A MULTIVARIATE APPROACH TO MAPPING FOREST
VEGETATION AND FUELS USING GIS DATABASES, SATELLITE
IMAGERY, AND FOREST INVENTORY PLOTS

Michael C. Wimberly *
University of Georgia, Athens, Georgia

Janet L. Ohmann, Kenneth B. Pierce Jr.
USDA Forest Service, Pacific Northwest Research Station, Corvallis, Oregon

Matthew J. Gregory
Oregon State University, Corvallis, Oregon

Jeremy S. Fried
USDA Forest Service, Pacific Northwest Research Station, Portland, Oregon

1. INTRODUCTION

Knowing the types and amounts of fuels at a site is an important prerequisite to evaluating fire risk, predicting fire behavior, and assessing potential fire effects. When these assessments are expanded to larger extents, the spatial configuration of the landscape fuel mosaic must also be considered (Keane et al. 2000: Keane et al. 1998). Spatial patterns of fuels, topography, and wind all interact to influence fire spread and intensity. Fuels are of particular interest from a management perspective because they can be modified through mechanical treatments or prescribed burning. Accurate maps of existing fuel patterns can help managers develop improved assessments of fire risk and target fuel reduction efforts more efficiently. In order for fuel maps to be useful to land managers they must accurately represent the spatial pattern of fuels across large, heterogeneous landscapes, while simultaneously and consistently mapping multiple fuel and vegetation variables.

Aerial photographs and satellite imagery have been used extensively to map forest land cover. However, the limited number of vegetation types mapped in classified images may not be sufficient to capture the continuous spatial variability of multiple forest attributes. An alternative to

traditional image classification methods is the predictive vegetation mapping approach (Franklin 1995). In this approach, statistical models are used to predict vegetation characteristics measured in the field using remotely sensed imagery and GIS layers describing climate, topography, and other biophysical variables. These mixed models provide a more detailed characterization of forest vegetation patterns than can be obtained independently from either remotely sensed imagery or biophysical data. Other types of spatial variables may also indirectly contribute to predicting vegetation and fuels. For example, ownership may be a predictive variable if different landowners implement different forest management regimes that produce distinctive fuel patterns.

Emerging imputation-based statistical methods can generate simultaneous spatial predictions of multiple vegetation attributes (Moeur and Stage 1995: Ohmann and Gregory 2002: Tuominen et al. 2003). In these approaches, a statistical model is used to choose a representative field plot from a set of sampled locations to characterize each unsampled area in the landscape. Multiple vegetation attributes from the selected plot are then assigned to each unsampled location. The advantages of this method are that it maintains the true correlation structure among the predicted variables, and also retains the full range of variability for each of the predicted variables (Moeur and Stage 1995).

The Gradient Nearest Neighbor (GNN) method, which combines a multivariate imputation

---

* *Corresponding author address*: Michael C. Wimberly, Univ. of Georgia, Warnell School of Forest Resources, Athens, GA 30602; email: wimberly@forestry.uga.edu

model with the predictive vegetation mapping strategy described previously, has been successfully applied in the Oregon Coast Range to map multiple forest structure variables measured on regional forest inventory plots (Ohmann and Gregory 2002). The main objective of this research was to explore the feasibility of mapping forest fuels in the Coast Range using the GNN approach. We used GNN to generate preliminary spatial models of several forest vegetation attributes related to fire risk, fire spread, and fire effects. An accuracy assessment was carried out to quantify the uncertainty of these spatial predictions. Patterns of forest structure and fuels were mapped to explore the spatial variability of the predicted vegetation attributes.

2. METHODS

The Oregon Coastal Province encompasses 3 million ha in western Oregon (Figure 1). Physiography is characterized by highly dissected terrain with steep slopes and high stream densities. Climate is generally wet and mild, with most precipitation falling between October and March. Major conifer species include Douglas-fir (*Pseudotsuga menziesii*), western hemlock (*Tsuga heterophylla*), and Sitka spruce (*Picea sitchensis*). Red alder (*Alnus rubra*) and bigleaf maple (*Acer macrophyllum*) often dominate on moist sites and in riparian areas. Ownership patterns reflect a legacy of large catastrophic wildfires, with many federal and state forests located on the sites of historical burns. In other areas, ownerships are distributed in a checkerboard pattern that is a relic of historical railroad land grants.

Vegetation data was collected between 1994 and 1997, and included 1300 field plots from the Natural Resources Inventory, encompassing Bureau of Land Management lands; the Current Vegetation Survey, encompassing the Siskiyou and Siuslaw National Forests; the Forest Inventory and Analysis Program, encompassing nonfederal lands; and an old growth survey that was carried out on federal lands. These data included dbh, height, and crown ratio measurements for all live trees; diameter, height/length, and decay class measurements for snags and down wood; and estimates of understory vegetation cover. Live tree structure was summarized as total basal area by species- and size-class. Dead tree structure was summarized as the total standing dead and down dead biomass. These structural indices were used as the vegetation response variables to develop the GNN model.

Spatial predictor variables from a variety of sources were incorporated into an integrated raster GIS database with a 30 m spatial grain. Satellite imagery was obtained from Landsat TM scenes collected in 1996, and included raw band values and tasseled cap transformations. Locations of recent timber harvests derived from multi-date analysis of satellite imagery (Cohen et al. 2002) were also incorporated into the database. Spatial climate data produced by the DayMet model were used to compute mean conditions and seasonal variability in climate (Thornton et al. 1997). Digital elevation models were used to compute elevation, slope, aspect, and solar radiation. A land ownership map was used to identify privately- versus publicly-owned lands.

The following steps were used to develop a predictive map of forest structure using the GNN method (Ohmann and Gregory 2002):

1. Direct gradient analysis was conducted using stepwise canonical correspondence analysis (CCA) (ter Braak and Prentice 1988) to develop a multivariate model that quantified the relationship between forest structure indices measured on the field plots and the mapped explanatory variables. Based on this CCA model, each plot received a set of axis scores that identified its location in an abstract, multidimensional spectral and environmental space.

2. For each pixel, scores for statistically significant CCA axes were predicted by applying coefficients from the model developed in step (1) to the mapped values for the explanatory variables. Thus, each pixel was projected into the same multidimensional space as the ground plots.

3. For each mapped pixel, the single plot that was nearest in eight-dimensional gradient space was identified. Distances were Euclidean and axis scores were weighted by their eigenvalues. Note that this distance was an abstract gradient distance, not the geographic distance between the plots. This step effectively matches each pixel with the ground plot that has the most similar suite of spectral and environmental characteristics. The specific spectral and environmental variables used in this analysis have already been selected and weighted in step 1 for maximum correlation with forest structure.

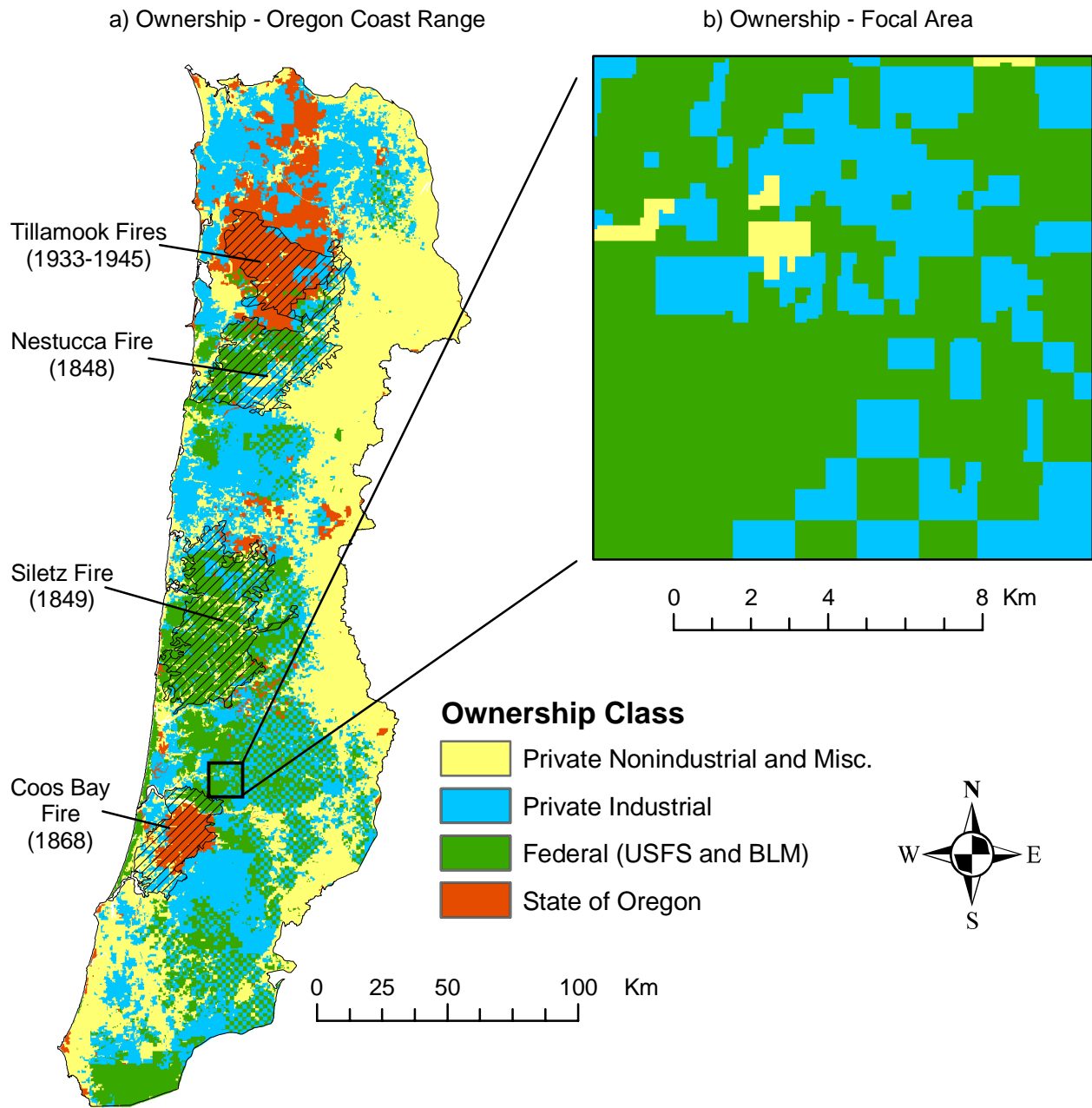4. The ground attributes of the nearest-neighbor plot were imputed to each mapped pixel.

a) Ownership - Oregon Coast Range

Tillamook Fires
(1933-1945)

Nestucca Fire
(1848)

Siletz Fire
(1849)

Coos Bay
Fire
(1868)

b) Ownership - Focal Area

0   2   4   8   Km

Ownership Class

Private Nonindustrial and Misc.

Private Industrial

Federal (USFS and BLM)

State of Oregon

0   25   50   100   Km

N
W   E
S

Figure 1: Land ownership maps of a) the Oregon Coast Range, and b) the focal area in the Central Coast Range. Hatched areas delineate the approximate boundaries of several large historical wildfires.

Each pixel is thus linked to all of the measured field data from the associated ground plot

After the GNN model was used to assign a representative inventory plot to each pixel in the landscape, fuel variables were computed from the imputed field data. Four of the forest canopy attributes required for input into the FARSITE model (Finney 1998) were computed for each pixel: canopy bulk density, height to crown base, stand height and canopy cover. Canopy bulk density was computed using methods similar to those of Cruz et al. (2003), based on published crown weight equations for Pacific Northwest tree species (Brown 1978: Snell and Brown 1980: Snell and Anholt 1981: Snell and Little 1983). Down dead fuel and snag fuel loadings in the > 1000 hour (> 20 cm diameter) time-lag class were also calculated. Although these larger fuels do not directly influence the behavior and intensity of the active fire front, they may contribute to fire effects and emissions through residual smoldering combustion (Bertschi et al. 2003).

Cross validation was carried out by comparing observed values for each plot with the values from its second-nearest neighbor in the ordination space. This approach has been shown to generate similar results to more traditional cross-validation approaches where individual plots or subsets of plots are eliminated from the dataset and then predicted using the model (Ohmann and Gregory 2002) Spatial patterns of mapped fuels variables were graphically displayed for a 150 km^2 focal area located in the central Coast Range just north of the Siuslaw River (Figure 1).

3. Results

The underlying CCA model explained 34.7% of the total variability in forest structure based on satellite imagery, biophysical variables, and land ownership (public versus private). Variables derived from Landsat TM imagery accounted for the majority (20.6%) of the explained variation, predicting spatial variability in live tree size and density. Land ownership was also related to tree size and density. Climate and topography were predominantly associated with spatial variability in tree species composition. Recent timber harvest activity was related to the biomass of snags and down wood.

Accuracy of the predicted canopy variables was highest for stand height, followed by canopy cover, height to base of canopy, > 1000 hour snag

fuels, canopy bulk density, and > 1000 hour down dead fuels (Figure 2). The prediction errors for stand height and height to base of canopy were relatively consistent across the range of values examined. Plots with crown bulk density greater than 0.3 kg/m^3 were under predicted, whereas several plots with crown bulk density between 0.1 and 0.2 kg/m^3 were over predicted. Errors in the prediction of crown cover were much higher for plots with less than 50% than for plots with higher crown cover. Errors in the predictions of large down wood and snag fuel loads were high across all levels of these variables.

When the predictive maps were viewed at a landscape scale within the 150 km^2 focal area, two different types of spatial variability could be discerned (Figure 3). Height to base of canopy, stand height, and large snag fuel loads exhibited coarse-scale spatial patterns that reflected the boundaries of major land ownership types, and the distribution of past forest management activities within these ownerships. In contrast, canopy bulk density, canopy cover, and large down wood fuel loads exhibited finer-grained spatial patterns in which the land ownership mosaic was much less apparent.

4. Discussion

This preliminary assessment demonstrated the feasibility of using the GNN method to develop predictive maps of forest fuels based on satellite imagery, biophysical data, and ownership data. The distinctive patterns generated for the different fuel variables emphasize the importance of making continuous, independent predictions of multiple vegetation attributes, rather than linking all the variables to a single classification scheme.

The accuracy of the predictions varied considerably among the variables tested. It is important to note that the preliminary CCA model used to generate these maps was a general model based on live tree basal area and dead tree biomass. Predictions of canopy fuel loads are therefore based on second-order relationships between the modeled variables and canopy fuel variables that were derived post-hoc from the imputed plot attributes. We expect that prediction accuracy can be improved by incorporating canopy variables as response variables in the underlying multivariate model, thereby optimizing the model for predictions of canopy fuels and canopy structure. Improving predictions of dead fuel loads will be more difficult. Down logs and
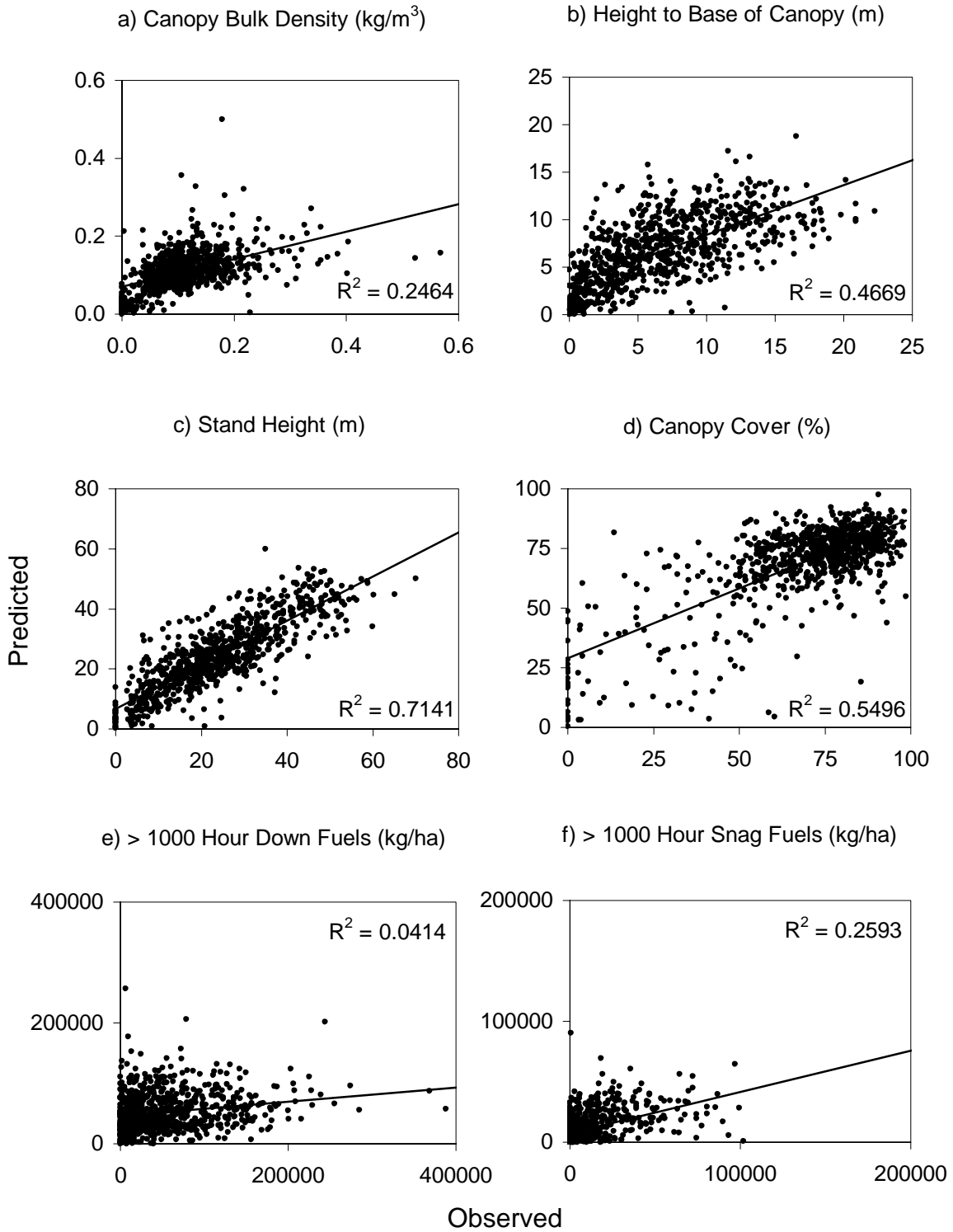
Figure 2: Cross-validation of forest structure and fuel variables mapped using the GNN method. Observed values measured on the field plots(x-axis) are plotted against predicted values generated through a second nearest-neighbor analysis of the GNN model (y-axis).
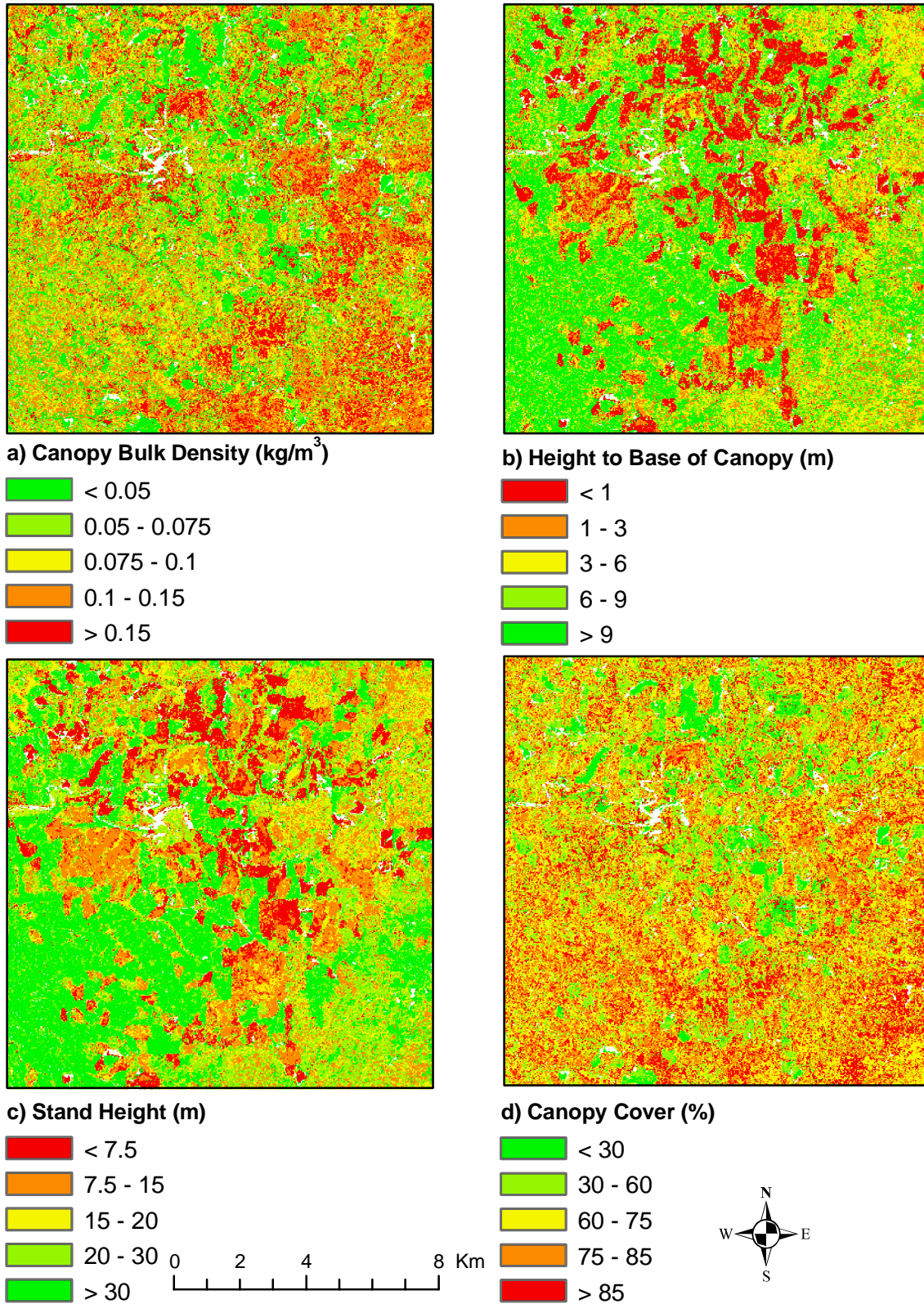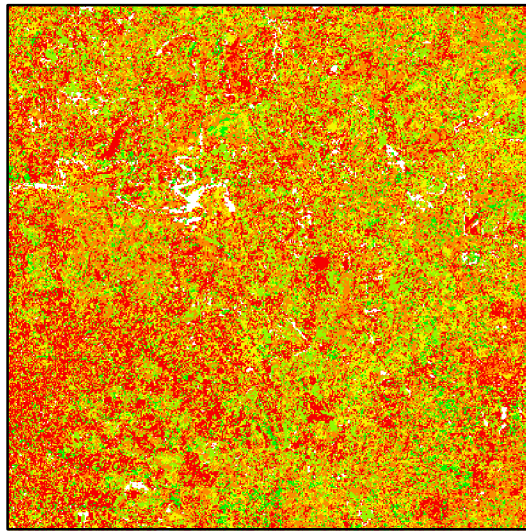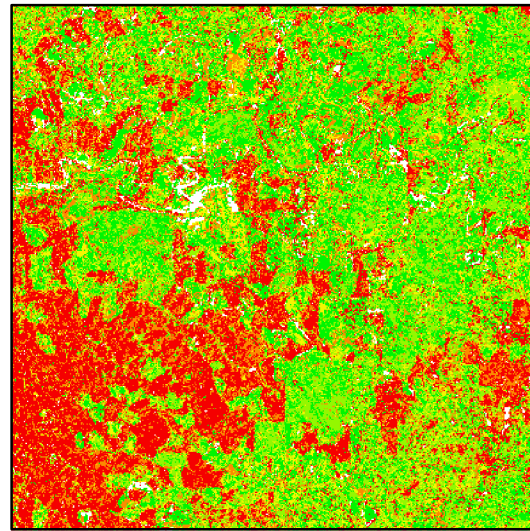
a) Canopy Bulk Density (kg/m^3)

🟩	< 0.05
🟨	0.05 - 0.075
⬜	0.075 - 0.1
🟧	0.1 - 0.15
🟥	> 0.15

b) Height to Base of Canopy (m)

🟥	< 1
🟧	1 - 3
🟨	3 - 6
🟩	6 - 9
🟩	> 9

c) Stand Height (m)

🟥	< 7.5
🟧	7.5 - 15
🟨	15 - 20
🟩	20 - 30
🟩	> 30

d) Canopy Cover (%)

🟩	< 30
🟨	30 - 60
⬜	60 - 75
🟧	75 - 85
🟥	> 85

0   2   4   8 Km

Figure 3: Spatial patterns of fuels and forest structure within a 150 km^2 landscape located in the Oregon Coast Range. Predictive vegetation maps were generated using the Gradient Nearest Neighbor model. White patches represent non-forested areas.

e) >1000-hr Down Fuels (kg/ha)

- < 2,500
- 2,500 - 15,000
- 15,000 - 30,000
- 30,000 - 70,000
- > 70,000

f) >1000-hour Snag Fuels (kg/ha)

- 0
- 0 - 2,000
- 2,000 - 5,000
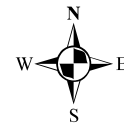- 5,000 - 15,000
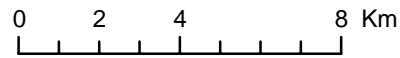- > 15,000

0   2   4   8 Km

Figure 3 (Continued)

snags cannot be directly measured using satellite imagery, and their regional patterns may be primarily related to disturbance history rather than environmental setting.

Although prediction accuracies can almost certainly be improved, some level of uncertainty will undoubtedly remain in the fuel maps generated using the GNN method. The implications of this uncertainty for the utilization of fuel maps in management and planning are unclear. Although fuel maps embody considerable uncertainty at the individual pixel level, they may still generate reliable aggregate predictions for stands, watersheds, or other larger landscape units. Even though error is unavoidable, predictive spatial models provide critical information about broad scale fuel patterns that cannot be obtained from any other data source. Future accuracy assessments will need to focus on evaluating map error in a spatially explicit context, and understanding how this error propagates through models of fire risk and fire behavior.

Ongoing work on the application of the GNN method to predictive fuels mapping includes refinement of the multivariate statistical model to improve prediction accuracy, and exploration of methods for quantifying spatial uncertainty and error propagation. The GNN mapping approach is also being applied in eastern Washington and in the Sierra Nevada in California to assess the potential for predictive fuel mapping across a variety of different ecoregions.

5. References

Bertschi, I., R. J. Yokelson, D. E. Ward, R. E. Babbitt, R. A. Susott, J. G. Goode, and W. M. Hao, 2003: Trace gas and particle emissions from fires in large diameter and belowground biomass fuels. *Journal of Geophysical Research-Atmospheres*, **108**.

Brown, J. K., 1978: Weight and Density of Crowns of Rocky Mountain Conifers. INT-RP-197, 56 pp.

Cohen, W. B., T. A. Spies, R. J. Alig, D. R. Oetter, T. K. Maiersperger, and M. Fiorella, 2002: Characterizing 23 years (1972-95) of stand replacement disturbance in western Oregon forests with Landsat imagery. *Ecosystems*, **5,** 122-137.

Cruz, M. G., M. E. Alexander, and R. H. Wakimoto, 2003: Assessing canopy fuel stratum characteristics in crown fire prone fuel types of western North America. *International Journal of Wildland Fire*, **12,** 39-50.

Finney, M. A., 1998: FARSITE: Fire Area Simulator - Model Development and Evaluation. RMRS-RP-4, 47 pp.

Franklin, J., 1995: Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, **19,** 474-499.

Keane, R. E., S. A. Mincemoyer, K. M. Schmidt, D. G. Long, and J. L. Garner, 2000: Mapping Vegetation and Fuels for Fire Management on the Gila National Forest Complex, New Mexico. RMRS-GTR-46-CD.

Keane, R. E., J. L. Garner, K. M. Schmidt, D. G. Long, J. P. Menakis, and M. A. Finney, 1998: Development of Input Data Layers for the FARSITE Fire Growth Model for the Selway-Bitterroot Wilderness Complex, USA. RMRS-GTR-3.

Moeur, M. and A. R. Stage, 1995: Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science*, **41,** 337-359.

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, USA. *Canadian Journal of Forest Research*, **32,** 725-741.

Snell, J. A. and J. K. Brown, 1980: Handbook for Predicting Residue Weights of Pacific Northwest Conifers. PNW-GTR-103, 43 pp.

Snell, J. A. and B. F. Anholt, 1981: Predicting Crown Weight of Coast Douglas-fir and Western Hemlock. PNW-RP-281, 13 pp.

Snell, J. A. and S. N. Little, 1983: Predicting Crown Weight and Bole Volume of Five Western Hardwoods. PNW-GTR-151, 37 pp.

ter Braak, C. J. F. and I. C. Prentice, 1988: A Theory of Gradient Analysis. *Advances in Ecological Research*, **18,** 271-317.

Thornton, P. E., S. W. Running, and M. A. White, 1997: Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, **190,** 214-251.

Tuominen, S., S. Fish, and S. Poso, 2003: Combining remote sensing, data from earlier inventories, and geostatistical interpolation in multisource forest inventory. *Canadian Journal of Forest Research*, **33,** 624-634.