

QUANTIFYING THE ABUNDANCE OF CO-OCCURRING CONIFERS ALONG INLAND NORTHWEST (USA) CLIMATE GRADIENTS

GERALD E. REHFELDT,¹ DENNIS E. FERGUSON, AND NICHOLAS L. CROOKSTON

Rocky Mountain Research Station, USDA Forest Service, Forestry Sciences Laboratory, 1221 S. Main, Moscow, Idaho 83843 USA

Abstract. The occurrence and abundance of conifers along climate gradients in the Inland Northwest (USA) was assessed using data from 5082 field plots, 81% of which were forested. Analyses using the Random Forests classification tree revealed that the sequential distribution of species along an altitudinal gradient could be predicted with reasonable accuracy from a single climate variable, a growing-season dryness index, calculated from the ratio of degree-days $>5^{\circ}\text{C}$ that accumulate in the frost-free season to the summer precipitation. While the appearance and departure of species in an ascending altitudinal sequence were closely related to the dryness index, the departure was most easily visualized in relation to negative degree-days (degree-days $<0^{\circ}\text{C}$). The results were in close agreement with the works of descriptive ecologists. A Weibull response function was used to predict from climate variables the abundance and occurrence probabilities of each species, using binned data. The fit of the models was excellent, generally accounting for $>90\%$ of the variance among 100 classes.

Key words: altitudinal distributions; climatic distributions; conifers; Inland Northwest, USA; Random Forests; realized climate niche; species abundance; Weibull regressions.

INTRODUCTION

Climatic gradients in mountainous systems are pronounced and steep (see Baker 1944, Geiger 1950). Climate, moreover, is generally recognized as the primary factor controlling the distribution of plants (e.g., Woodward 1987). Tree limits indeed are one of the most sensitive paleoecologic indicators of past climates (see Kullman 1995). Zonate patterns of vegetation tend to be common and prominent in the phytosociological descriptions of mountainous ecosystems. It is baffling, therefore, that the complex of environmental factors responsible for zonate distribution of the vegetation as a whole and their component species in particular is largely the subject of conjecture, observation, and intuition (see Brown and Gibson 1983). For the Inland Northwest of the United States, for instance, climate variables purported to control the distribution of plant species have included low temperatures for the upper altitudinal margins and moisture stress at the lower (Larsen 1930), the balance between temperature and precipitation (Haig et al. 1941), and summer droughts (Daubenmire 1956).

To be sure, general climatic gradients along altitudinal transects are superimposed on microclimatic effects associated with topography, aspect, wind, and soils (see Rosenberg 1974). Both, moreover, interact with disturbance history to determine vegetation composition (see Bonan 2002). Nonetheless, despite a potential

complex of interacting environmental factors, altitudinal distributions of species tend to be well defined (Fig. 1; Daubenmire 1952, 1966).

Daubenmire's (1966) conceptual ordination of conifers in the Inland Northwest (Fig. 1) depicts the relative entrance and departure of species in an altitudinal sequence. In his view, the sequence extends along an environmental gradient from warm and dry at low elevation to cold and wet at high elevation. Although ecological amplitudes were represented by a bar, Daubenmire recognized that species distributions tended to be Gaussian. Successive treatments of plant sociology that expand on Daubenmire's work (Pfister et al. 1977, Steele et al. 1981, Cooper et al. 1991) have used similar bar charts, but neither Daubenmire nor his successors have attempted to elucidate and quantify niche space, even though most believed that summer droughts limit distributions at lower elevations and summer heat deficiencies limit distributions at higher elevations (Daubenmire 1956, Daubenmire and Daubenmire 1968, Pfister et al. 1977).

Our goals began as an objective assessment of the climate variables relevant to the distribution of conifers depicted in Fig. 1. Positive results led us toward a quantitative estimate of the realized climatic niche and toward predicting abundance and occurrence probabilities.

The concept of the niche has been a focal point of ecological research since its formulation by Hutchinson (1957). Yet, only with the relatively recent availability of immense data sets and powerful computers and their software has the quantification of niche dimensions become feasible (see Thuiller et al. 2004). Recent attention to the realized niche, for example, has

Manuscript received 7 December 2006; revised 4 December 2007; accepted 5 December 2007. Corresponding Editor: J. A. Jones.

¹ E-mail: jrehfeldt@gmail.com

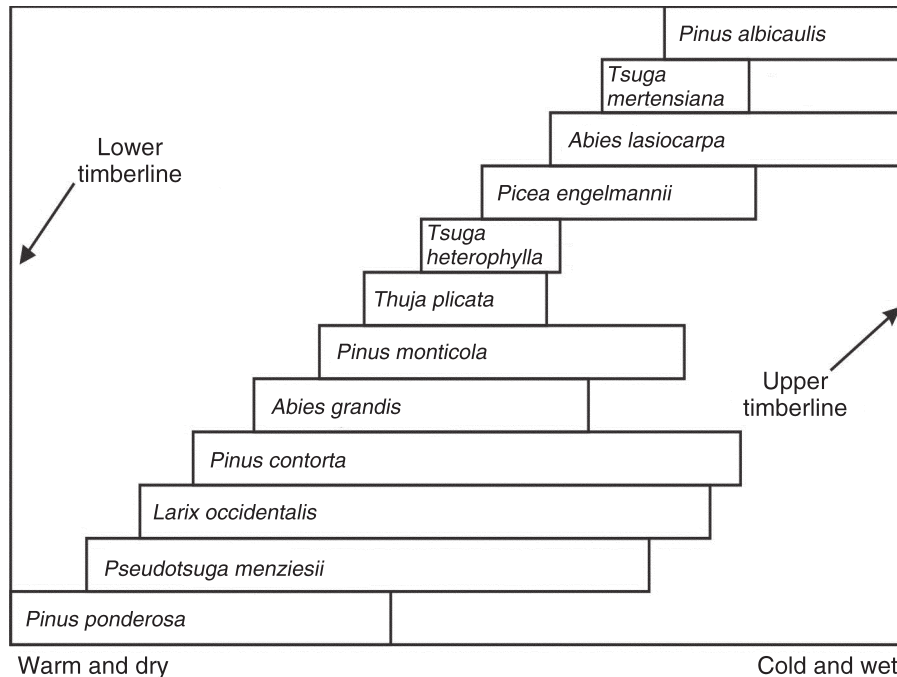


FIG. 1. Daubenmire's conceptual ordering of conifers along an altitudinal gradient in the forests of the Inland Northwest (redrawn to scale from Daubenmire [1966]).

considered organisms ranging from marine vertebrates (Huff et al. 2005) to species of *Eucalyptus* (Austin et al. 1990). Niche dimensions, moreover, have been used to interpret paleoecologic responses to Holocene climates (e.g., Brunelle et al. 2008) as well as projected future responses to global warming (e.g., Rehfeldt et al. 2006). Our approach is at the community level, relating shifts in composition to regional climatic gradients.

METHODS

We use data from permanent sample plots maintained by the Forest Inventory and Analysis, USDA Forest Service. The plots are systematically located to sample the vegetation on forested and non-forested lands throughout the United States (see Alerich et al. 2006, Bechtold and Patterson 2005). We use presence/absence data for 12 species for 4573 plots from the Inland Northwest, the region within which Daubenmire's classification system is applicable (45°30'–49° N and 114°–118° W). Plots are ordinarily established with four subplots, but for our analyses, data from subplots were collated. A screening of the inventory data revealed 210 plots at an elevation <300 m that do not occur within our geographic window. These data points were considered to be compromised and were eliminated. Inventory plots were augmented with the addition of 719 sample plots of Cooper et al. (1991), producing 5082 plots for our analyses (Table 1), 81.2% of which were forested.

The climate surfaces of Rehfeldt (2006) were used to estimate normalized (1961–1990) monthly means of total precipitation and average, maximum, and minimum temperature of each plot. These surfaces, built with the thin plate splines of Hutchinson (1991, 2000), make point predictions of climate from geographic input (latitude, longitude, and elevation). Although the precise geographic location of Forest Inventory plots is proprietary, we were allowed to generate estimates of climate variables directly from their databases.

Our analyses employ 36 variables derived from the monthly estimates (see Rehfeldt 2006, Rehfeldt et al. 2006). These variables include simple expressions of average temperature and precipitation (e.g., mean annual temperature, mean annual precipitation), temperature sums (e.g., degree-days >5°C, degree-days <0°C), freezing dates (e.g., the day of the year of the last freeze of spring), and simple interactions of temperature and precipitation (e.g., the ratio of degree-days >5°C to mean annual precipitation). Eighteen of the climate variables were derived from the monthly estimates, while the remainder involved interactions. Of these variables, note that (1) summer precipitation is defined as the April–September sum, (2) the growing-season dryness index is the ratio of the square root of the degree-days >5°C that accumulate between the last frost of spring and the first frost of autumn to summer precipitation, and (3) the summer–winter temperature differential is the difference between the mean temper-

TABLE 1. Occurrence and co-occurrence of the 12 tree species comprising the forest communities of the Inland Northwest.

Species	Occurrence	Co-occurrence (%)											
		PIPO	PSME	LAOC	PICO	ABGR	PIMO	THPL	TSHE	PIEN	ABLA	TSME	PIAL
PIPO	1046	...	86	31	24	35	5	9	4	9	3	0	<1
PSME	2854	32	...	45	39	53	13	25	14	27	24	2	2
LAOC	1553	21	83	...	52	56	20	36	23	39	31	4	1
PICO	1513	17	73	53	...	42	15	21	16	45	45	4	5
ABGR	1882	19	80	46	34	...	21	42	24	32	23	2	<1
PIMO	559	10	71	56	40	69	...	55	47	40	36	9	2
THPL	1015	9	69	54	31	77	30	...	43	33	24	2	<1
TSHE	612	7	66	59	39	72	43	72	...	38	32	2	1
PIEN	1360	7	57	45	50	45	17	25	17	...	70	8	7
ABLA	1343	2	52	36	51	32	15	18	15	71	...	11	11
TSME	183	0	36	31	32	22	27	10	5	56	84	...	7
PIAL	168	<1	26	7	45	6	5	2	2	55	91	7	...

Notes: Data are summarized from Alerich et al. (2006) and Bechtold and Patterson (2005). Species abbreviations are: PIPO, *Pinus ponderosa*; PSME, *Pseudotsuga menziesii*; LAOC, *Larix occidentalis*; PICO, *Pinus contorta*; ABGR, *Abies grandis*; PIMO, *Pinus monticola*; THPL, *Thuja plicata*; TSHE, *Tsuga heterophylla*; PIEN, *Picea engelmannii*; ABLA, *Abies lasiocarpa*; TSME, *Tsuga mertensiana*; and PIAL, *Pinus albicaulis*.

ature in the warmest month and that of the coldest month.

Identifying relevant climate variables

Inventory data were reassembled such that the occurrence of each species was linked to the climate of the plot within which the species occurred. This produced a data set with 14088 observations (Table 1). On average, the climate of a plot was represented 3.4 times within this data set, with each repetition coding the occurrence of a different species.

The Random Forests classification tree of Breiman (2001) was used to predict the species best suited to the climate variables describing an observation. This algorithm, available in R (Liaw and Wiener 2002, R Development Core Team 2004), builds a set of independent trees, called a forest. The process begins with the drawing of a bootstrap sample consisting of approximately two-thirds of the total number of observations. This sample is used to build a tree, while those omitted are used to compute classification errors. For each node of a tree, a random sample of the predictor variables is selected. The number ordinarily equals the square root of the number of predictors. Of these, the variable that minimizes the classification error is identified. Nodes are further split until no more improvement can be achieved. Predictions are made by running an observation down all trees in all forests. A "vote" concerning the classification of that observation is then available from each tree; prediction ordinarily is based on the voting majority. Because errors of prediction converge to a limit as the number of trees in the forest becomes large, overfitting is avoided.

For classification trees, the algorithm in R produces two measures for evaluating the relative importance of independent variables: the mean decrease in accuracy and the mean decrease in the Gini index of class purity (see Breiman and Cutler 2004). The first relies on an iterative process of randomly permuting (noising up) a

predictor variable to determine the effect of a variable on the classification error. The second, also known in ecological research as the Gini-Simpson index (see Pranab 2005), expresses the reduction in node purity attributable to a variable when it is used to split a node. Although correlative models ordinarily are not desirable for assessing mechanistic interactions, these indices are useful in assorting a large number of predictor variables or in culling superfluous variables (Breiman 2001).

Classification errors are ordinarily quantified as the errors of omission and errors of commission that can be computed by comparing predictions with observations in the sample not used to build the trees. For us, however, the co-occurrence of species in the same plot would inflate classification errors unrealistically; the algorithm would accept as being correct only one prediction when two or more predictions could be correct, depending on the number of a species in a plot. To obtain a more realistic estimate of the errors of prediction, we summarized the votes cast by the trees. Errors of omission were arbitrarily defined as occurring if a species received <12% of the votes cast within a plot when that species was present, and the errors of commission were defined as occurring if a species received >2% of the votes when absent. These levels were chosen so that the errors of omission and commission were contributing equally to the combined error.

Breiman (2001) notes that classification errors are reduced when the number of observations within groups is approximately equal. To prevent inequalities from becoming excessive, we arbitrarily limited the number of observations within species to 1500. This was done by discarding a random sample of 48% of the observations for *Pseudotsuga menziesii*, 11% for *Abies grandis*, 5% for *Larix occidentalis*, and 1% for *Pinus contorta*. All but four species, therefore, were represented by 1000–1500 observations (Table 1). Two samples were obtained,

each containing ~12 200 observations. Each sample was used to develop a forest of 100 trees.

Analyses began by using a full complement of the 36 predictors for building the trees in each forest. Importance values calculated from either the Gini index or the mean decrease in accuracy then were used to remove the least important variables in a stepwise iterative process designed to identify the most important variables. Although separate analyses were made using the alternative measures of variable importance, models using the Gini index to select relevant variables consistently provided the least error; those using the mean decrease in accuracy are not presented.

Quantifying niche space

Data points were ordered with regard to the relevant climate variables. The ordinations were compared visually to Daubenmire's conceptual view, which was then recast quantitatively. In recasting, limits of distribution were equated to the mean of the 10 most distal data points. Although Euclidean distances are used frequently to eliminate outlying data points (e.g., Lawesson 2000), outliers in our distributions should be in error only in degree, largely because plots with fallacious locations had been eliminated previously. Using the mean of the distal 10 thus required discarding no additional data but nonetheless reduced the impact of the outliers.

Predicting abundance

Inventory data were arranged into 100 classes of equal breadth for each of the relevant climate variables. A class was added to the beginning and end of the sequence within which presence of all species was assigned a value of zero. Because distributions tended to be skewed, a Weibull response function was used to predict for each species and for all data (forested and nonforested plots) the number of plots in each class. The Weibull function was originally conceived as a probability density function suited to normal or skewed normal distributions (Johnson et al. 1994). We used the formulation of Fang and Xu (1987) (see Rehfeldt et al. 2003). In this formulation, response, R , is the number of plots in which a species occurs in class X :

$$R = c(X' - \delta)^{(\alpha-1)} \exp\left[-\frac{(X' - \delta)^\alpha}{\beta}\right], \quad (1)$$

where $X' = (X - X_{\min})(X_{\max} - X_{\min})^{-1}$; c is the performance capacity, >0 ; α is a shape parameter, >1 ; β is a scale parameter, >0 ; and δ is a position parameter, $<X'$;

$$\log R = \log c + (\alpha - 1)\log(X' - \delta) - \frac{(X' - \delta)^\alpha}{\beta}. \quad (2)$$

It follows that

$$R' = c' + a_1 \log(X' - \delta) + a_2 (X' - \delta)^{(a_1+1)}. \quad (3)$$

Let

$$X'_1 = \log(X' - \delta)$$

and

$$X'_2 = (X' - \delta).$$

Then,

$$R' = c' + a_1 X'_1 + a_2 X'_2^{(a_1+1)} \quad (4)$$

where $c' = \log(c)$; $a_1 = \alpha - 1$, which is >-1 ; and $a_2 = -(1/\beta)$, which is <0 .

Regressions were performed iteratively using the NLIN algorithm of SAS (2003) to find a value of $(X' - \delta)$ that minimized the residual mean square. In this analysis, X is scaled to values between zero and one and, therefore, could be represented by the class number or the median value of the climate variable for the class.

Relative abundance of species was assessed from ratios of the area under the species-specific curves to that for the curve describing the climatic distribution of all forested and nonforested plots within the geographic window.

Occurrence probabilities

The probability that a species occurs within any of the 102 classes defined above is estimated by the ratio of the number of plots in which a species occurs within a class to the total number of forested and nonforested plots for that class. The Fung-Xu formulation of the Weibull function was used to model these probabilities using the relevant climate variables as predictors. We predict, therefore, conditional probabilities, the probability of a species occurring given the occurrence of a specific climate.

RESULTS

Relevant climate variables

Classification errors for the two forests were essentially the same. As defined by us, the errors averaged 11.4% for the 36-variable model and remained essentially unchanged through the stepwise elimination process until six variables remained. The six-variable model produced an error of 11.7%, which rose to 12.0% with the three-variable model, 12.6% for the two-variable model, and 13.3% for the one-variable model. A two- or three-variable model seemingly would be the most parsimonious. Species tended to contribute equally to the errors of commission, but ~56% of the errors of omission were traceable to the most common species, *Pseudotsuga menziesii* (Table 1).

Multiple representation of plots within the data set made difficult an assessment of the statistical fit of the classification tree. To be sure, using $<12\%$ of the votes to define errors of omission and $>2\%$ to define errors of commission was arbitrary. Yet, when the independent variables of the three-variable model were used to predict presence or absence of individual species (see

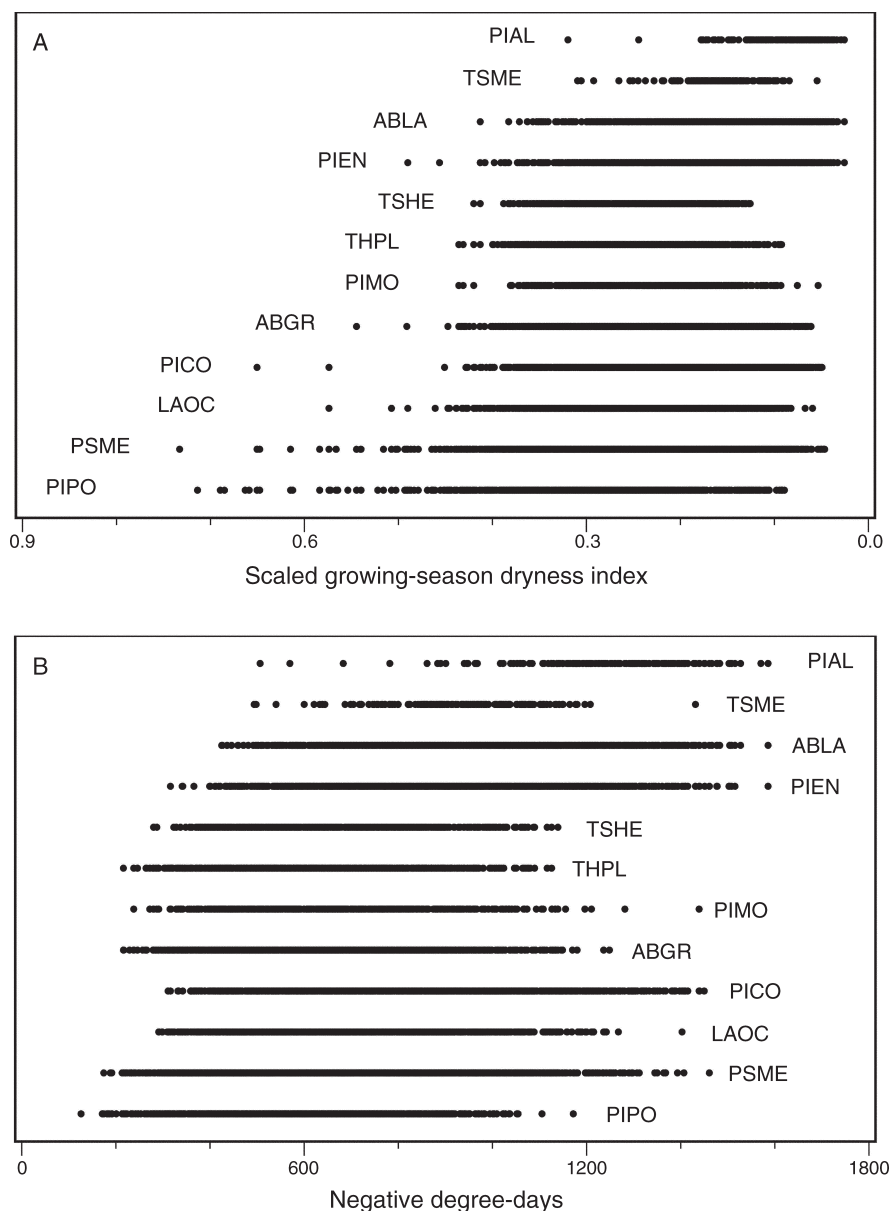


FIG. 2. Distribution of observations in an ordination of conifer species (arranged according to Fig. 1) along (A) gradients in the growing-season dryness index (the ratio of the square root of degree-days $>5^{\circ}\text{C}$ accumulating in the frost-free period to summer precipitation) scaled to values between 0 and 1 and (B) negative degree-days (degree-days $<0^{\circ}\text{C}$). See Table 1 for an explanation of species abbreviations.

Rehfeldt et al. 2006), out-of-bag errors ranged from 10% (*A. lasiocarpa*) to 28% (*P. menziesii*). It would seem, therefore, that the limits chosen by us to define classification errors are reasonable and that our model is statistically sound.

According to the Gini index, the growing-season dryness index was the most important climate variable to both forests. The product of mean annual precipitation with (1) degree-days $>5^{\circ}\text{C}$ and (2) the summer–winter temperature differential each were identified by one of the forests as the variables second in importance

to the dryness index. These two variables, however, not only accounted for little classification error in addition to that accounted by the dryness index but also contributed little toward quantifying and interpreting Fig. 1; they will not be considered further.

Realized climate niche

An ordination of observations with respect to the growing-season dryness index (Fig. 2A) shows that the index effectively reproduces the order in which species appear and depart the altitudinal sequence of Fig. 1,

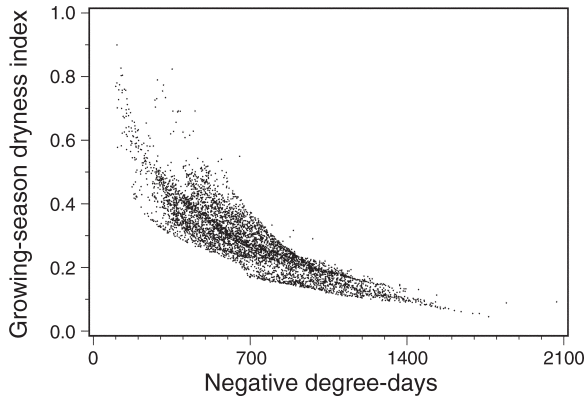


FIG. 3. Scatter of data points relating the scaled growing-season dryness index (the ratio of the square root of degree-days $>5^{\circ}\text{C}$ accumulating in the frost-free period to summer precipitation) and negative degree-days (degree-days $<0^{\circ}\text{C}$).

particularly when outlying data are ignored. Outliers are disjunct data points in an otherwise continuous distribution. Because we eliminated obvious location errors at the outset, outliers in our data set most likely are due to the occurrence of a species in favorable microsites in an otherwise inhospitable climate, misidentification of species, or errors in estimating climate. Note that the dryness index presented in Fig. 2 and subsequent figures has been scaled to values between zero and one, the values used for the Weibull regressions.

Correspondence between Figs. 1 and 2 is most obvious for the entrance of species into the ascending

sequence. Although the correspondence is equally strong for the departure, the differences among species are small at low values of the index. As a result, pronounced ecological effects are concentrated between values of 0 and 0.1. A scanning of the data revealed that the exit of species from the sequence was more easily visualized in relation to negative degree-days (Fig. 2B). This variable is closely related to the dryness index in our study area (Fig. 3), particularly for cold climates. In fact, the relationship is so strong for cold climates that one could serve as a surrogate for the other.

Daubenmire's ordination (Fig. 1) was recast quantitatively by (1) estimating the limits of distribution from the mean of the 10 most distal data points to reduce the impact of outliers and (2) overlaying Fig. 2B on Fig. 2A (Fig. 4). The point of overlay is at the climate values associated with the exit of the first species, *Pinus ponderosa*, from the sequence. This point occurs at a scaled growing-season dryness index of 0.19 and at negative degree-days equaling 1056. The result is an ordination (Fig. 4) that is much the same as that of Daubenmire (Fig. 1).

Abundance

Abundance, expressed as the number of plots within a class, was modeled using the growing-season dryness index and negative degree-days as predictors. An approximate *F* test using the ratios of mean squares (SAS 2003) suggested that all Weibull regression models were statistically significant ($P < 0.0001$). The models in

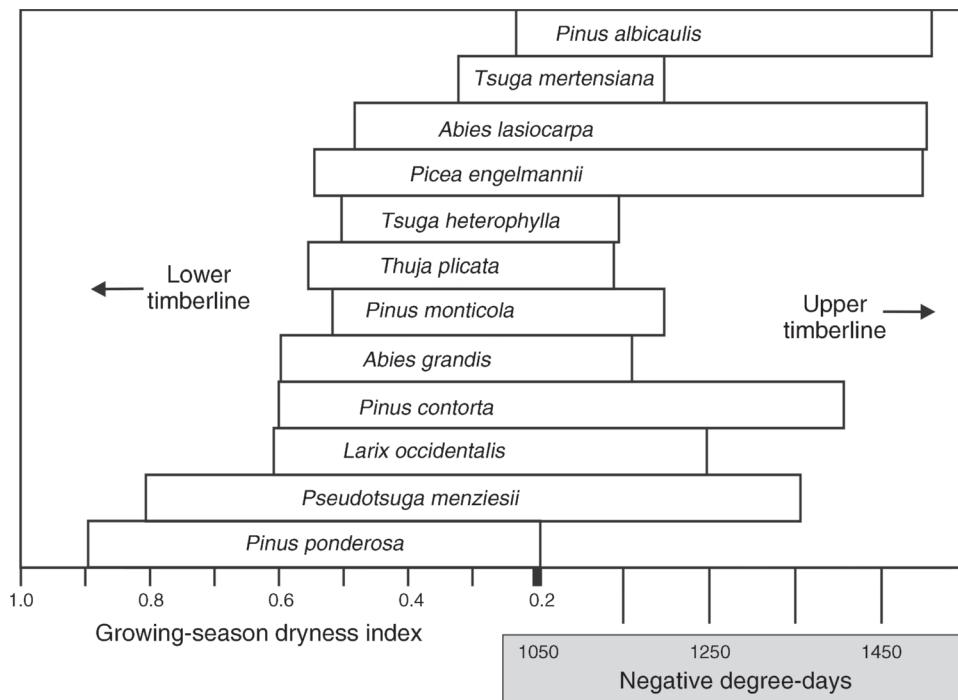


FIG. 4. Ordination of conifer species (arranged as in Fig. 1) using Fig. 2A to determine the entrance of a species into the sequence and Fig. 2B to determine the departure.

TABLE 2. Proportion of variance explained by Weibull regressions of abundance (number of plots) and occurrence probabilities (proportion of plots to the total number of plots) on the growing-season dryness index and negative degree-days.

Species	Models predicting abundance		Models predicting occurrence probabilities	
	Dryness index	Negative degree-days	Dryness index	Negative degree-days
<i>Pinus ponderosa</i>	0.94	0.70	0.96	0.94
<i>Pseudotsuga menziesii</i>	0.98	0.90	0.98	0.98
<i>Larix occidentalis</i>	0.98	0.97	0.96	0.95
<i>Pinus contorta</i>	0.95	0.91	0.96	0.94
<i>Abies grandis</i>	0.95	0.96	0.98	0.97
<i>Pinus monticola</i>	0.91	0.91	0.92	0.85
<i>Thuja plicata</i>	0.96	0.95	0.95	0.95
<i>Tsuga heterophylla</i>	0.96	0.97	0.92	0.90
<i>Picea engelmannii</i>	0.96	0.97	0.93	0.89
<i>Abies lasiocarpa</i>	0.98	0.96	0.94	0.93
<i>Tsuga mertensiana</i>	0.93	0.82	0.79	0.70
<i>Pinus albicaulis</i>	0.89	0.81	0.86	0.81

general accounted for >90% of the variance in the dependent variable, although those for the two least common species, *Tsuga mertensiana* and *P. albicaulis*, tended to be somewhat less (Table 2). Transforming the growing-season dryness index by taking the cube root improved the fit of the resulting models; no transformation was necessary when negative degree-days was the predictor.

The fit of models using the dryness index as a predictor was generally outstanding (Fig. 5, Table 2), as was that for predicting the total number of forested and non-forested plots (Fig. 6). For the latter, the regression model accounted for 98% of the variance. Nonetheless, close inspection of these figures reveals that this function tended to slightly overestimate data points near either the upper or lower limits of the climatic niche where the curve rises from the x-axis. Although not contributing greatly to the residual mean square, this deficiency is apparent in Fig. 5 for *P. contorta*, *L. occidentalis*, *P. monticola*, *T. heterophylla*, *Picea engelmannii*, and *A. lasiocarpa*.

When the area under the curves in Fig. 5 is expressed as a proportion to the total area of forested and non-forested land (Fig. 6), the relative abundance of these species in the region as a whole becomes apparent (Table 3). These proportions suggest, for instance, that *Pseudotsuga menziesii* occurs on more than half of the land base of the region. Because forests occupy ~81% of the land base, one could expect, therefore, for *P. menziesii* to be present on nearly 70% of the forested lands; *T. mertensiana* and *Pinus albicaulis*, by contrast, would occur on only 3.6%.

Models predicting abundance from negative degree-days also were statistically robust (Table 2). The modeled effects (Fig. 7) illustrate the preference of seven of the species for moderately warm winters; the predominant occurrence of *P. contorta*, *T. mertensiana*, *A. lasiocarpa*, and *Picea engelmannii* in the subalpine

climates with cool winters; and the restriction of *Pinus albicaulis* to climates with the coldest winters.

Occurrence probabilities

Regressions predicting occurrence probabilities from the growing-season dryness index and negative degree-days also were of excellent fit (Table 2). Probabilities predicted from the dryness index (Fig. 8) were as high as 0.8 for *Pseudotsuga menziesii* and *A. lasiocarpa* but reached a maximum of only 0.2 for *T. mertensiana*. These probabilities are consistent, with the abundance of the respective species (Table 1).

The probability functions, moreover, provide for an understanding of plant occurrences that are not available from the abundance alone. The predictions of abundance (Fig. 5) show, for instance, that the amount of *Pinus ponderosa* that occurs at a dryness index >0.5 is negligible. Yet Fig. 8A indicates that for sites with index values between 0.5 and 0.8, there is a reasonable probability of *P. ponderosa* occurring. The probabilities, however, are conditional; sites with a scaled dryness index of 0.6, for instance, are infrequent (Fig. 6), but given their occurrence, the probability of *P. ponderosa* being present would be about 0.2. Quite similarly, the occurrence of *P. albicaulis* is limited to a small range of index values and is never expected to be abundant (Fig. 5). Yet, at a scaled value of the dryness index of ~0.09, the relatively high probability of the species occurring (0.6, Fig. 8C) seems incongruous with its lack of abundance. Also noteworthy in Fig. 8B are the occurrence probabilities of *P. monticola*, *Thuja plicata*, and *Tsuga heterophylla*, which peak at probabilities <0.4; these species, therefore, are expected to be absent more than one-half of the time even when the climate is ideal.

Fig. 8 demonstrates that seemingly small changes in the dryness index can result in large changes in composition. At a scaled dryness index of 0.2, for instance, the probability of *Pseudotsuga menziesii*

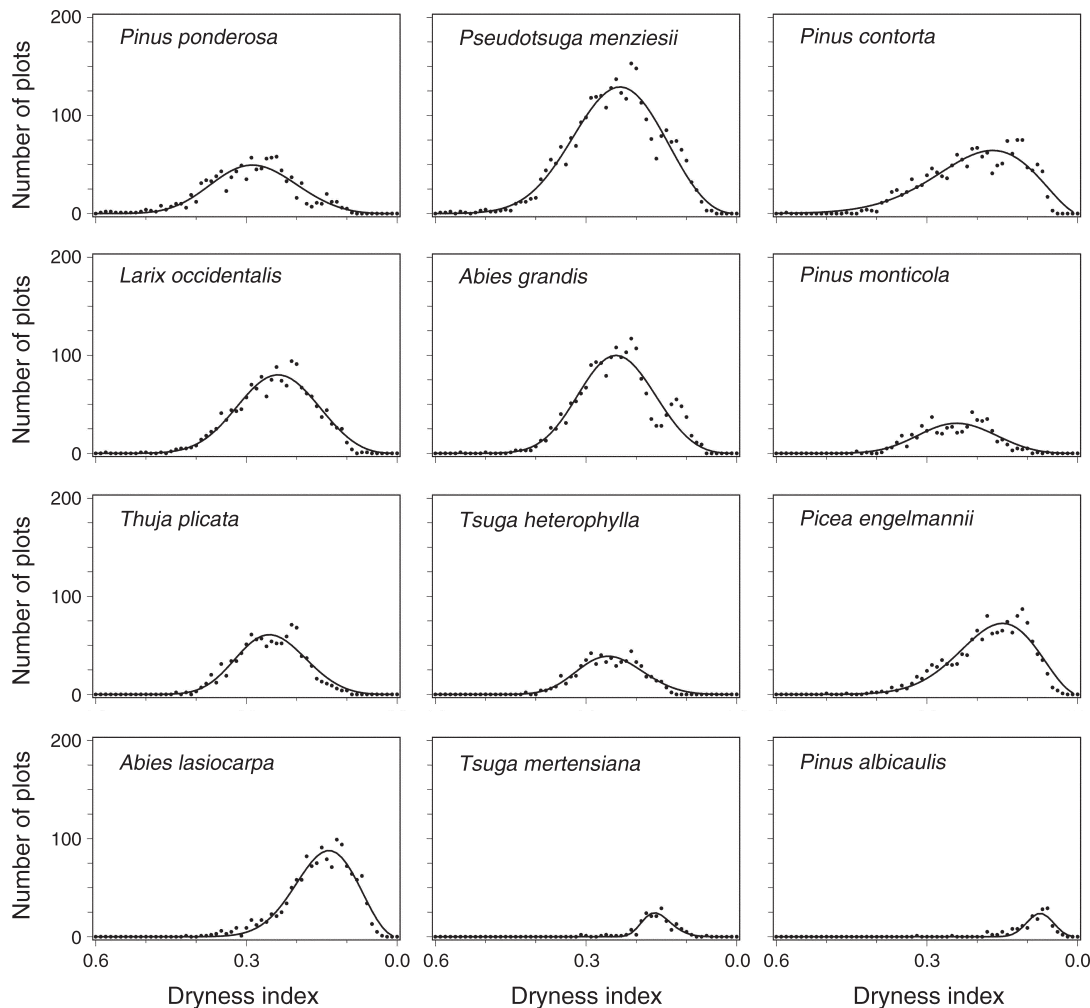


FIG. 5. Plots of observed data points (dots) against those predicted (lines) from species-specific regressions using a Weibull response function of the number of plots within each of 102 discrete classes on scaled values of the growing-season dryness index (the ratio of the square root of degree-days $>5^{\circ}\text{C}$ accumulating in the frost-free period to summer precipitation).

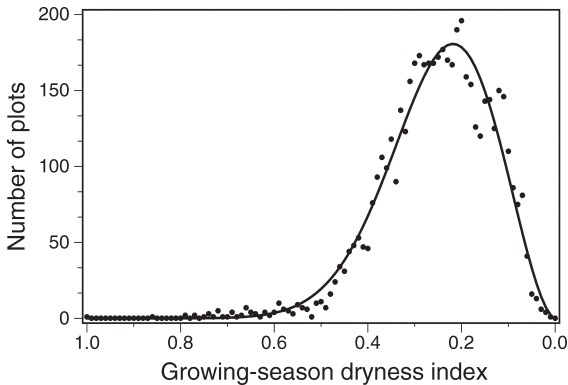


FIG. 6. Plots of observed data points (dots) against those predicted (lines) from a Weibull response function of the total number of forested and non-forested plots within each of 102 discrete classes on the scaled growing-season dryness index (the ratio of the square root of degree-days $>5^{\circ}\text{C}$ accumulating in the frost-free period to summer precipitation).

TABLE 3. Ratio of the area under species-specific curves (Fig. 5) to that for all forested and non-forested lands (Fig. 6).

Species	Proportion of total area
<i>Pinus ponderosa</i>	0.20
<i>Pseudotsuga menziesii</i>	0.56
<i>Larix occidentalis</i>	0.31
<i>Pinus contorta</i>	0.31
<i>Abies grandis</i>	0.36
<i>Pinus monticola</i>	0.11
<i>Thuja plicata</i>	0.20
<i>Tsuga heterophylla</i>	0.12
<i>Picea engelmannii</i>	0.27
<i>Abies lasiocarpa</i>	0.26
<i>Tsuga mertensiana</i>	0.03
<i>Pinus albicaulis</i>	0.03

occurring reaches a peak value near 0.7. Yet at only 0.1 units lower on this index, *A. lasiocarpa*, a species of the subalpine climates (Fig. 7), would be expected to occur with a similarly high probability. Also, at a scaled index value of 0.25, all species except *Pinus albicaulis* could be present at probabilities >0.2, with six, in fact, at their peak probabilities. All species common to the subalpine forests, moreover, achieve highest occurrence probabilities near values of 0.1. The extensive ecological responses that occur at low values of the dryness index scalar point again to the value of using negative degree-days for interpreting interspecific interactions in cold climates.

Occurrence probabilities for negative degree-days (Fig. 9) illustrate the affinity for mild winters for many species that was apparent in Fig. 7. Fig. 9 shows also that *Pseudotsuga menziesii* would be the most likely species occurring where winters are mild, while *A. lasiocarpa* would be the most probable where winters are cold. Yet all species except *Pinus albicaulis* and *T. mertensiana* can occur with probabilities >0.15 where winters average ~700 degree-days <0°C. However, as shown in both Figs. 8 and 9, *P. albicaulis* reaches its highest probability of occurrence at the coldest sites. Note also, as illustrated in Fig. 9, that predicted probabilities for *Picea engelmannii*, *A. lasiocarpa*, *Pinus contorta*, and *P. albicaulis* fail to reach the x-axis in cold environments. This failure is consistent with the general distributions of these species, which exist elsewhere in climates colder than the coldest of the Inland Northwest.

DISCUSSION

The Random Forests multiple regression tree identified one climate variable of overriding importance for segregating 12 conifers along the altitudinal gradients in Inland Northwest forests. This variable, here called “the growing-season dryness index,” is the ratio of the growing degree-days that accumulate during the frost-free season to summer (April–September) precipitation. High values denote arid climates, and low values reflect abundant moisture, which, for our region, tends to be associated with cold climates. With this definition of a dryness index, however, the upper altitudinal distributions of all species tended to occur at index values near zero. Consequently, interspecific relationships for the cold climates were visualized most easily along gradients in negative degree-days. For the middle and upper altitudes, the dryness index and negative degree-days are nearly synonymous.

Quantification of phytosociological relationships

Our results illustrate the same ordering of species along altitudinal gradients that were described conceptually long ago (Daubenmire 1966). We show that the entrance of a species into the sequence is closely related to the dryness index, while the departure correlates with negative degree-days. Our results not only support the phytosociological descriptions of Daubenmire but also

support the intuitive views of many descriptive ecologists who believed that the lower altitudinal limits of species were determined by moisture stress, while the upper limits were controlled by either winter cold or lack of growing-season warmth, the two of which are difficult to separate (Fig. 3).

Although generally minor, most of the discrepancies between Daubenmire’s conceptual view and our quantitative ordination (Fig. 1 vs. Fig. 4) are resolved in studies that were extensions of Daubenmire’s work. Relative to Fig. 1, for instance, our ordination (Fig. 4) shows wider ecological amplitudes for *P. ponderosa*, *Picea engelmannii*, and *A. lasiocarpa*. Ours also disagrees with the upper altitudinal limit of *L. occidentalis* relative to that of *Pseudotsuga menziesii*. Studies conducted within or peripheral to Daubenmire’s core region, by Pfister et al. (1977) for western Montana, Steele et al. (1981) for central Idaho, Cooper et al. (1991) for northern Idaho, and Lillybridge et al. (1995) for northeastern Washington, support our relationships in Fig. 4.

Unresolved, however, would be conflicts regarding the entrance of *Pinus monticola* into the altitudinal sequence relative to that of *Thuja plicata*. Descriptive ecologists (e.g., Pfister et al. 1977, Cooper et al. 1991, Williams et al. 1995) unanimously agree with Daubenmire (Fig. 1). It is likely, however, that our results in particular have been compromised by mortality to *P. monticola* from white pine blister rust (*Cronartium ribicola*), which has reduced the western white pine cover type to 10% of its former range on millions of hectares (O’Laughlin et al. 1993), most of which was within the area studied by Daubenmire.

In quantifying Daubenmire’s ordination, relationships among species became prominent that are not apparent in Fig. 1. Daubenmire presented the entrance of species into the sequence as a regular stair step that obviously was not meant for quantitative interpretation. Fig. 4, however, suggests fine-scaled differences among many species for their lower altitudinal thresholds. Eight species enter the sequence at scaled values of the dryness index between 0.6 and 0.5 (Fig. 4).

However, in considering dissonance between Figs. 1 and 4, it is pertinent to note that one of the goals of the descriptive ecologists was to depict “vegetation in essentially pristine and stable condition” on upland sites that were homogeneous (see Daubenmire 1966). The data we used, however, were from inventory plots that were systematically distributed and contained, therefore, seral as well as successional advanced vegetation. The plots, moreover, occurred not only on homogeneous upland sites but also in ecotones and riparian zones not sampled by the descriptive ecologists. Our results, moreover, stem from summaries of four geographically proximal but nonetheless discrete subplots, collated because the climate surfaces of Rehfeldt (2006) do not address fine-scale topographic effects, particularly aspect. The correspondence between our ordination and

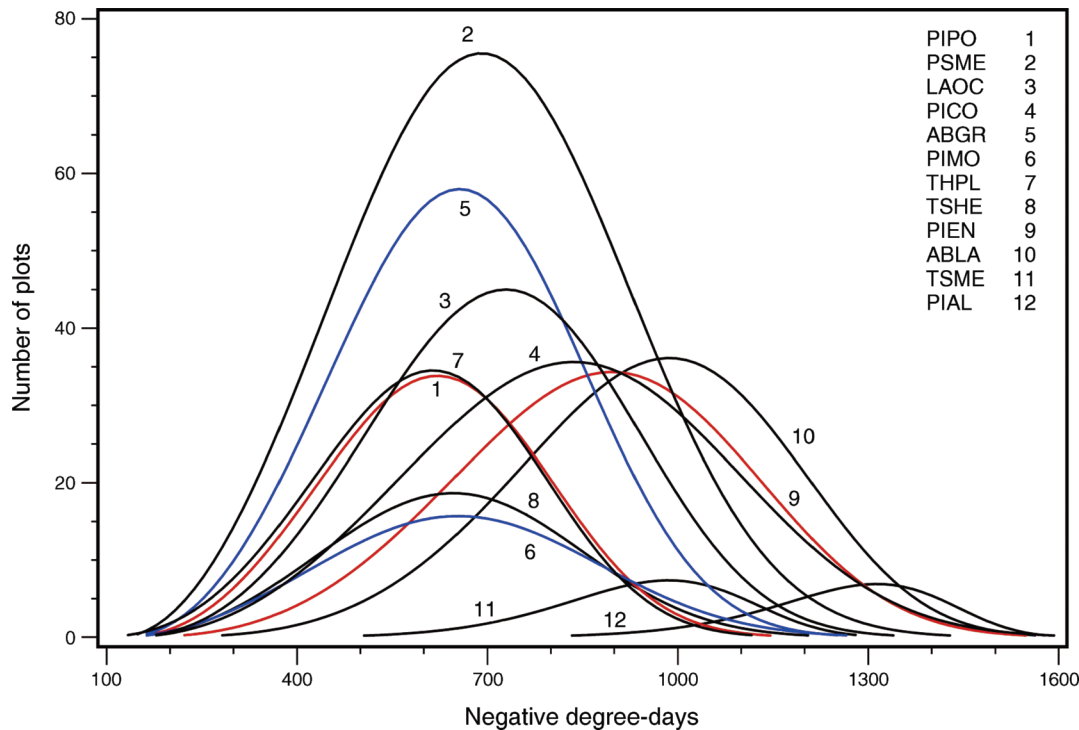


FIG. 7. Predicted abundance of 12 species from species-specific Weibull response functions of the number of plots within each of 102 discrete classes on degree-days $<0^{\circ}\text{C}$. See Table 1 for an explanation of species abbreviations.

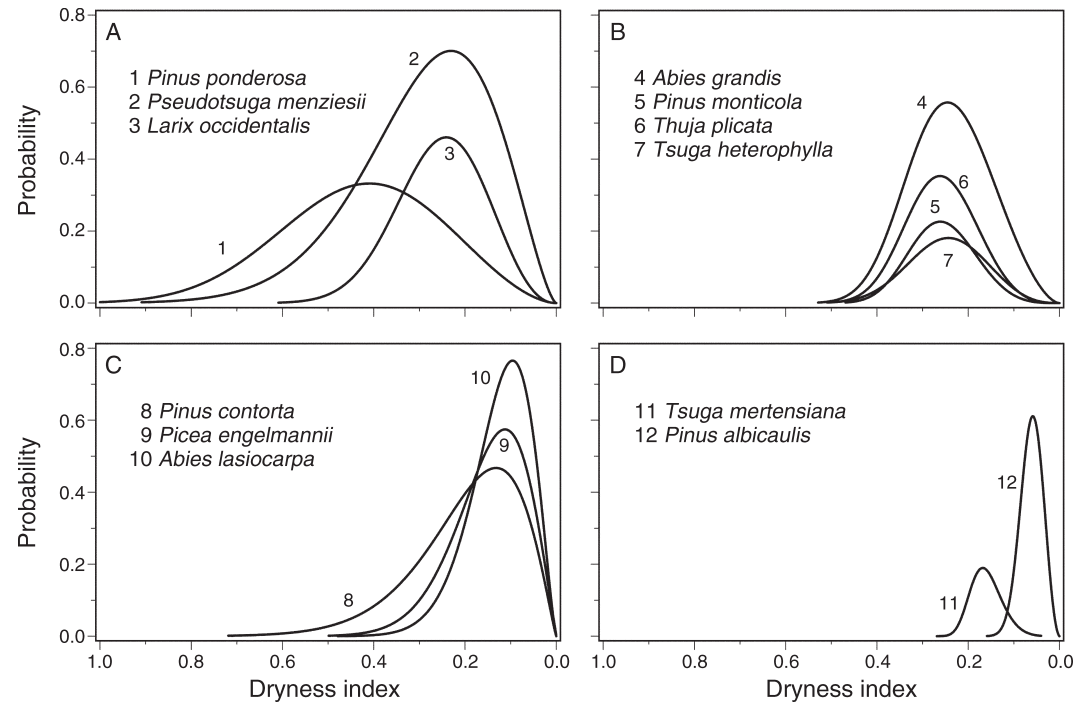


FIG. 8. Predicted occurrence probability of 12 species from Weibull models driven by the growing-season dryness index scalar (the ratio of the square root of degree-days $>5^{\circ}\text{C}$ accumulating in the frost-free period to summer precipitation).

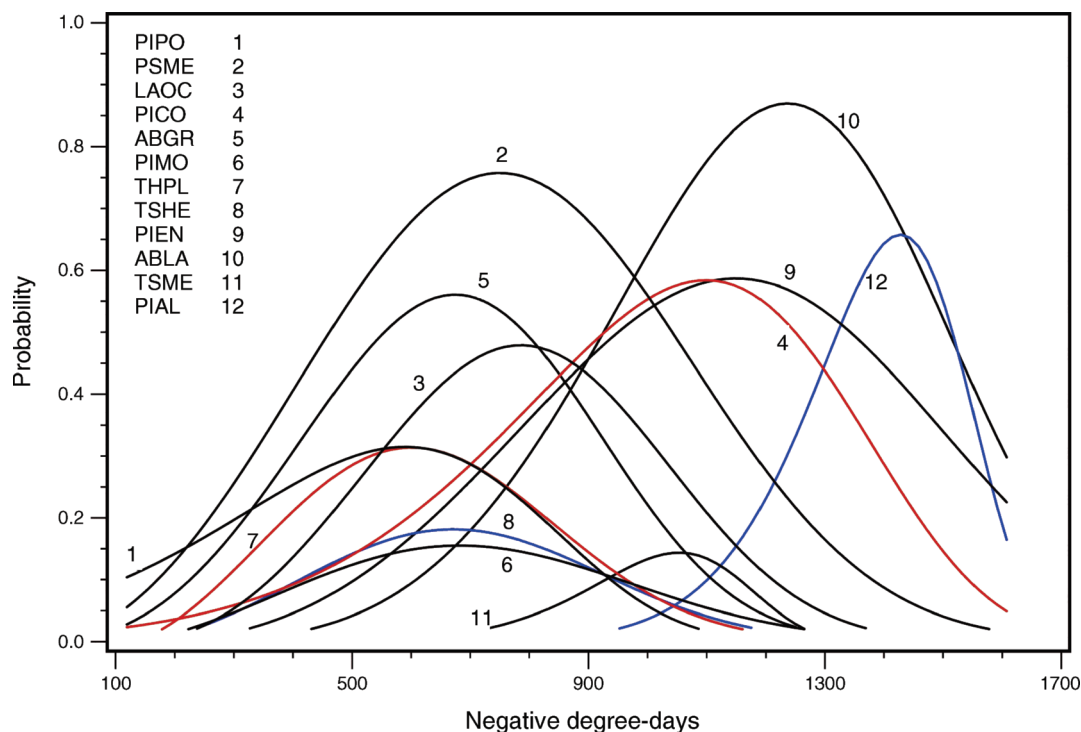


FIG. 9. Predicted occurrence probabilities of 12 species from Weibull models driven by degree-days $<0^{\circ}\text{C}$. See Table 1 for an explanation of species abbreviations.

that of the descriptive ecologists despite these potential sources of dissonance provides further evidence of the close relationship between the altitudinal distribution of species and climate.

Abundance and occurrence

We used a formulation of the Weibull function that provided an excellent fit (Table 2, Fig. 5) to models predicting the abundance and occurrence probabilities from distributions that tended to be skewed. Of the 12 species, the distributions of seven were skewed toward high dryness indices, although for only two, *P. contorta* and *Picea engelmannii*, was the skew pronounced. Occurrence of the remaining species was skewed slightly toward cold climates. Although transformations of the dryness index reduced the skew, the distributions support Austin et al. (1990), who concluded that realized niche space is rarely Gaussian.

The models show that occurrence probabilities can approach 0.7 (*Pseudotsuga menziesii* and *A. lasiocarpa*; Figs. 8 and 9), but for most species tend to be less than 0.5. This suggests that even when the climate is optimal, the probability tends to be high that a species will not be present. Among the factors other than climate tending to influence occurrence of species, disturbance history seems particularly important in the Inland Northwest, where, for instance, the Great Fire of 1910 devastated $\sim 12,000 \text{ km}^2$ (see Pyne 1982). Another factor related to the disturbance regime would be the successional status

of the vegetation. *Pseudotsuga menziesii*, for instance, occurs on $\sim 56\%$ of the area within our geographic window (Table 3), while *A. grandis* and *L. occidentalis* occur on 35% and 31%, respectively. All three are seral to the successional advanced *T. plicata* and *Tsuga heterophylla* (Daubenmire 1966, Daubenmire and Daubenmire 1968) and therefore would be rare in successional advanced forests. Quite similarly, the low occurrence probabilities (Figs. 8 and 9) and relative scarcity (Table 3) of *Pinus monticola* would reflect the epidemic of blister rust that apparently will have a similar devastating effect on *P. albicaulis* (Campbell and Antos 2000).

The occurrence probabilities (Figs. 8 and 9) illustrate the co-occurrence of species that are not readily apparent in the bar-chart ordinations of Figs. 1 and 4. Fig. 1 and Table 1, for instance, suggest that *P. ponderosa* and *T. mertensiana* should not co-occur; yet Figs. 2, 4, 5, and 7 suggest that the realized climatic niche of these species overlap. Fig. 8, however, shows that at a dryness index of 0.2, a climate in which co-occurrence should be the highest, the probability of *P. ponderosa* occurring is ~ 0.15 , while that for *T. mertensiana* is ~ 0.1 . Differences in successional position notwithstanding (see Daubenmire 1966), co-occurrence at this dryness index should be $\sim 3\%$, a rate much more compatible with actual occurrences (Table 1) than implied by Fig. 2. Likewise, the bar charts suggest that *P. albicaulis* and *T. mertensiana* should co-occur $\sim 50\%$

of the time, but Figs. 8 and 9 imply that probability of co-occurrence is low. Table 1 indeed shows that co-occurrence of these species is only ~7%.

*Contributions toward an understanding
of the realized niche*

Our results imply that the segregation of species into the realized niches characteristic of Inland Northwest forests may be under relatively simple climatic control: balance between growing-season temperatures and summer precipitation and, perhaps, the severity of winters, particularly in the high elevations. For this region, separation of the effects of a summer heat deficiency from the coldness of winters is not possible with correlative models. The results, however, apply to a region that encompasses only a portion of the geographic distributions of these 12 conifers. In regions where the array of competitors differs or where the balance between growing-season precipitation and annual precipitation differs, the realized niche of any one of these species could be defined by a different assortment of variables.

When distributions from throughout the western United States are considered, for instance, two variables, the ratio of growing degree-days to annual precipitation and negative degree-days, are required for a climatic separation of *P. contorta* and *P. albicaulis* (Brunelle et al. 2008). Bioclimatic models, moreover, that deal with much of the geographic distributions of four of the species considered herein, *P. ponderosa*, *Pseudotsuga menziesii*, *Picea engelmannii*, and *L. occidentalis*, require five to seven climate variables (Rehfeldt et al. 2006) to describe the realized climate niche, and for none of these species did the array of climate variables include growing-season dryness index. To be sure, in these studies as well as in the present, variables of importance in predicting contemporary distributions always included a ratio involving degree-days >5°C and precipitation and frequently included a variable describing the coldness of winters. Nevertheless, the determinants and dimensions of the realized niche will depend on the size of the geographic distribution being considered. When considering niche space for a regional flora, the results indeed may lead to simple interpretations, but niche space across broad geographic/climatic tracts may be much more complex.

Conclusions

Our results quantify relationships among species for the forests of the Inland Northwest that long have been recognized by descriptive ecologists: (1) the tolerance of *Pinus ponderosa* to the highest moisture stresses; (2) the broad ecological amplitude of *Pseudotsuga menziesii*, whose apparent indifference to winter cold and moisture stress combine to make it the most abundant conifer in the region's forests; (3) the similarities in climatic tolerances of *Pinus monticola*, *L. occidentalis*, *A. grandis*, *Thuja plicata*, and *Tsuga heterophylla*, all of which are

components of the plant associations that recur with mild winters and low potentials for moisture stresses; and (4) the occurrence of species of the subalpine associations (*P. contorta*, *Picea engelmannii*, *A. lasiocarpa*, and *Pinus albicaulis*) in climates of short growing season, cold winters, and adequate precipitation, where moisture stresses rarely develop.

A Weibull response function provided a remarkably good fit to models predicting the abundance of species and their occurrence probabilities from climate. Because the function was fit to general climate gradients across a heterogeneous region, modeled effects reflect a biogeographic perspective. The models, therefore, may seem unrealistic to physiologists seeking fine-scale predictions of vegetation on discrete landscapes. Consequently, using our approach in landscape ecology would require topographic subroutines (e.g., Hungerford et al. 1989) for the spline climate model. Additional climate variables, such as evapotranspiration and water balance (see Stephenson 1990), also may be suited for differentiating topographic effects. As the technology becomes available, effects of drainages on mountain temperatures could be incorporated using GIS-based systems. The resulting tool would be of use in either the fine- or coarse-scale work of ecophysicists or plant geographers in predicting the occurrence of plant associations and their component species on forested landscapes.

Our objective analyses have supported and quantified the conceptual views of descriptive ecologists. In the process, the Random Forests classification tree proved effective in sifting through an array of intercorrelated climate variables to select the most relevant for segregating species into their respective climatic niches. The results also demonstrated that the skewed-normal distributions relating abundance of species and the probability of their occurrence to climate can be fit with high precision using Weibull response functions. In the forests of the Inland Northwest, allocation of conifer species to their characteristic realized niches is driven by a single climate variable, the growing-season dryness index, although visualizing niche space is facilitated by considering a second variable, negative degree-days.

ACKNOWLEDGMENTS

We are indebted to Forest Inventory and Analysis, USDA Forest Service, Ogden, Utah and Portland, Oregon for the use of their data. For thoughtful comments on various versions of this manuscript, we thank Andy Hudak, John Marshall, Dennis Joyce, and Rudy King. We are particularly grateful for the thorough and stimulating criticism of Xiaoli Fern and an anonymous referee.

LITERATURE CITED

- Alerich, C. A., L. L. Bednar, G. J. Brand, K. Campbell, L. Klevgard, K. Nimerfro, L. Royer, M. E. Rubey, G. Sendhil, R. Wanek, C. Washington, and S. Waters. 2006. The forest inventory and analysis database: database description and user's guide. Version 2.1. (http://ncrs.fs.fed.us/4801/tools-data/data/doc/FIADB_v2-1_07_05_06.pdf)
- Austin, M. P., A. O. Nicholls, and C. R. Margules. 1990. Measurement of the realized qualitative niche: environmental

- niches of five *Eucalyptus* species. *Ecological Monographs* 60: 161–177.
- Baker, F. S. 1944. Mountain climates of the western United States. *Ecological Monographs* 14:223–254.
- Bechtold, W. A., and P. L. Patterson. 2005. The enhanced forest inventory and analysis program—national sampling design and estimation procedures. General Technical Report SRS-80. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, North Carolina, USA.
- Bonan, G. B. 2002. *Ecological climatology*. Cambridge University Press, New York, New York, USA.
- Breiman, L. 2001. Random forests. *Machine Learning* 45:5–32.
- Breiman, L., and A. Cutler. 2004. Random Forests. (http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#varimp)
- Brown, J. H., and A. C. Gibson. 1983. *Biogeography*. C. V. Mosby, St. Louis, Missouri, USA.
- Brunelle, A., G. E. Rehfeldt, B. Bentz, and A. S. Munson. 2008. Holocene records of *Dendroctonus* bark beetles in high elevation pine forests of Idaho and Montana, USA. *Forest Ecology and Management* 255:836–846.
- Campbell, E. M., and J. A. Antos. 2000. Distribution and severity of white pine blister rust and mountain pine beetle on whitebark pine in British Columbia. *Canadian Journal of Forest Research* 30:1051–1059.
- Cooper, S. V., K. E. Neiman, and D. W. Roberts. 1991. Forest habitat types of northern Idaho: second approximation. General Technical Report INT-236. U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah, USA.
- Daubenmire, R. 1952. Forest vegetation of northern Idaho and adjacent Washington and its bearing on concepts of vegetation classification. *Ecological Monographs* 22:301–330.
- Daubenmire, R. 1956. Climate as a determinant of vegetation distribution in eastern Washington and northern Idaho. *Ecological Monographs* 26:131–154.
- Daubenmire, R. 1966. Vegetation: identification of tygal communities. *Science* 151:291–298.
- Daubenmire, R., and J. B. Daubenmire. 1968. Forest vegetation of eastern Washington and northern Idaho. Washington Agriculture Experiment Station Technical Bulletin 60. Washington State University, Pullman, Washington, USA.
- Fang, K. T., and J. L. Xu. 1987. Statistical distributions. Science Press, Beijing, China. [In Chinese.]
- Geiger, R. 1950. *The climate near the ground*. Harvard University Press, Cambridge, Massachusetts, USA.
- Haig, I. T., K. P. Davis, and R. H. Weidman. 1941. Natural regeneration in the western white pine type. Technical Bulletin Number 767. U.S. Department of Agriculture, Washington, D.C., USA.
- Huff, D. D., S. L. Huber, and A. N. Borisenko. 2005. Using field data to estimate the realized thermal niche of aquatic vertebrates. *North American Journal of Fisheries Management* 25:346–360.
- Hungerford, R. D., R. R. Nemani, S. W. Running, and J. C. Coughlan. 1989. MTCLIM: a mountain microclimate simulation model. Research Paper INT-414. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, Utah, USA.
- Hutchinson, G. E. 1957. Concluding remarks: population studies: animal ecology and demography. Cold Spring Harbor Symposia on Quantitative Biology 22:415–427.
- Hutchinson, M. F. 1991. The application of thin plate smoothing splines to continent-wide data assimilation. Pages 104–113 in J. D. Jasper, editor. *Data assimilation systems*. Research Report 27. Bureau of Meteorology, Melbourne, Victoria, Australia.
- Hutchinson, M. F. 2000. ANUSPLIN. Version 4.1. User's guide. Australian National University, Centre for Resource and Environmental Studies, Canberra, A.C.T., Australia.
- Johnson, N. L., S. Kotz, and N. Balakrishnan. 1994. *Continuous univariate distributions*. Volume 1. John Wiley and Sons, New York, New York, USA.
- Kullman, L. 1995. Holocene tree-limit and climate history from the Scandes Mountains, Sweden. *Ecology* 76:2490–2502.
- Larsen, J. A. 1930. Forest types of the northern Rocky Mountains and their climatic controls. *Ecology* 11:631–672.
- Lawesson, J. E. 2000. Danish deciduous forest types. *Plant Ecology* 151:199–221.
- Liaw, A., and M. Wiener. 2002. Classification and regression by Random Forests. *R News* 2:18–22.
- Lillybridge, T. R., B. L. Kovalchik, C. K. Williams, and B. G. Smith. 1995. Field guide for forested plant associations of the Wenatchee National Forest. General Technical Report 359. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.
- O'Laughlin, J., J. G. MacCracken, D. L. Adams, S. C. Bunting, K. A. Blatner, and C. E. Keegan, III. 1993. Forest health conditions in Idaho. Report Number 11. Idaho Forest, Wildlife, and Range Policy Analysis Group, College of Natural Resources, University of Idaho, Moscow, Idaho, USA.
- Pfister, R. D., B. L. Kovalchik, S. F. Arno, and R. C. Presby. 1977. Forest habitat types of Montana. General Technical Report 34. U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah, USA.
- Pranab, K. S. 2005. Gini diversity index, Hamming distance, and curse of dimensionality. *Metron International Journal of Statistics* 63:329–349.
- Pyne, S. J. 1982. *Fire in America: a cultural history of wildland and rural fire*. University of Washington Press, Seattle, Washington, USA.
- R Development Core Team. 2004. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rehfeldt, G. E. 2006. A spline model of climate for western United States. General Technical Report 165. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ft. Collins, Colorado, USA.
- Rehfeldt, G. E., N. L. Crookston, M. V. Warwell, and J. S. Evans. 2006. Empirical analyses of plant–climate relationships for the western United States. *International Journal of Plant Science* 167:1123–1150.
- Rehfeldt, G. E., N. M. Tchebakova, L. I. Milyutin, E. I. Parfenova, and W. R. Wykoff. 2003. Assessing population responses to climate in *Pinus sylvestris* and *Larix* spp. of Eurasia with climate-transfer models. *Eurasian Journal of Forest Research* 6:82–93.
- Rosenberg, N. J. 1974. *Microclimate: the biological environment*. John Wiley and Sons, New York, New York, USA.
- SAS. 2003. *SAS user's guide*. Release 9.1 edition. SAS Institute, Cary, North Carolina, USA.
- Steele, R., R. D. Pfister, R. A. Ryker, and J. A. Kittams. 1981. Forest habitat types of central Idaho. General Technical Report 114. U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah, USA.
- Stephenson, N. L. 1990. Climatic control of vegetation distribution: the role of the water balance. *American Naturalist* 135:649–670.
- Thuiller, W., S. Lavorel, G. Midgley, S. Lavergne, and T. Tebelo. 2004. Relating plant traits and species distributions along bioclimatic gradients for 88 *Leucadendron* taxa. *Ecology* 85:1688–1699.
- Williams, C. K., B. F. Kelley, B. G. Smith, and T. R. Lillybridge. 1995. Forested plant associations of the Colville National Forest. General Technical Report 360. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.
- Woodward, F. I. 1987. *Climate and plant distribution*. Cambridge University Press, New York, New York, USA.