

# Dynamics and Classification of Forest Tree Species Composition in North America

Akane O Abbasi<sup>1</sup>, Rajeev Sahay<sup>2</sup>, Ann Christine Catlin<sup>3</sup>, Christopher G. Brinton<sup>2</sup>, Alain Paquette<sup>4</sup>, Han Y. H. Chen<sup>5,6</sup>, Jingjing Liang<sup>1</sup>

1. Department of Forestry and Natural Resources, Purdue University 2. School of Electrical and Computer Engineering, Purdue University 3. Rosen Center for Advanced Computing, Purdue University 4. Centre for Forest Research, Université du Québec à Montréal, Département des sciences biologiques 5. Faculty of Natural Resources Management, Lakehead University 6. Key Laboratory for Humid Subtropical Eco-Geographical Processes of the Ministry of Education, Institute of Geography, Fujian Normal University

## Introduction

Forest type is commonly the fundamental unit for forest research and management, identified by similar tree species composition and structural characteristics. While historical forest classifications were based solely on experts' subjective opinions, some recent studies attempted quantitative forest classification. However, a majority of them were limited in spatial coverage and represented only one point in time despite the dynamic nature of forest types. Here, we present the first and most updated forest classification map for North America. Underpinned by the spatial ecology cyberinfrastructure (SEC), a new cyber-environment that supports real-time forest classification based on autoencoder neural network and K-means cluster analysis. We created a spatially continuous map of forest types and identified forest types with high conservation priority based on tree species vulnerability to climate change and insect and disease. Furthermore, we present the dynamics of forest classification by comparing current forest types to those in the past.

## Methods

**Data:** We used ground-sourced tree species data recorded in and after 2000 from the Forest Inventory and Analysis, the Cooperative Alaska Forest Inventory, permanent sample plot networks of Canada, and Canada's National Forest Inventory ground plot network. For each plot and species, we calculated importance value, which equally weighs relative abundance and relative dominance. The plot data was then aggregated into a forest area grid map. **Cluster analysis:** We used an autoencoder neural network to reduce data dimensions, and conducted a K-means cluster analysis to define forest types. **Mapping:** We estimated forest type using random forest models with 38 geoclimatic, environmental and anthropogenic predictor variables. **Conservation priority:** We calculated climate change vulnerability score and insect and disease vulnerability score for each forest type and grid based on species scores and species composition in each forest type as well as the proportion of forest types in each grid. **Forest type dynamics:** We defined and mapped forest types of 1970-1999 and compared the spatial changes of forest types over time using randomization tests.

## Results and Discussion

We identified a total of 46 forest types (Fig. 1). Overall mean silhouette width (a measure of between-cluster heterogeneity) was 0.25, indicating an effective separation of forest types. Balanced accuracy of random forest classifiers ranged from 0.53 to 0.95, and F1 score ranged from 0.08 to 0.81. The overall clustering and mapping performance were comparable with similar studies.

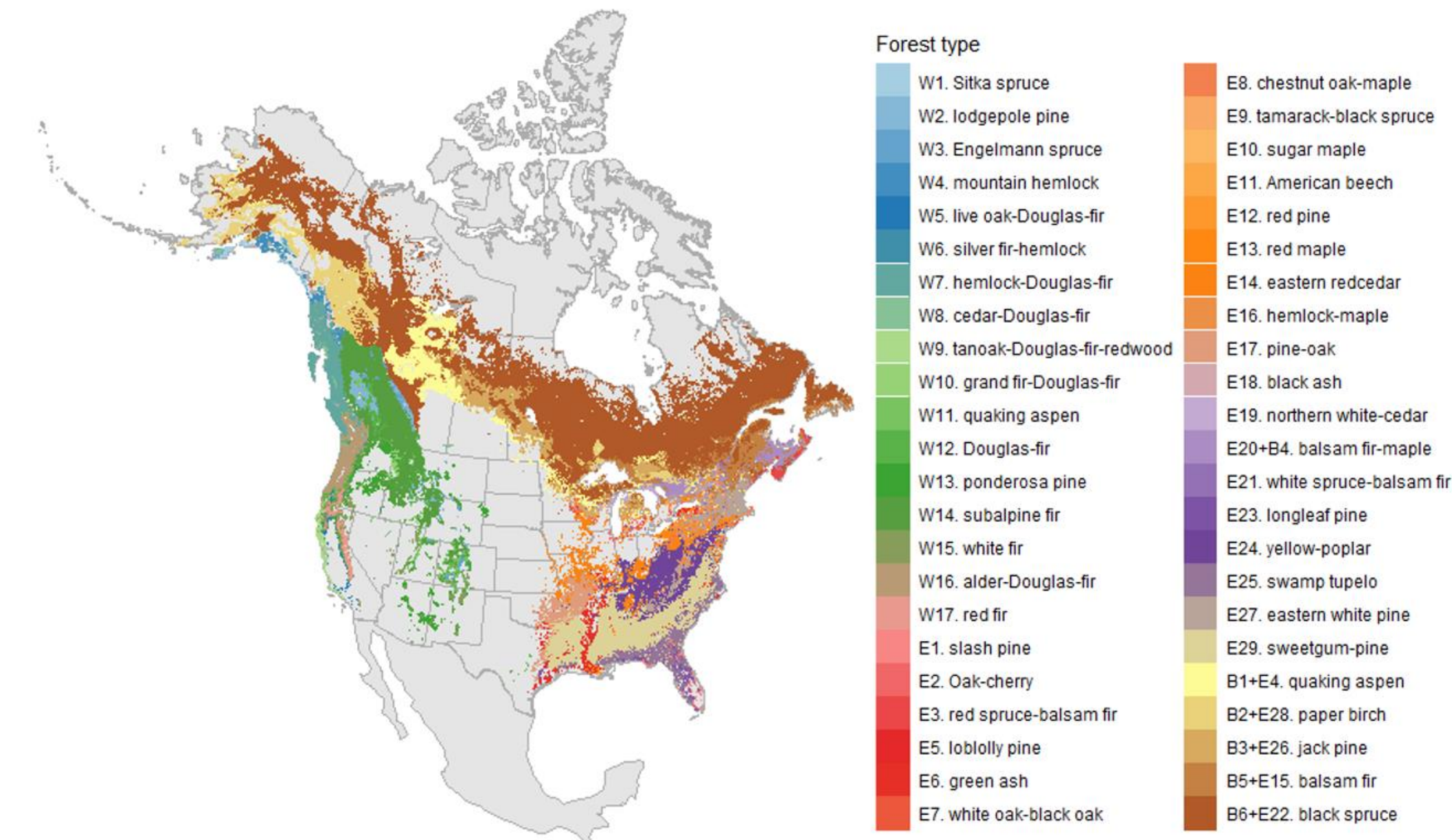


Figure 1. A map of 46 forest types from 2000-2019.

Boreal coniferous forest types ranked high for climate change vulnerability along with some eastern forest types (Fig. 2a). Since rare tree species with highest vulnerability scores were mainly excluded from this study due to small sample size, the species in our vulnerable forest types are potentially

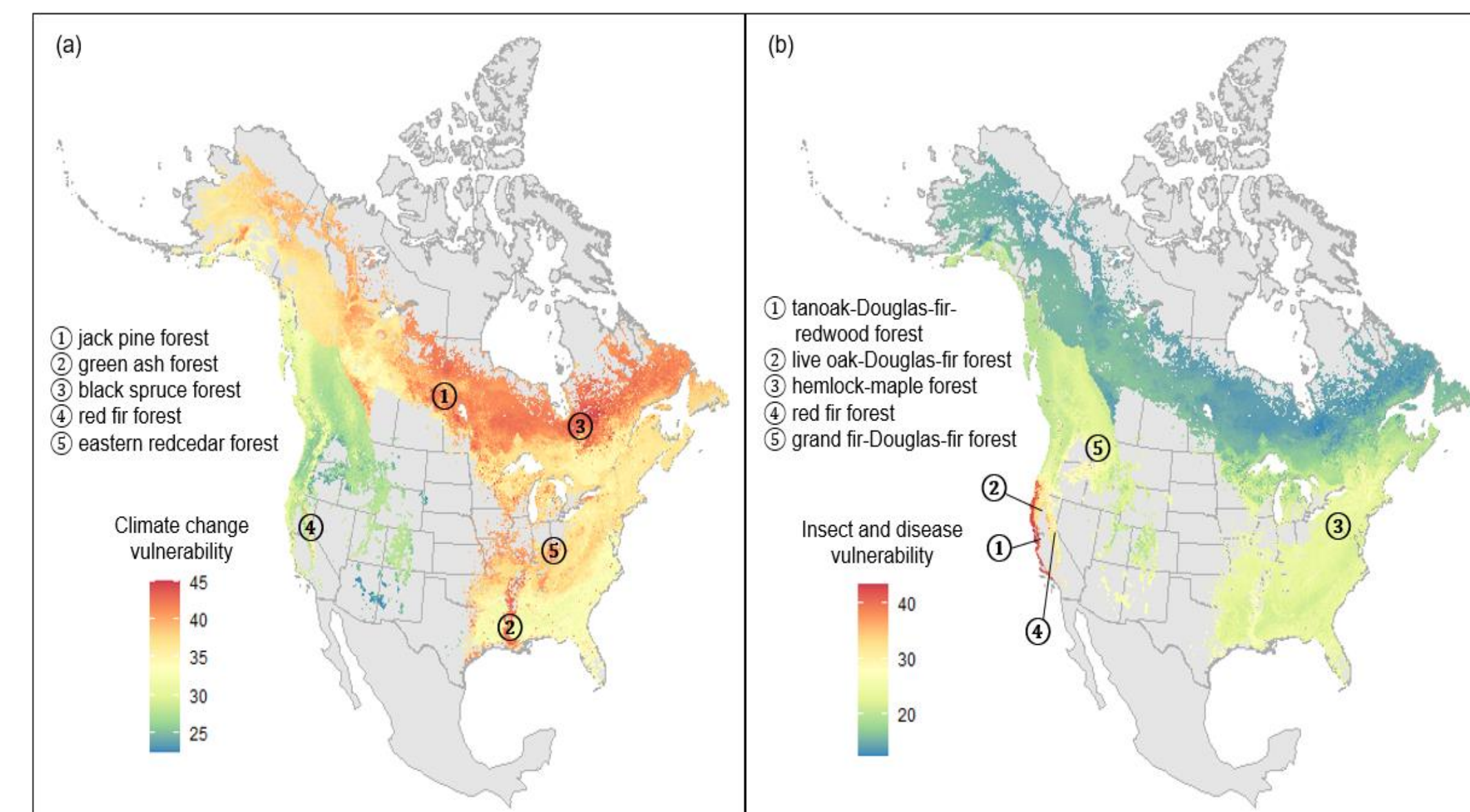


Figure 2. (a) Climate change vulnerability and (b) insect and disease vulnerability. Higher values indicate more vulnerability. The five most vulnerable forest types are marked and named with its ranking.

capable of withstanding climate change due to low sensitivity. However, these species are predicted to migrate northward in the future, changes need to be closely monitored and updated. Many forest types in the west were recognized vulnerable to insect and disease (Fig. 2b). Tanoak-Douglas-fir-redwood forest (ranked 1<sup>st</sup>) and live oak-Douglas-fir forest (2<sup>nd</sup>) have been experiencing sudden oak death, and true fir forests, including red fir forest and grand fir forest, are threatened by root diseases. In the eastern forests, hemlock woolly adelgid and emerald ash borer are the major threats to eastern hemlock and ash trees, respectively. Given that effectiveness of control approaches depends on forest composition and some treatments could alter forest composition, consistent monitoring and updates are critical.

Comparison of the 1970-1999 and 2000-2019 forest types revealed some significant differences in the spatial distribution of corresponding forest types over time (Fig. 3). Westward shifts were largely observed in the eastern forests, and northward shifts in the west. This indicates that the conventional classifications, many of which are still widely in use today, no longer represent the real-time forest type distributions, and future update is undoubtedly in need

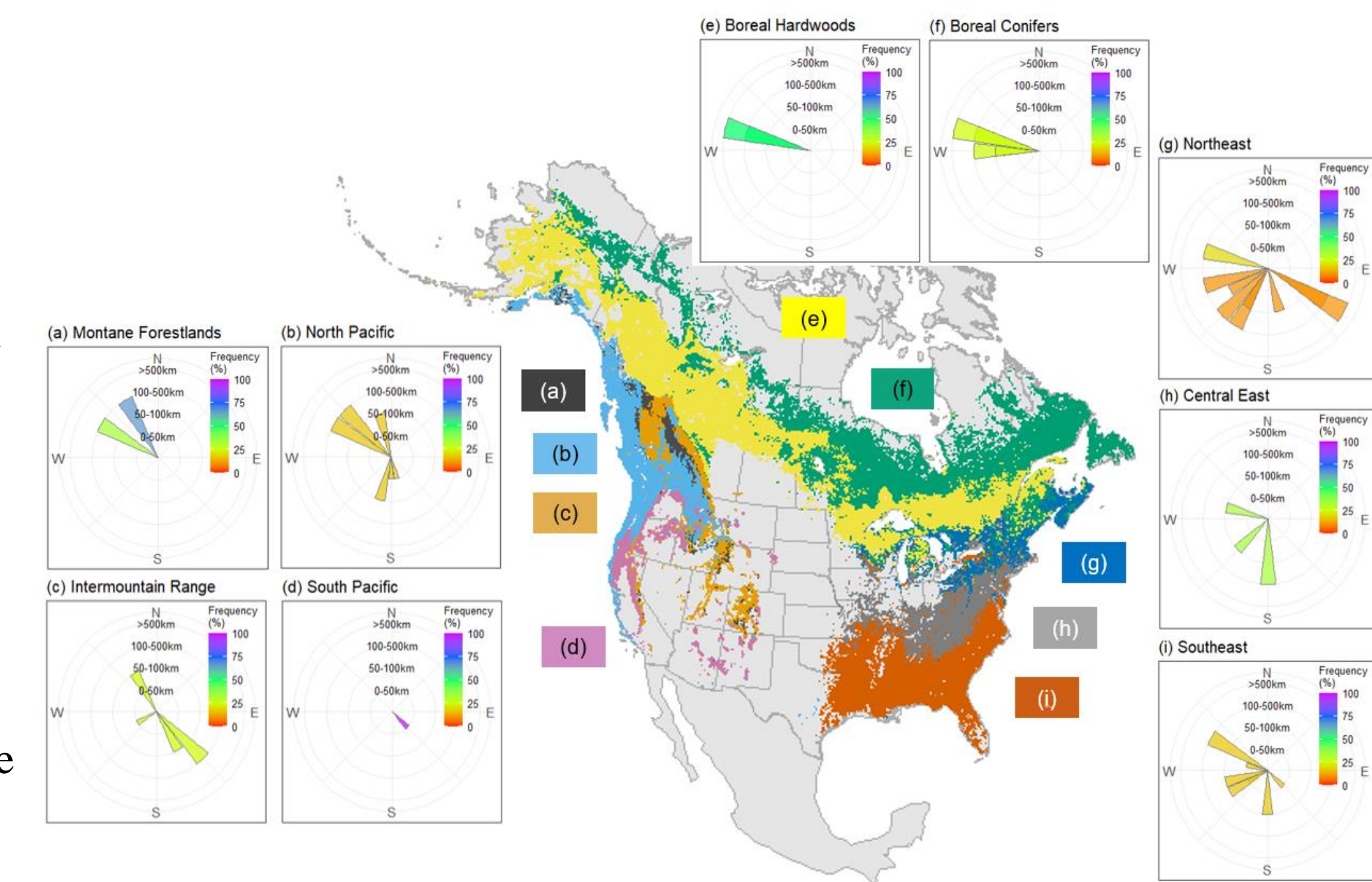


Figure 3. Summary of estimated spatial shifts of forest types within each of nine forest groups from 1970-1999

## Conclusion

Forest classification has long been the fundamental baseline for forest scientists and management communities. As data availability and computational power increase, the quality of quantitative forest classification increases. Compared to the conventional forest classifications, which are mostly available as a paper print, digitized forest classification holds a tremendous amount of details (precise geographic location, species composition, proportion, etc.) and is easier to manipulate for further analyses and applications. Forests are exposed to numerous threats and change their spatial distribution, species composition, and dynamics in the future. Our study highlights that quantitative forest classification is a useful tool to keep track of such ever-changing forests. We hope that more and more forest scientists and foresters use our classification method, available through the spatial ecology cyberinfrastructure (SEC), to understand their target forests as a part of their research and management.

## Acknowledgment

We thank the Global Forest Biodiversity Initiative for facilitating the international research collaboration that leads to the development of this study and J. Knott for providing information on the randomization tests. This study was funded by the Department of Forestry and Natural Resources at Purdue University and Takenaka Scholarship Foundation.

## Contact

albca.liang@gmail.com