How does spatial resolution affect Gradient Nearest Neighbor vegetation maps?

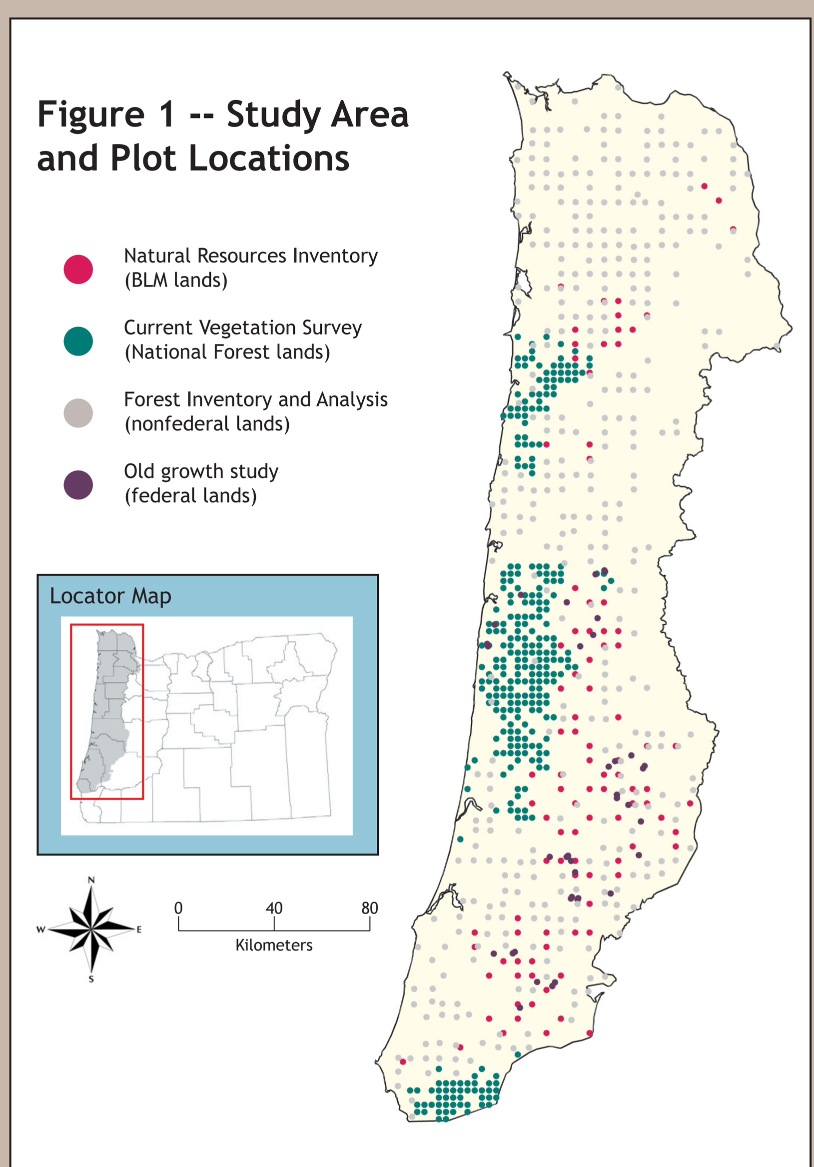
Janet L. Ohmann, Pacific Northwest Research Station, USDA Forest Service, Corvallis, OR 97331 USA
Matthew J. Gregory, Department of Forest Science, Oregon State University, Corvallis, OR 97331 USA

Introduction

Landscape ecologists now have access to tools and data to predictively map forest vegetation across large regions at fine spatial resolution and with great floristic and physiognomic detail. However, the information often must be aggregated and summarized for analysis and interpretation at broad geographic scales. We examined the effects of spatial resolution on vegetation mapped with the Gradient Nearest Neighbor (GNN) method for three million hectares in the coastal province of Oregon, USA (fig. 1).

Our **objectives** were to:

(1) understand how vegetation-environment relationships change with spatial resolution,
(2) quantify the effects of increasing grain-size on regional estimates and local prediction accuracy, and
(3) discuss implications for users of GNN vegetation maps.



The Gradient Nearest Neighbor Method (GNN)

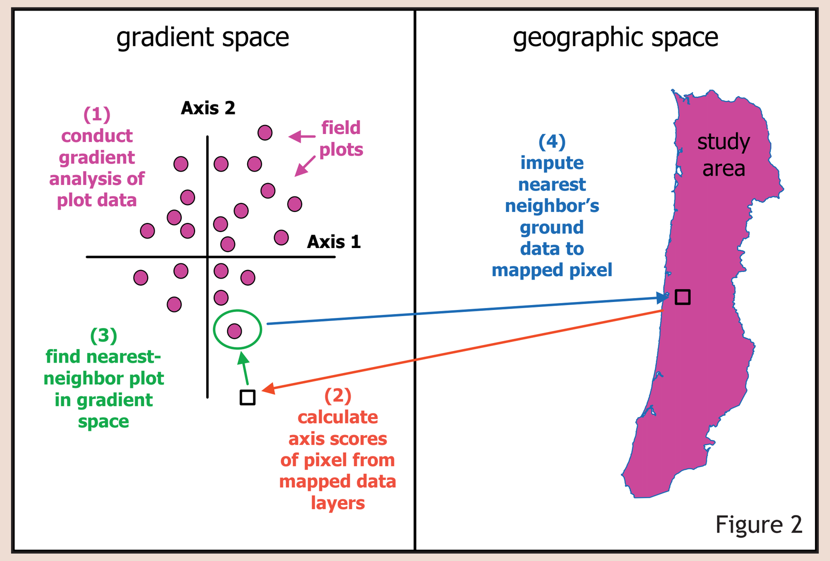


Figure 2

The Gradient Nearest Neighbor (GNN) method integrates vegetation measurements from regional networks of field plots, mapped environmental data, and Landsat TM data to characterize forest vegetation across a region. The method applies direct gradient analysis (canonical correspondence analysis, CCA) and nearest-neighbor imputation to ascribe detailed ground attributes of vegetation to each patch in a regional landscape. Steps in GNN are (fig. 2): (1) Quantify relations between ground and mapped data for plots using CCA. (2) For each mapped pixel, predict scores on the first eight CCA axes from the mapped explanatory variables. (3) For each pixel, identify the single plot that is nearest in eight-dimensional gradient space (Euclidean distances with scores weighted by eigenvalues). (4) Impute the ground attributes of the nearest-neighbor plot to the pixel. Maps then can be constructed for any vegetation attribute measured on the plots.

Methods

- We compared GNN maps at four spatial resolutions: 0.1 ha (25m x 25m), 1 ha (100m x 100m), 9 ha (300m x 300m), 100 ha (1,000m x 1,000m).
- Vegetation data were from 823 field plots established by regional forest inventories (fig. 1). Response variables were basal area by tree species and size-class (0-25, 25-50, 50-75, 75-100, and >100cm DBH).
- Explanatory variables were from GIS grids representing topography, geology, climate, and Landsat imagery (Table 1). We resampled the original 0.1-ha grids to 1 ha, 9 ha, and 100 ha (fig. 3). Larger gridcells were assigned mean values of the 0.1-ha gridcells for continuous variables, and majority values for categorical variables. Values from the resampled grids were assigned to each plot location for analysis.
- We developed stepwise CCA models and GNN vegetation maps at each spatial resolution using the resampled input grids.
- We evaluated the effects of increasing grain-size on: (1) vegetation-environment associations, (2) landscape proportions of predicted vegetation classes, (3) representation of the full range of variability in the predictions, and (4) prediction accuracy at local sites (using an independent set of field plots).

Figure 3 -- Resampling effects on two explanatory variables

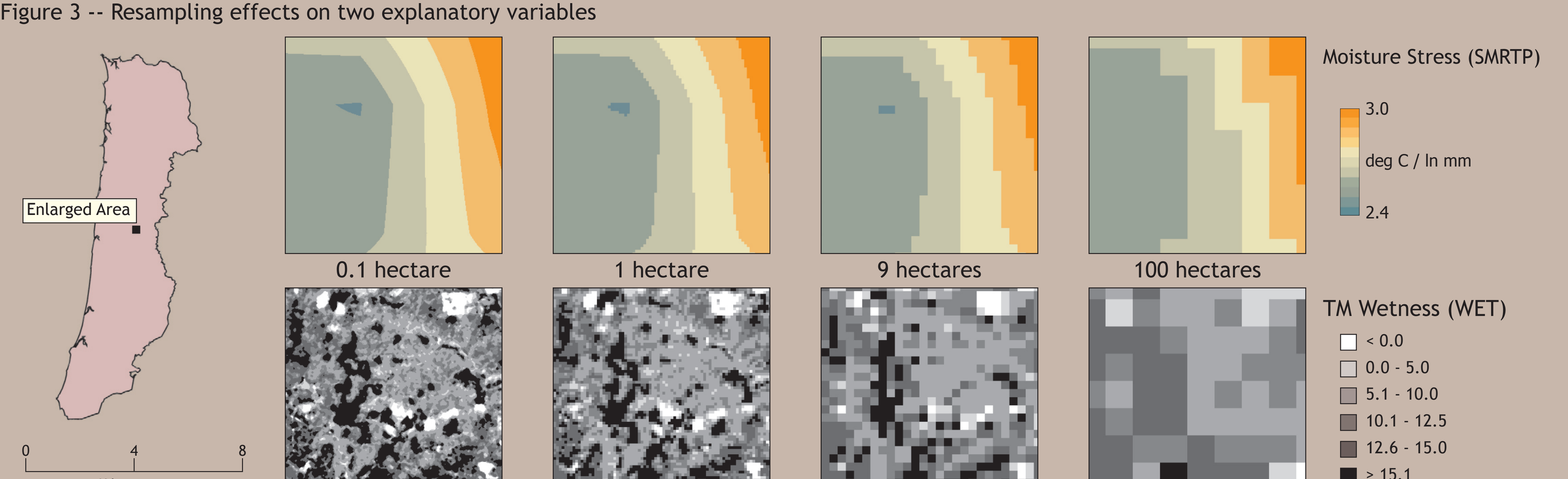


Table 1 -- Explanatory variables selected in stepwise canonical correspondence analysis models

Class	Code	Definition	Class	Code	Definition
Ownership	PUB	Ownership (public or private)	Topography (30-m DEM)	ELEV SLOPE SLPOS SOLAR	Elevation (m) Slope (percent) Slope position, from 0 (bottom) to 100 (ridgetop) Solar radiation (cal/cm²), from program SOLARRAD
Geology	VOLC MAFO	Igneous: volcanic and intrusive rocks Igneous: mafic rocks--miocene and older			
Climate (PRISM model)	SMRPRE CVPRE SMRTP AUGMAXT DIFTMP STRATUS	Mean precipitation from May-September (nat. log, mm) Coefficient of variation of December and July precipitation Growing-season moisture stress (SMRTMP/SMRPRE) Mean maximum temperature in August (C) August max. temperature - December min. temperature (C) Marine stratus ceiling <1,524 m and visibility <8 km (%)	Landsat TM	B2 B3 B4 BRT WET R43 R57 DISTURB	Band 2 (blue) Band 3 (red) Band 4 (near-infrared) Brightness (axis 1 from tasseled cap transformation) Wetness (axis 3 from tasseled cap transformation) Ratio of band 4 to band 3 Ratio of band 5 to band 7 Disturbance (yr), from multitemporal Landsat (Cohen et al.)
Location	X Y	Longitude (decimal degrees) Latitude (decimal degrees)			

Results

Explanatory power of gradient models

- Total variation explained held constant in the 0.1-m, 1-ha, and 9-ha CCA models, but declined in the 100-ha model (fig. 4).
- Contributions of classes of variables to explained variation in CCA changed only slightly or not at all with spatial resolution (fig. 4). The importance of Landsat declined in the 100-ha model.

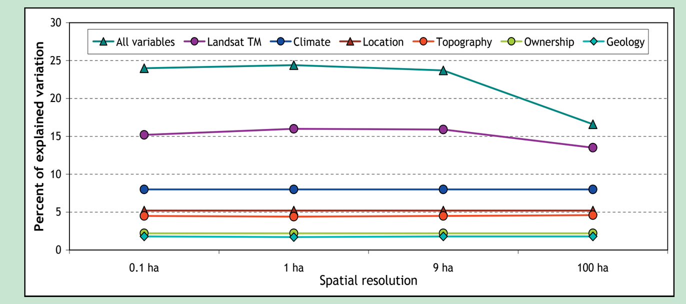


Figure 4 -- Variation explained in partial CCA (see Table 1 for variable definitions)

Dominant gradients in vegetation and environment

- Dominant gradients were the same in the 0.1-ha, 1-ha, and 9-ha CCA models, but axis 1 and 2 were switched in the 100-ha model.
- For all but the 100-ha model, CCA axis 1 reflected variation in forest structure (tree size) and was associated with Landsat TM and ownership (fig. 5). Axis 2 was a gradient in species composition, associated with the maritime climatic gradient.

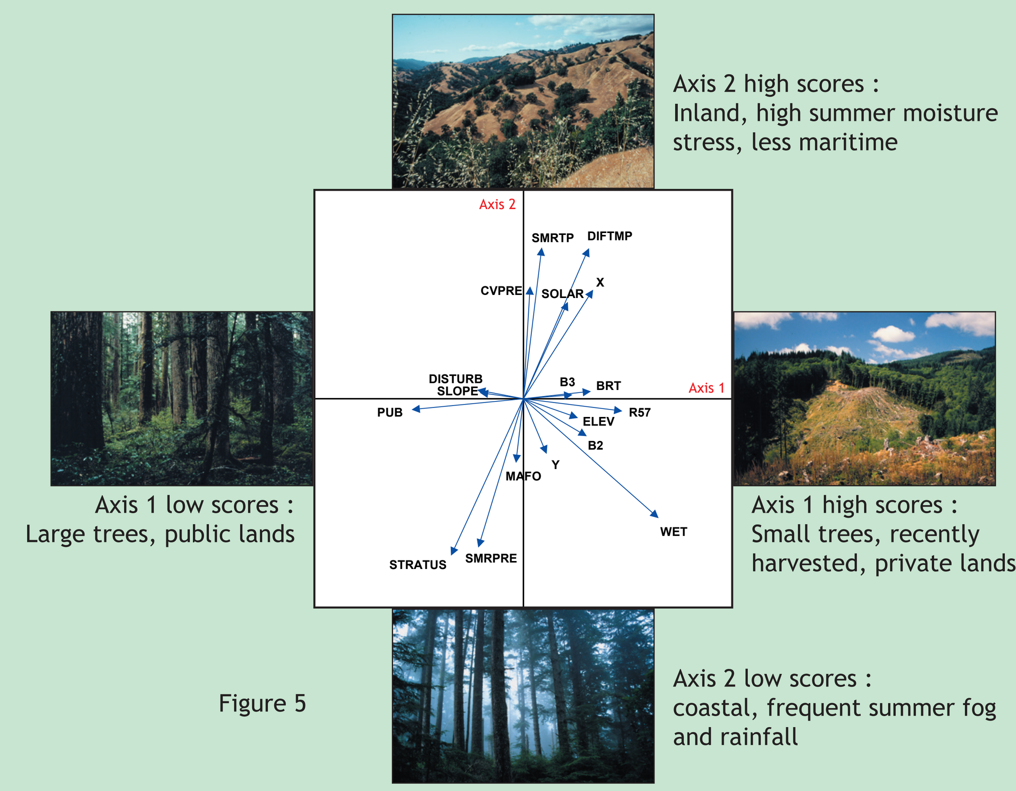


Figure 5

Table 2 -- Descriptive statistics for observed (n = 823 plots) and predicted (mapped) vegetation at four spatial resolutions

Vegetation attribute	Spatial resolution	Mean	Standard deviation
Total basal area (BA) (m^2/ha)	Plots 0.1 ha 1 ha 9 ha 100 ha	33.9 31.0 31.3 33.3 34.8	20.6 22.3 22.1 21.1 21.0
Broadleaf proportion (proportion of BA)	Plots 0.1 ha 1 ha 9 ha 100 ha	0.27 0.26 0.26 0.32 0.27	0.32 0.32 0.31 0.31 0.32
Quadratic mean diameter (cm)	Plots 0.1 ha 1 ha 9 ha 100 ha	34.6 33.2 32.9 34.0 35.3	22.4 24.6 23.8 23.4 23.6
Trees per hectare > 100 cm dbh	Plots 0.1 ha 1 ha 9 ha 100 ha	3.0 3.0 2.9 3.0 3.2	7.5 7.7 7.5 7.7 8.1

Regional landscape characteristics

- At the aggregate, regional level, the mapped GNN predictions very closely matched the means and ranges of variability present in the plot data at all grain-sizes (Table 2).
- As grain-size increased, the landscape proportion of 'open' and 'small conifer' decreased, 'medium mixed' and 'medium conifer' increased, and other classes were unchanged (fig. 6).
- The resemblance of landscape composition as predicted by GNN to estimates from a semi-independent sample of field plots varied among the vegetation classes (fig. 6). Resemblance declined with increasing grain-size for the 'open' class. Resemblance was greatest at 1 ha for 'medium mixed' and 'medium conifer,' and at 9 ha for 'small conifer.' Resemblances of other classes were unaffected by resolution.

Figure 6 -- Landscape proportions by vegetation classes* for plot-based estimates and GNN predictions

* Open: < 1.5 m^2/ha basal area; broadleaf: 65-100% of basal area is broadleaf; mixed: 20-64% broadleaf; conifer: 0-20% broadleaf; small: 0-25 cm DBH; medium: 25-50 cm DBH; large: 50-75 cm DBH; very large: >75 cm DBH.

Local prediction accuracy

- Prediction accuracy at the site level declined with increasing grain-size for all vegetation measures (fig. 7). Pearson correlation coefficients between ground-measured and GNN-predicted vegetation were 33-54% worse for the 100-ha model than for the 0.1-ha model.
- Prediction accuracy for 10 vegetation classes (see fig. 6) declined with increasing grain-size (fig. 7).
- In contrast, prediction accuracy for seven tree species was worst at the 0.1-ha resolution (fig. 8). The relationship between accuracy and grain-size differed among species.

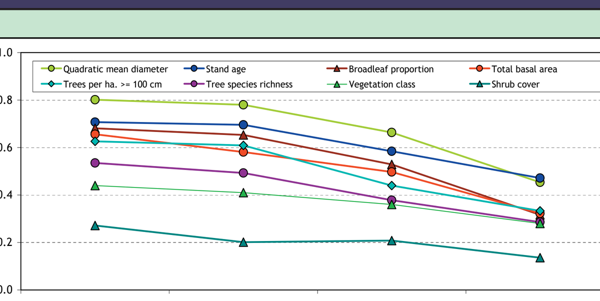


Figure 7 -- GNN prediction accuracy for vegetation attributes for n=823 plots (correlation between observed (ground data) and GNN-predicted values, except vegetation class shows proportion correctly classified -- see fig. 6 for class definitions)

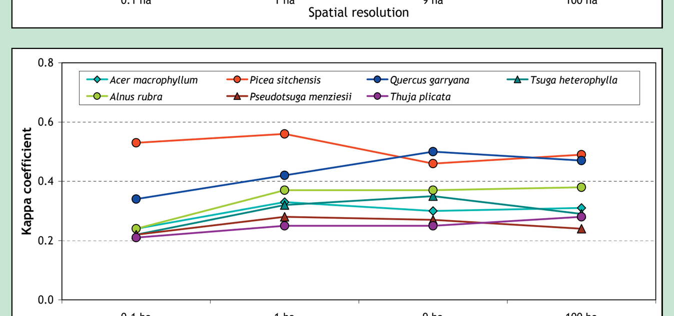
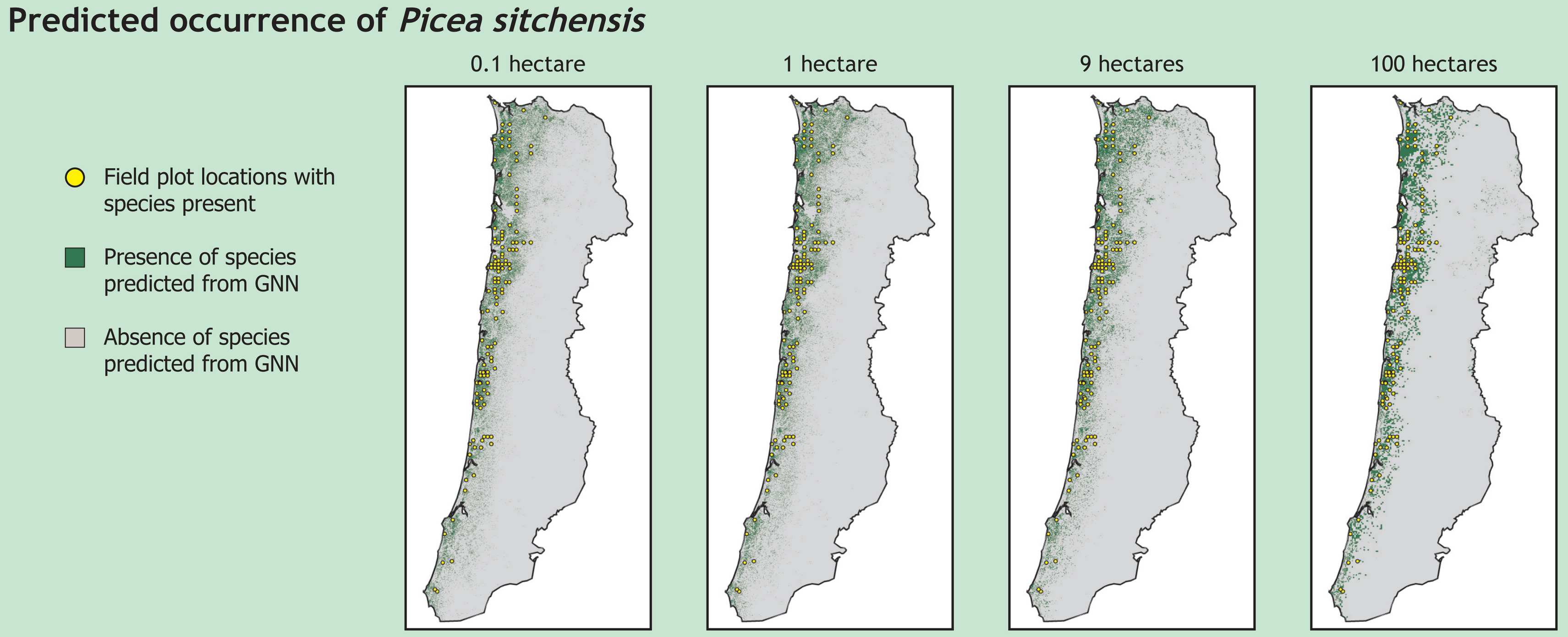


Figure 8 -- GNN prediction accuracy for seven tree species for n=823 plots (Kappa coefficient of agreement between observed (ground data) and GNN-predicted values)

Predicted vegetation classes

Predicted occurrence of *Picea sitchensis*



Conclusions

- The direct gradient analysis models underlying the GNN method are robust to changes in spatial resolution up to 9 ha. At the coarsest resolution examined (100 ha), Landsat data became less important in the model and gradient interpretation changed slightly to favor species composition over forest structure.
- Users concerned primarily with general regional patterns in vegetation can feel confident using GNN maps developed at coarser resolutions, at least up to 9-ha. The coarser-resolution maps represented the range of variability present in the ground sample, but landscape proportions of some vegetation types were altered.
- If site-level accuracy is an important objective for a vegetation map, the 0.1-ha GNN model should be used for most vegetation attributes. Users interested solely in mapping species ranges could consider the intermediate grain-sizes.
- The GNN method applied at coarser resolutions executes much more quickly and requires less disk storage, and can be more readily run over broader geographic extents.
- Our findings on effects of spatial resolution on GNN should be generalizable to other forested regions where similar spatial and plot data are available. Choice of an appropriate spatial resolution will depend on the user's objectives, accuracy requirements, and budget.

For more information on GNN . . .

Ohmann, J.L.; Gregory, M.J. In review. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in the coastal province of Oregon, USA. Canadian Journal of Forest Research. (Submitted February 2001)

Website : http://www.fsl.orst.edu/clams/gnn/index.html

Acknowledgements

This research was funded by the USDA Forest Service, Pacific Northwest Research Station, as part of the Coastal Landscape Analysis and Modeling Study (CLAMS).