# CHAPTER 6.

## Links between Land Cover and Lichen Species Richness at Large Scales in Forested Ecosystems across the United States

Susan Will-Wolf

Randall S. Morin

Mark J. Ambrose

Kurt Riitters

Sarah Jovan

## INTRODUCTION

Lichen community composition is well known for exhibiting response to air pollution, and to macroenvironmental and microenvironmental variables. Lichens are useful indicators of air quality impact, forest health, and forest ecosystem integrity across the United States (McCune 2000, reviews in Nimis and others 2002, USDA Forest Service 2007). Recent studies suggest lichen composition of a forested area can also be affected by the proportion of forest in the nearby landscape, at least at relatively small scales (Stofer and others 2006; Will-Wolf and others 2002, 2010, 2011b). A lichen is a close symbiotic relationship of a fungus (mostly Ascomycota) with green algae, cyanobacteria ("blue-green algae"), or both. The fungus provides a stable environment for the algae to live, while the algae provide the fungus with energy from photosynthesis. A macrolichen can be detached from its substrate and is large enough to see easily. Lichens grow on tree trunks and branches, on rocks, and on soil where vascular plants are sparse.

The Forest and Inventory Analysis (FIA) Program monitors the status of forests nationwide on a national grid of permanent plots (Woodall and others 2011). Lichen data were collected on an interspersed subset of this grid; they are suitable primarily for evaluating large-scale patterns and trends in forest health. Lichen data are collected using standard protocols (USDA Forest Service 2011) that

have remained unchanged since they were first implemented in 1994 under the Forest Health Monitoring (FHM) Program. The number of macrolichen species found at a forested plot, referred to here as Lichen S, is an index for lichen species richness available from FIA data. Lichen S has been recommended as an indicator for the condition of the surrounding forested ecosystem across both large and small geographic areas (McCune 2000).

Lichen S is a general indicator for the condition of forested ecosystems. Because it does not depend on a particular species composition (that changes between different geographic regions), Lichen S is a consistent variable across the entire country that can be compared among regions. FIA lichen data are available for public download (at http://apps.fs.fed.us/fiadb-downloads/datamart.html), representing almost two-thirds of forested areas across the conterminous United States, probably the most extensive (but with relatively low spatial intensity) quantitative lichen dataset in the world.

Will-Wolf and others (2011a) found that Lichen S has potential to indicate broad-scale response of forests to climate (stronger in the Western United States) and to air quality (stronger in the Eastern United States) in both East and West regions and in several subregions of each large region. In that earlier study, we did not consider the influence of nearby land cover pattern. We later hypothesized that air

pollution and climate variables may well be correlated with land cover pattern near plots when evaluated across these large regions. If such relationships occur, information on Lichen S variation with nearby land cover is needed to support accurate interpretation of the relationship between Lichen S and pollution or climate.

The purpose of this study was to investigate whether the inclusion of land cover variables improved models for explaining variation in Lichen S as compared with the previous study. We investigated this possibility by exploring with the same data used by Will-Wolf and others (2011a) the relationships of Lichen S, climate, and air quality with new land cover variables for a subset of six geographic areas from the earlier study. We selected our analysis tools to maximize comparability of results among regions and robustness of conclusions, at the cost of lower statistical power for individual results.

Our general questions for this project were:

- How are land cover variables related to Lichen S across large regions?
- How are land cover variables related to environmental variables and each other?
- What implications do these relationships have for interpreting effects of both environment and nearby land cover on Lichen S? Can general recommendations be suggested for analyses?
- Does inclusion of land cover variables improve models to explain Lichen S?

## METHODS

### Study area

For this project, we selected six of the regions defined in Will-Wolf and others (2011a) for the conterminous United States based on Bailey's ecoregions (Bailey 1989, Cleland and others 2005) (fig. 6.1). We include the entire East and West regions, plus two subregions within each region. In each region, one subregion is mountainous and the other subregion has less relief. The region described hereafter as "West all" includes all plots from west of the area in figure 6.1 with no data; "W Sierras" is the Sierras/Coast Mountains subregion, including lowlands; "W CO Plateau" is the Colorado Plateau/S Dry Mountains subregion with less relief. The region described hereafter as "East all" includes all plots from east of the area in figure 6.1 with no data; "E Adiron" is the Adirondacks mountainous subregion; "E Decid" is the Eastern Deciduous subregion with less relief.

### Lichen data

We used lichen data compiled by Will-Wolf and others (2011a) for the six regions described above—2,482 unique plots. These data were the single most recent FIA (or FHM) sample for each plot surveyed 1994–2002 in the conterminous United States; sample year thus varied by plot. Lichen S is the number (count data) of species on a plot from a timed (up to 2-hour) survey of macrolichens on all easily accessible woody substrates in a 0.379-ha (0.937-acre) plot by trained nonspecialists, with species identification by specialists. Lichen data are from the "forest
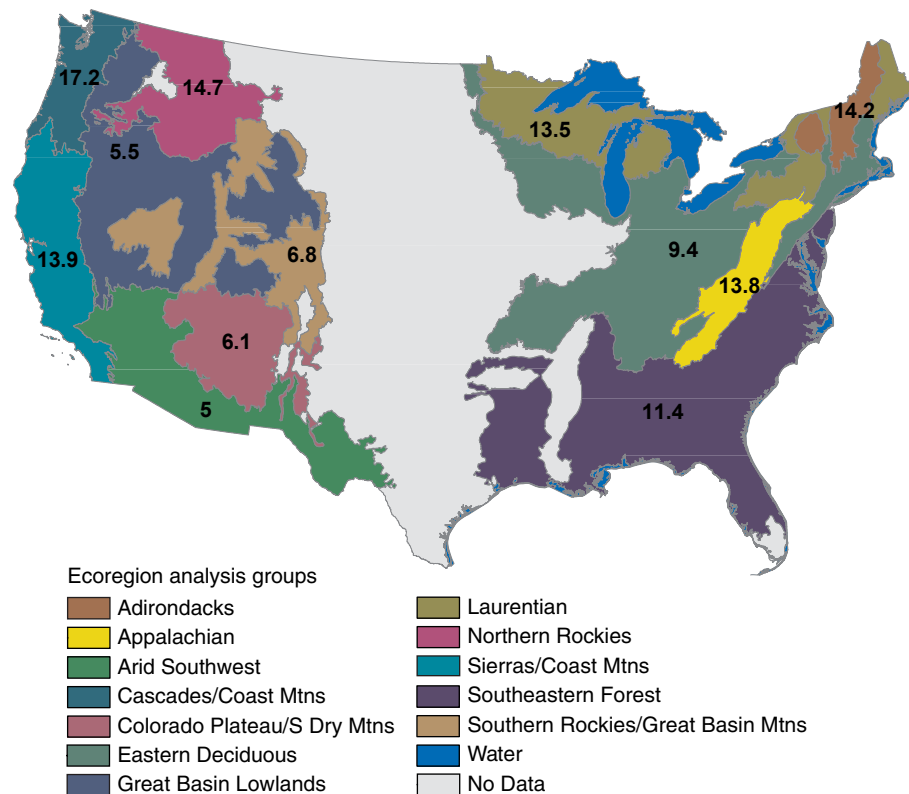
*Figure 6.1—Ecoregion groups from Will-Wolf and others (2011a), based on Bailey's ecoregion provinces (Bailey 1989, Cleland and others 2005). For this project, the "West all" region includes all ecoregion groups west of the large pale gray area with no data, and the "East all" region includes all ecoregion groups east of this same area. Some ecoregion groups include parts of States with no Forest Inventory and Analysis lichen data. Printed numbers are the average for each ecoregion of Lichen S, the number of macrolichen species found per plot. Note the greater variation in average Lichen S among groups in the West.*

health" (phase 3) subset of FIA plots (Woodall and others 2010), whose average density is 1 plot/39 072 ha (per 96,000 acres).

## Explanatory variables

Five land cover variables were developed for neighborhoods of five sizes for each plot (table 6.1). A neighborhood is defined as a square geographic area centered on the approximate coordinates (see next paragraph) for each plot with sides facing the cardinal directions. Within each neighborhood, land cover measurements were taken from the 2001 National Land Cover Database (NLCD) land cover map, which portrays 16 land cover types at a spatial resolution of 30 m (0.09 ha/square pixel; Homer and others 2007). We evaluated forest density ($Fden$), forest connectivity ($Fcon$), and the percentage of pixels in three types of generalized land cover—natural (including seminatural; $pctNt$), agricultural ($pctAg$), and developed ($pctDv$)—to represent the type and degree of human modification within a neighborhood. $Fcon$ for a plot is determined by counting the total number of pixel edges (within the neighborhood) that have a forest pixel on at least one side, then calculating the percentage of those edges that have forest pixels on both sides (Riitters and others 2007); higher values of $Fcon$ indicate less forest fragmentation. A neighborhood with no forest cover pixels is assigned $Fcon = 0$; connectivity cannot be calculated there, but it is certainly less than for a neighborhood with one forest cover pixel, where $Fcon = 1$. Plot neighborhoods with no forest cover pixels do occur in our data; a clearcut plot

in forest land use (FIA definition) may coincide with a pixel interpreted in NLCD coverage as having no forest cover. Both *Fden* and *Fcon* were measured as continuous variables, then were converted to ordinal indices for our analyses. The three land cover variables related to human modification (*pctNt*, *pctAg*, and *pctDv*) were derived from previously prepared national maps of the categorical "landscape mosaic" tripolar classification model (Riitters 2011, Riitters and others 2007). For that model, "natural" land cover is any land cover not specifically classified as developed (human structures) or agricultural. Ordinal variables with eight numerical values representing the range 0 to 100 percent of each target cover class in the landscape were created from the model categories. Thus, all five land cover metrics are unitless, ordinal indices with a higher index value for more cover of that type or more forest connectivity. There are correlations among the five variables for a given plot; for instance, *Fden* represents a subset of the pixels represented by *pctNt* for the same area, and values for the same variable at different neighborhood sizes are correlated. We added size code (table 6.1) as a suffix to a variable name to indicate the size of the neighborhood. For instance, *Fden4* is forest density evaluated across a neighborhood size of 590.5 ha. Size code 1 has been reserved for a smaller neighborhood size of 4.41 ha for a planned later study to be compared with this project. Each of the three largest neighborhood sizes is 9 times larger than the next lower size. The largest neighborhood size might share a substantial minority of its pixels with the neighborhood of a plot in an adjacent

FIA grid cell, making values for plots in adjacent cells potentially not strictly independent at this size. We estimate up to 20 percent of plots in our data are in a grid cell immediately adjacent to another grid cell with a plot.

We used data compiled for Will-Wolf and others (2011a) for the environmental and plot variables defined in that study. Geographic location for each plot was represented by approximate longitude (*long*) and approximate latitude (*lat*) as in the public online FIA databases, and by elevation (*elev*) from field Global Positioning System coordinates (Woudenberg and others 2010). Use of the public location data allows our study to serve as a model broadly relevant to scientists outside FIA; exact plot locations are private by law and are available only through special agreements. The geographic variables are indirect indicators of climate at a plot as well as of other possible factors linked to geographic location; relationship to direct climate variables and other possible factors differs for each region and subregion. Forest structure within a plot was represented by total live tree basal area (*tBA*), and percentage of *tBA* in softwoods, conifers, or both (*pctSoft*); the latter is also a general indicator of variability in lichen substrates at a plot. Location and forest structure variables were extracted from FIA and legacy databases.

Air pollution is represented by 1998–2004 average annual wet deposition values for sulfur dioxide ($SO_2^{--}$), nitrate ($NO_3^-$), and ammonium ($NH_4^+$) (geographic pattern of $NO_3^-$ deposition illustrated in Will-Wolf and others 2011a)

**Table 6.1—Definition of land cover variables and the neighborhoods across which they were evaluated**

Landscape variables

| Name (variables) | Definition | National Land Cover Database classes | Range of values |
|---|---|---|---|
| Forest density *(Fden)* | Index for percentage of pixels in the neighborhood that were forest cover | 41, 42, 43, 90 | $1-22^a$ |
| Forest connectivity *(Fcon)* | Index for percentage of all forest cover pixel edges that were forest-to-forest edges | 41, 42, 43, 90 | $0-22^a$ |
| Percentage natural *(pctNt)* | Index for percentage of natural to seminatural land cover | 11, 12, 31, 41, 42, 43, 52, 71, 90, 95 | $0-22^b$ |
| Percentage agricultural *(pctAg)* | Index for percentage of agricultural land cover | 81, 82 | $0-22^b$ |
| Percentage developed *(pctDv)* | Index for percentage of developed land cover | 21, 22, 23, 24 | $0-22^b$ |

Neighborhoods

| Size | 4.41 ha | 15.21 ha | 65.61 ha | 590.5 ha | 5314 ha | 47 830 ha |
|---|---|---|---|---|---|---|
| Size code | 1 | 2 | 3 | 4 | 5 | 6 |

Note: All values are ordinal, unitless numbers representing ranges of percentages.

[a] Value 1 represents exactly 0 percent; 22 represents exactly 100 percent. The 20 values (from 2 through 21) represent approximately equal divisions of the 1- through 99-percent range. For *Fcon*, the value 0 represents a landscape with no forest cover pixels; see text.

[b] Value 0 represents exactly 0 percent; 22 represents exactly 100 percent. The 6 intermediate values represent roughly equal divisions of the 1- to 99-percent range.

interpolated for each plot from models using National Atmospheric Deposition Program data (Coulston and others 2004). These modeled variables represent average background pollution in the general region; they do not capture variability over time or spatial variation from either local point sources or diffuse semilocal pollution "hot spot." $SO_2^{--}$ and $NO_3^-$ are strongly correlated with each other in our datasets (Will-Wolf and others 2011a), so $NO_3^-$ represents both acidic pollution variables for most analyses (tables 6.2, 6.3, 6.4, and 6.5). Our pollution variables adequately represent total background air pollution in Eastern States, but they underestimate total background pollution in Western States, where dry deposition is an important contributor (Fenn and others 2003). Wet deposition estimates represent relative air pollution loads well within East regions and reasonably well within West regions (better in the wetter parts of the West). They do not well represent the relative air pollution loads between the East and West, because dry deposition contributes so much more to total pollution in the West.

**Table 6.2—West region and groups: Spearman correlations of land cover variables with Lichen S at the neighborhood size with strongest correlation**

| Variables | West all n = 1,397 | W Sierras n = 254 | W CO Plateau n = 151 |
|---|---|---|---|
| *Fden* | max rho = **+ 0.416** with *Fden6* | max rho = + 0.187 with *Fden6* | max rho = **+ 0.572** with *Fden6* |
| *Fcon* | max rho = + 0.262 with *Fcon6* | NS | max rho = + 0.518 with *Fcon6* |
| *pctNt* | max rho = − 0.204 with *pctNt4* | max rho = − 0.294 with *pctNt4* | NS |
| *pctAg* | max rho = + 0.205 with *pctAg4* | max rho = + 0.294 with *pctAg5* | NS |
| *pctDv* | max rho = + 0.204 with *pctDv4* | max rho = **+ 0.301** with *pctDv4* | NS |

n = sample size, "max" = "maximum." NS = nonsignificant. See text and Table 6.1 for definitions of variable acronyms.
Note: The rho value for the strongest correlation in each region is in bold. Correlations with rho < ± 0.40 are considered weak. For the smallest n, rho = ± 0.30 is significant at $p \leq 0.0005$, though the correlation represents less than 10 percent of the variation in Lichen S. For rho > 0.447, at least 20 percent of variation is represented by the correlation.

**Table 6.3—West region and subregions: intercorrelations of variables featuring *Fden* and *pctAg***

| Regions/subregions | *Fden* and similar variables | *pctAg* and similar variables | Environmental variables |
|---|---|---|---|
| **West all** | n = 1,397 | | |
| Correlations[a] | *Fden*: rho = + 0.611 with *precip*, − 0.453 with *mxJul* at *Fden6*. (*Fcon* similar pattern but weaker.) | *pctAg*: weak negative correlation with *elev* max at *pctAg5*. (*pctNt* similar patterns but weaker, signs opposite.) | Intercorrelations of environmental variables: all rho < 0.600. |
| **W Sierras** | n = 254 | | |
| Correlations[a] | *Fden*: weak negative correlation with pollution, max at *Fden6*; rho = + 0.742 with *precip*, + 0.509 with *tBA* at *Fden6*. (*Fcon* similar pattern but weaker.) | *pctAg*: weak positive correlation with pollution, max at *pctAg4*; rho = − 0.494 with *elev* at *pctAg4*. (*PctNt* similar pattern but weaker, signs opposite.) | Pollutants intercorrelated, correlated with *long, lat*, all rho > 0.600; *elev, mxJul, pctSoft* all intercorrelated rho > 0.600. |
| **W CO Plateau** | n = 151 | | |
| Correlations[a] | *Fden*: rho = + 0.417 with $SO_2^{--}$, + 0.562 with $NO_3^{-}$, + 0.633 with *elev*, + 0.765 with *precip*, − 0.752 with *mxJul*, + 0.427 with *tBA*, − 0.431 with *pctSoft*, all at *Fden6*. (*Fcon* similar pattern but weaker.) | *pctAg*: weak positive correlation with pollution, max at *pctAg4*; rho = − 0.494 with *elev* at *pctAg4*. (*PctNt* similar pattern but weaker, signs opposite.) | $NO_3^{-}$, *elev, mxJul* intercorrelated rho > 0.600. |

n = sample size, "max" = "maximum." See text and Table 6.1 for definitions of variable acronyms.
Note: The strongest Spearman correlations of *Fden* and *pctAg* with environmental variables are listed when absolute value of rho $\geq$ 0.40 ($p < 0.0005$; at least 16 percent of variation explained). Weaker results are described; "much weaker" indicates correlations were > 0.150 weaker. Selected correlations between environmental variables (rho > 0.60) are summarized in parentheses in the far right column from Will-Wolf and others (2011a). Weaker correlations are not mentioned.

[a] For absolute value of rho > 0.447, at least 20 percent of variation is represented by the correlation; for absolute value of rho $\leq$ 0.60, 36 percent or less of variation is represented. Correlations not mentioned are very weak to random.

**Table 6.4—East region and subregions: Spearman correlations of land cover variables with Lichen S at the neighborhood size with strongest correlation**

| Variables | East all<br>n = 1,085 | E Adiron<br>n = 152 | E Decid<br>n = 264 |
|---|---|---|---|
| *Fden* | max rho = + 0.370 with *Fden6* | NS | max rho = + 0.428 with *Fden6* |
| *Fcon* | max rho = + 0.311 with *Fcon6* | NS | max rho = + 0.383 with *Fcon5* |
| *pctNt* | max rho = **+ 0.410** with *pctNt6* | max rho = **+ 0.333** with *pctNt4* | max rho = **+ 0.440** with *pctNt6* |
| *pctAg* | max rho = − 0.355 with *pctAg6* | max rho = **− 0.332** with *pctAg4* | max rho = − 0.342 with *pctAg6* |
| *pctDv* | max rho = − 0.264 with *pctDv4* | max rho = − 0.307 with *pctDv4* | max rho = − 0.282 with *pctDv6* |

n = sample size, "max" = "maximum." NS = nonsignificant. See text and Table 6.1 for definitions of variable acronyms.
Note: The rho value for the strongest correlation in each region is in bold, plus a second almost equal for one region. Correlations with rho < ± 0.40 are considered weak For the smallest n, rho = ± 0.30 is significant at $p \leq 0.0005$, though the correlation represents less than 10 percent of the variation in Lichen S. For rho > 0.447, at least 20 percent of variation is represented by the correlation.

**Table 6.5—East region and subregions: intercorrelations of variables featuring *Fden* and *pctAg***

| Regions/subregions | *Fden* and similar variables | *pctAg* and similar variables | Environmental variables |
|---|---|---|---|
| East all | n = 1,085 | | |
| Correlations[a] | *Fden*: weak negative correlations with $NH_4^+$. (*Fcon* similar pattern but weaker.) | *pctAg*: rho = + 0.456 with $NH_4^+$, weak positive with $SO_2^{--}$, $NO_3^-$, all max at *pctAg6*. (*pctNt* similar pattern slightly weaker, signs opposite.) | Intercorrelations of environmental variables: all rho < 0.600. |
| E Adiron | n = 152 | | |
| Correlations[a] | *Fden*: rho = + 0.558 with *elev* at *Fden4*. (*Fcon* similar pattern but weaker.) | *pctAg*: rho = + 0.447 with $SO_2^{--}$, − 0.498 with *elev*, max at *pctAg5*; rho = − 0.521 with *lat*, + 0.536 with *minJan*, max at *pctAg4*. (*pctNt* similar pattern slightly weaker, signs opposite.) | Pollutants strongly intercorrelated, correlated with *long*, *lat*, *minJan*, all rho > 0.600; *long*, *lat*, *minJan*, all intercorrelated rho > 0.600. |
| E Decid | n = 264 | | |
| Correlations[a] | *Fden*: rho = − 0.601 with $NH_4^+$ at *Fden6*, weak positive correlations with *long*. (*Fcon* similar pattern but weaker.) | *pctNt*: rho = − 0.585 with $NH_4^+$ at *pctNt6*. *pctAg*: rho = + 0.684 with $NH_4^+$ at *pctAg6*. | $NH_4^+$ correlated with *long*, rho > 0.60.0 |

n = sample size, "max" = "maximum." See text and Table 6.1 for definitions of variable acronyms.

Note: Spearman correlations of *Fden* and *pctAg* with environmental variables are listed when absolute values of rho $\geq$ 0.40 ($p < 0.0005$; at least 16 percent of variation explained). Weaker results are described; "much weaker" indicates correlations were > 0.150 weaker. Strength (not sign) of selected intercorrelations between environmental variables (absolute value of rho > 0.60) summarized from Will-Wolf and others (2011a) in the far right column. Weaker correlations are not mentioned.

[a] For absolute value of rho > 0.447, at least 20 percent of variation is represented by the correlation; for absolute value of rho $\leq$ 0.60, 36 percent or less of variation is represented. Correlations not mentioned are very weak to random.

Climate is represented by average annual precipitation (*precip*), average minimum January temperature (*minJan*), and average maximum July temperature (*mxJul*); each is the 1971–2000 30-year average interpolated for each plot from the Climate Source model (PRISM: Daly and Taylor 2000). Pollution modeled as wet deposition is much higher and varies more widely in the East, and climate varies much more widely in the West (appendices in Will-Wolf and others 2011a).

The full suite of explanatory variables as described above was available for each plot included in the analyses. The numbers of plots for each analysis are reported in tables 6.2 through 6.5.

## Analysis methods

Our analyses were organized to evaluate these preliminary hypotheses:

- Nearby land cover is correlated with Lichen S across large regions.
- In the West, land cover variables are correlated with climate variables.
- In the East, land cover variables are correlated with pollution variables.
- The relationships of land cover variables to environmental variables, and thus the interpretations of their relationships with Lichen S, vary by region.
- Regression models to explain variation in Lichen S are improved when land cover variables are included.

As with Will-Wolf and others (2011a), we found that few of the pairwise scatterplots we examined for variables showed linear relationships. We did not quantitatively evaluate whether parametric assumptions were met for each analysis or what specific data transformations might be most appropriate for each variable. Instead we applied a single data transformation as needed (see next paragraph) to address all such issues.

We calculated Pearson product-moment (parametric) correlations (r) and Spearman rank (nonparametric) correlations (rho) to evaluate the first four hypotheses, and we developed linear regression models with both original and ranked data to evaluate the fifth hypothesis. Rank transformation, the recommended option for standard analysis of FIA lichen data (Will-Wolf 2010), compensates for many different data distribution issues with count, ordinal, and continuous data (Yandell 1997) that affect analysis. It also equalizes data ranges for independent variables, which is useful for avoiding possible bias in regression models from differently scaled variables. Data were always ranked so that higher rank corresponded to higher original value for the variable. All analyses were performed in SPSS for Windows (Release 16.0.1 © SPSS, Inc. 1989–2007).

Many correlations between explanatory variables, as well as between those and Lichen S, were examined to determine which variables were most informative. We report the strongest correlations of Lichen S and environmental

variables with land cover variables and briefly summarize correlations between environmental variables. We discuss only correlations with absolute value of r or rho ≥ 0.40 ($p < 0.0005$ for the smallest sample size used) to highlight the potentially most ecologically important (beyond merely statistically significant) relationships. For correlation strength (sign ignored) > 0.447, a minimum of about 20 percent of variation is represented by the correlation; for correlation strength ≤ 0.60, about 36 percent or less of variation is explained by the correlation.

Many linear regression models were developed for each geographic region, primarily with hand selection of variables and forced simultaneous entry; all reported models were developed this way. Each land cover variable was entered into a particular model at only one neighborhood size. All land cover variables were tested for each region, but a final model was not required to include a land cover variable. Multiple independent variables that had correlations stronger than 0.60 were not entered into the same regression model; instead alternate models were developed. Multiple alternate regression models were examined. Only variables that were significant at $p < 0.05$ were retained in a model. We made these conservative choices for model development to avoid overestimating the significance of a model; the lower power of our analyses was an acceptable cost. We report the single strongest (highest $R^2$) regression model that met our development criteria, unless there were two similarly strong alternate models. All of our

"independent" variables including land cover variables are estimated and subject to error, not exact and fixed as assumed for regression models. This is likely to be the case in all large-scale ecological analyses for practical reasons, and should be considered when interpreting predictive models.

## RESULTS

Spearman correlations were often stronger than Pearson correlations, suggesting data may not meet assumptions for parametric statistical tests in those cases. Similarly, many regression models using ranked data were stronger than equivalent models using raw data. For consistency, we report Spearman correlations and regression models using ranked data, except as specifically noted. Our conservative criteria for reporting correlations and for entering variables into regression models mean the patterns we do report are quite robust.

Among land cover variables, percentage of forest cover (*Fden*) in West regions and percentage of natural to seminatural land cover (*pctNt*) in East regions usually had the strongest correlations (+/-) with Lichen S (tables 6.2 and 6.4). Percentage of a neighborhood in agricultural land (*pctAg*) or developed land (*pctDv*) each had the strongest (+/-) correlation with Lichen S in one region. Connectivity of forest cover (*Fcon*) gave results similar to percentage of forest cover, but weaker. As we expected, percentage of forest, percentage of natural to seminatural land cover, and connectivity of forest were positively correlated

in each region, whereas all three were negatively correlated with proportion of agricultural land (statistics not reported). Percentage in developed land usually had weak positive correlations with percentage in agricultural land and was mostly random regarding other land cover variables. We focused on percentage in forest cover and percentage in agricultural land to report relationships of land cover with environmental variables (tables 6.3 and 6.5).

In the West region and both West subregions, land cover variables had relatively strong correlations with climate and related environmental variables (table 6.3). For the East, in contrast, only in the mountainous Adirondacks subregion did land cover variables have relatively strong correlations with climate and related variables (table 6.5). Correlations between land cover variables and pollution variables were relatively weak in the West and slightly stronger in the East. The strongest such correlation, of *pctAg6* with $NH_4^+$ in the Eastern Deciduous subregion, explained almost 50 percent of variation (table 6.5). Correlations between environmental variables that affected interpretation of other statistical relationships varied quite widely among regions; some areas had few or no strong correlations, but others had many strong relationships (tables 6.3 and 6.5).

We found two different patterns for effect of neighborhood size on land cover variables. For percentage of agricultural land and developed land, the strength of several correlations with Lichen S and environmental variables peaked at neighborhood sizes of 590.5 ha or 5314 ha. For percentage of forest and related forest cover variables (*Fcon* and *pctNt*), this happened only twice. Identification of a neighborhood size at which a correlation is strongest helps focus the search for a mechanism causing the inferred effect. For *Fden* and related land cover variables most of the time and for all land cover variables at least some of the time the strength of correlations increased continuously with neighborhood size, lending little help to the search for underlying causes.

All four West regression models included climate variables; only two included land cover variables. Land cover variables were included in the best regression model for region West all:

R-Lichen S = 946.322 - 0.569*R-*elev* + 0.316*R-*Fden6* - 0.195*R-*mxJul* + 0.077*R-*pctAg5*       (1)

where

$R^2 = 0.472$
$p < 0.0005$
n = 1,397.

In this region, the land cover variables were more strongly linked to climate-related factors than to pollution variables (table 6.3). Model (1) suggests that climate and human-modified landscape pattern have more influence than pollution on Lichen S in region West all. Land cover variables were not included in the best regression model for the W Sierras subregion:

R-Lichen S = 207.119 - 0.519*R-*elev* - 0.268*R-$NH_4^+$ + 0.162*R- *tBA*  (2)

where

$R^2$ = 0.330
$p < 0.0005$
n = 254.

Land cover variables were much more strongly linked to climate-related variables than to pollution variables in this region (table 6.3). Model (2) indicates that climate, pollution, and forest structure have more influence than landscape pattern on Lichen S in the W Sierras subregion. There were two equally strong regression models for the W CO Plateau subregion; one included a land cover variable and the other did not:

R-Lichen S = 42.092 + 0.374*R-$NO_3^-$ + 0.292*R-*precip* - 0.176*R-*pctSoft*  (3)

where

$R^2$ = 0.404
$p < 0.0005$
n = 151.

R-Lichen S = 48.461 + 0.327*R-$NO_3^-$ + 0.298*R-*Fden6* - 0.210*R-*pctSoft*  (4)

where

$R^2$ = 0.403
$p < 0.0005$
n = 151.

*Fden* and $NO_3^-$ were both strongly linked to climate-related variables (table 6.3), and pollution is relatively low in this subregion (Will-Wolf and others 2011a). It is possible that neither pollution nor landscape configuration has an independent influence on Lichen S here; separating their impact from climate will require more detailed studies. Models (3) and (4) suggest Lichen S increases with more pollution. This unexpected result is another indication that in the W CO Plateau subregion, variable $NO_3^-$ may indirectly represent some climate factor, rather than the direct influence of pollution.

All three East models included both pollution and land cover variables; two also included variables linked to climate. Despite the stronger single correlations of land cover variable *pctNt*

with Lichen S in East regions (table 6.5), *Fden* (two regions) and *pctAg* (one region) were the land cover variables entered into the best East regression models. The best regression model for region East all was:

R-Lichen S = 591.290 - 0.435*R-NO$_3^-$ + 0.345*R-*Fden6*  (5)

where

$R^2$ = 0.324
$p < 0.0005$
n = 1,085.

Land cover variables were not strongly correlated with either climate-related or pollution variables in this region (table 6.5). Model (5) suggests that pollution and landscape pattern have more influence than climate on Lichen S in the region East all. For the E Adiron subregion, models with original data were stronger than those with ranked data (strongest model with ranked data had $R^2$ = 0.324); the only one of our models for which this was the case:

Lichen S = 30.920 - 0.506*NO$_3^-$ - 0.204*pctAg4 - 0.180*elev  (6)

where

$R^2$ = 0.375
$p < 0.0005$
n = 152.

Model (6) suggests that pollution, climate, and landscape pattern all influence Lichen S in the E Adiron subregion. Geographic variables *lat* and *long* were strongly correlated with both climate and NO$_3^-$ in this dataset (table 6.5). The effect of urban and industrial pollution on climate may thus be overestimated, with NO$_3^-$ indirectly representing a climate variable in model (6). This E Adiron model, developed using original rather than ranked data, is more subject to bias from data distribution and other data issues than are other models. The best regression model for subregion E Decid was:

R-Lichen S = 137.761 + 0.534*R-*Fden6* - 0.297*R-NO$_3^-$ - 0.278*R-*long*  (7)

where

$R^2$ = 0.323
$p < 0.0005$
n = 264.

Neither longitude nor land cover variables were strongly correlated with climate variables

in this subregion (table 6.5). Thus, model (7) clearly suggests that landscape pattern and urban and industrial air pollution have more influence than does climate on Lichen S here. Because $NH_4^+$ is strongly correlated with *long* as well as with *Fden6*, the influence of agricultural air pollution is probably represented indirectly in the model.

We found that most of our final linear regression models developed using our very conservative data treatment and analysis choices were stronger ($R^2$ higher by 0.04 to 0.10) than the comparable models reported in Will-Wolf and others (2011a) that used original data and less conservative model development choices, even when a land cover variable was not included in our final model. This is additional evidence that nonstandard distributions of original data probably reduced the power of regression models using the original data. Only for the E Adiron subregion was our best model no stronger than the model for that subregion in the earlier study (Will-Wolf and others 2011a). Best regression models are somewhat stronger for the West (explaining 33 to 47 percent of variation) than for the East (explaining 32 to 38 percent of variation).

## DISCUSSION

Our first hypothesis, that pattern of nearby land cover is linked to Lichen S, is supported at least minimally; a minimum of about 10 percent of the variation in Lichen S is explained by correlation with at least one estimate for nearby land cover in each region. Lichen S is a general index for the condition of a forest lichen community; it is likely that variation in lichen species composition will be even more strongly correlated with neighboring land cover. Our results suggest more intensive studies of the relationship between neighboring land cover and the condition of forests are likely to improve our understanding of impacts on forest health. Regression models for the East support the importance of landscape variables to explain Lichen S more clearly than do models for the West, even though East models are weaker overall than are West models. The land cover variables that we measured have the greatest potential usefulness in the three East regions, and the least potential usefulness in the West Sierras subregion.

Our second hypothesis, that nearby land cover strongly reflects the influence of climate in the West, is supported. In the two West subregions, land cover variables do not clearly improve regression models, while for "West all" they do appear to improve regression models.

These results reflect that local landscape patterns are strongly correlated with climate on forests in the West. More intensive studies should be designed at smaller spatial scales to separate the effects of neighboring land cover from the effects of climate on forest lichens. Such studies will probably use data on abundance of individual species at each site rather than just number of lichen species at a site. They will be required to support conclusions about whether landscape pattern independently affects the condition of forest lichen communities in the West.

Our third hypothesis, that nearby land cover strongly reflects the influence of pollution in the East, is not supported. Correlations of land cover variables with modeled background pollution in the East are mostly moderate to weak. The one exception is for the mountainous Adirondacks region, where pollution is also correlated with climate; possible indirect links of land cover with climate could mean the link between pollution and land cover is overestimated for this region. In the East, landscape variables do improve regression models for all three regions, though interpretation is sometimes difficult. It thus appears that in the East landscape variables contribute information to explaining patterns of lichen species richness, potentially independent of pollution and climate. Lichen S does appear to have some potential as an indicator for condition of forests in the Eastern Deciduous Forest subregion when land cover variables are included, in contrast to our earlier analyses

(Will-Wolf and others 2011a) that did not include neighboring land cover.

Our fourth hypothesis, that relationships of land cover variables with Lichen S and with environmental variables differ strongly between regions, is supported. No single landscape variable appears the best to use in all cases. Indices for percentage of forest cover or percentage of natural to seminatural land cover in a neighborhood were often most strongly correlated with Lichen S. The index for percentage of forest cover was most frequently included in the best regression model, even when it was not the land cover variable having the strongest correlation with Lichen S. The indices for percentage of agricultural land or of developed land also sometimes had strong relationships with Lichen S. Correlations of land cover variables with individual climate and pollution variables differed enough between regions that no general recommendations are made; several standard land cover variables should be tested in each new region. We do conclude from this study that land cover composition (represented by indices for percentage of forest, natural and seminatural, or agricultural) may be more important than forest fragmentation (represented by our index of forest connectivity) to explain variation in Lichen S. We also observed that at these large spatial scales our index for percentage of developed land was usually not as useful as other land cover variables.

Our fifth hypothesis, that inclusion of land cover variables improves models to explain variation in Lichen S, is supported for the East but not for the West. In the East, including land cover improves regression models for all three regions tested, and thus helps to explain patterns in lichen species richness independent of pollution and climate. Our failure to demonstrate the usefulness of land cover variables in the West is, however, not conclusive. More intensive studies across smaller geographic areas in West regions more affected by human land use, as well as the use of data reflecting lichen species composition rather than merely species counts, might find that nearby land cover independently affects the condition of forest lichen communities in those circumstances.

For most of our land cover variables, we could not identify neighborhood sizes associated with the strongest correlations. Strength of correlation often increased continuously to the maximum size, at which neighborhoods for plots in adjacent sample grid cells may overlap and possibly inflate the strength of the correlation. The neighborhood sizes at which correlations of Lichen S with our index for percentage of agricultural land peaked are often smaller than the maximum, suggesting a particular scale of impact. That area is still much larger than the area across which a similar land cover variable had the strongest correlations with lichen community composition in a more intensive study (Will-Wolf and others 2005). This means

our study provides limited support to focus on possible mechanisms even as it highlights the importance of exploring further the impact of land cover pattern on forest lichen communities. The impact of land cover pattern can itself be considered a proxy for more direct causes such as altered disturbance regime, dispersal limitations, competition from invaders, or alterations in other regional ecological processes (Sillett and others 2000, Stofer and others 2006, Werth and others 2006).

Several additional studies are suggested by our results. One is to evaluate the importance of calculating land cover variables using exact as opposed to approximate plot locations. If the latter approach is found to be adequate, use of public FIA data with approximate plot locations is supported and analyses can be conducted much more easily by a wide variety of investigators. Use of exact plot locations in an additional study would also allow evaluation of a smaller neighborhood size that better corresponds with small scales at which effects of landscape pattern on lichens have been found in other studies (Werth and others 2006, Will-Wolf and others 2005). Another useful addition would be to include an evapotranspiration variable to represent climate in future analyses; this could be particularly helpful for the East, where our simpler climate variables are poorly linked to Lichen S. Yet another useful addition to future studies would be to compare correlations with pollution using dry, wet, and total pollution

deposition, to evaluate which of the pollution deposition variables is most strongly correlated with Lichen S and with land cover variables in different regions. Dry deposition has been shown to be a very important contributor to ground-level pollution in Western States (Fenn and others 2003).

Our study clearly suggests that impact of land cover pattern on forest health indicators should be considered further for analysis of FIA data from Eastern States. The results of this study of large areas using Lichen S, a very general indicator of the condition of lichen communities, suggests likely stronger links between land cover and forest response from investigations with more precise forest health indicators in Eastern States. The strong relationship between land cover pattern and climate at large spatial scales in Western States suggests more research is needed to decide whether independent effects of land cover pattern on forest lichen communities can be identified in the West.

## ACKNOWLEDGMENTS

## LITERATURE CITED

Bailey, R.G. 1989. Explanatory supplement to ecoregions map of the continents. Environmental Conservation. 16: 307-309.

Cleland, D.T.; Freeouf, J. A.; Keys, J.E., Jr. [and others]. 2005. Ecological subregions: sections and subsections for the conterminous United States. June 14, 2005 version. Washington, DC: U.S. Department of Agriculture Forest Service. [Map, presentation scale 1:3,500,000; Albers equal area projection; colored].

Cleland, D.T.; Freeouf, J.A.; Keys, J.E., Jr. [and others]. 2007. Ecological subregions: sections and subsections for the conterminous United States. 1:3,500,000; Albers equal area projection; colored. In: Sloan, A.M., tech. ed. Gen. Tech. Rep. WO-76. Washington, DC: U.S. Department of Agriculture Forest Service. Map, presentation scale. Also as a GIS coverage in ArcINFO format on CD-ROM or at http://fsgeodata.fs.fed.us/other_resources/ecosubregions.html. [Date accessed: March 18, 2011].

Coulston, J.W.; Riitters, K.H.; Smith, G.C. 2004. A preliminary assessment of Montreal Process indicators of air pollution for the United States. Environmental Monitoring and Assessment. 95: 57-74.

Daly, C.; Taylor, G. 2000. United States average monthly or annual precipitation, temperature, and relative humidity 1961–90. Arc/INFO and ArcView coverages. Corvallis, OR: Oregon State University, Spatial Climate Analysis Service.

Fenn, M.E.; Haeuber, R.; Tonnesen, G.S. [and others]. 2003. Nitrogen emissions, deposition, and monitoring in the Western United States. BioScience. 53(4): 391-403.

Homer, C.; Dewitz, J.; Fry, J. [and others]. 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. Photogrammetric Engineering and Remote Sensing. 73: 337-341.

McCune, B. 2000. Lichen communities as indicators of forest health. Bryologist. 103: 353-356.

Nimis, P.L.; Scheidegger, C.; Wolseley P., eds. 2002. Monitoring with lichens—monitoring lichens. In: Earth and Environmental Sciences, vol. 7. NATO Science Series. Dordrecht, The Netherlands: Kluwer Academic Publishers. 412 p.

Riitters, K. 2011. Spatial patterns of land cover in the United States: a technical document supporting the Forest Service 2010 Resources Planning Act (RPA) Assessment. General Technical Report SRS-136. Asheville, NC: US Department of Agriculture Forest Service, Southern Research Station. 64 p.

Riitters, K.H.; Wickham, J.D.; Wade, T.G. 2007. An indicator of forest dynamics using a shifting landscape. Mosaic. Ecological Indicators. 9: 107-117.

Sillett, S.C.; McCune, B.; Peck, J.E. [and others]. 2000. Dispersal limitations of epiphytic lichens result in species dependent on old-growth forests. Ecological Applications. 10(3): 789-799.

Stofer, S.; Bergamini, A.; Aragon, G. [and others]. 2006. Species richness of lichen functional groups in relation to land use intensity. Lichenologist. 38(4): 331-353.

U.S. Department of Agriculture (USDA) Forest Service. 2007. Forest Inventory and Analysis (FIA) fact sheet series: lichens community indicator. http://www.fia.fs.fed.us/library/fact-sheets/p3-factsheets/lichen.pdf. [Date accessed: May 8, 2014].

U.S. Department of Agriculture (USDA) Forest Service. 2011. Forest Inventory and Analysis (FIA) Phase 3 field guide – lichens community, version 5.1. http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2012/field_guide_p3_5-1_sec21_10_2011.pdf. [Date accessed: May 8, 2014].

Werth, S.; Wagner, H.H.; Gugerli, F. [and others]. 2006. Quantifying dispersal and establishment limitation in a population of an epiphytic lichen. Ecology. 87(8): 2037-2046.

Will-Wolf, S. 2010. Analyzing lichen indicator data in the Forest Inventory and Analysis Program. Gen. Tech. Rep. PNW-GTR-818. Portland, OR: U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station. 61p.

Will-Wolf, S.; Ambrose, M.J.; Morin, R.S. 2011a. Relationship of a lichen species diversity indicator to environmental factors across continental United States. In: Conkling, B.L.; Ambrose, M.J., eds. Forest health monitoring 2007 national technical report. Gen. Tech. Rep. SRS-147. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station: 25-63.

Will-Wolf, S.; Esseen, P.-A.; Neitlich, P. 2002. Monitoring biodiversity and ecosystem function: forests. In: Nimis, P.L.; Scheidegger, C.; Wolseley P.A., eds. Monitoring with lichens—monitoring lichens. Dordrecht, The Netherlands: Kluwer Academic Publishers: 203-222.

Will-Wolf, S., Makholm, M.M., Roth, J. [and others]. 2005. Lichen bioaccumulation and bioindicator study near Alliant Energy - WPL Columbia Energy Center. Energy Environmental Research Program Final Report. Madison, WI: Division of Energy, Department of Admininistration, State of Wisconsin, and Wisconsin Department of Natural Resources. 57 p. https://focusonenergy.com/sites/default/files/research/willwolflichenlowres_summaryreport.pdf [Accessed May 8, 2014].

Will-Wolf, S.; Nelsen, M.P.; Trest, M.T. 2010. Responses of small foliose lichen species to landscape pattern, light regime, and air pollution from a long-term study in upper Midwest USA. In: Nash, T.H. III; Geiser, L.; McCune, B. [and others], eds. Biology of lichens—symbiosis, ecology, environmental monitoring, systematics, cyber applications. Bibliotheca Lichenologica. 105: 167-182.

Will-Wolf, S.; Nelsen, M.P.; Trest; M.T. 2011b. Does morphological response of four common lichen species to pollution, shade, and landscape pattern predict long-term changes in distribution? In: Bates, S.; Bungartz, F.; Elix, J. [and others], eds. Biomonitoring, ecology, and systematics of lichens: festschrift Thomas H. Nash III. Bibliotheca Lichenologica. 106: 375-386.

Woodall, C.W.; Amacher, M.C.; Bechtold, W.A. [and others]. 2011. Status and future of the forest health indicators program of the USA. Environmental Monitoring and Assessment. 177: 419-436.

Woodall, C.W.; Conkling, B.L.; Amacher, M.C. [and others]. 2010. The Forest Inventory and Analysis database: database description and users manual version 4.0 for phase 3. Gen. Tech. Rep. NRS-61. Newtown Square, PA: U.S. Department of Agriculture Forest Service, Northern Research Station. 180 p.

Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M. [and others]. 2010. The Forest Inventory and Analysis database: database description and users manual version 4.0 for phase 2. Gen. Tech. Rep. RMRS-GTR-245. Fort Collins, CO: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station. 336 p.

Yandell, B.S. 1997. Practical data analysis for designed experiments. New York: Chapman & Hall. 437 p.