

ARTICLE

Extracting gradients from community data with multiple stressors and empty plots: A case study and simulations

Adrienne Kovasi¹  | Bruce McCune² | Sarah Jovan³

¹USDA Forest Service, Six Rivers National Forest, Eureka, California, USA

²Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA

³USDA Forest Service, PNW Research Station, Portland, Oregon, USA

Correspondence

Adrienne Kovasi

Email: adriennekovasi@gmail.com

Funding information

U.S. Forest Service, Grant/Award Number: 19-JV-11261979-116

Handling Editor: Michael P. Perring

Abstract

Ecological community data are used to infer levels of environmental stressors; for example, epiphytic lichen communities can be used to estimate levels of air pollutants. If stressors are so severe that the community of interest disappears altogether, then “empty” plots are recorded. The problem compounds when multiple environmental stressors strongly influence the community. We explore this problem using simulated data and a case study of epiphytic lichens in the Sierra Nevada mountains of the United States. This area experiences relatively high amounts of nitrogen (N) deposition that can suppress lichen communities, but this signal is confounded at high elevations where epiphytic lichens become absent. This combination of severe stressors interferes with extracting a lichen community gradient that corresponds with the air quality gradient for the full elevation range in this study area. We evaluated three general approaches: (1) methods of nonmetric multidimensional scaling that allow inclusion or exclusion of empty plots, (2) the “dummy species” method that populates empty plots with a consistent low abundance, and (3) bypassing problematic distance measures by nonparametric regression of the environmental parameter of interest (N deposition) against total abundance. In the simulated gradient, we added stressors at one or both ends of a dataset with a dominant environmental gradient to demonstrate the effect of extreme suppression of communities by environmental conditions. With lichen and simulated data that are “stressed” at one or both ends of the community gradients, Euclidean and Gower distances distorted the primary gradients in a circular manner that weakened relationships with environmental variables. This is similar to the “horseshoe effect” that is recognized as a problem in ordinations of ecological communities. If one excluded empty plots, ordination axes represented the underlying gradients better with quantitative Sørensen (Bray–Curtis) distance than with Euclidean and Gower distances, but this limits the scope of its application in bioindication. The distortion of environmental gradients can be revealed with ordination phase plots, a new method of projecting environmental gradients, whether linear or nonlinear, onto any community ordination space. The regression-based alternative to ordination

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Ecosphere* published by Wiley Periodicals LLC on behalf of The Ecological Society of America. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

extracted the strongest relationship between N deposition and lichen community data, and included empty plots.

KEYWORDS

bioindication, CMAQ, distance-based methods, dummy species, nitrogen deposition, nonmetric multidimensional scaling, nonparametric multiplicative regression, ordination phase plot, southern Sierra Nevada, TDEP

INTRODUCTION

Bioindication uses organisms to indicate environmental conditions. One common form of this is to use ecological community data to infer or estimate levels of environmental stressors. For example, one can use ordinations of epiphytic lichen communities to determine relationships between axis scores and levels of air pollutants, then use that relationship to estimate levels of air pollutants where the community has been measured but the pollutants were not (e.g., Jovan, 2008; Jovan & McCune, 2005; Will-Wolf et al., 2015). While this strategy is often successful, sometimes the stressors can become so severe that the community of interest disappears altogether. This produces “empty” sample units (e.g., plots) where no species are recorded, presenting analytical challenges to methods that are based on species composition. The empty plot problem becomes particularly severe when more than one environmental stressor suppresses the target community under different circumstances within the same dataset.

Empty plots are rarely encountered in many ecosystems, for example, the Pacific Northwest of North America (Jovan, 2008; Smith et al., 2017) and the Northeastern United States (Cleavitt et al., 2015; Smith et al., 2020; Will-Wolf et al., 2015). These areas have favorable climate regimes resulting in virtually no forested landscapes that lack epiphytic lichens. In these regions, nonmetric multidimensional scaling (NMS; Kruksal, 1964) using common proportional distance measures, such as quantitative Sørensen (Bray–Curtis) distance, effectively extracts air quality gradients from lichen community data.

Other regions, such as the Southwestern United States, are so climatically harsh that forested plots lacking epiphytic lichens are encountered, rendering the proportional Sørensen distance unusable. That region has a single dominant stressor (aridity), and NMS can continue to solve this problem with Euclidean distance, taking advantage of the signal from overall abundance and allowing empty plots in the analysis (McCune et al., 2022).

Another distance measure option is Gower distance (Gower, 1971; McCune & Mefford, 2018), which is one of

the few distance measures that allow both categorical and quantitative variables and can be applied to empty plots, though we could not find examples of this application in the literature. Gower distance is a universal distance measure, in that it can be applied to any data matrix. For quantitative data, Gower distance is a city-block distance, since it is based on the absolute differences between values. It is also proportional: two objects have maximum distance (1) when they have nothing in common, and minimum distance (0) when they are identical, regardless of the variable types. Sørensen and Jaccard are also proportional city-block distances, but they are proportionalized by sample unit totals, rather than Gower's proportionalizing by absolute ranges of the individual variables. This makes Gower's distance relatively sensitive to minor species in the data, rather than being controlled by a few dominant species. Because of the universality of Gower's distance, we include it here as a potential solution to the empty plot problem.

Here, we explore the empty plot problem with a dataset in which it is particularly severe. In the southern Sierra Nevada, epiphytic lichens at high elevations are suppressed by climatic stress and a lack of trees. At the other (low) end of the elevation gradient, a second environmental stressor, very poor air quality, can virtually eliminate epiphytic lichens. The effect of two stressors at opposite ends of the elevation gradient is that the air quality and climatic signals are confounded by the extreme simplification or loss of communities at both ends of the gradient. Most lichen community analyses partially deal with this problem by excluding empty plots in order to use quantitative Sørensen (Bray–Curtis) distances for ordinations (Jovan & McCune, 2004, 2005, 2006). But even strongly suppressed communities with just one or a few species can be problematic for Sørensen distances (Clarke et al., 2006). So, we sought a method to effectively represent all the available data, in our case, including both extreme high and low elevations.

Although the most common solution has been to exclude empty plots, other options exist. Another solution has been to augment the data matrix with a “dummy” species with a constant, low abundance, and then to proceed with ordination based on Sørensen distances

(Clarke et al., 2006). To use this method requires the assumption that there is a single likely cause for empty plots (Clarke et al., 2006). The “dummy” species method has been used in studies of rocky intertidal communities, which often encounter empty plots (Weitzman et al., 2021) as well as communities that experience reductions in diversity and abundance from pollution stress (Sales & Ballesteros, 2009) and seasonal changes in climate (Sobocinski et al., 2013).

In the context of estimating beta diversity and nestedness, the function `beta.pair_zerosppecies` in R package `allanecology/BetaDivMultifun` (unpublished, R documentation 9 November 2023) introduced ad hoc rules to avoid division by zero in calculating proportional distances:

1. If two plots are empty, then dissimilarity = 0
2. If one plot is empty and one plot has species, then dissimilarity = 1.

In the context of ordination with multiple factors suppressing communities, both rules clearly will degrade recovery of environmental gradients, so we did not evaluate them further. For example, Rule 1 degrades the case of two plots stressed by the same factor, but one plot has only one species and one plot has none. Despite similar environment and communities, the dissimilarity is assigned a maximum value. On the other hand, Rule 2 degrades the case when plots have different stressors responsible for the absence of species.

Environmental gradients usually have nonlinear relationships to community gradients (e.g., McCune et al., 2022). NMS deals effectively with nonlinear relationships among species and between species and the environment, but practitioners commonly view community relationships with environmental variables graphically with linear biplots and corresponding linear statistics (R^2 or correlation coefficients). In contrast, phase plots have been used commonly in other fields, such as chemistry and physics (Schnell & Mendoza, 2004), for studying nonlinear dynamical systems. Traditional phase plots graph a series of points in a coordinate system defined by variables that have been recorded through time. For example, if you measured the temperature and humidity of a system through time, a time domain graph shows temperature and humidity as a function of time as the x axis, while the phase plot shows a series of points in time in a coordinate system defined by axes of temperature and humidity. It thus shows the trajectory of the system in the state space defined by the variables of interest. Applied to ordinations of ecological communities, we broaden that definition to include the path of either

temporal or environmental gradients in the state space as represented by an ordination, that is, reduced species space. A phase plot shows the path of an environmental gradient in the community ordination space. These plots can show how key environmental and community gradients are related to each other in species space and can reveal structure that is otherwise unclear. A method for creating gradient phase plots in ordination space (B. McCune and Heather T. Root, unpublished) is applied for the first time here.

Our primary objective was to evaluate alternative solutions to the problem of analyzing community responses to the environment in the face of severe environmental stressors on two ends of a community gradient, stressors so severe that communities effectively disappear at both ends. If there are stressors that are severe enough to produce empty plots at only one end of the community gradient, this still poses analytical problems, but not nearly as severe as for stressors at two ends. To address these, we used both real lichen data and simulated community data. Our second objective was to demonstrate a new method, ordination phase plots, to show how methodological choices may misrepresent underlying gradients, again using both real and simulated datasets.

STUDY AREA

Our study area falls along a nitrogen (N) deposition and elevational gradient (50–3700 m) that extends north-eastward from the agricultural San Joaquín Valley into the granitic southern Sierra Nevada mountains of California (Figure 1). Mean annual temperature decreases and mean annual precipitation increases along this gradient. A strong rainshadow on the eastern edge of our study area starts in local valleys around 2500 m in elevation on the west side and continues eastward (Figure 2).

Like the adjacent interior southwest, the southern Sierra Nevada mountains have a paucity of epiphytic lichens in many areas (Jovan et al., 2021; Jovan & McCune, 2006). Nevertheless, much of the higher elevation epiphyte-poor area falls in Class I Wilderness that requires, under the Clean Air Act, monitoring air quality as an agent of ecological change (Geiser et al., 2021; Jovan et al., 2021).

The San Joaquín Valley has large urban population centers among large expanses of agricultural land, thus being a major source of N emissions (Almaraz et al., 2018; Bytnerowicz et al., 2019). We included four plots located in city parks to capture the high end of N deposition; other plots were on rural federal- or

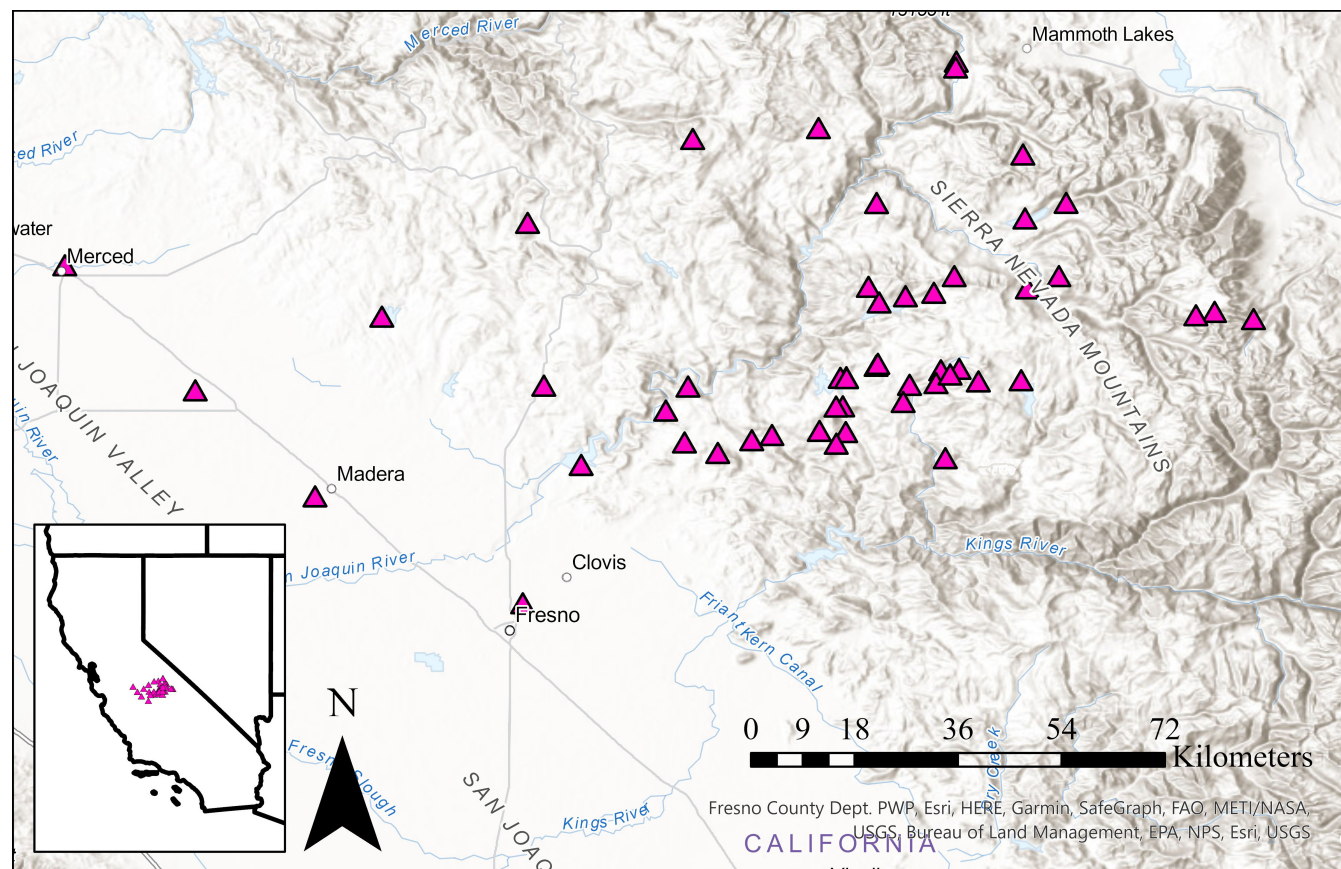


FIGURE 1 Topographic map of study area and plot locations in central California. Solid pink triangles represent plot locations.

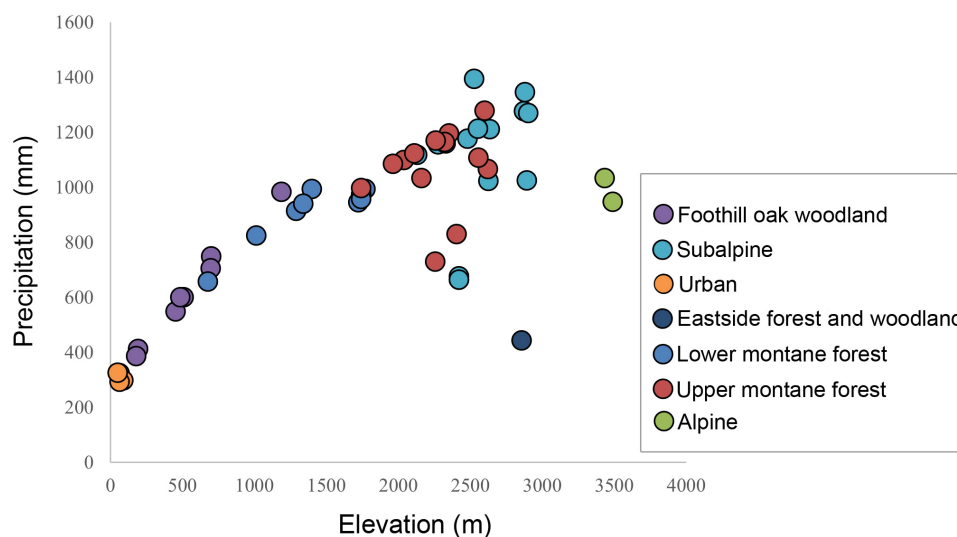


FIGURE 2 Relationship between mean annual precipitation and elevation with ecological zones color coded for each lichen community plot.

state-managed land from the oak woodland foothills up into the alpine high elevations of the Sierra Nevada.

Plots in high N environments favor nitrophiles; the proportional abundance of nitrophiles within the

epiphytic lichen community indicates higher N deposition (Jovan & McCune, 2005, 2006; van Herk, 1999). More sensitive lichen species can tolerate only low-to-moderate levels of N deposition so that plots have

lower species richness and lower abundance of sensitive species at high levels of N deposition.

Sulfur is another common pollutant to which lichens respond (Nash III, 2002), but deposition is low throughout California and thus has little ecological effect in that area, including on lichen communities (Bytnerowicz et al., 2016; Fenn et al., 2010; Jovan & McCune, 2006).

METHODS

Lichen community survey protocol

All existing lichen community data within the study area were extracted from the US Forest Service Air Resources Management program (Jovan et al., 2020) and were collected by both the US Forest Service Air Resources Management program and the Forest Inventory and Analysis program (FIA; Bechtold & Patterson, 2005) over the last 25 years. We added more low elevation and urban plots between the years 2019 and 2021, for a total of 52 plots (Kovasi, 2023). Five plots (9% of the dataset) lacked epiphytic macrolichens.

The epiphytic macrolichen community was sampled in circular 34.7 m radius plots using the standardized FIA protocol (Jovan et al., 2020; McCune et al., 1997). Every macrolichen species found above 0.5 m from the ground on wood as well as in litterfall was collected on the plots over the course of 2 h or less. Each species was assigned a pseudo-logarithmic abundance rating in each plot: 1 = one to three individuals, 2 = four to 10 individuals, 3 = more than 10 individuals but present on less than half of allowable substrate, and 4 = more than 10 individuals and present on at least half of allowable substrate. Elevation, total basal area, and conifer and hardwood basal area were recorded for each plot.

Plot environmental variables

Three-year rolling averages of modeled N deposition estimates came from the Community Multiscale Air Quality model v. 5.0.2 (CMAQ; Bash et al., 2013; Byun & Schere, 2006) produced by the Environmental Protection Agency and from the Total Deposition model v. 2018.02 (TDEP; Schwede & Lear, 2014) produced by the National Atmospheric Deposition Program Total Deposition Science Committee. Model output resolution is 12 × 12 km. CMAQ estimates deposition by modeling chemical transport, meteorology, and emissions inventory data (Byun & Schere, 2006), while TDEP adjusts the CMAQ model by empirical measurements of actual

deposition (Schwede & Lear, 2014). Thirty-year climate normals for environmental variables, such as mean annual precipitation and temperature (Appendix S1: Table S3), were extracted from the Parameter elevation Regressions on Independent Slopes Model (PRISM; Daly et al., 2000). We extracted values from CMAQ and TDEP in ArcMap 5.1 (Environmental Systems Research Institute, 2018) using plot coordinates.

Simulated dataset

Using the program Compas (Minchin, 1987) in Windows NT DOS simulator, we first generated species response functions for two gradients and 100 species. We chose to use simulated data with known underlying structure because it allowed us to cleanly test alternative methods for dealing with the empty plot problem. If the analytical method can recover the underlying gradients, then the method was successful. Use of two gradients simulates the usual case of empirical community datasets being driven by more than one important factor. We used Compas defaults except that the second gradient was weakened compared to the first gradient by giving species broader ranges and sampling a narrower range on the second gradient, relative to the first. Minchin's method uses a generalized beta function to generate species response curves, with default distributions for alpha and gamma applying various levels of skewness and kurtosis in the responses to the gradients. We used the option to adjust modes of major species to have more even spacing along the gradients; this applied to 50% of the species. Next, we generated a 30 × 4 sampling grid of 120 plots. Lastly, we generated quantitative community datasets from the noiseless species response functions along those two gradients. These two gradients are represented by the variables Env1 and Env2.

DATA ANALYSIS

Lichen community data

We chose NMS as a common framework that will accept any distance measure and that excels at extracting nonlinear patterns from species space (McCune & Grace, 2002). The full lichen species matrix (52 plots × 60 species) contained the abundance scores of each species at a given plot, including empty rows (no species present). The accompanying environmental matrix (52 plots × 15 variables) contained the environmental and two lichen community variables discussed elsewhere (Appendix S1: Table S3). A second version of both

matrices excluded the empty plots. The reduced species matrix consisted of 47 plots $\times$ 60 species and the environmental matrix was 47 plots $\times$ 15 variables. The dust bunny index was 0.84, which indicates a high degree of departure from multivariate normality (McCune & Root, 2015). The dust bunny index indicates for a community dataset the degree of departure from multivariate normality toward the usual zero-rich condition, with zero the expected value for multivariate normality and 1.0 the expected value for a perfect dust bunny (all points fall on $p-1$ -dimensional corners of the p -dimensional space; McCune & Root, 2015). Whittaker's beta diversity = 4.1 when empty plots were excluded, 86% zeros, and an average of eight species per plot. A third matrix was created from the full matrix using the Clarke et al. (2006) method; we added a dummy species with an abundance of one for all the plots (52 plots $\times$ 61 species). "Air scores" were calculated using the following methods from Jovan and McCune (2006). The purpose of air scores is to utilize lichen biomonitoring data to tell us something about air quality in a streamlined way. The proportion of nitrophile abundance (PNA) is:

$$\text{PNA}_i = \frac{\sum_{j=1}^s x_{ij} w_j}{\sum_{j=1}^s x_{ij}}, \quad (1)$$

where s is the number of species, x_{ij} is the abundance of species j in plot i , and w_j is a vector of binary values indicating whether each species is a nitrophile ($w_j = 1$) or not ($w_j = 0$). The expected PNA, Y , for a given elevation (in meters) was calculated with the following equation derived from nonlinear regression (Jovan & McCune, 2006):

$$Y = \frac{0.48}{1 + (\text{elevation}/1689)^{5.91}}. \quad (2)$$

Air scores were then generated by multiplying the unstandardized residuals by 100:

$$\text{Air score} = 100 (\text{PNA} - Y). \quad (3)$$

We referred to Jovan and McCune (2004, 2005, 2006) to classify species as nitrophilous or not. Because PNA is proportional to plot totals, it cannot be calculated for empty plots, which prevents calculating air scores. To bypass this problem, we set $\text{PNA} = 0$ for empty plots because empty plots were at high elevations where no nitrophiles were expected. This assumption was supported by observations of the non-nitrophile *Letharia vulpina* below 0.5 m (excluded by sampling rules) in the otherwise empty plots. Plots that had lichens, but no

nitrophiles, also had $\text{PNA} = 0$. We ignored the dummy species when calculating the air scores.

Species present in two or fewer plots were removed to reduce noise and improve relationships to environmental gradients. This reduced the total number of species from 60 to 42; 43 with the dummy species. No plots were identified as outliers in the reduced or dummy datasets using an average Sørensen distance from other plots of three SDs beyond the grand mean of distances (McCune & Grace, 2002). Similarly, no outliers were identified for the full dataset using Euclidean and Gower distances.

Simulated community data

The quantitative community abundance values were rounded to the nearest integer (range 0–53), then $\log(x + 1)$ transformed. Empty columns (species) resulting from rounding very small values were deleted. The rounded and log-transformed data had 120 plots $\times$ 31 species remaining, and a moderate departure from multivariate normality, as measured by the dust bunny index = 0.71 (McCune & Root, 2015), Whittaker's beta diversity = 1.6, 62% zeros, and an average of 12 species per plot.

To introduce additional environmental stressors, we suppressed communities at one or both ends of the main gradient by multiplying each raw abundance by a proportion from 0 to 1. For a single stressor, we multiplied abundances at one end of the gradient by 0, the other end by 1, and a linear series of proportions between 0 and 1 at each of the 29 steps between the ends. For two stressors, we used similar multipliers from both ends of the gradient, with 0 at the ends and 1 in the middle, and evenly ramping up and down around the middle. We then adjusted each dataset by rounding to the nearest integer, $\log(x + 1)$ transforming, and deleting empty species that had an abundance of zero for all plots after the sampling.

This yielded a full matrix of 120 plots $\times$ 24 species with one stressor, including four empty plots, and a full matrix of 120 plots $\times$ 21 species with two stressors, including eight empty plots. This adjustment increased the dust bunny index only slightly to 0.76 with one stressor and 0.73 with two stressors. The small range in the dust bunny index among the three variants of the dataset indicates that the degree of departure from multivariate normality was similar among all three. To apply the dummy species method of avoiding empty plots, we also augmented each log-transformed matrix with a dummy species containing a constant abundance of 0.301, the smallest nonzero value in each matrix, increasing the species count for each matrix by one.

Nonmetric multidimensional scaling

We summarized community relationships in species space with NMS (PC-ORD v. 7.07; McCune & Mefford, 2018), using both real and simulated datasets. We used Euclidean and Gower distance on matrices that had empty plots and Sørensen distance on all matrices that did not. For equations and properties of these measures, see Beals (1984), McCune and Grace (2002), and Clarke et al. (2006). We did not use Euclidean distance or Gower distances on the datasets without empty plots, because we wanted to retain the most complete dataset allowed by each method. For the lichen data, a 2-D solution was recommended based on 100 runs with both real and randomized data from random starting configurations, tie penalization, and a maximum of 200 iterations per run. Stable solutions were verified by checking plots of stress by iteration number and final instability. We calculated coefficients of determination between original plot distances and distances in the final ordination solutions to determine how much variability in lichen community composition was represented by the NMS axes. The same NMS procedure was used with the simulated data, but forcing a 2D solution, because two environmental factors completely determined this species matrix.

Ordination phase plots

A new phase plot method displayed selected environmental gradients in relationship to ordinations of plots in species space. Details of the conceptual steps and calculations are in Appendix S2. The gradients can be either temporal or environmental. Although standard biplots or joint plots also superimpose environmental gradients on the ordination space, they assume that the relationship is linear (McCune & Grace, 2002), while phase plots can show both linear and nonlinear gradients. Furthermore, monotonicity is not required for either the species relationships or ordination axis relationships to the environmental variable. Phase plots can be applied to any community ordination method and distance measure. Phase plots can reveal how environmental gradients are represented by different ordination methods and different distance measures.

We constructed phase plots by first fitting 100-step nonparametric regression (NPMR) curves (e.g., Bowman & Azzalini, 1997, chapter 3) in HyperNiche v. 2.30 (McCune & Mefford, 2009) of Axis 1 and 2 NMS ordination scores separately against key environmental variables using a local mean with a Gaussian kernel, the smoothing parameters optimized by maximizing the leave-one-out cross-validated R^2 . We then used these regressions to calculate the fitted values of the ordination scores for 100 incremental steps along the environmental gradient, which allowed us to

superimpose the environmental gradient in the same coordinate system as the ordination of plots in species space. We applied the same method to ordinations of the simulated datasets to show how the known underlying environmental gradients, those used to generate the species responses, were represented in the community ordinations. An ideal, simple, easily interpreted outcome for an environmental gradient in a phase plot on a community ordination is a long straight line, but phase plots will also reveal strongly nonlinear relationships between an environmental gradient and the ordination space.

Nonparametric multiplicative regression

Finally, we addressed empty plots a different way using nonparametric multiplicative regression (NPMR; McCune, 2006) as implemented in HyperNiche, regressing modeled nitrogen deposition variables (CMAQ and TDEP N) against a pool of predictors consisting of lichen community and environmental variables. The predictors were the same 13 environmental variables used in the ordinations, in addition to the lichen variables PNA, total abundance, and richness (Appendix S1: Table S3). Abundance was the sum of the abundance of each lichen species at a given plot, and richness was the number of species in a given plot. NPMR fits a response surface to the predictors using a kernel smoother (local mean, Gaussian kernel). Smoothing parameters and predictors are selected to maximize fit, subject to overfitting controls. Fit was evaluated by leave-one-out cross-validation to calculate $\times R^2$ (cross-validated coefficient of determination). We chose the “conservative” software setting to force stiffer response surfaces to help bridge gaps in the data along the primary environmental gradients. We set the minimum average neighborhood size to 20% of available plots. Smoothing parameters were expressed as “tolerance,” the range of a predictor covered by one SD of the Gaussian kernel function. We evaluated the importance of individual parameters in the model by sensitivity analysis. Sensitivity is the root mean squared difference in response resulting from nudging the predictors, expressed as a proportion of the range of the response variable. For parsimony, we required a minimum gain of $\times R^2$ of 0.08 to add a predictor to the model.

RESULTS

Distance measure comparison

Community patterns had stronger relationships to environmental variables when using Sørensen distance versus

the nonproportional Euclidean and Gower distances, with Gower performing worse than Euclidean. The Sørensen NMS ordination had a final stress of 16.8, less than expected by chance ($p = 0.02$). Stress measures departure from monotonicity between the distance matrices in ordination space and the original space; the lower the stress, the better the representation of the distance matrix by the ordination. Final instability was adequate at 0.001. The ordination was rotated to align Axis 1 with mean temperature, because it was the environmental variable with the strongest linear relationship to the community ordination (Figure 3). Two axes captured 79% of the variation in lichen communities (Axis 1, 54.8%; Axis 2, 24.1%). Axis 1 represented a community gradient with increasing abundance of *Candelaria pacifica*, *Melanelixia californica*, and *Physconia isidiigera* and decreasing abundance of *Letharia vulpina* and *Ahtiana sphaerosporella*. This community gradient had strong linear relationships ($R^2 > 0.45$) to numerous environmental variables (Appendix S1: Table S1): air scores, CMAQ N, PNA, temperature, and maximum vapor pressure deficit increased along Axis 1, while total basal area, elevation, and precipitation decreased. TDEP N had only a weak relationship with Axis 1 ($R^2 = 0.09$). Axis 2 had no strong linear relationship to any measured environmental variable.

Using Euclidean distance, the NMS 2-D ordination had a final minimum stress of 12.2, lower than expected by chance ($p = 0.02$). We rotated it to align with mean temperature, to aid comparison to the other ordinations (Figure 3). Final instability was 0.00029; 87.9% of the variation in the lichen community was captured by the two NMS axes (Axis 1, 50.5%; Axis 2, 37.4%). Species relationships to Axis 1 were similar to the Sørensen ordination, except that *Physconia perisidiosa* rather than *M. californica* increased strongly along it. Positive Axis 1 scores are most strongly related ($R^2 > 0.2$) to decreasing elevation and precipitation, and increasing maximum vapor pressure deficit, PNA, air scores, and temperature. Both models of N deposition correlated weakly with Axis 1 ($R^2 < 0.2$). Many of these variables also had strong correlations in the Sørensen ordination, although overall the R^2 values were lower using Euclidean distances. Axis 2, as in the Sørensen ordination, was not correlated strongly with any environmental variables ($R^2 < 0.07$).

Using Gower distance, the NMS 2-D ordination yielded slightly poorer results than Euclidean distance, with final stress of 16.9 and weaker correlations with environmental variables (Appendix S1: Table S1). Gower and Euclidean distances brought the urban and high-elevation empty plots very close to each other in species space, while the two groups of plots should ideally have been placed at opposite ends of the environmental gradient (Figure 3).

The contrast between ordinations based on nonproportional distance measures (Euclidean and Gower) and a proportional distance measure (Sørensen distance) is illustrated by the shift in position of the four urban plots (Figure 3). Using Sørensen distance, the urban plots were close to plots with much higher lichen richness and abundance than would be expected based on the limited diversity in the urban plots. This derived from the fact that the few species present at the urban plots were most commonly found in the high richness and abundance sites. Moreover, this positioning strengthened correlations of Axis 1 with climate and other environmental variables because the urban plots are similar climatically to foothill plots that had higher lichen richness and abundance. With Euclidean and Gower distances, however, urban plots fell close to high-elevation plots with zero to low abundance of lichens, even though they had no species in common. Euclidean and Gower distances considered the plots with overall low abundance of lichens to be close in species space, even though they were at opposite ends of the elevation, CMAQ N, and climatic gradients.

Dummy species analysis

Using Sørensen distance for NMS on the dummy species matrix yielded a 2-D solution with final stress of 11.0, significantly lower than expected by chance. As for the other ordinations, this one was also rotated by mean temperature (Figure 3). Final instability was 0.04; 80.7% of the variation in the lichen community was captured (Axis 1, 53.7%; Axis 2, 27.0%). The species strongly related to Axis 1 were the same as in the Sørensen ordination of the nondummy species matrix. Axis 1 had strong linear relationships ($R^2 > 0.45$) to longitude, elevation, temperature, maximum vapor pressure deficit, and PNA (Appendix S1: Table S1). All of the correlations were stronger than for the ordinations using nonproportional distance measures, but weaker than for the Sørensen ordination. Axis 2, as in the other Sørensen ordination, was at most weakly correlated with the environmental variables. The urban and empty plots are somewhat close to each other in species space, but further away than when using Gower and Euclidean distances (Figure 3).

Ordination phase plots

The Sørensen ordination phase plots had straighter elevation (Appendix S1: Figure S1), CMAQ N deposition (Figure 4), and precipitation (Appendix S1: Figure S1)

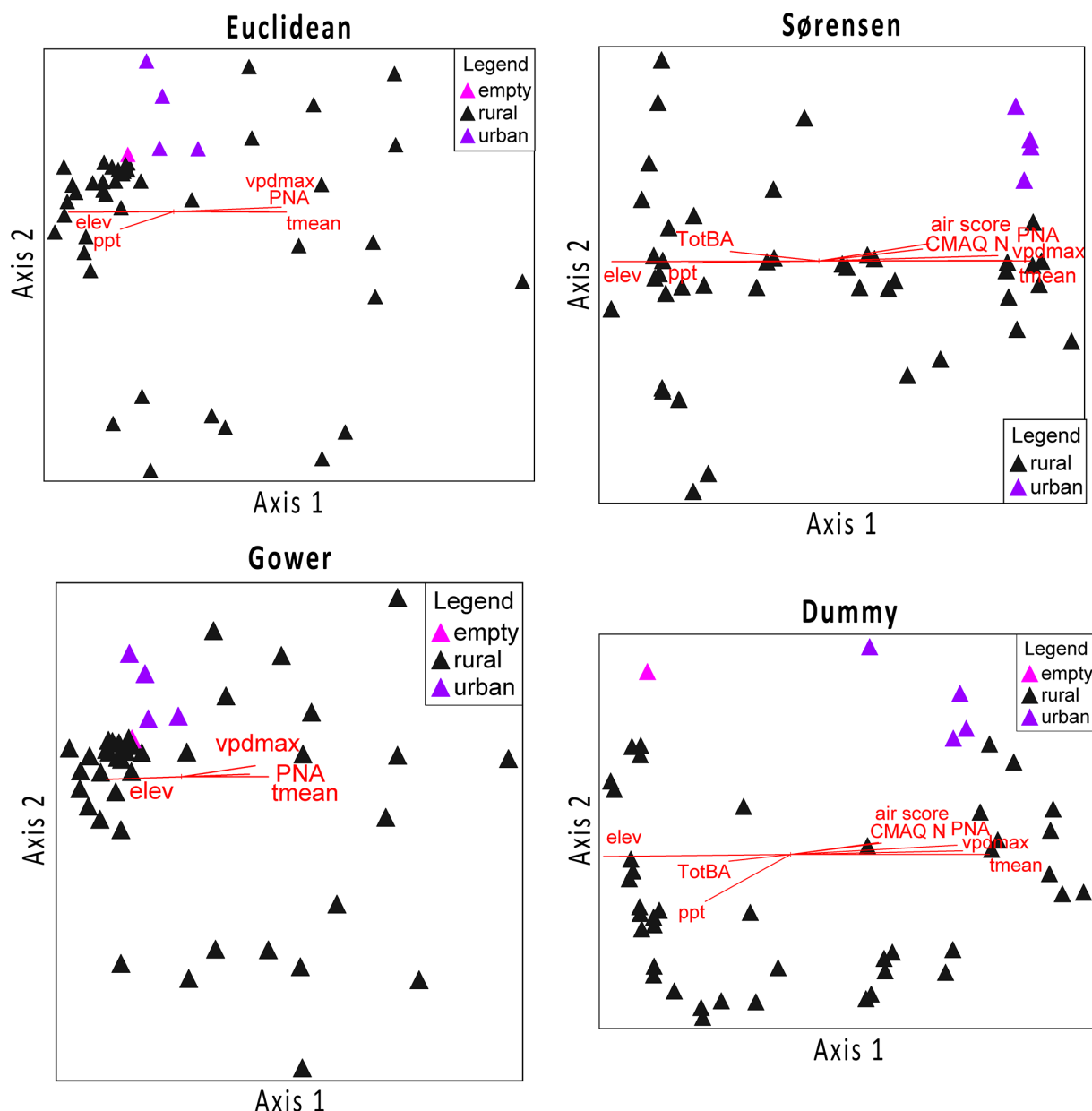


FIGURE 3 Euclidean, Gower, and Sørensen (Bray–Curtis) distance nonmetric multidimensional scaling ordinations and the dummy variable ordination of plots (solid black triangles) in species space overlaid with red vectors of environmental and lichen community variables ($R^2 > 0.2$). Latitude and longitude were removed to alleviate crowding of vector labels, but correlations can be seen in Appendix S1: Table S1. Abbreviations for environmental vectors shown are: elev, elevation; ppt, mean annual precipitation; tmean, mean annual temperature; vpdmax, maximum annual vapor pressure deficit; TotBA, total basal area; CMAQ N, Community Multiscale Air Quality model v. 5.0.2 modeled N deposition estimates; PNA, proportion of nitrophile abundance; and air score, Jovan & McCune (2006) air scores.

gradient lines through species space than did the Euclidean, Gower, and dummy ordination phase plots (Figure 4; Appendix S1: Figure S1). The phase plots also exposed sampling gaps on the environmental gradients. The Euclidean, Gower, and dummy phase plots' environmental gradient lines are generally rather circular, with the elevation and CMAQ N gradient lines being more circular than those for precipitation. The plots are

compressed at one end of the environmental gradient line and become spread out at the other end. The straight segments of the gradient lines in the Euclidean elevation and CMAQ N phase plots show where we had to manually connect the gradient lines because of gaps in the elevational and CMAQ N gradients; that is, we did not have plots at those missing elevations and CMAQ N deposition amounts.

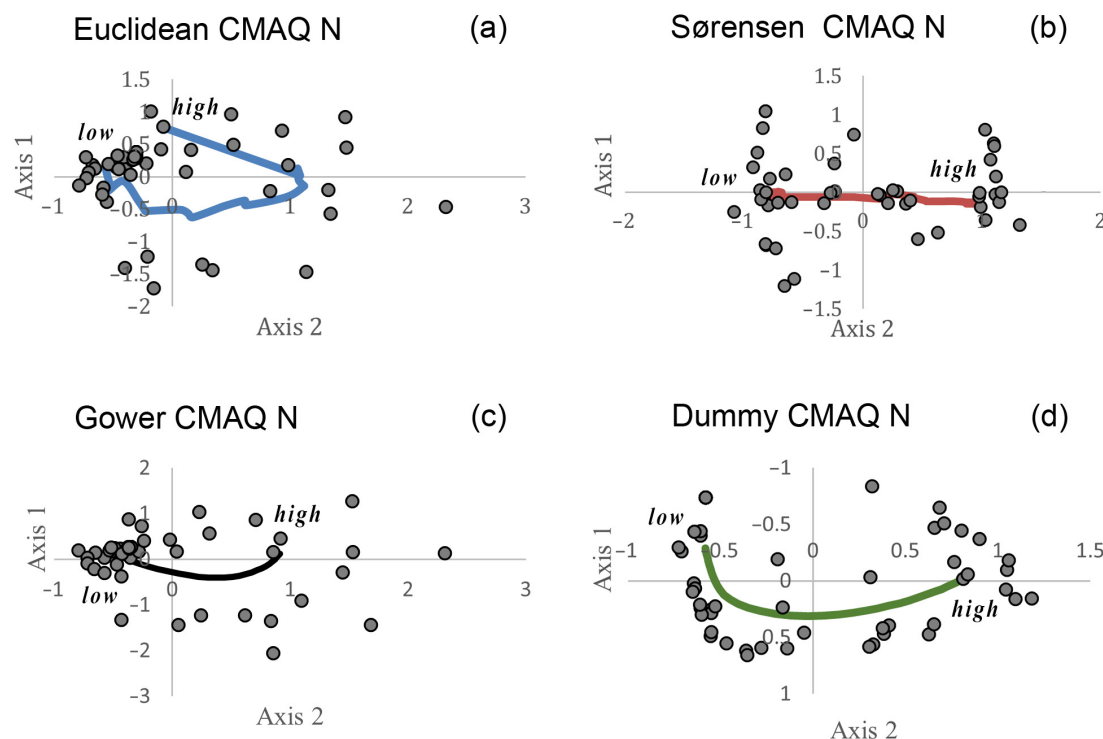


FIGURE 4 Euclidean, Sørensen, Gower, and dummy nonmetric multidimensional scaling (NMS) ordination phase plots showing environmental gradient lines of Community Multiscale Air Quality model v. 5.0.2 (CMAQ) N deposition. A point on a phase plot line represents the average position of plots near that point on the environmental gradient in the ordination space. Environmental gradient lines are labeled “High” at the high end of the deposition gradient and “Low” at the low end of the gradient.

Distance measure comparison for simulated data

All NMS ordinations on the simulated data had a final stress ≤ 0.2 , found stable solutions, and were rotated to align with Env1. The results of NMS ordinations on the original simulated dataset with no stressors, and thus no empty plots, were quite different depending on the choice of distance measure (Figure 5a,d,g,j). Correlations between the main environmental gradient, Env1, and Axis 1 were, however, high for all ordinations ($R^2 = 0.97$). Correlations between Env1 and Axis 1 using Gower distance were slightly lower ($R^2 = 0.94$). The Env1 gradient is expressed as a straight line parallel to Axis 1 using Sørensen distance (Figure 5d), while with Euclidean and Gower distance the Env1 gradient was arched (Figure 5a,g). Even worse, using Euclidean and Gower distances, Axis 2 had almost no relationship with the secondary environmental gradient, Env2, while with Sørensen distance the community gradient expressed by Axis 2 was closely related to Env2 ($R^2 = 0.9$).

With a single stressor added to one end of the community gradient, all ordinations curved the main gradient, but more extremely so with Gower and Euclidean

distances (Figure 5b,h). Axis 1 using Euclidean distance had a weaker, but still strong relationship with Env1 ($R^2 = 0.82$). The correlation involving Gower distance was weaker ($R^2 = 0.7$). Axis 1 using Sørensen distance maintained a strong relationship with Env1 ($R^2 = 0.97$). Axis 2 with Gower or Euclidean distance had no relationship to Env2. Application of the single stressor along the main gradient destroyed the relationship between Axis 2 and Env2 using Sørensen distance ($R^2 = 0.05$).

Adding the second stressor at the opposite end of the community gradient made the environmental gradients even harder to detect. Ordination with Euclidean and Gower distances only weakly found the major gradient, Env1 ($R^2 = 0.5$ and 0.43 , respectively), and there were no relationships between Axis 2 and Env2. Ordination with Sørensen distance did effectively detect a relationship between Axis 1 and Env1 ($R^2 = 0.95$) but detected no relationship between Axis 2 and Env2. Ordination with Euclidean and Gower distances and two stressors represented the main gradient as a circle with the opposing ends of the community gradient nearly touching each other in species space (Figure 5c,i). The Sørensen ordination became strongly arched (Figure 5f).

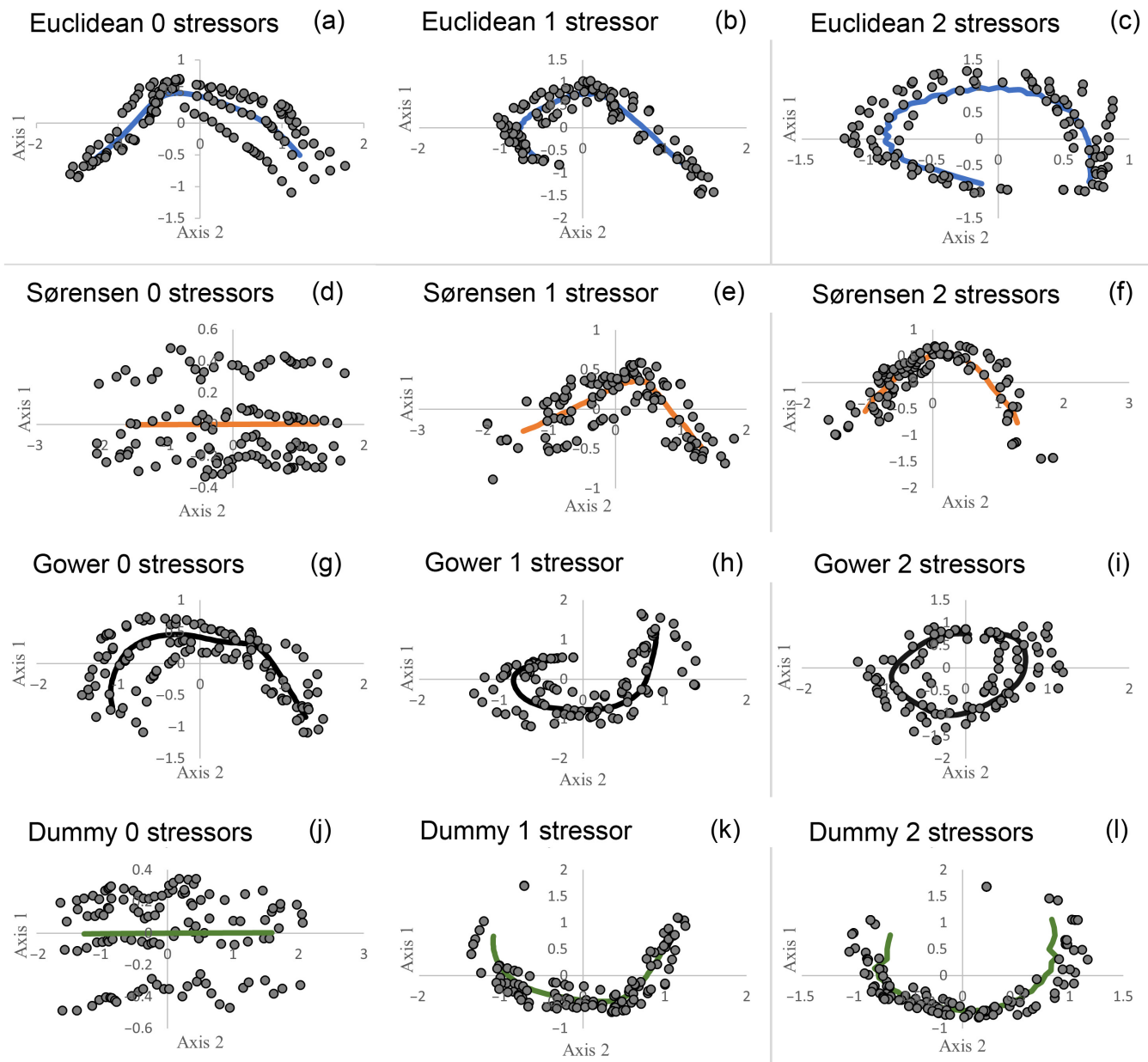


FIGURE 5 Nonmetric multidimensional scaling (NMS) ordination phase plots of simulated datasets with the addition of stressors and using Euclidean, Gower, or Sørensen distances, as well as the dummy method. Points are plotted on a sampling grid along one major and one minor environmental gradient, the minor gradient independent of the major gradient. The phase plot overlay shows the main environmental gradient, Env1, as lines through ordination space. In all cases, the species have noiseless response curves to the two environmental factors. The ideal representation in the ordination would be a straight line phase plot and a rectangular grid of plots that is elongate along Axis 1.

Simulated data dummy species analysis

NMS with Sørensen distance of the dummy species dataset exhibited similar results to the Euclidean and Gower ordinations of the one- and two-stressor datasets, with opposing ends of the community gradient becoming close in species space (Figure 5c,i,l). Correlations between Env1 and Axis 1 were strong for

the simulated dummy dataset with no stressors ($R^2 = 0.97$), while correlations with Env2 were low ($R^2 = 0.003$). Correlations with Env1 weakened as stressors were added, with R^2 decreasing slightly to 0.95 for the one-stressor dataset, then decreasing sharply to $R^2 = 0.20$ for the two-stressor dataset. Correlations with Env2 stayed low throughout the analyses.

Applying NPMR to the empty plot problem

A NPMR model effectively estimated nitrogen deposition (CMAQ N) from two predictors, total lichen abundance and elevation (NPMR: $\times R^2 = 0.81$, $p = 0.01$). Tolerances for predictors are smoothing parameters in the model, with smaller smoothing parameters for more influential variables. Raw tolerances, expressed in the original units of the predictors, were 15.5 for total abundance and 344 m for elevation; relativized into a percentage of the range of those predictors, 25% for abundance and 10% for elevation. Sensitivities express the influence of predictors by nudging the input values and recording the magnitude of the response. The sensitivity was 0.13 for abundance and 0.37 for elevation, showing elevation to be the more influential variable.

The nonlinear relationship of CMAQ N to elevation and abundance is illustrated by a 3-D response surface (Figure 6). At high elevations, CMAQ N is low regardless of lichen abundance, but between about 1500 and 2500 m, lichen abundance has a counterintuitive pattern, with more lichen abundance indicating higher N levels. At low elevations, N is high and lichen abundance is informative as to just how high and with the expected pattern: more N means less lichen abundance. At middle elevations (1000–1500 m) lichen abundance is moderate to high, but not informative about N levels. Our model could be used to estimate local N deposition for future plots with known lichen abundance and elevation using HyperNiche's prediction tool. In general, modeled CMAQ N ranges from 2.5 to 12.5 kg ha⁻¹ year⁻¹, total lichen species abundance from 0 to 62, and elevation ranges from 53 to 3494 m (Appendix S1: Table S4). Modeled estimates of CMAQ N in relationship to lichen abundance and elevation are in Appendix S1: Table S4.

The best model for TDEP N deposition ($\times R^2 = 0.67$) was a three-predictor model involving lichen species richness, longitude, and precipitation. In this model, TDEP N decreases as longitude increases (becomes more easterly) and lichen richness increases (Figure 6; precipitation not shown).

DISCUSSION

A primary goal of community analysis is to be able to indicate or estimate environmental variables based on species composition. In this case, we wished to estimate total N deposition from lichen community data. Most commonly, this has been done with community ordination techniques (Jovan, 2008; Smith et al., 2020; Will-Wolf et al., 2015).

Bioindication from species composition is challenged when stressors are so severe that the community of interest may be completely absent, producing “empty plots.”

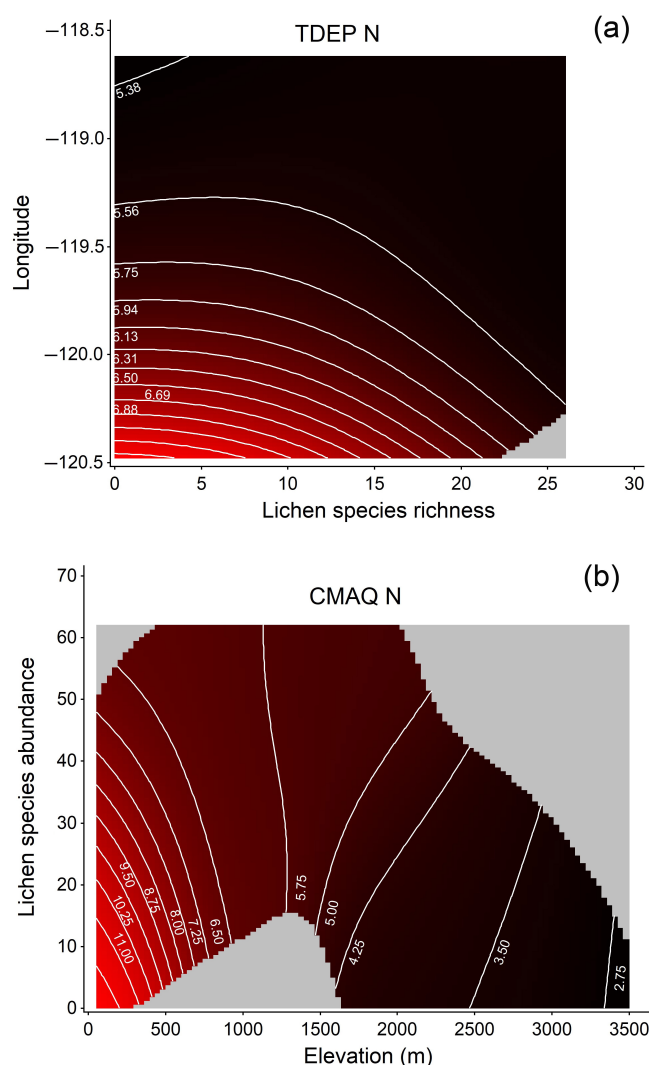


FIGURE 6 Nonparametric multiplicative regression model contour plots of response surfaces. (a) Isolines of Total Deposition model v. 2018.02 (TDEP) N in relationship to longitude and lichen species richness. (b) Isolines of Community Multiscale Air Quality model v. 5.0.2 (CMAQ) N response surface is in relationship to elevation and total lichen species abundance. Gray areas have insufficient data to estimate the response.

How can we effectively recover an environmental gradient under these conditions? Our trials with various methods designed to overcome this community analytical problem proved ineffective with both real and simulated data. These fixes included applying nonproportional distance measures (Euclidean, Gower), deleting empty plots, and introducing dummy species.

Distance measure comparison

With both real lichen data and simulated data that are “stressed” at one or both ends of the community

gradients, Euclidean and Gower distances distorted the primary gradients in species space in a circular manner that weakened relationships with environmental variables. This is similar to, but more extreme than, the “horseshoe effect” or “arch effect” that has long been recognized as a problematic phenomenon in ordinations of ecological communities; the arch effect is predictably severe with some ordination methods (Austin & Noy-Meir, 1971; Beals, 1973). However, the problem becomes especially severe when both ends of the environmental gradient have suppressed communities and the two ends of the arch are pulled together by the empty or near-empty sample units into a circular pattern. On the other hand, the use of Sørensen distances represented the environmental gradient more linearly in the ordination space, but required excluding empty plots, thus shortening the environmental gradient.

Ordination phase plots showed how environmental gradients ran through species space, whether straight or curved, highlighting how the community ordinations related to key environmental gradients. While one can also capture nonlinear gradients with contour overlays, only one gradient can be displayed at once, while multiple environmental gradients can be shown in a single phase plot. Furthermore, the reduction of the pattern to a single curved or straight line more clearly connects with the concept of a gradient than a contour overlay or other response surface.

Dummy species analysis

With the real data, adding a dummy species to enable use of Sørensen distance for all plots distorted the main environmental gradient into a strongly arched pattern, as shown by the phase plots. This is because the near-empty urban plots and empty plots at high elevations were both dominated by the dummy species, placing them near each other in species space. Similarly, for simulated data, the dummy species method distorted the main environmental gradient into an arched pattern for one stressor and a nearly circular pattern for two stressors.

Ecological explanations for empty plots are often more complicated than a single stressor, so exclusion of the empty plots often makes more sense than using the dummy species method (Elahi et al., 2013; Jovan & McCune, 2005). The dummy species method should only be used when there is a single cause for empty plots (Clarke et al., 2006).

Applying NPMR to the empty plot problem

NPMR bypassed the limitations of the distance measures underlying community ordinations and effectively predicted

CMAQ N. The strongest model represented 81% of the variance in N deposition based on total lichen species abundance and elevation, while retaining empty plots, even though using total abundance ignores the information present in species composition.

Our regression model was better able to predict CMAQ N than TDEP N. The TDEP N deposition model has been found to have weaker relationships than the CMAQ N model to lichen elemental concentration data in another study (Kovasi et al., 2022). Likewise, our Sørensen NMS ordination found that Axis 1 scores had a strong relationship with CMAQ modeled estimates of N, but not TDEP N. The CMAQ and TDEP models did not exist at the time of the Jovan & McCune (2006) study, but the relationship of the Jovan & McCune (2006) air scores to CMAQ N is also stronger than to TDEP N. Because a subset of our plots ($n = 9$) was monitored for N deposition in 2012 and 2013 (Bytnerowicz et al., 2019), we compared their measurements to CMAQ and TDEP N estimates from that time period and found CMAQ related more strongly to direct N measurements than did TDEP (data not shown).

While our regression model used total lichen species abundance and elevation to predict CMAQ N deposition, this is not a causal model. In fact, both N deposition and elevation affect lichen abundance. Recognizing that all three variables have error, we used lichen abundance and elevation as predictors of N deposition (in kilograms per hectare per year) to facilitate practical application of the lichen data.

Geiser et al. (2021) also used lichen community data to create estimates of CMAQ N deposition using community weighted means of species deposition optima for the entire United States. While this method is similar to our NPMR approach, it does not work well in the southern Sierra Nevada given the low species diversity and inability to make estimates for empty or near-empty plots; their method requires at least five species with calculated deposition optima to calculate an air score. This would have excluded almost half of our plots (23 of 52). The Jovan & McCune (2006) method allows calculation of air scores with just one epiphytic lichen species present, but it has other drawbacks. First, these air scores are not estimates of deposition. Instead, they are unstandardized residuals from regression of PNA against elevation. Their air scores are thus the differences between the observed and expected PNA values, with a possible range of -100 to $+100$. Second, because the regression is based on elevation, and climate change will surely alter climate relationships with elevation (Rangwala & Miller, 2012), the applicability of the method will degrade over time.

Our regression of CMAQ against elevation and lichen abundance provides the most effective method to date for

estimating total N deposition from lichen community data in the study area, even though it does not take advantage of information on species composition. Future work should evaluate this regression method with expanded geographic scope, by drawing on the extensive lichen community surveys outside of our study area.

The lichen community and simulated datasets showed us that ordinations of lichen community data based on Euclidean, Gower, or Sørensen distance each had their own problems with representing the strong air quality and climatic gradients of the southern Sierra Nevada. Sørensen distance was more effective than Euclidean and Gower distances when used on datasets that are stressed on both ends of the community gradient, but this distance measure requires exclusion of empty plots such that the full range of the underlying environmental gradients is not represented. NPMR of N deposition against lichen species abundance and elevation solved this problem for the lichen community data, effectively predicting modeled N deposition while still retaining empty plots. This investigation provides more evidence that epiphytic lichen communities in the southern Sierra Nevada have strong relationships to both climate and N deposition.

AUTHOR CONTRIBUTIONS

Adrienne Kovasi: Conceptualization; data curation; formal analysis; investigation; methodology; resources; visualization; roles/writing—original draft; writing—review and editing. **Bruce McCune:** Conceptualization; supervision; writing—review and editing. **Sarah Jovan:** Conceptualization; funding acquisition; supervision; writing—review and editing.

ACKNOWLEDGMENTS

This research was funded by the US Forest Service Pacific Northwest Research Station through the Joint Venture Agreement 19-JV-11261979-116 between the US Forest Service and Oregon State University. The USFS Region 5 Air Resources Management Program paid for fieldwork expenses. The Forest Inventory and Analysis program provided critical information for completing fieldwork. Rob Smith and Rick Shory supplied data and answered questions. Daphne Stone and Mike Haldeman assisted with lichen identification. Don Schweizer, Lupe Amezcuita, and Trent Procter helped with fieldwork logistics. Don Schweizer also reviewed an early draft of this paper. Evan Kovasi and Peri Lee Pipkin assisted with field data collection. Julie Pett-Ridge, Steve Perakis, Pat Muir, and two anonymous reviewers provided valuable feedback on the manuscript. Thanks to Eastman Lake (Army Corps of Engineers), Millerton Lake State Recreation Area, and the San Joaquín River Gorge for granting collection permits.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data (Kovasi, 2023) are available from the Oregon State University ScholarsArchive@OSU: <https://doi.org/10.7267/pv63g710j>.

ORCID

Adrienne Kovasi  <https://orcid.org/0000-0002-4439-7818>

REFERENCES

- Almaraz, M., E. Bai, C. Wang, J. Trousdell, S. Conley, I. Faloona, and B. Z. Houlton. 2018. "Erratum: Agriculture Is a Major Source of NO_x Pollution in California (Science Advances)." *Science Advances* 4: 1–9. <https://doi.org/10.1126/SCIADV.AAU2561>.
- Austin, M. P., and I. Noy-Meir. 1971. "The Problem of Non-Linearity in Ordination: Experiments with Two-Gradient Models." *Journal of Ecology* 59: 763–774. <https://doi.org/10.2307/2258138>.
- Bash, J. O., E. H. Cooter, R. L. Dennis, J. T. Walker, and J. E. Pleim. 2013. "Evaluation of a Regional Air-Quality Model with Bidirectional NH₃ Exchange Coupled to an Agroecosystem Model." *Biogeosciences* 10: 1635–45. <https://doi.org/10.5194/bg-10-1635-2013>.
- Beals, E. W. 1973. "Ordination: Mathematical Elegance and Ecological Naivete." *Journal of Ecology* 61: 23–34. <https://doi.org/10.2307/2258914>.
- Beals, E. W. 1984. "Bray-Curtis Ordination: An Effective Strategy for Analysis of Multivariate Ecological Data." *Advances in Ecological Research* 14: 1–55. [https://doi.org/10.1016/S0065-2504\(08\)60168-3](https://doi.org/10.1016/S0065-2504(08)60168-3).
- Bechtold, W. A., and P. L. Patterson. 2005. *The Enhanced Forest Inventory and Analysis Program - National Sampling Design and Estimation Procedures. General Technical Report (Issue SRS-80)*. Asheville, NC: USFS, Southern Research Station.
- Bowman, A. W., and A. Azzalini. 1997. *Applied Smoothing Techniques for Data Analysis*. Oxford: Clarendon Press.
- Bytnerowicz, A., M. E. Fenn, E. B. Allen, and R. Cisneros. 2016. "Atmospheric Chemistry." In *Ecosystems of California*, edited by H. Mooney and E. Zavaleta. Berkeley, CA: University of California Press.
- Bytnerowicz, A., M. E. Fenn, R. Cisneros, D. Schweizer, J. Burley, and S. L. Schilling. 2019. "Nitrogenous Air Pollutants and Ozone Exposure in the Central Sierra Nevada and White Mountains of California – Distribution and Evaluation of Ecological Risks." *Science of the Total Environment* 654: 604–615. <https://doi.org/10.1016/j.scitotenv.2018.11.011>.
- Byun, D., and K. L. Schere. 2006. "Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System." *Applied Mechanics Reviews* 59: 51–76. <https://doi.org/10.1115/1.2128636>.
- Clarke, K. R., P. J. Somerfield, P. J. Somerfield, and M. G. Chapman. 2006. "On Resemblance Measures for Ecological Studies, Including Taxonomic Dissimilarities and a Zero-Adjusted Bray-Curtis Coefficient for Denuded

- Assemblages.” *Journal of Experimental Marine Biology and Ecology* 330: 55–80. <https://doi.org/10.1016/j.jembe.2005.12.017>.
- Cleavitt, N. L., J. W. Hinds, R. L. Poirot, L. H. Geiser, A. Dibble, B. Leon, R. Perron, and L. H. Pardo. 2015. “Epiphytic Macrolichen Communities Correspond to Patterns of Sulfur and Nitrogen Deposition in the Northeastern United States.” *Bryologist* 118: 304–324. <https://doi.org/10.1639/0007-2745-118.3.304>.
- Daly, C., G. H. Taylor, W. P. Gibson, T. W. Parzybok, G. L. Johnson, and P. Pasteris. 2000. “High-Quality Spatial Climate Datasets for the United States and beyond.” *Transactions of the American Society of Agricultural and Biological Engineers* 43: 1957–62.
- Elahi, R., C. Birkeland, K. P. Sebens, K. R. Turner, and T. R. Dwyer. 2013. “Limited Change in the Diversity and Structure of Subtidal Communities over Four Decades.” *Marine Biology* 160: 3209–19. <https://doi.org/10.1007/s00227-013-2308-3>.
- Environmental Systems Research Institute. 2018. “ArcMap version 10.5.1.”
- Fenn, M. E., E. B. Allen, S. B. Weiss, S. Jovan, L. H. Geiser, G. S. Tonnesen, R. F. Johnson, et al. 2010. “Nitrogen Critical Loads and Management Alternatives for N-Impacted Ecosystems in California.” *Journal of Environmental Management* 91: 2404–23. <https://doi.org/10.1016/j.jenvman.2010.07.034>.
- Geiser, L. H., H. T. Root, R. J. Smith, S. Jovan, L. St Clair, and K. L. Dillman. 2021. “Lichen-Based Critical Loads for Deposition of Nitrogen and Sulfur in US Forests.” *Environmental Pollution* 291: 118187. <https://doi.org/10.1016/j.envpol.2021.118187>.
- Gower, J. C. 1971. “Statistical Methods of Comparing Different Multivariate Analyses of the Same Data.” In *Mathematics in the Archaeological and Historical Sciences*, edited by F. R. Hodson, D. G. Kendal, and P. Tautu, 138–149. Edinburgh: Edinburgh University Press.
- Jovan, S. 2008. *Lichen Bioindication of Biodiversity, Air Quality, and Climate: Baseline Results from Monitoring in Washington, Oregon, and California*. General Technical Report-PNW-GTR-737. Portland, OR: USDA Forest Service, Pacific Northwest Research Station.
- Jovan, S., M. E. Fenn, M. Buhler, A. Bytnerowicz, A. Kovasi, M. Hutten, E. Di Meglio, and D. Schweizer. 2021. “Challenges Characterizing N Deposition to High Elevation Protected Areas: A Case Study Integrating Instrument, Simulated, and Lichen Inventory Datasets for the Devils Postpile National Monument and Surrounding Region, USA.” *Ecological Indicators* 122: 107311. <https://doi.org/10.1016/j.ecolind.2020.107311>.
- Jovan, S., and B. McCune. 2005. “Air-Quality Bioindication in the Greater Central Valley of California, with Epiphytic Macrolichen Communities.” *Ecological Applications* 15: 1712–26. <https://doi.org/10.1890/03-5368>.
- Jovan, S., and B. McCune. 2006. “Using Epiphytic Macrolichen Communities for Biomonitoring Ammonia in Forests of the Greater Sierra Nevada, California.” *Water, Air, and Soil Pollution* 170: 69–93. <https://doi.org/10.1007/s11270-006-2814-8>.
- Jovan, S., and B. McCune. 2004. “Regional Variation in Epiphytic Macrolichen Communities in Northern and Central California Forests.” *Bryologist* 107: 328–339. [https://doi.org/10.1639/0007-2745\(2004\)107\[0328:RVIEMC\]2.0.CO;2](https://doi.org/10.1639/0007-2745(2004)107[0328:RVIEMC]2.0.CO;2).
- Jovan, S., S. Will-Wolf, L. Geiser, K. Dillman, M. Haldeman, and R. Shory. 2020. “User Guide for the National Forest Inventory and Analysis Lichen Database (Version 1.0).” General Technical Report PNW-GTR-988. USDA Forest Service, Pacific Northwest Research Station. https://www.fs.usda.gov/pnw/pubs/pnw_gtr988.pdf.
- Kovasi, A., B. McCune, and S. Jovan. 2022. “The Potential Use of Elemental Content of Saxicolous Lichens as Bioindicators of Nitrogen Deposition in the Central and Southern California Mountains.” *Ecological Indicators* 144(October): 109541. <https://doi.org/10.1016/j.ecolind.2022.109541>.
- Kovasi, A. I. 2023. “Lichen Biomonitoring from Valley to Alpine in Central and Southern California (Version 2).” Oregon State University ScholarsArchive@OSU. https://doi.org/10.7267/PV63G710J_V2.
- Kruskal, J. 1964. “Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis.” *Psychometrika* 29: 1–27. <https://doi.org/10.1007/BF02289565>.
- McCune, B. 2006. “Nonparametric Habitat Models with Automatic Interactions.” *Journal of Vegetation Science* 17: 819–830.
- McCune, B., J. P. Dey, J. E. Peck, D. Cassell, K. Heiman, S. Will-Wolf, and P. N. Neitlich. 1997. “Repeatability of Community Data: Species Richness Versus Gradient Scores in Large-Scale Lichen Studies.” *Bryologist* 100: 40–46. <https://doi.org/10.2307/3244385>.
- McCune, B., and J. B. Grace. 2002. *Analysis of Ecological Communities*. Corvallis, OR: MjM Software Design.
- McCune, B., and M. J. Mefford. 2009. *HyperNiche Version 2.30*. Corvallis, OR: MjM Software Design.
- McCune, B., and M. J. Mefford. 2018. *PC-ORD Version 7.07*. Corvallis, OR: MjM Software Design.
- McCune, B., and H. T. Root. 2015. “Origin of the Dust Bunny Distribution in Ecological Community Data.” *Plant Ecology* 216: 645–656. <https://doi.org/10.1007/s11258-014-0404-1>.
- McCune, B., S. Yang, S. Jovan, and H. Root. 2022. “Climate and Epiphytic Macrolichen Communities in the Four Corners Region of the USA.” *Bryologist* 125: 70–90. <https://doi.org/10.1639/0007-2745-125.1.070>.
- Minchin, P. R. 1987. “Simulation of Multidimensional Community Patterns: Towards a Comprehensive Model.” *Vegetatio* 71: 145–156. <http://www.jstor.org/stable/20038177>.
- Nash, T. H., III. 2002. “Lichens as Bioindicators of Sulfur Dioxide.” *Symbiosis* 33: 1–21.
- Rangwala, I., and J. R. Miller. 2012. “Climate Change in Mountains: A Review of Elevation-Dependent Warming and Its Possible Causes.” *Climate Change* 114: 527–547. <https://doi.org/10.1007/s10584-012-0419-3>.
- Sales, M., and E. Ballesteros. 2009. “Shallow *Cystoseira* (Fucales: Ochrophyta) Assemblages Thriving in Sheltered Areas from Menorca (NW Mediterranean): Relationships with Environmental Factors and Anthropogenic Pressures.” *Estuarine, Coastal and Shelf Science* 84: 476–482. <https://doi.org/10.1016/j.ecss.2009.07.013>.
- Schnell, S., and C. Mendoza. 2004. “The Condition for Pseudo-First-Order Kinetics in Enzymatic Reactions Is Independent of the Initial Enzyme Concentration.” *Biophysical Chemistry* 107: 165–174. <https://doi.org/10.1016/j.bpc.2003.09.003>.
- Schwede, D. B., and G. G. Lear. 2014. “A Novel Hybrid Approach for Estimating Total Deposition in the United States.”

- Atmospheric Environment* 92: 207–220. <https://doi.org/10.1016/j.atmosenv.2014.04.008>.
- Smith, R. J., S. Jovan, and B. McCune. 2017. “Lichen Communities as Climate Indicators in the U.S. Pacific States. General Technical Report PNW-GTR-952.” Pacific Northwest Research Station, USDA Forest Service. <https://research.fs.usda.gov/treearch/54532>.
- Smith, R. J., S. Jovan, D. Stanton, and S. Will-Wolf. 2020. “Epiphytic Macrolichen Communities Indicate Climate and Air Quality in the U. S. Midwest.” *Bryologist* 123: 517–533. <https://doi.org/10.1639/0007-2745-123.3.516>.
- Sobocinski, K. L., R. J. Orth, M. C. Fabrizio, and R. J. Latour. 2013. “Historical Comparison of Fish Community Structure in Lower Chesapeake Bay Seagrass Habitats.” *Estuaries and Coasts* 36: 775–794. <https://doi.org/10.1007/s12237-013-9586-3>.
- van Herk, K. 1999. “Mapping of Ammonia Pollution with Epiphytic Lichens in The Netherlands.” *The Lichenologist* 31: 9–20. <https://doi.org/10.1006/lich.1998.0138>.
- Weitzman, B., B. Konar, K. Iken, H. Coletti, D. Monson, R. Suryan, T. Dean, D. Hondolero, and M. Lindeberg. 2021. “Changes in Rocky Intertidal Community Structure during a Marine Heatwave in the Northern Gulf of Alaska.” *Frontiers in Marine Science* 8: 1–18. <https://doi.org/10.3389/fmars.2021.556820>.
- Will-Wolf, S., S. Jovan, P. Neitlich, J. E. Peck, and R. Rosentreter. 2015. “Lichen-Based Indices to Quantify Responses to Climate and Air Pollution across Northeastern U.S.A.” *Bryologist* 118: 59–82. <https://doi.org/10.1639/0007-2745-118.1.059>.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kovasi, Adrienne, Bruce McCune, and Sarah Jovan. 2025. “Extracting Gradients from Community Data with Multiple Stressors and Empty Plots: A Case Study and Simulations.” *Ecosphere* 16(10): e70394. <https://doi.org/10.1002/ecs2.70394>