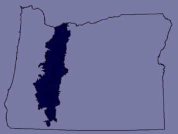
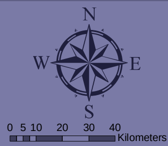
A Tale of Two Techniques:
Gradient Nearest Neighbor and Random Forest
For Mapping Ecological Systems in Oregon's Western Cascades





Emilie Grossmann[1], Janet Ohmann[2], Jimmy Kagan[3], Kenneth Pierce[2], Heather May[1] and Matthew Gregory[1]

[1]Oregon State University, Department of Forest Science, [2]US Forest Service, Corvallis Forestry Sciences Laboratory, [3]Oregon State University, Institute for Natural Resources

Contact: Emilie.Grossmann@oregonstate.edu

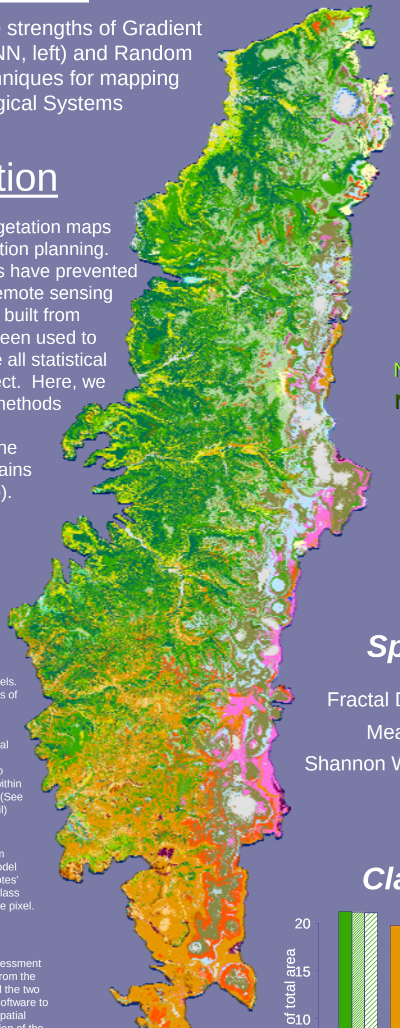
http://www.fsl.orst.edu/lemma/gap

Objective

To assess the relative strengths of Gradient Nearest Neighbor (GNN, left) and Random Forest (RF, right) techniques for mapping Nature Serve's Ecological Systems classification.

Introduction

Large-scale, detailed vegetation maps are needed for conservation planning. However, data limitations have prevented their construction from remote sensing techniques alone. Maps built from statistical models have been used to help fill the need, but like all statistical models, they are imperfect. Here, we compare two statistical methods for building predictive vegetation maps within the western Cascade Mountains of Oregon (shown above).



Methods

Data

We used 6099 inventory and analysis plots within the study area, classified to ecological system, and an array of predictor variables describing climate, topography, disturbance history, and soil parent material to build our statistical models. Models were then projected to spatial grids of the predictor variables.

Gradient Nearest Neighbor

Defines ecological gradients with Canonical Correspondence Analysis. Assigns pixel values (systems, in this case) according to their proximity to the original data points within the space defined by the ordination axes. (See Ohmann and Gregory 2002 for more detail)

Random Forest

Builds many classification trees, each from a subset of the plot and predictor data. Model predictions are generated as each tree 'votes' on a classification for a given pixel. The class with the majority of votes is assigned to the pixel. (See Breiman 2004 for more detail)

Map Assessment

We used error-matrix based accuracy assessment on each map, using 3947 plots withheld from the model-building phase. We also compared the two maps with an error-matrix. We used IAN software to calculate landscape statistics describing spatial patterns. We estimated areal representation of the ecological systems from FIA inventory plots which are a systematic sample of the landscape.

Results

Classification

	GNN	RF
Accuracy	40%	57%
Fuzzy Accuracy	87%	89%
Matrix Systems		
North Pacific Maritime Dry-Mesic Douglas-fir	38%	43%
Mediterranean California Mesic Mixed Conifer	53%	70%
Rare Systems		
North Pacific Dry Douglas-fir	5%	0%
North Pacific Mesic Western Hemlock – Silver Fir	42%	0%

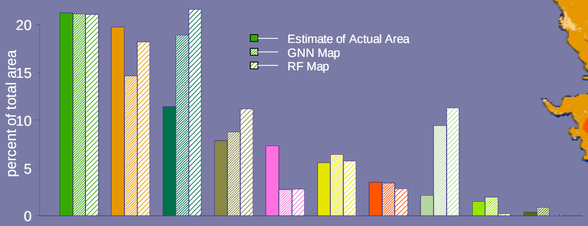
Map Agreement

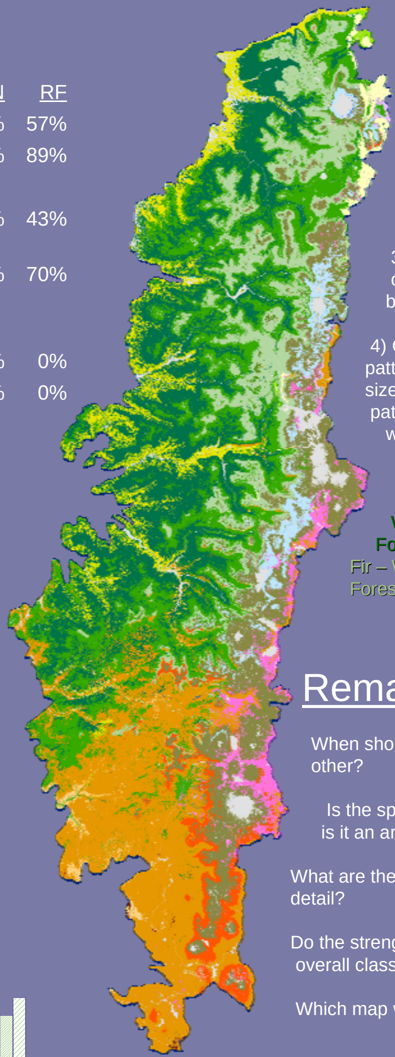
Absolute Agreement	50%
Fuzzy Agreement	93%

Spatial Patterns

	GNN	RF
Fractal Dimension (Box)	1.95	1.95
Mean Polygon Area	26.85	74.81
Shannon Weaver Diversity	2.05	1.65

Class Representation





Conclusions

1) The Random Forest map mapped ecological systems more accurately than the GNN map.

2) GNN represented the minor ecological systems more effectively, while RF consistently showed higher accuracy for the most abundant systems.

3) Although map agreement was only 50%, the disagreements between the two maps were minor.

4) GNN yielded a finer-grained pattern, (smaller patch-size and higher diversity), although patch complexity (fractal dimension) was similar between the two maps

5) Both models over-represented the North Pacific Maritime Mesic Wet Douglas-fir/Western Hemlock Forest, and the North Pacific Silver Fir – Western Hemlock – Douglas-fir Forest.

Remaining Questions

When should one map be preferred to the other?

Is the spatial detail of the GNN map real, or is it an artifact?

What are the values and hazards of spatial detail?

Do the strengths of GNN outweigh its lower overall classification accuracy?

Which map would you use? Why?

References

Breiman, L. 2004. Random Forests. Machine Learning. 45:5-32.

IAN Software: http://landscape.forest.wisc.edu/projects/IAN/

Nature Serve: http://www.natureserve.org/explorer/

Ohmann, J.L. and M.J. Gregory, 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, U.S.A. Canadian Journal of Forest Research 32:725-741.

Acknowledgements

Thanks to the USGS Gap Analysis Program for the funding that enabled this research.

- California Montane Jeffrey Pine-(Ponderosa Pine) Woodland
- Columbia Plateau Western Juniper Woodland and Savanna
- East Cascades Mesic Montane Mixed-Conifer Forest and Woodland
- East Cascades Oak-Ponderosa Pine Forest and Woodland
- Klamath-Siskiyou Lower Montane Serpentine Mixed Conifer Woodland
- Mediterranean California Dry-Mesic Mixed Conifer Forest and Woodland
- Mediterranean California Lower Montane Black Oak-Conifer Forest and Woodland
- Mediterranean California Mesic Mixed Conifer Forest and Woodland
- Mediterranean California Mixed Evergreen Forest
- Mediterranean California Red Fir Forest
- North Pacific Dry Douglas-fir Forest and Woodland
- North Pacific Dry-Mesic Silver Fir-Western Hemlock-Douglas-fir Forest
- North Pacific Hardwood-Conifer Swamp
- North Pacific Lowland Mixed Hardwood Conifer Forest and Woodland
- North Pacific Maritime Dry-Mesic Douglas-fir-Western Hemlock Forest
- North Pacific Maritime Mesic Subalpine Parkland
- North Pacific Maritime Mesic-Wet Douglas-fir-Western Hemlock Forest
- North Pacific Mesic Western Hemlock-Silver Fir Forest
- North Pacific Mountain Hemlock Forest
- North Pacific Oak Woodland
- Northern California Mesic Subalpine Woodland
- Northern Rocky Mountain Ponderosa Pine Woodland and Savanna
- Rocky Mountain Aspen Forest and Woodland
- Rocky Mountain Lodgepole Pine Forest
- Rocky Mountain Poor-Site Lodgepole Pine Forest
- Rocky Mountain Subalpine Dry-Mesic Spruce-Fir Forest and Woodland
- Rocky Mountain Subalpine-Montane Limber-Bristlecone Pine Woodland
- Sierra Nevada Subalpine Lodgepole Pine Forest and Woodland