876.2011.01030.x.

Freeman E a., Moisen GG. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol Modell* 2008;217:48–58. doi:10.1016/j.ecolmodel.2008.05.015.

Frescino TS, Edwards TC, Moisen GG. Modeling spatially explicit forest structural attributes using generalized additive models 2001;12:15–26.

Friedrichs F, Igel C. Evolutionary tuning of multiple SVM parameters. *Neurocomputing* 2005;64:107–17. doi:10.1016/j.neucom.2004.11.022.

Frost C, Thompson SG. Correcting for regression dilution bias: Comparison of methods for a single predictor variable. *J R Stat Soc Ser A* 2000;163:173–89.

Fuller AK, Harrison DJ. Influence of partial timber harvesting on American martens in North-Central Maine. *J Wildl Manage* 2005;69:710–22.

Gavier-Pizarro GI, Kuemmerle T, Hoyos LE, Stewart SI, Huebner CD, Keuler NS, et al. Monitoring the invasion of an exotic tree (*Ligustrum lucidum*) from 1983 to 2006 with Landsat TM/ETM+ satellite data and Support Vector Machines in Córdoba, Argentina. *Remote Sens Environ* 2012;122:134–45. doi:10.1016/j.rse.2011.09.023.

Ghoggali N, Melgani F, Bazi Y. A multiobjective genetic SVM approach for classification problems with limited training samples. *IEEE Trans Geosci Remote Sens* 2009;47:1707–18. doi:10.1109/TGRS.2008.2007128.

Gislason PO, Benediktsson JA, Sveinsson JR. Random Forests for land cover classification. *Pattern Recognit Lett* 2006;27:294–300. doi:10.1016/j.patrec.2005.08.011.

Goldberg DE. *Genetic Algorithms in Search, Optimisation and Machine Learning*. New York: Addison-Wesley; 1989.

Goldstein A, Kapelner A, Bleich J, Pitkin E. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *J Comput Graph Stat* 2014;24:44–65. doi:10.1080/10618600.2014.907095.

Guisan A, Weiss SB, Weiss AD. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecol* 1999;143:107–22. doi:10.1023/A:1009841519580.

Guo Q, Kelly M, Graham CH. Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecol Modell* 2005;182:75–90. doi:10.1016/j.ecolmodel.2004.07.012.

Gustafson EJ. When relationships estimated in the past cannot be used to predict the future: Using mechanistic models to predict landscape ecological dynamics in a changing world. *Landsc Ecol* 2013;28:1429–37. doi:10.1007/s10980-013-9927-4.

Gustafson EJ, Shifley SR, Mladenoff DJ, Nimerfro KK, He HS. Spatial simulation of forest succession and timber harvesting using LANDIS. *Can J For Resour* 2000;30:32–43.

Hagan JM, Irland LC, Whitman AA. *Changing timberland ownership in the Northern Forest and implications for biodiversity*. Brunswick, Maine, USA: Manomet Center for Conservation Sciences; 2005.

Häntzschel J, Goldberg V, Bernhofer C. GIS-based regionalisation of radiation, temperature and coupling measures in complex terrain for low mountain ranges. *Meteorol Appl* 2005;12:33–42. doi:10.1017/S1350482705001489.

- He H, Garcia E a. Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 2009;21:1263–84. doi:10.1109/TKDE.2008.239.
- Healey SP, Cohen WB, Spies TA, Moeur M, Pflugmacher D, Whitley MG, et al. The relative impact of harvest and fire upon landscape-level dynamics of older forests: Lessons from the Northwest Forest Plan. *Ecosystems* 2008;11:1106–19. doi:10.1007/s10021-008-9182-8.
- Henderson EB, Ohmann JL, Gregory MJ, Roberts HM, Zald H. Species distribution modelling for plant communities: Stacked single species or multivariate modelling approaches? *Appl Veg Sci* 2014;17:516–27. doi:10.1111/avsc.12085.
- Hennigar CR, Wilson JS, MacLean DA, Wagner RG. Applying a spruce budworm decision support system to Maine: Projecting spruce-fir volume impacts under alternative management and outbreak scenarios. *J For* 2011;109:332–42.
- Hepinstall JA, Sader SA, Krohn WB, Boone RB, Bartlett RI. Development and testing of a vegetation and land cover map of Maine. Orono, Maine, USA: Maine Agricultural and Forest Experiment Station, University of Maine; 1999.
- Herms DA, McCullough DG. Emerald ash borer invasion of North America: History, biology, ecology, impacts, and management. *Annu Rev Entomol* 2014;59:13–30. doi:10.1146/annurev-ento-011613-162051.
- Holland JH. *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan: University of Michigan Press; 1975.
- Homyack JA, Harrison DJ, Krohn WB. Effects of precommercial thinning on snowshoe hares in Maine. *J Wildl Manage* 2007;71:4–13. doi:10.2193/2005-481.
- Huang C-L. ACO-based hybrid classification system with feature subset selection and model parameters optimization. *Neurocomputing* 2009;73:438–48. doi:10.1016/j.neucom.2009.07.014.
- Huang C, Song K, Kim S, Townshend JRG, Davis P, Masek JG, et al. Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sens Environ* 2008;112:970–85. doi:10.1016/j.rse.2007.07.023.
- Huang C, Thomas N, Goward SN, Masek JG, Zhu Z, Townshend JRG, et al. Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *Int J Remote Sens* 2010;31:5449–64. doi:10.1080/01431160903369642.
- Huang CL, Wang CJ. A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst Appl* 2006;31:231–40. doi:10.1016/j.eswa.2005.09.024.
- Hubert M, Engelen S. Robust PCA and classification in biosciences. *Bioinformatics* 2004;20:1728–36. doi:10.1093/bioinformatics/bth158.
- Hubert M, Rousseeuw P, Verdonck T. Robust PCA for skewed data and its outlier map. *Comput Stat Data Anal* 2009;53:2264–74. doi:10.1016/j.csda.2008.05.027.

Hudak AT, Crookston NL, Evans JS, Hall DE, Falkowski MJ. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sens Environ* 2008;112:2232–45. doi:10.1016/j.rse.2007.10.009.

Irland LC, Dimond JB, Stone JL, Falk J, Baum E. The spruce budworm outbreak in Maine in the 1970's - assessment and directions for the future. Orono, Maine, USA: Maine Agricultural and Forest Experiment Station, University of Maine; 1988.

Jackson ST, Sax DF. Balancing biodiversity in a changing environment: extinction debt, immigration credit and species turnover. *Trends Ecol Evol* 2010;25:153–60. doi:10.1016/j.tree.2009.10.001.

James PMA, Fortin M, Fall A, Kneeshaw D, Messier C. The effects of spatial legacies following shifting management practices and fire on boreal forest age structure. *Ecosystems* 2007;10:1261–77. doi:10.1007/s10021-007-9095-y.

Jiménez-Valverde A, Acevedo P, Barbosa AM, Lobo JM, Real R. Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. *Glob Ecol Biogeogr* 2013;22:508–16. doi:10.1111/geb.12007.

Jin S, Sader SA. Effects of forest ownership and change on forest harvest rates, types and trends in northern Maine. *For Ecol Manage* 2006;228:177–86. doi:10.1016/j.foreco.2006.03.009.

Jin S, Sader SA. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sens Environ* 2005;94:364–72. doi:10.1016/j.rse.2004.10.012.

Jin Y, editor. *Multi-Objective Machine Learning*. vol. 16. Berlin: Springer-Verlag; 2006.

Kennedy RE, Yang Z, Cohen WB, Pfaff E, Braaten J, Nelson P. Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sens Environ* 2012;122:117–33. doi:10.1016/j.rse.2011.09.024.

Kim JH. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Comput Stat Data Anal* 2009;53:3735–45. doi:10.1016/j.csda.2009.04.009.

Kohm KA, Franklin JF, editors. *Creating a forestry for the 21st century*. Washington D.C.: Island Press; 1997.

Konak A, Coit DW, Smith AE. Multi-objective optimization using genetic algorithms: A tutorial. *Reliab Eng Syst Saf* 2006;91:992–1007. doi:10.1016/j.ress.2005.11.018.

Kubat M, Matwin S. Addressing the curse of imbalanced training sets: one-sided selection. *Proc Fourteenth Int Conf Mach Learn* 1997:179–86. doi:10.1017/CBO9781107415324.004.

Kuemmerle T, Chaskovskyy O, Knorn J, Radeloff VC, Kruhlov I, Keeton WS, et al. Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007. *Remote Sens Environ* 2009;113:1194–207. doi:10.1016/j.rse.2009.02.006.

Kyriakidis P. Geostatistical models of uncertainty for spatial data. In: Hunsaker CT, Goodchild MF, Friedl MA, Case TJ, editors. *Spat. Uncertain. Ecol. Implic. Remote Sens. GIS Appl.*, New York, NY: Springer-Verlag; 2001, p. 175–213.

Langford WT, Gergel SE, Dietterich TG, Cohen W. Map misclassification can cause large errors in landscape pattern indices: examples from habitat fragmentation. *Ecosystems* 2006;9:474–88. doi:10.1007/s10021-005-0119-1.

Lechner AM, Langford WT, Bekessy SA, Jones SD. Are landscape ecologists addressing uncertainty in their remote sensing data? *Landsc Ecol* 2012;27:1249–61. doi:10.1007/s10980-012-9791-7.

Legaard KR, Sader SA, Simons-Legaard EM. Evaluating the impact of abrupt changes in forest policy and management practices on landscape dynamics: analysis of a Landsat image time series in the Atlantic Northern Forest. *PLoS One* 2015;10:e0130428. doi:10.1371/journal.pone.0130428.

Li S, Tan M. Tuning SVM parameters by using a hybrid CLPSO–BFGS algorithm. *Neurocomputing* 2010;73:2089–96. doi:10.1016/j.neucom.2010.02.013.

Li S, Wu H, Wan D, Zhu J. An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine. *Knowledge-Based Syst* 2011;24:40–8. doi:10.1016/j.knosys.2010.07.003.

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

Likens GE, Franklin JF. Ecosystem thinking in the Northern Forest - and beyond. *Bioscience* 2009;59:511–3.

Lin S-W, Ying K-C, Chen S-C, Lee Z-J. Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert Syst Appl* 2008;35:1817–24. doi:10.1016/j.eswa.2007.08.088.

Loiselle B a., Howell C a., Graham CH, Goerck JM, Brooks T, Smith KG, et al. Avoiding pitfalls of using species distribution models in conservation planning. *Conserv Biol* 2003;17:1591–600. doi:10.1111/j.1523-1739.2003.00233.x.

Lorena AC, De Carvalho ACPLF. Evolutionary tuning of SVM parameter values in multiclass problems. *Neurocomputing* 2008;71:3326–34. doi:10.1016/j.neucom.2008.01.031.

Lorimer CG, White AS. Scale and frequency of natural disturbances in the northeastern US: implications for early successional forest habitats and regional age distributions. *For Ecol Manage* 2003;185:41–64. doi:10.1016/S0378-1127(03)00245-7.

Maine Forest Service. Silvicultural activities report. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 2014.

Maine Forest Service. Forest Trees of Maine. 14th ed. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 2008.

Maine Forest Service. Silvicultural activities report. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 2005.

Maine Forest Service. Silvicultural activities report. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 2000.

Maine Forest Service. Silvicultural activities report. Augusta, Maine, USA: Maine Forest Service, Department of Conservation; 1994.

Maine Natural Areas Program. Natural Community Fact Sheet - Black Ash Swamp n.d.
http://www.maine.gov/dacf/mnap/features/communities/black_ash_swamp.pdf (accessed May 2, 2016).

Martens D, Baesens B, Van Gestel T, Vanthienen J. Comprehensible credit scoring models using rule extraction from support vector machines. *Eur J Oper Res* 2007;183:1466–76.
doi:10.1016/j.ejor.2006.04.051.

Mas JF, Flores JJ. The application of artificial neural networks to the analysis of remotely sensed data. *Int J Remote Sens* 2008;29:617–63. doi:10.1080/01431160701352154.

McCaskill GL, McWilliams WH, Barnett CJ, Butler BJ, Hatfield MA, Kurtz CM, et al. Maine's Forests 2008. Newtown Square, PA: USDA Forest Service, Northern Research Station; 2011.

McGarigal K, Cushman SA, Ene E. Fragstats v4: Spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. 2012. <http://www.umass.edu/landeco/research/fragstats/fragstats.html>.

McGarigal K, McComb WC. Relationships between landscape structure and breeding birds in the Oregon coast range. *Ecol Monogr* 1995;65:235–60. doi:10.2307/2937059.

McGwire KC, Fisher P. Spatially variable thematic accuracy: Beyond the confusion matrix. In: Hunsaker CT, Goodchild MF, Friedl MA, Case TJ, editors. *Spat. Uncertain. Ecol. Implic. Remote Sens. GIS Appl.*, New York, NY: Springer-Verlag; 2001, p. 308–29.

McInerney GJ, Purves DW. Fine-scale environmental variation in species distribution modelling: Regression dilution, latent variables and neighbourly advice. *Methods Ecol Evol* 2011;2:248–57.
doi:10.1111/j.2041-210X.2010.00077.x.

McRoberts RE, Bechtold WA, Patterson PL, Scott CT, Reams GA. The Enhanced Forest Inventory and Analysis Program of the USDA Forest Service: historical perspective and announcement of statistical documentation. *J For* 2005;103:304–8.

McRoberts RE, Cohen WB, Næsset E, Stehman S V., Tomppo EO. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scand J For Res* 2010;25:340–67. doi:10.1080/02827581.2010.497496.

Mladenoff DJ. LANDIS and forest landscape models. *Ecol Modell* 2004;180:7–19.
doi:10.1016/j.ecolmodel.2004.03.016.

Morin H, Jardon Y, Gagnon R. Relationship between spruce budworm outbreaks and forest dynamics in eastern North America. In: Johnson EA, Miyanishi K, editors. *Plant Disturb. Ecol.*, Burlington, Massachusetts, USA: Elsevier; 2007, p. 555–77.

Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: A review. *ISPRS J Photogramm Remote Sens* 2011;66:247–59. doi:10.1016/j.isprsjprs.2010.11.001.

Mouton AM, De Baets B, Goethals PLM. Ecological relevance of performance criteria for species distribution models. *Ecol Modell* 2010;221:1995–2002. doi:10.1016/j.ecolmodel.2010.04.017.

Nightingale JM, Fan W, Coops NC, Waring RH. Predicting Tree Diversity Across the United States As a Function of Modeled Gross Primary Production. *Ecol Appl* 2008;18:93–103. doi:10.1890/07-0693.1.

Ohmann JL, Gregory MJ. Predictive mapping of forest composition and structure with direct gradient analysis and nearest- neighbor imputation in coastal Oregon, U.S.A. *Can J For Res* 2002;32:725–41. doi:10.1139/x02-011.

Ohmann JL, Gregory MJ, Roberts HM. Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. *Remote Sens Environ* 2014;151:3–15. doi:10.1016/j.rse.2013.08.048.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. Package “vegan” - Community Ecology Package, version 2.4-3 2017.

Ollinger S V, Aber JD, Federer CA, Lovett GM, Ellis JM. Modeling physical and chemical climate of the Northeastern United States for a geographic information system. Radnor, Pennsylvania, USA: USDA Forest Service, Northern Research Station; 1995.

Olofsson P, Foody GM, Stehman S V., Woodcock CE. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens Environ* 2013;129:122–31. doi:10.1016/j.rse.2012.10.031.

Pal M, Mather PM. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens Environ* 2003;86:554–65. doi:10.1016/S0034-4257(03)00132-9.

Pasolli L, Notarnicola C, Bruzzone L, Bertoldi G, Della Chiesa S, Niedrist G, et al. Polarimetric RADARSAT-2 imagery for soil moisture retrieval in alpine areas. *Can J Remote Sens* 2011;37:535–47.

Pickett STA. Space-for-time substitution as an alternative to long-term studies. In: Likens GE, editor. *Long-term Stud. Ecol. Approaches Altern.*, New York: Springer-Verlag; 1989, p. 110–35.

Powell SL, Cohen WB, Healey SP, Kennedy RE, Moisen GG, Pierce KB, et al. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sens Environ* 2010;114:1053–68. doi:10.1016/j.rse.2009.12.018.

Preisendorfer RW. *Principal component analysis in meteorology and oceanography*. Amsterdam: Elsevier; 1988.

Price B, McAlpine CA, Kutt AS, Ward D, Phinn SR, Ludwig JA. Disentangling how landscape spatial and temporal heterogeneity affects Savanna birds. *PLoS One* 2013;8:e74333. doi:10.1371/journal.pone.0074333.

Puertas OL, Brenning A, Meza FJ. Balancing misclassification errors of land cover classification maps using support vector machines and Landsat imagery in the Maipo river basin (Central Chile, 1975-2010). *Remote Sens Environ* 2013;137:112–23. doi:10.1016/j.rse.2013.06.003.

Quinn P, Beven K, Chevallier P, Planchon O. The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrol Process* 1991;5:59–79.

R Core Team. R: A language and environment for statistical computing. R Found Stat Comput 2017.

Radford JQ, Bennett AF, Cheers GJ. Landscape-level thresholds of habitat cover for woodland-dependent birds. *Biol Conserv* 2005;124:317–37. doi:10.1016/j.biocon.2005.01.039.

Ranco D, Arnett A, Latty E, Remsburg A, Dunckel K, Quigley E, et al. Two Maine forest pests: A comparison of approaches to understanding threats to hemlock and ash trees in Maine. *Maine Policy Rev* 2012;21:76–89.

Rehfeldt GE. A spline model of climate for the western United States. Fort Collins, Colorado, USA: USDA Forest Service, Rocky Mountain Research Station; 2006.

Rejou-Mechain M, Muller-Landau HC, Detto M, Thomas SC, Le Toan T, Saatchi SS, et al. Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. *Biogeosciences* 2014;11:6827–40. doi:10.5194/bg-11-6827-2014.

Riemann R, Wilson BT, Lister A, Parks S. An effective assessment protocol for continuous geospatial datasets of forest characteristics using USFS Forest Inventory and Analysis (FIA) data. *Remote Sens Environ* 2010;114:2337–52. doi:10.1016/j.rse.2010.05.010.

van Rijsbergen CJ. Information Retrieval. 2nd ed. London: Butterworths; 1979.

Robinson C, Saatchi S, Neumann M, Gillespie T. Impacts of spatial variability on aboveground biomass estimation from L-band radar in a temperate forest. *Remote Sens* 2013;5:1001–23. doi:10.3390/rs5031001.

Robinson L. Ecological relationships among partial harvesting, vegetation, snowshoe hares, and Canada lynx in Maine. M.Sc. Thesis. The University of Maine, 2006.

Rondinini C, Wilson K a, Boitani L, Grantham H, Possingham HP. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecol Lett* 2006;9:1136–45. doi:10.1111/j.1461-0248.2006.00970.x.

Sader SA, Bertrand M, Wilson EH. Satellite change detection of forest harvest patterns on an industrial forest landscape. *For Sci* 2003;49:341–53.

Sader SA, Jin S, Metzler JW, Hoppus M. Exploratory analysis of forest harvest and regeneration pattern among multiple landowners. *For Chron* 2006;82:203–10.

Sader SA, Legaard KR. Inclusion of forest harvest legacies, forest type, and regeneration spatial patterns in updated forest maps: A comparison of mapping results. *For Ecol Manage* 2008;255:3846–56. doi:10.1016/j.foreco.2008.03.047.

Sader SA, Winne JC. RGB-NDVI colour composites for visualizing forest change. *Int J Remote Sens* 1992;13:3055–67.

Salcedo-Sanz S, Rojo-Alvarez JL, Martinez-Ramon M, Camps-Valls G. Support vector machines in engineering: an overview. *Wiley Interdiscip Rev Min Knowl Discov* 2014;4:234–67. doi:10.1002/Widm.1125.

Samadzadegan F, Hasani H, Schenk T. Determination of optimum classifier and feature subset in hyperspectral images based on ant colony system. *Photogramm Eng Remote Sens* 2012;78:1261–73.

Scheller RM, Domingo JB, Sturtevant BR, Williams JS, Rudy A, Gustafson EJ, et al. Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecol Model* 2007;201:409–19.

Scheller RM, Van Tuyl S, Clark K, Hayden NG, Hom J, Mladenoff DJ. Simulation of forest change in the New Jersey Pine Barrens under current and pre-colonial conditions. *For Ecol Manage* 2008;255:1489–500. doi:10.1016/j.foreco.2007.11.025.

Schrott GR, With KA, King AW. On the importance of landscape history for assessing extinction risk. *Ecol Appl* 2005;15:493–506.

Seymour RS. The northeastern region. In: Barrett JW, editor. *Reg. Silv. United States*. 3rd ed., New York: Wiley; 1995, p. 31–79.

Seymour RS. The red spruce-balsam fir forest of Maine: evolution of silvicultural practice in response to stand development patterns and disturbances. In: Kelty MJ, Larson BC, Oliver CD, editors. *Ecol. Silv. Mix. For.*, Norwell, Massachusetts, USA: Kluwer Academic Publishers; 1992, p. 217–44.

Seymour RS, Hunter MLJ. Principles of ecological forestry. In: Hunter MLJ, editor. *Maint. Biodivers. For. Ecosyst.*, Cambridge, UK: Cambridge University Press; 1999, p. 22–61.

Shao G, Wu J. On the accuracy of landscape pattern analysis using remote sensing data. *Landsc Ecol* 2008;23:505–11. doi:10.1007/s10980-008-9215-x.

Simons EM. Influences of past and future forest management on the spatiotemporal dynamics of habitat supply for Canada lynx and American martens in northern Maine. Ph.D. Dissertation. The University of Maine, 2009.

Small C, Elvidge CD. Night on Earth: Mapping decadal changes of anthropogenic night light in Asia. *Int J Appl Earth Obs Geoinf* 2013;22:40–52. doi:10.1016/j.jag.2012.02.009.

Smith WB. Forest inventory and analysis: a national inventory and monitoring program. *Environ Pollut* 2002;116 Suppl:S233–42.

Soenen SA, Peddle DR, Coburn CA. SCS+C: A modified sun-canopy-sensor topographic correction in forested terrain. *Geosci Remote Sensing, IEEE Trans* 2005;43:2148–59. doi:10.1109/TGRS.2005.852480.

Stehman S V, Czaplewski RL. Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sens Environ* 1998;64:331–44.

Story M, Congalton RG. Accuracy assessment: a user's perspective. *Photogramm Eng Remote Sens* 1986;52:397–9. doi:10.1111/j.1530-9290.2010.00257.x.

Sturtevant BR, Miranda BR, Wolter PT, James PMA, Fortin M, Townsend PA. Forest recovery patterns in response to divergent disturbance regimes in the Border Lakes region of Minnesota (USA) and Ontario (Canada). *For Ecol Manage* 2014;313:199–211. doi:10.1016/j.foreco.2013.10.039.

Suttorp T, Igel C. Multi-objective optimization of support vector machines. *Stud Comput Intell* 2006;16:199–220.

Tang Y, Zhang YQ, Chawla N V., Krasser S. SVMs modeling for highly imbalanced classification. *IEEE Trans Syst Man, Cybern Part B Cybern* 2009;39:281–8. doi:10.1109/TSMCB.2008.2002909.

Thomas NE, Huang C, Goward SN, Powell S, Rishmawi K, Schleeweis K, et al. Validation of North American Forest Disturbance dynamics derived from Landsat time series stacks. *Remote Sens Environ* 2011;115:19–32. doi:10.1016/j.rse.2010.07.009.

Üstün B, Melssen WJ, Oudenhuijzen M, Buydens LMC. Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization. *Anal Chim Acta* 2005;544:292–305. doi:10.1016/j.aca.2004.12.024.

Václavík T, Meentemeyer RK. Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecol Modell* 2009;220:3248–58. doi:10.1016/j.ecolmodel.2009.08.013.

Vapnik VN. *Statistical Learning Theory*. New York, NY: John Wiley & Sons; 1998.

Vapnik VN. *The Nature of Statistical Learning Theory*. New York: Springer; 1995.

Verboven S, Hubert M. MATLAB library LIBRA. *Wiley Interdiscip Rev Comput Stat* 2010;2:509–15. doi:10.1002/wics.96.

Wang L, Liu H. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *Int J Geogr Inf Sci* 2006;20:193–213. doi:10.1080/13658810500433453.

Ward JH. Hierarchical groups to optimize an objective function. *J Am Stat Assoc* 1963;58:236–44.

Willmott C. On the validation of models. *Phys Geogr* 1981;2:184–94.

Wilson BT, Lister AJ, Riemann RI. A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *For Ecol Manage* 2012;271:182–98. doi:10.1016/j.foreco.2012.02.002.

Wilson EH, Sader SA. Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sens Environ* 2002;80:385–96.

Wilson KA, Westphal MI, Possingham HP, Elith J. Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biol Conserv* 2005;122:99–112. doi:10.1016/j.biocon.2004.07.004.

With KA. Invoking the ghosts of landscapes past to understand the landscape ecology of the present... and the future. In: Bissonette JA, Storch I, editors. *Temporal Dimens. Landsc. Ecol.*, New York: Springer; 2007, p. 43–58.

Wu C-H, Tzeng G-H, Lin R-H. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. *Expert Syst Appl* 2009;36:4725–35. doi:10.1016/j.eswa.2008.06.046.

Xi W, Coulson RN, Birt AG, Shang Z-B, Waldron JD, Lafon CW, et al. Review of forest landscape models: Types, methods, development and applications. *Acta Ecol Sin* 2009;29:69–78. doi:10.1016/j.chnaes.2009.01.001.

Xu Y, Dickson BG, Hampton HM, Sisk TD, Palumbo JA, Prather JW. Effects of mismatches of scale and location between predictor and response variables on forest structure mapping. *Photogramm Eng Remote Sensing* 2009;75:313–22.

Zäpfel G, Braune R, Bögl M. *Metaheuristic Search Concepts: A Tutorial with Applications to Production and Logistics*. New York: Springer; 2010.

Zevenbergen LW, Thorne CR. Quantitative analysis of land surface topography. *Earth Surf Process Landforms* 1987;12:47–56.

Zhu Z, Woodcock CE. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens Environ* 2012;118:83–94. doi:10.1016/j.rse.2011.10.028.

APPENDIX A: CHAPTER 1 LANDSAT IMAGE PROCESSING

Forest harvest and composition maps were assembled from a time series of Landsat Multispectral Scanner (MSS), Thematic Mapper (TM), and Enhanced Thematic Mapper Plus (ETM+) images acquired during summer leaf-on conditions (Table 1.1). Consecutive images were spaced 1-4 years apart, as determined by the availability of high quality, predominantly cloud-free imagery. Images were either obtained from the U.S. Geological Survey (USGS) Earth Resources Observation and Science Center or available for use through other programs (Hepinstall et al., 1999; Lunetta et al., 1998).

Change detection and composition mapping procedures were applied to forested pixels as identified by the 1993 Maine Gap Analysis Program (GAP) land cover map. The GAP map represents conditions near the midpoint of our time series, and discriminated forest from non-forest with an estimated 100% accuracy within our study area (Hepinstall et al., 1999). All images were geo-referenced to a previously rectified 1991 image that was used to produce the GAP map. TM and ETM+ images acquired 1988-2007 were rectified using a second-order polynomial transformation applied to 30-35 well distributed ground control points, with nearest neighbor resampling (RMSE <15 m). The 2010 TM image was obtained from the USGS with Level 1T Standard Terrain Correction and close inspection indicated that no further geocorrection was necessary. MSS images were rectified using a second-order polynomial transformation applied to 25-30 ground control points (RMSE <30 m) and resampled to 30 m by cubic convolution to match the spatial resolution of the TM/ETM+ imagery.

For each of the MSS and TM/ETM+ image sequences, a subset of image bands was selected for change detection and forest type mapping. TM/ETM+ red band 3, near-infrared band 4, and mid-infrared band 5 were retained, a combination that provides most of the image information content for northern temperate and boreal forests (Häme, 1991; Horler and Ahern, 1986; Sader, 1990). MSS green band 1, red band 2, and near-infrared band 4 were retained following the observation that near-infrared band 3 was less comparable to TM/ETM+ data (Crist and Cicone, 1984). Clouds and cloud shadows were

delineated and masked using an on-screen digitization procedure. Cloud cover typically affected a small fraction of forestland (Table 1.1). Extensive cloud cover on 17 June 2007 was mitigated by substituting cloud-contaminated areas with TM image data acquired on 22 August 2007. The substitution of cloud-free data was not possible for images acquired in 1993 and 1997.

To facilitate visual interpretation, images were transformed to a common radiometric scale using a relative radiometric normalization procedure applied separately to MSS and TM/ETM+ imagery. A preliminary change detection procedure known as multivariate alteration detection was first applied to consecutive image pairs to identify pixels whose spectral characteristics had not changed (Canty et al., 2004). Band values were extracted from a random sample of 5000 no-change pixels and linear normalization parameters were estimated using Theil-Sen regression (Olthof et al., 2005). Normalization parameters were used to derive a common radiometric scale for each band, preserving the full radiometric resolution of all images (Du et al., 2002, 2001). Normalization was performed to enhance visual consistency between images, and to reduce image-to-image differences in the impact of atmospheric effects on derived vegetation index values. However, the classification procedures used to produce forest harvest and composition maps (unsupervised classification guided by visual interpretation of Landsat images and ancillary data) do not assume a common radiometric scale across images. Normalization was therefore not a requirement (Song et al., 2001), and normalization outcomes were accordingly evaluated by qualitative visual assessment only.

References

- Canty MJ, Nielsen AA, Schmidt M. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sens Environ* 2004;91:441–51. doi:10.1016/j.rse.2003.10.024.
- Crist EP, Ciccone RC. Comparisons of the dimensionality and features of simulated Landsat-4 MSS and TM data. *Remote Sens Environ* 1984;14:235–46. doi:10.1016/0034-4257(84)90018-X.
- Du Y, Cihlar J, Beaubien J, Latifovic R. Radiometric normalization, compositing, and quality control for satellite high resolution image mosaics over large areas. *IEEE Trans Geosci Remote Sens* 2001;39:623–34.

Du Y, Teillet PM, Cihlar J. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sens Environ* 2002;82:123–34. doi:10.1016/S0034-4257(02)00029-9.

Häme T. Spectral interpretation of changes in forest using satellite scanner images. *Acta For Fenn* 1991;222:1–111.

Hepinstall JA, Sader SA, Krohn WB, Boone RB, Bartlett RI. Development and testing of a vegetation and land cover map of Maine. Orono, Maine, USA: Maine Agricultural and Forest Experiment Station, University of Maine; 1999.

Horler DNH, Ahern FJ. Forestry information content of Thematic Mapper data. *Int J Remote Sens* 1986;7:405–28.

Lunetta RS, Lyon JG, Guindon B, Elvidge CD. North American Landscape Characterization Dataset Development and Data Fusion Issues. *Photogramm Eng Remote Sensing* 1998;64:821–9.

Olthof I, Pouliot D, Fernandes R, Latifovic R. Landsat-7 ETM+ radiometric normalization comparison for northern mapping applications. *Remote Sens Environ* 2005;95:388–98. doi:10.1016/j.rse.2004.06.024.

Sader SA. Multispectral and seasonal characteristics of northern hardwood and boreal forest types in Maine. In: Elvidge C, editor. *Image Process. '89 Sparks*, Nevada, 23 May 1989, Bethesda, Maryland, USA: The Society; 1990, p. 109–16.

Song C, Woodcock CE, Seto KC, Lenney MP, Macomber SA. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sens Lett* 2001;75:230–44.

APPENDIX B: VALIDATION OF CHAPTER 1 MAPS

Validation Overview

The U.S. Forest Service Forest Inventory and Analysis (FIA) Program provides quality-assured measurements of forest attributes from a national network of field plots adhering to a statistically rigorous systematic sampling design (McRoberts et al., 2005; Smith, 2002). We made extensive use of FIA data for validation of Landsat-derived forest harvest and type maps. Since 1999, 20% of FIA plots within Maine have been surveyed annually during 5-year inventory cycles. Earlier inventories were conducted at irregular intervals using different designs (McRoberts et al., 2005). Under the modern inventory design, FIA plots consist of a center subplot with three satellite subplots, each 0.017 ha in size (McRoberts et al., 2005). Subplots will generally fall within an area defined by a 3x3 pixel block (90x90 m), and subplots constitute 8% of that area. The suitability of FIA data for map validation is affected by mismatches in location and scale between FIA plots and pixel neighborhoods, uncertainty of field measurements, and for maps spanning long time periods, changes in inventory design. Validation using FIA data should be considered an assessment of agreement with an accepted and widely utilized source of information on forest conditions, rather than an assessment of accuracy against ground truth. The FIA program maintains the confidentiality of plot locations to protect the privacy of landowners and to preserve plot integrity (Coulston et al., 2006; Smith, 2002). True plot locations were made available for our use through a collaborative agreement with the USFS Northern Research Station FIA Program.

Harvest Time Series Validation Procedure

Visual interpretation of TM/ETM+ imagery has been established as a credible source of reference data for the validation of forest disturbance maps (Cohen et al., 2010, 1998; Sader et al., 2003; Sader and Legaard, 2008). Even very light partial harvests are discernible due to characteristic spectral response, texture, and clearly visible access roads. Stand-replacing and heavy partial harvests are similarly apparent in MSS imagery. However, visual discrimination of disturbance intensity classes is

difficult and subjective. Forest inventory data provided by the FIA program provide a valuable alternative source of reference data. Forest age estimates at FIA plots represent the average age of overstory trees and, assuming they were established at the time of disturbance, the year of disturbance should be given by the year of field measurement minus stand age. However, if the overstory cohort was established from a seed source following disturbance or as advance regeneration prior to disturbance, the date of stand origin may be over- or underestimated, respectively. Field methods introduce additional uncertainty. Typically 2-3 trees are subjectively selected for aging. Age is estimated by coring and ring count at 1.37 m with the addition of a constant to approximate time elapsed between germination and growth to the height of coring (U.S. Forest Service, 2012). Thomas et al. (2011) validated forest disturbance time series (occurrence maps) by independent analyses of TM/ETM+ image interpretations and FIA forest age estimates. They concluded that validation is improved through the use of both reference sources, but they did not integrate the two into a single assessment. We found the visual discrimination of harvest intensity classes to be highly subjective, and FIA age to be an uncertain measure of time since disturbance. However, FIA plot data provide an objective basis for the discrimination of harvest intensity classes and image interpretation provides accurate identification of harvest dates. We adopted an approach to reference class labeling that leveraged the strength of one against the weakness of the other.

Our validation approach was based on the visual interpretation of satellite imagery over FIA plot locations to obtain reference class labels. Image interpretation was used to date harvest events; FIA plot data were used to discriminate stand-replacing and partial harvests. A trained image interpreter recorded the occurrence of a harvest provided spectral changes, image texture, and other contextual cues were consistent with harvesting in the vicinity of the plot and provided harvest operations appeared to have affected the majority of pixels within a 3x3 neighborhood surrounding plot center. Use of a 3x3 majority is consistent with FIA plot configuration (Cooke, 2000). A harvest recorded by

visual interpretation was labeled stand-replacing provided FIA age dated stand origin to 1970 or later (allowing for advance regeneration established prior to 1973) and the field-assigned stand size class was either sapling or poletimber. A substantial fraction of plots sampled multiple forest condition classes, as identified by FIA field crews (U.S. Forest Service, 2012). Age and stand size criteria were required of all sampled conditions. Age estimates were unreliable indicators of disturbance intensity for harvests that occurred after 1999, because age estimates frequently corresponded to a few remaining large stems rather than the new cohort which will eventually dominate the canopy. For plots harvested after 1999, an alternate approach was used to assign reference classes based on repeated plot measurements made during the 1999-2003 and 2004-2008 inventory cycles. Where the 5-year period between plot measurements included a harvest recorded by visual interpretation, that harvest was labeled stand-replacing if plot basal area (cross-sectional area of stems measured at 1.37 m) had been reduced by at least 70%. This basal area removal threshold was identified as that for which errors between stand-replacing and partial harvest classes were best balanced and mapped class extents least biased.

Of 671 FIA plots, we excluded 111 that contained non-forest cover types or field condition classes. We excluded 51 samples where a harvest was interpreted to have occurred after 1999 but the timing of plot measurements did not allow for both pre-harvest and post-harvest assessment of forest conditions. This occurred when the harvest either preceded plot measurement during the 1999-2003 inventory cycle or followed plot measurement during the 2004-2008 inventory cycle. The latter case included all harvests that occurred 2008-2010. The validation sample size of 509 was insufficient to produce reasonably precise estimates of class accuracy for individual time series intervals. We therefore aggregated intervals into the following six harvest validation classes: 1973-1988 stand-replacing harvest, 1988-1999 stand-replacing harvest, 1988-1999 partial harvest, 1999-2010 stand-replacing harvest, 1999-2010 partial harvest, and intact mature forest (no history of harvest, 1973-2010). Map and reference validation class labels for each sample were assigned in a manner consistent with the construction of

cumulative harvest maps. Where multiple entries were mapped or interpreted to have occurred, the corresponding map or reference label was assigned based on the date of the first stand-replacing disturbance. If multiple partial harvests occurred, a validation class label of either 1988-1999 stand-replacing or 1999-2010 stand-replacing was assigned based on the date of the second entry. Reference labels were compared to the pixel locations coincident with plot centers, resulting in a per-pixel validation of harvest data.

Map and reference labels were compiled into an error matrix. Overall accuracy, user accuracy (the complement of class commission error), producer accuracy (the complement of class omission error), and corresponding standard error estimates were calculated using poststratified estimators (Card, 1982; Zhu et al., 2000). Mapped pixel counts were calculated for each of the validation classes, and the validation sample was treated as a random sample stratified by validation class.

Poststratification produces more efficient estimates of overall and producer accuracy than those obtained using formulae for a simple random sample; user accuracy estimates are equivalent (Card, 1982; Stehman, 2009). Additionally, we evaluated the accuracy of our 2010 cumulative harvest map by further aggregating validation classes into regenerating, partially harvested, and intact mature forest.

Forest Type Validation Procedure

FIA plot measurements of coniferous and deciduous live tree basal area were used to derive reference class labels for validation of the 1975 and 2004 forest type maps. Reference labels were compared to pixel locations coincident with plot centers. The 2004 map was validated with data collected during the 1999-2003 inventory cycle. The 1975 map was validated using FIA data collected during the 1980-1982 inventory, which included a large proportion of samples consisting of a single 0.08 ha plot (U.S. Forest Service, 1981). At these locations, multiple forest conditions within a 3x3 pixel neighborhood were less likely to have been sampled than under the modern plot design. To improve the comparability of validation data sets, we removed multiple-condition plots from the 2004 validation

sample. Differences in dates between field data collection and satellite image acquisition resulted in cases where intervening harvests altered forest conditions. For 2004 map validation, we excluded samples where pixels within the 3x3 neighborhood surrounding plot center were mapped as harvested 1999-2004. For 1975 map validation, we excluded samples where neighborhood pixels were harvested 1975-1982. A total of 445 samples remained for validation of the 2004 map. Because accurate plot coordinates are known for only a small subset of plots sampled during the 1982 inventory, only 70 samples were available for validation of the 1975 map.

We identified coniferous-dominant and deciduous-dominant class thresholds for which errors were best balanced and mapped class extents least biased. To do so, we varied coniferous and deciduous threshold values from 50-95% in increments of 5%, assigned reference class labels based on threshold values, and calculated omission and commission error rates. We iteratively refined the maps and reevaluated error rates until a reasonable balance was achieved at the same threshold for both maps, facilitating meaningful comparisons of class extent between maps. Although we initially defined coniferous-dominant and deciduous-dominant classes using a 75% basal area threshold, we were better able to balance commission and omission error after adjusting class thresholds to 80% and 70%, respectively. An error matrix was compiled for each map based on these selected threshold values. Estimates of overall, user, and producer accuracy were calculated by poststratification (Card, 1982; Zhu et al., 2000).

Interpretation of Validation Outcomes

The overall agreement between map and reference harvest validation classes was high (Tables 1.2 and 1.3). The largest source of disagreement was confusion between harvest intensities, rather than confusion between harvest periods. Errors between stand-replacing and partial harvests were balanced for 1999-2010, indicating that the stand-replacing harvest class consistently represented harvests where >70% of basal area had been removed. Errors were similarly well balanced for 1973-1988, but not for

1988-1999. Use of field age and stand size criteria to infer stand-replacing disturbance may have been more appropriate for the clearcutting practices of the 1970s and 1980s than the partial harvest practices of the 1990s. Alternatively, harvest intensity may have been systematically under-represented for the 1990s, perhaps due to cloud cover in several 1990s images (Table 1.1). We adapted the three-date classification method to detect change in cloudy areas using preceding and succeeding images. As a consequence, harvest intensity may have been under-estimated due to regrowth during the longer periods between clear observations. We cannot verify this due to insufficient validation sample sizes. Possible impacts to regenerating forest metrics are unknown, but the alternative of wholly missing harvests due to cloud cover would certainly have affected cumulative harvest area time series and intact mature forest metrics.

Forest type classes for both 1975 and 2004 were mapped with reasonably high overall accuracies (Tables 1.4 and 1.5). Overall accuracy and individual class accuracy estimates were higher for 1975 than for 2004. Although differences between FIA inventory designs and sample sizes complicate comparison, lower accuracies for the 2004 map probably reflect more heterogeneous forest landscape conditions. Off-diagonal entries in both error matrices indicated confusion between the mixed class and both coniferous- and deciduous-dominant classes. There was little confusion between coniferous and deciduous classes. Using coniferous-dominant and deciduous-dominant class thresholds of >80% and >70% basal area, respectively, errors were very well balanced for 2004 forest type classes and reasonably well balanced for 1975. User and producer accuracies for the 1975 map suggested under-representation of coniferous forest area and over-representation of mixed forest under these same class definitions, but the relatively small validation sample and correspondingly large standard error estimates made this inconclusive. Available validation data suggested that user and producer accuracies were best balanced under these class definitions. Note that had more historic field plot locations been available, perhaps we could have balanced errors using the original class thresholds of 75%.

Nonetheless, our validation procedure served the purpose of ensuring that forest type classes in both maps represented the same forest conditions, so that map comparisons were meaningful.

References

- Card DH. Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogramm Eng Remote Sensing* 1982;48:431–9.
- Cohen WB, Fiorella M, Gray J, Helmer E, Anderson K. An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. *Photogramm Eng Remote Sensing* 1998;64:293–300.
- Cohen WB, Yang Z, Kennedy R. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. *Remote Sens Environ* 2010;114:2911–24. doi:10.1016/j.rse.2010.07.010.
- Cooke WH. Forest/non-forest stratification in Georgia with Landsat Thematic Mapper data. In: McRoberts RE, Reams GA, Van Deusen PC, editors. *Proc. First Annu. For. Invent. Anal. Symp.*, St. Paul: US Forest Service, North Central Research Station; 2000, p. 28–30.
- Coulston JW, Riitters KH, McRoberts RE, Reams GA, Smith WD. True versus perturbed forest inventory plot locations for modeling: a simulation study. *Can J For Res* 2006;36:801–7. doi:10.1139/X05-265.
- McRoberts RE, Bechtold WA, Patterson PL, Scott CT, Reams GA. The Enhanced Forest Inventory and Analysis Program of the USDA Forest Service: historical perspective and announcement of statistical documentation. *J For* 2005;103:304–8.
- Sader SA, Bertrand M, Wilson EH. Satellite change detection of forest harvest patterns on an industrial forest landscape. *For Sci* 2003;49:341–53.
- Sader SA, Legaard KR. Inclusion of forest harvest legacies, forest type, and regeneration spatial patterns in updated forest maps: A comparison of mapping results. *For Ecol Manage* 2008;255:3846–56. doi:10.1016/j.foreco.2008.03.047.
- Smith WB. Forest inventory and analysis: a national inventory and monitoring program. *Environ Pollut* 2002;116 Suppl:S233–42.
- Stehman S V. Model-assisted estimation as a unifying framework for estimating the area of land cover and land-cover change from remote sensing. *Remote Sens Environ* 2009;113:2455–62. doi:10.1016/j.rse.2009.07.006.
- Thomas NE, Huang C, Goward SN, Powell S, Rishmawi K, Schleeweis K, et al. Validation of North American Forest Disturbance dynamics derived from Landsat time series stacks. *Remote Sens Environ* 2011;115:19–32. doi:10.1016/j.rse.2010.07.009.
- U.S. Forest Service. Forest inventory and analysis national core field guide, volume1: field data collection procedures for phase 2 plots, Version 6.0. vol. I. Washington DC, USA: US Department of Agriculture, Forest Service, Forest Inventory and Analysis; 2012.

U.S. Forest Service. Resources evaluation field instructions for Maine: 1980-1981. Broomall, Pennsylvania, USA: US Department of Agriculture, Forest Service, Northeastern Experiment Station, Renewable Resources Evaluation Work Unit; 1981.

Zhu Z, Yang L, Stehman SV, Czaplewski RL. Accuracy assessment for the U.S. Geological Survey regional land-cover mapping program: New York and New Jersey Region. *Photogramm Eng Remote Sensing* 2000;66:1425–35.

APPENDIX C: PREPARATION OF SPATIAL PREDICTOR DATA USED IN CHAPTER 3

Our primary source of spatial predictor data for species occurrence mapping was Landsat TM and ETM+ imagery acquired at different times throughout the growing season (late April through early October) (Table C.1). Frequent cloud cover necessitated the collection of imagery across multiple years. We selected eight relatively cloud- and snow-free images spanning a roughly 5-year observation period (2001-2006) to match a full forest inventory field measurement cycle (McRoberts et al., 2005). Landsat images were obtained from the Multi-Resolution Land Characteristics consortium and U.S. Geological Survey Earth Resources Observation and Science Center at 30 m resolution with standard terrain correction applied. Clouds and cloud shadows were masked using a semi-automated procedure developed in-house. Masks were inspected and errors were corrected by on-screen digitization. Bands 1-5 and 7 (visible and reflective infrared) were extracted for further processing as spatial covariates. Visible snow cover in early-season imagery was masked by unsupervised classification using an ISODATA algorithm and visual interpretation of snow-covered classes. Images were converted to top-of-atmosphere reflectance and then corrected for topographic illumination effects using the SCS+C algorithm (Soenen et al., 2005), with slope and aspect calculated from the 1 arc-second (30 m) National Elevation Dataset (NED).

Several additional factors confounded associations between multi-temporal imagery and field plot data. Harvesting during the 5-year observation period dissociated image characteristics from field measurements at affected reference locations. Additionally, rapid vegetation growth following harvests that preceded the observation period introduced spectral variation across images that was due primarily to changes in canopy cover rather than canopy phenology. We therefore masked locations of apparent canopy cover change using available summer, leaf-on images acquired in 2001, 2004, and 2007 (Table C.1). The iteratively-reweighted multivariate alteration detection transformation (IR-MAD; Canty and Nielsen, 2008) was applied to 2001-2004 and 2004-2007 image pairs to estimate a probability of spectral

change during each interval. Intervals were combined by selecting the maximum probability of change, and a probability threshold was selected that resulted in identification of 20% of forest pixels as change pixels. Threshold selection was arbitrary, but visual inspection of the resulting 2001-2007 change mask indicated close correspondence with contemporary disturbance and visible regrowth in previously disturbed stands. Forest and non-forest pixels were differentiated using the 1993 Maine Gap Analysis Program (GAP) land cover map, augmented with the agricultural classes of the 2001 National Land Cover Database (NLCD). The 1993 GAP map differentiated forest from non-forest with an estimated 100% accuracy in our study area (Hepinstall et al., 1999), but incorporation of the 2001 NLCD agricultural classes was necessary to account for a small amount of apparent land cover change.

Table C.1. Landsat images used to model and map tree species occurrence (ca. 2004) and canopy disturbance (2004-2007). Images were acquired over Landsat Worldwide Reference System-2 path 12, row 28. Unless otherwise indicated, images were obtained from the U.S. Geological Survey Earth Resources Observation and Science Center.

Acquisition date	Landsat sensor	Landsat satellite	% forest under cloud/shadow	% forest under snow
Species occurrence:				
April 29, 2006	TM	5	-	1.0
May 12, 2005	TM	5	1.7	0.8
May 25, 2001	ETM+	7	1.2	<0.1
June 10, 2004 ²	TM	5	0.4	-
July 20, 2001 ^{1,2}	TM	5	0.9	-
Sept. 14, 2004	TM	5	0.3	-
Sept. 30, 2001	ETM+	7	-	-
Oct. 6, 2006	TM	5	3.0	-
Canopy disturbance:				
June 10, 2004 ²	TM	5	0.4	-
June 19, 2007 ²	TM	5	9.0	-

¹Available from the Multi-Resolution Land Characteristics consortium.

²Images used to mask spectral change resulting from disturbance and regrowth over the observation period used for species occurrence modeling and mapping. For this purpose only, cloud-contaminated data in the June 19, 2007 image were replaced with data from a Landsat 5 image acquired on Aug. 22, 2007.

Landsat images used for canopy disturbance mapping were acquired during summer leaf-on conditions. Specific images were selected based on patterns of cloud cover. Landsat TM images acquired

in 2004 and 2007 (Table C.1) were obtained from the U.S. Geological Survey Earth Resources Observation and Science Center at 30 m resolution with standard terrain correction applied. No additional steps were taken to prepare imagery for disturbance mapping. Cloud and cloud shadow masks were produced for each image as a basis for evaluating the accuracy of MOSVC in affected areas.

Additional spatial covariates for species occurrence mapping included climate and terrain attributes thought to be relevant to tree establishment or growth. Terrain data included 10 morphometry, 8 lighting/visibility, and 11 hydrology variables (Table C.2) calculated from the 1 arc-second (30 m) NED and the National Hydrography Dataset (NHD) using the freely distributed System for Automated Geoscientific Analyses GIS software, version 2.1.4 (Conrad et al., 2015). The NED was lightly smoothed with a Gaussian filter to reduce the effects of random error and systematic artifacts (circular filter element, radius = 90 m, $\sigma = 1.5$). Terrain slope, aspect, and curvature were calculated from a second-order polynomial fit (Zevenbergen and Thorne, 1987). Direct insolation was calculated at mid-month, April-September, by assuming a uniform 65% atmospheric transmittance, a value that produced insolation estimates in good agreement with a previously published regional climate model (Ollinger et al., 1995). Hydrology variables including catchment area, flow path length, and distance to stream channel were calculated using a bidimensional flow routing algorithm (Quinn et al., 1991) after filling sinks in the NED (Wang and Liu, 2006). Synthetic stream channel networks were derived from the catchment area raster after masking and dilating NHD water bodies using a 5x5 filter element. The dilated water body mask reduced the tendency for channels to initiate near the edges of water bodies, where the flow routing algorithm produced large estimates of flow accumulation. Climate data were obtained from the USDA Forest Service Rocky Mountain Research Station, Moscow Forestry Sciences Laboratory, and included 17 variables (Table C.2) derived from monthly temperature and precipitation surfaces interpolated from weather station data for the climate normal period of 1961-1990 (Rehfeldt, 2006). Climate data were available at approximately 1 km spatial resolution.

Table C.2. Terrain and climate variables used to model and map tree species occurrence (ca. 2004).

Terrain variables were calculated using the System for Automated Geoscientific Analyses (SAGA) GIS software (Conrad et al., 2015) with default settings unless otherwise specified. Climate variables were obtained directly from the USDA Forest Service Rocky Mountain Research Station, Moscow Forestry Sciences Laboratory.

Terrain morphometry (10 variables)		
Elevation		
Slope	Local terrain slope, from fit of second-order polynomial	
Aspect	Local terrain aspect, from fit of second-order polynomial	$\cos(\text{aspect} - 45^\circ) + 1$ (Beers et al., 1966)
Curvature	Local terrain curvature, from fit of second-order polynomial	Tangential, profile, and plan curvature
Topographic position index	Difference between elevation and mean elevation of circular neighborhood (Guisan et al., 1999)	150 m, 300 m, 1000 m, 2000 m neighborhood radii
Lighting/visibility (8 variables)		
Visible sky	Proportion of hemisphere unobstructed by terrain (Häntzschel et al., 2005)	10,000 m search radius
Sky view factor	Ratio of diffuse irradiance to that of an unobstructed horizontal surface (Häntzschel et al., 2005)	10,000 m search radius
Direct insolation	Potential incoming solar radiation	single day estimate at mid-month, April-September; 65% atmospheric transmittance
Hydrology (11 variables)		
Catchment area	Upslope area or flow accumulation	\log_{10} transformed
Catchment height	Difference between elevation and mean elevation of upslope pixels	
Catchment slope	Mean slope of upslope pixels	
Catchment aspect	Mean aspect of upslope pixels	$\cos(\text{aspect} - 45^\circ) + 1$ (Beers et al., 1966)
Flow path length	Mean distance of flow from upslope pixels	
Distance to stream channel	Shortest distance (or distance component) to synthetic stream channel network calculated by flow routing algorithm	overland distance and horizontal, vertical distance components; stream networks from 10 ha and 50 ha flow initiation thresholds

Table C.2 continued

Climate (17 variables)	
d100	Julian date on which the sum of degree-days >5°C reaches 100
dd0	Degree-days <0°C (from mean monthly temperatures)
dd5	Degree-days >5°C (from mean monthly temperatures)
fday	Julian date of the first autumn freeze
ffp	Length of the frost-free period (days)
gsdd5	Degree-days >5°C accumulated over the frost-free period
gsp	Growing season precipitation (April-September)
map	Mean annual precipitation
mat_tenths	Mean annual temperature
mmax_tenths	Mean maximum temperature of warmest month
mmindd0	Degree-days <0°C (from mean minimum monthly temperatures)
mmin_tenths	Mean minimum temperature of coldest month
mtcm_tenths	Mean temperature of coldest month
mtwm_tenths	Mean temperature of warmest month
sday	Julian date of last spring freeze
smrpb	Summer precipitation balance (July+Aug.+Sept. / April+May+June)
smrsprpb	Summer/spring precipitation balance (July+Aug. / April+May)

References

- Beers TW, Dress PE, Wensel LC. Aspect transformation in site productivity research. *J For* 1966;64:691–2.
- Canty MJ, Nielsen AA. Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation. *Remote Sens Environ* 2008;112:1025–36. doi:10.1016/j.rse.2007.07.013.
- Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L, et al. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci Model Dev* 2015;8:1991–2007. doi:10.5194/gmd-8-1991-2015.
- Guisan A, Weiss SB, Weiss AD. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecol* 1999;143:107–22. doi:10.1023/A:1009841519580.
- Häntzschel J, Goldberg V, Bernhofer C. GIS-based regionalisation of radiation, temperature and coupling measures in complex terrain for low mountain ranges. *Meteorol Appl* 2005;12:33–42. doi:10.1017/S1350482705001489.
- Hepinstall JA, Sader SA, Krohn WB, Boone RB, Bartlett RI. Development and testing of a vegetation and land cover map of Maine. Orono, Maine, USA: Maine Agricultural and Forest Experiment Station, University of Maine; 1999.
- McRoberts RE, Bechtold WA, Patterson PL, Scott CT, Reams GA. The Enhanced Forest Inventory and Analysis Program of the USDA Forest Service: historical perspective and announcement of statistical documentation. *J For* 2005;103:304–8.

Ollinger S V, Aber JD, Federer CA, Lovett GM, Ellis JM. Modeling physical and chemical climate of the Northeastern United States for a geographic information system. Radnor, Pennsylvania, USA: USDA Forest Service, Northern Research Station; 1995.

Quinn P, Beven K, Chevallier P, Planchon O. The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrol Process* 1991;5:59–79.

Rehfeldt GE. A spline model of climate for the western United States. Fort Collins, Colorado, USA: USDA Forest Service, Rocky Mountain Research Station; 2006.

Soenen SA, Peddle DR, Coburn CA. SCS+C: A modified sun-canopy-sensor topographic correction in forested terrain. *Geosci Remote Sensing, IEEE Trans* 2005;43:2148–59. doi:10.1109/TGRS.2005.852480.

Wang L, Liu H. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *Int J Geogr Inf Sci* 2006;20:193–213. doi:10.1080/13658810500433453.

Zevenbergen LW, Thorne CR. Quantitative analysis of land surface topography. *Earth Surf Process Landforms* 1987;12:47–56.

BIOGRAPHY OF THE AUTHOR

Kasey Reed Legaard was born in Twin Falls, Idaho on April 26, 1978. He grew up in Idaho and graduated from Pocatello High School in 1995. He attended Idaho State University and graduated with B.S. degrees in Mathematics and Biology in 2001. After moving to Maine with his wife, Kasey completed a M.S. degree in Oceanography from the University of Maine in 2004, with a thesis titled “Temporal variability of satellite-derived chlorophyll and sea surface temperature in the California Current.” This work was subsequently published in the Journal of Geophysical Research. Kasey entered the Doctoral Program in Forest Resources at the University of Maine in 2005. In 2008, he accepted a full-time position within the University of Maine School of Forest Resources as Associate Scientist, and he has since maintained this position through external awards. Kasey is a candidate for the Doctor of Philosophy degree in Forest Resources from the University of Maine in May 2018.