Chapter 6. Modeling Ozone Bioindicator Injury with Microscale and Landscape-Scale Explanatory Variables: A Logistic Regression Approach

John W. Coulston

Introduction

Tropospheric ozone occurs at phytotoxic levels in the United States (Lefohn and Pinkerton 1988). Several plant species, including commercially important timber species, are sensitive to elevated ozone levels. Exposure to elevated ozone can cause growth reduction and foliar injury and make trees more susceptible to secondary stressors such as insects and pathogens (Chappelka and Samuelson 1998). In response to this threat, the Forest Service, U.S. Department of Agriculture, maintains a national ozone biomonitoring program.

The goal of ozone biomonitoring is to identify geographic areas where the risk of ozone injury is high and the forest community is sensitive. These areas may then become candidate areas for followup investigation through the evaluation monitoring tier of the National Forest Health Monitoring (FHM) Program of the Forest Service (see the definition of evaluation monitoring in chapter 1). Information about plant injury from ozone is collected at biomonitoring plots by examining bioindicator species. In general, biomonitoring plots are located in relatively open areas within or near to forests, and biomonitoring species are both tree and nontree species (table 6.1). To achieve the goal of the biomonitoring program (to identify geographic areas with sensitive forest communities and high risk of ozone injury), spatial models, e.g., kriging, are used to predict the likelihood of ozone injury (based on biomonitoring data) at plot locations from the Forest Inventory and Analysis (FIA) Program of the Forest Service. The plot data are then used to identify the sensitivity of the forest community. Smith and others (2007) provide guidance on how to select an appropriate spatial interpolation model, but they also note that future research will attempt to improve the precision of the estimates from the spatial models. The objective of this chapter is to identify appropriate ancillary data and the appropriate spatial scale of those data for use in spatial modeling of risk of ozone injury.

Ozone injury to plants is a function of the sensitivity of the plant species, the ambient ozone concentration, and environmental conditions (McCool 1998). Here, we examine the importance of microscale variables recorded on the biomonitoring plots and landscape-scale variables available as Geographic Information System maps for predicting the likelihood of ozone injury to biomonitoring plants. The microscale variables examined are aspect, terrain position, soil depth, soil wetness, and soil drainage. The landscape-scale variables are SUM06 ozone, Palmer Drought Severity Index (PDSI), aspect, available water capacity, terrain relative moisture index, and population density (a more complete description of each variable is presented in the following section) (table 6.2). A logistic regression model is used to identify the significant variables for predicting the likelihood of ozone injury on biomonitoring plots.

Table 6.1—Common and scientific names of bioindicator species

Common name	Scientific name
Blackberry	*Rubus allegheniensis*
Black cherry	*Prunus serotina*
Common and tall milkweed	*Asclepias spp.*
Yellow poplar	*Liriodendron tulipifera*
White ash	*Fraxinus americana*
Sassafras	*Sassafras albidum*
Spreading dogbane	*Apocynum androsaemifolium*
Big leaf aster	*Aster macrophylum*
Sweetgum	*Liquidambar styraciflua*
Pin cherry	*Prunus pensylvanica*
Ponderosa pine	*Pinus ponderosa*
Jeffrey pine	*Pinus jeffreyi*
Blue elderberry	*Sambucus cerulea*
Quaking aspen	*Populus tremuloides*
Scouler's willow	*Salix scouleriana*
Red alder	*Alnus rubra*
Skunk bush	*Rhus trilobata*
Ninebark	*Physocarpus malvaceus*
Mountain snowberry	*Symphoricarpos oreaphilus*
Western wormwood	*Artemesia ludoviciana*
Red elderberry	*Sambucus racemosa*
Huckleberry	*Vaccinium membranaceum*
Evening primrose	*Oenothera elata*
Mugwort	*Artemesia douglasiana*
California black oak	*Quercus kellogii*
Pacific ninebark	*Physiocarpus capitatus*

Data

The data used for model development were acquired from several sources. The FIA program collects information on bioindicator plant injury at biomonitoring sites (fig. 6.1). At each ozone biomonitoring plot, the amount and severity of ozone injury on bioindicator species was collected. Bioindicator species include ozone-sensitive species such as black cherry (*Prunus serotina*) in the Eastern United States and Scouler's willow (*Salix scouleriana*) in the Western United States (table 6.1). We used the biomonitoring plot data as the binary response variable in our model (0 = no injury detected, 1 = injury detected) for each year 2003 through 2005. We also used microscale variables collected on biomonitoring sites during the 2003 through 2005 field seasons. Landscape-scale variables were obtained from the U.S. Environmental Protection Agency, U.S. Geological Survey, and other sources (table 6.2).

Table 6.2—Potential explanatory variables used in logistic regression

Explanatory variable	Type	Categories	Variable name	Spatial scale	Reference
Microscale					
Aspect	Continuous		AspP[a]	Plot level	U.S. Department of Agriculture Forest Service 2002[b]
Terrain position	Categorical		TerrPos	Plot level	U.S. Department of Agriculture Forest Service 2002[b]
	1	Ridgetop or upperslope			
	2	Bench or level area along slope			
	3	Lower slope			
	4	Flat and unrelated to slope			
	5	Bottomland with occasional flooding			
Soil depth	Binary		SoilDpt	Plot level	U.S. Department of Agriculture Forest Service 2002[b]
	1	Bedrock is not exposed			U.S. Department of Agriculture Forest Service 2002[b]
	2	Bedrock is exposed. Soil generally shallow			
Soil drainage (Eastern United States)	Categorical		SoilDrn	Plot level	U.S. Department of Agriculture Forest Service 2002[b]
	1	Well drained			
	2	Generally wet			
	3	Excessively dry			
Soil wetness (Western United States)	Categorical		SoilWt	Plot level	U.S. Department of Agriculture Forest Service 2002[b]
	1	Wet			
	2	Moderately dry			
	3	Very dry			
Landscape scale					
Ambient ozone	Continuous		SUM06[c]	5-Km raster cells	U.S. Environmental Protection Agency 2004
Available water capacity	Continuous		Awc	1: 250,000	Miller and White 1998
Palmer drought severity	Continuous		PDSI[d]	U.S. Climate Division	National Climate Data Center 1994
Aspect	Continuous		AspG[a]	3 arc-second raster cells	U.S. Geological Survey 1993
Terrain relative moisture index	Continuous		TRMI[e]	3 arc-second raster cells	U.S. Geological Survey 1993
Population density	Continuous		Pden	U.S. counties	U.S. Census Bureau 2004

[a] Aspect was rescaled to a continuous variable denoting northernness scaled 0 to 2 by [cos(aspect)+1] to account for the fact that, for example, aspects of 15° and 345° have the same northernness.
[b] Forest Service, U.S. Department of Agriculture, 2002. Forest inventory and analysis national core field guide: field data collection procedures for phase 3 plots. Version 1.7. Internal report. Vol. 2. [Not paged]. On file with: Forest Service, U.S. Department of Agriculture, Forest Inventory and Analysis, Rosslyn Plaza, 1620 North Kent Street, Arlington, VA 22209.
[c] For each U.S. Environmental Protection Agency ozone monitoring station the sum of all hourly ozone concentration > 0.06 parts per million (ppm) were summarized from 8 a.m. to 8 p.m. for June, July, and August for 2003, 2004, and 2005. SUM06 (ppm-hours) ozone values were assigned to each biomonitoring plot by inverse distance weighting interpolation.
[d] Average June, July, and August Palmer Drought Severity Index (PDSI) was calculated for each climate division in the coterminous United States for each year (2003, 2004, 2005). PDSI is scaled from ~-7 to 7 where negative values indicate drought stress.
[e] TRMI is generated using a digital elevation model. The algorithm identifies topographic position, e.g., ridgetops and valley bottoms, to assign a moisture index scale from 0 (dry) to 60 (wet).

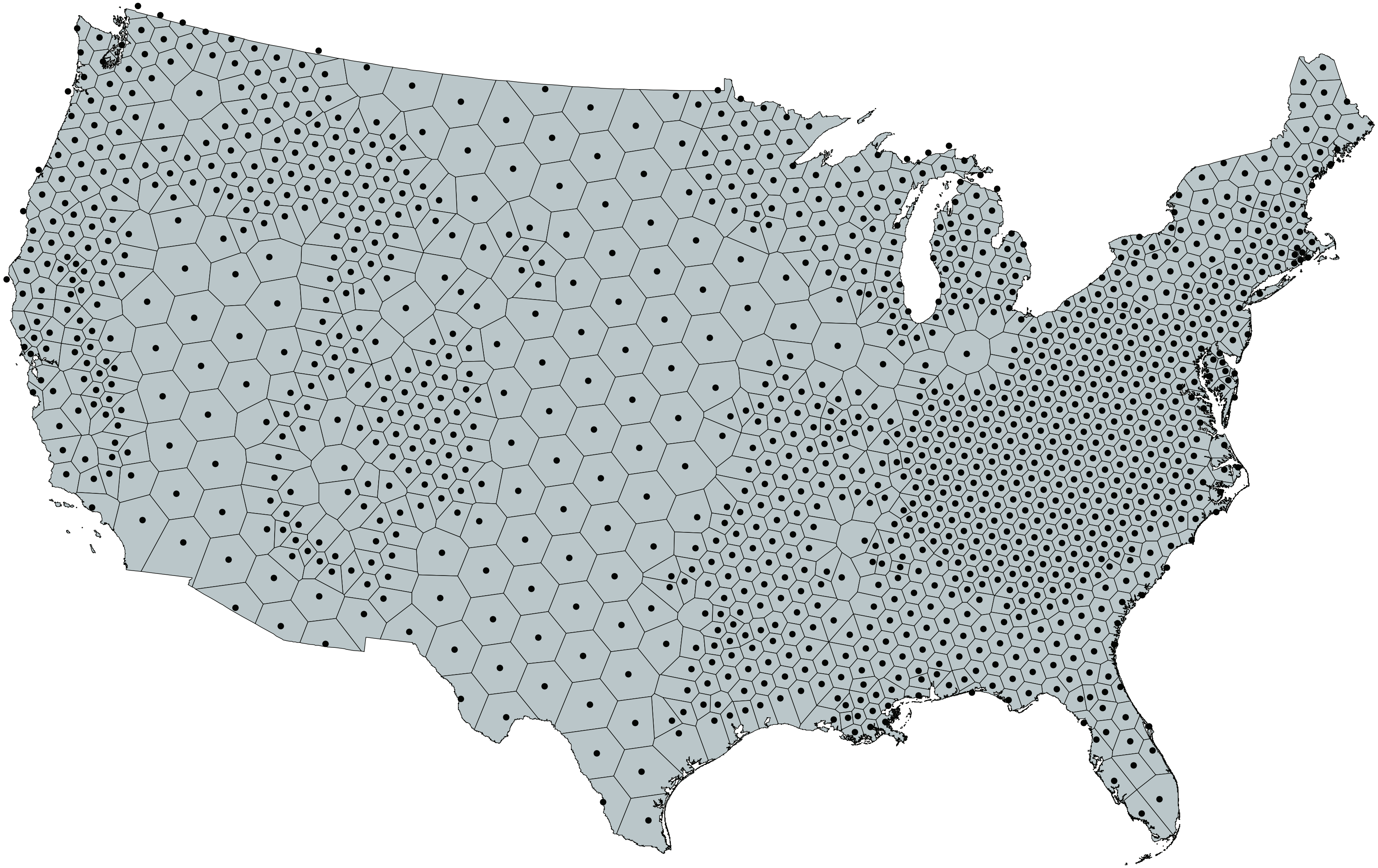

Figure 6.1—Ozone biomonitoring sampling grid for the coterminous United States. The points represent the approximate center of each sampling polygon.

Methods

Logistic regression (SAS 2004) was used to examine the relationship among biosite ozone injury, microscale variables, and landscape-scale variables. Because of different sampling intensities, each biosite was weighted by the area it represents (fig. 6.1). The general form of the linear logistic model is:

$$Log\,(\pi\ /\ (1-\pi) = a + B`x$$

where

Log = natural logarithm of (·)

π = the probability that the response equals 1 (ozone injury recorded) given the vector of explanatory variables x

a = intercept

B = vector of parameter estimates

The explanatory variables in x can be binary, categorical, ordinal, or continuous; and interactions among variables can also be examined. Because of potential regional differences in environmental conditions and ambient ozone concentrations, each FIA region [North, South, Interior West, and Pacific Northwest (which includes California)] was examined independently. Also, the North and South FIA regions collect information on soil drainage on biomonitoring plots while the Interior West and Pacific Northwest FIA regions collect information on soil wetness (table 6.2). Specific interactions were selected to examine the potential relationship among ozone injury, ambient ozone concentrations, terrain position, and moisture based on landscape and plot-level variables. In the East (North and South FIA regions), the full set of explanatory variables tested is denoted x_e and in the West (Interior West and Pacific Northwest FIA regions) the full set of explanatory variables tested is denoted x_w

where

$$x_e = \begin{bmatrix} AspP \\ TerrPos \\ SoilDpt \\ SoilDrn \\ SUM06 \\ Awc \\ PDSI \\ AspG \\ TRMI \\ Pden \\ SUM06 * TerrPos \\ SUM06 * SoilDrn \\ SUM06 * PDSI \\ SUM06 * TRMI \end{bmatrix} \quad \text{and } x_w = \begin{bmatrix} AspP \\ TerrPos \\ SoilDpt \\ SoilWt \\ SUM06 \\ Awc \\ PDSI \\ AspG \\ TRMI \\ Pden \\ SUM06 * TerrPos \\ SUM06 * SoilWt \\ SUM06 * PDSI \\ SUM06 * TRMI \end{bmatrix}$$

The null hypothesis for the full model (H_o: $B = 0$) was tested and each variable was examined. Variables that were not significant at the p = 0.10 level were removed from the models. The final model selected for each region was the model where $B \neq 0$, each variable was significant at $p < 0.10$, and the minimum Akaike Information Criterion value was minimized (SAS Institute 2004). The generalized coefficient of determination (pseudo-R^2) was used to examine the predictive power of each final model.

Results

From 2003 through 2005, ozone injury to bioindicator plants was recorded in every FIA region except for the Interior West (fig. 6.2). Injury occurrence was relatively constant in the North and Pacific Northwest FIA regions (2003 through 2005). However, in the South FIA region, the number of biomonitoring plots with injury tended to decrease even though the total number of plots examined increased between 2003 and 2005 (fig. 6.2). The Interior West FIA region was not examined using the logistic regression approach because of the lack of recorded ozone injury, but ambient ozone concentrations from 2003 through 2005 in the Interior West were comparable to the other FIA regions (fig. 6.3).

Statistically significant logistic regression models were developed for the North, Pacific Northwest, and South FIA regions (table 6.3).

The models for the three regions all contained SUM06 as a significant explanatory variable. Most of the interaction terms examined were not statistically significant. However, in the South FIA region, the SUM06*PDSI interaction was significant (table 6.4). TerrPos was the only microscale variable selected for the final model and only was used in the North FIA

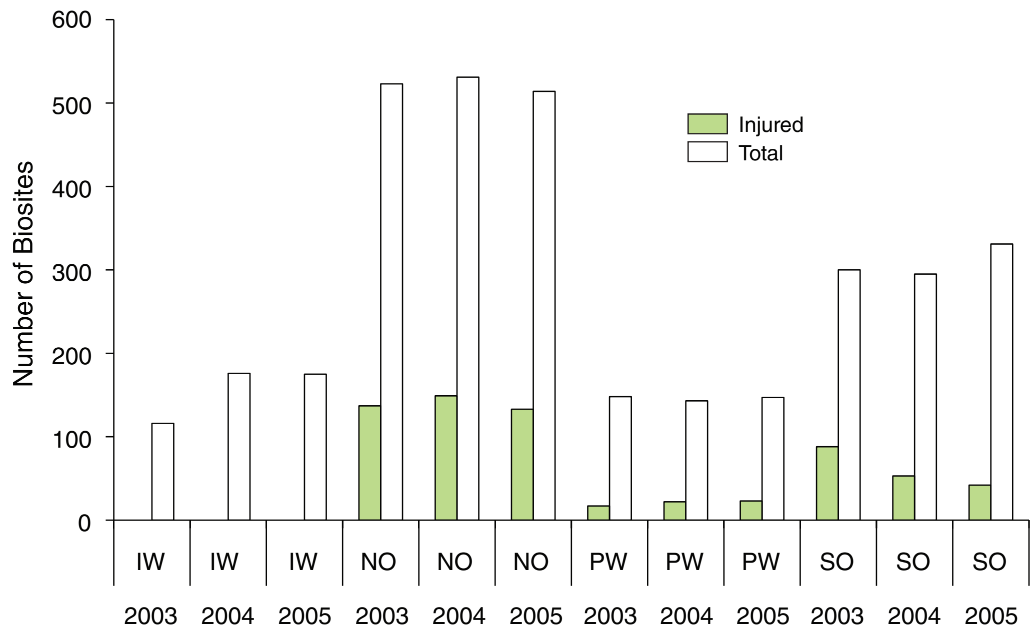

Figure 6.2—The number of total ozone biosites measured and the number of ozone biosites with injury, recorded by Forest Inventory and Analysis region (IW=Interior West, NO=North, PW=Pacific Northwest, SO=South) and year. (Data source: U.S. Department of Agriculture, Forest Service Forest Inventory and Analysis Program)

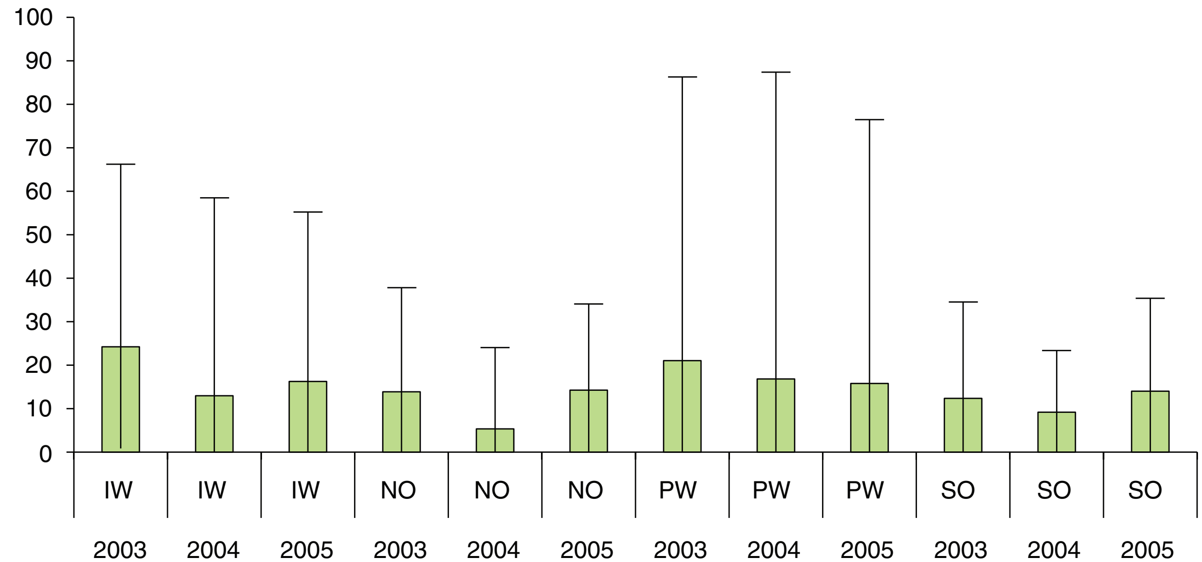

Figure 6.3—Average area weighted SUM06 ozone exposure for each Forest Inventory and Analysis region (IW=Interior West, NO=North, PW=Pacific Northwest, SO=South) and year. Error bars represent the minimum and maximum observed SUM06 value. (Data source: U.S. Environmental Protection Agency)

Table 6.3—Results from logistic regression model for each Forest Inventory and Analysis region

Region	Explanatory variables (x)	Chi-square	p-value	R^2	Percent of concordant pairs
					percent
North	SUM06, TerrPos, Pden	103.31	0.0001	0.0638	66.6
Pacific Northwest	PDSI, SUM06, Pden	94.96	0.0001	0.1949	84.1
South	PDSI, PDSI*SUM06	89.78	0.0001	0.0924	73.7

region model. Terrain position 1 (ridgetop or upperslope) and terrain position 4 (flatland unrelated to slope) had lower p-values (table 6.4) than the other terrain position categories. The best model for predicting the probability of ozone injury to bioindicator plants, based on R^2 and the percent of concordant pairs, was the Pacific Northwest model. The models for the South and North FIA regions had R^2 of 0.092 and 0.064, respectively.

Discussion

Ozone injury to plants is related to the ambient ozone concentration, plant species' sensitivity to ozone, soil moisture, and light, all of which influence ozone uptake by plants. The purpose of this analysis was to examine the relationship between ozone injury to bioindicator plants and microscale and landscape-scale explanatory variables. Generally, only landscape-scale explanatory variables were selected for the final model for each FIA region with terrain position as the exception in the North FIA region. However, the terrain relative moisture index landscape-scale variable can be used in place of the terrain position variable with minimal impact to the predictive power of the model. Smith and others (2007) suggested using either inverse distance weighting interpolation or kriging to predict the likelihood of ozone injury to bioindicator plants at unmeasured locations but also encouraged the development of spatially explicit models that include explanatory variables such as ambient ozone concentrations and moisture conditions. The results presented here indicate that landscape-scale variables were statistically significant most frequently, i.e., explained significant amounts of the variation in ozone injury to bioindicator plants.

Table 6.4—Significance of explanatory variables for each logistic regression model and Forest Inventory and Analysis region

Region	Explanatory variables (x)	Parameter estimate (B)	p-value
North	Intercept (a)	–5.55	0.0001
	SUM06	0.024	0.0033
	TerrPos		0.0032
	1	0.241	0.09
	2	–0.234	0.2404
	3	0.201	0.2519
	4	0.421	0.0002
	Pden	0.359	0.0001
Pacific Northwest	Intercept (a)	–6.326	0.0001
	PDSI	0.124	0.329
	SUM06	0.053	0.001
	Pden	0.292	0.0044
South	Intercept (a)	–1.79	0.0001
	PDSI	–0.325	0.0016
	PDSI*SUM06	0.045	0.0001

The significance of SUM06 ozone and PDSI from this analysis corresponds with results from Smith and others (2003) and Davis and Orendovici (2006). Davis and Orendovici (2006) found a statistically significant relationship between the incidences of ozone symptoms on vegetation in the Edwin B. Forsythe National Wildlife Refuge (New Jersey) and plant species, PDSI, and the interaction of two different ambient ozone statistics. Smith and others (2003) found that SUM06 ozone and PDSI were significant explanatory variables in a linear regression to predict a composite ozone bioindicator variable in the North FIA region. The statistical significance of county-level population density was not tested in the studies described above. In fact, population density is not a causal mechanism of ozone-induced foliar injury, and the correlation between population density and SUM06 ozone was < 0.23 in the North, South, and Pacific Northwest FIA regions. However, population density may serve as a surrogate for other explanatory variables not included in this analysis.

The results presented in this chapter identify key landscape-level variables that account for statistically significant amounts of the variance of ozone-induced plant injury from ozone. While statistically significant logistic models were developed for each FIA region where ozone injury occurred on biomonitoring sites, logistic regression provided little improvement, in respect to predictive power, over standard spatial interpolation techniques such as kriging and inverse distance weighting. However, the results from our analysis provide direction for future research:

1. Future modeling efforts should focus on using landscape-scale variables rather than microscale (plot-level) variables.
2. Predictive models vary regionally and perhaps subregionally. Future modeling efforts should examine the importance of using subregional areas for model development.
3. The SUM06 index was used to represent ambient ozone exposure. Based on suggestions from Davis and Orendovici (2006), other ambient ozone indices such as N100 (number of hours that ambient ozone is ≥ 100 parts per billion) may be more appropriate explanatory variables.
4. PDSI is derived by climate division, which may be too coarse for this kind of modeling. Other variables such as the ratio of precipitation to evapotranspiration should be examined as a potential fine-scale surrogate for PDSI.

Literature Cited

Chappelka, A.H.; Samuelson, L.J. 1998. Ambient ozone effects on forest trees of the Eastern United States: a review. New Phytologist. 139: 91–108.

Davis, D.D.; Orendovici, T. 2006. Incidence of ozone symptoms on vegetation within a national wildlife refuge in New Jersey, USA. Environmental Pollution. 143: 555–564.

Lefohn, A.S.; Pinkerton, J.E. 1988. High resolution characterization of ozone data for sites located in forested areas of the United States. Journal of the Air Pollution Control Association. 38: 1504–1511.

McCool, P.M. 1998. Introduction. In: Flager, R.B., ed. Recognition of air pollution injury to vegetation: a pictorial atlas. Pittsburgh: Air and Waste Management Association: 1.1–1.2.

Miller, D.A.; White, R.A. 1998. A conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling. Earth Interactions. 2: 1–15.

National Climate Data Center. 1994. Time bias corrected divisional temperature-precipitation-drought index. Documentation for dataset TD–9640. 12 p. On file with: Database Management Branch, National Climate Data Center, National Oceanic and Atmospheric Administration, Federal Building, 37 Battery Park Avenue, Asheville, NC 28801–2733.

SAS Institute Inc. 2004. SAS/STAT 9.1 user's guide. Cary, NC. SAS Institute Inc. 5,121 p.

Smith, G.; Coulston, J.; Jepsen, J.; Prichard, T. 2003. A national ozone biomonitoring program: results from field surveys of ozone sensitive plants in northeastern forest (1994–2000). Environmental Monitoring and Assessment. 87: 271–291.

Smith, G.C.; Smith, W.D.; Coulston, J.W. 2007. Ozone biomonitoring sampling and estimation. Gen. Tech. Rep. NRS–20. Newtown Square, PA: Forest Service, U.S. Department of Agriculture, Northern Research Station. 34 p.

U.S. Census Bureau. 2004. U.S. census database, 2000. Reston, VA: U.S. Geological Survey. http://www.census.gov/geo/www/gazetteer/places2k.html,. [Date accessed: June 6, 2007].

U.S. Environmental Protection Agency. 2004. Air quality system raw data summarization formulas. Draft version 1.1. http://www.epa.gov/ttn/airs/airsaqs/manuals/Precision%20and%20Accuracy%20Summary%20Formulas.pdf. [Date accessed: November 13, 2006].

U.S. Geological Survey. 1993. Digital elevation models—data users guide 5. Reston, VA: 48.