Spatial variation of climatic and non-climatic controls on species distribution: the range limit of *Tsuga heterophylla*

Daniel G. Gavin* and Feng Sheng Hu

**ABSTRACT**

**Aim** To assess which climatic variables control the distribution of western hemlock (*Tsuga heterophylla*), how climatic controls vary over latitude and between disjunct coastal and interior sub-distributions, and whether non-climatic factors, such as dispersal limitation and interspecific competition, affect range limits in areas of low climatic control.

**Location** North-western North America.

**Methods** We compared four bioclimatic variables [actual evapotranspiration (AET), water deficit (DEF), mean temperature of the coldest month (MTCO), and growing degree-days (GDD5)] with the distribution of *T. heterophylla* at a 2-km grid cell resolution. The distribution is based on a zonal ecosystem classification where *T. heterophylla* is the dominant late-successional species. For each bioclimatic variable and at each degree of latitude, we calculated the threshold that best defines the *T. heterophylla* distribution and assessed the extent to which *T. heterophylla* was segregated to one end of the bioclimatic gradient. We also fitted two forms of multivariate bioclimatic models to predict the *T. heterophylla* distribution: a simple threshold model and a complex Gaussian mixture model. Each model was trained separately on the coastal and interior distributions, and predicted areas outside of the *T. heterophylla* distribution (overprediction) were evaluated with respect to known outlier populations.

**Results** Actual evapotranspiration was the most accurate predictor across the *T. heterophylla* distribution; other variables were important only in certain areas. There was strong latitudinal variation in the thresholds of all variables except AET, and the interior distribution had wider bioclimatic thresholds than the coastal distribution. The coastal distribution was predicted accurately by both bioclimatic models; areas of overprediction rarely occurred > 10 km from the observed distribution and generally matched small outlier populations. In contrast, the interior distribution was poorly predicted by both models; areas of overprediction occurred up to 140 km from the observed distribution and did not match outlier populations. The greatest overprediction occurred in Idaho and Montana in areas supporting species that typically co-exist with *T. heterophylla*.

**Main conclusions** The high predictive capacity of AET is consistent with this species’ physiological requirements for a mild and humid climate. Spatial variation of MTCO, GDD5 and DEF thresholds probably reflects both the correlation of these variables with AET and ecotypic variation. The level of overprediction in portions of the interior suggests that *T. heterophylla* has not completely expanded into its potential habitat. *Tsuga heterophylla* became common in the interior 2000–3500 years ago, compared with > 9000 years ago in the coastal region. The limited time for dispersal, coupled with frequent fires at the margins of the distribution and competition with disturbance-adapted species, may have retarded range expansion in the interior. This study
INTRODUCTION

The distributions of plant species are affected by both climatic factors (e.g. temperature and water balance) and non-climatic factors (e.g. biotic interactions: Woodward, 1987; Huston, 2003). Climate affects species distributions directly by imposing physiological constraints on growth and reproduction, and indirectly by mediating competitive interactions (Shao & Halpin, 1995; Sykes et al., 1996), although some constraints may vary spatially because of local adaptation (Case & Taper, 2000). For some species, distribution limits correlate poorly with climatic gradients, especially at fine spatial resolutions, suggesting the importance of other controls such as dispersal limitation, disturbance and competition (Malanson et al., 1992; Loehle, 1998). The difficulty of distinguishing controls has led to several critiques that warn against directly mapping species distributions from climate (Woodward & Beerling, 1997; Davis et al., 1998; Hampe, 2004). However, no studies have attempted to distinguish the importance of climatic and non-climatic controls throughout the entire distribution of a species. A detailed examination of the distribution–climate correlation, especially in the light of a species’ autecology, synecology (interactions with commonly associated species), ecophysiology and historical distribution, may yield insights on the relative roles of climatic and non-climatic controls.

Western hemlock [Tsuga heterophylla Raf. (Sarg.)], a common late-successional tree in north-western North America, presents an excellent case for examining the climatic and non-climatic controls on distribution. On one hand, the distribution of T. heterophylla should strongly reflect regional climate because of its physiological requirement for mild and humid conditions (Packee, 1990). The species requires higher water potential to maintain cell turgor than most other conifers in the region (Jackson & Spomer, 1979; Lassoe et al., 1985), and it has a relatively high tendency for xylem cavitation (Bond & Kavanagh, 1999). The shallow roots of the species (Bennett et al., 2002) further exacerbate its vulnerability to water deficits. On the other hand, dispersal limitation and interspecific competition may reduce the importance of climatic controls in some areas more than others. Pollen records indicate the relatively recent arrival of T. heterophylla in a large disjunct interior distribution (2000–3500 yr BP: Mehringer, 1996; Rosenberg et al., 2003) compared with the main coastal distribution (mostly occupied by 9000 yr BP), thus T. heterophylla had less time to reach its climatically defined distributional limits in the interior than in the coastal region. Unlike in the coastal distribution, stand-replacing fire is more frequent in the interior, and T. heterophylla must compete with early successional tree species (e.g. Pinus contorta, Picea glauca × engelmannii, Betula papyrifera) in post-fire forests. These factors suggest that T. heterophylla may be subject to stronger non-climatic controls at its distribution limits in the interior than in the coastal distribution.

In this study we address spatial variation of climatic controls on the distribution of T. heterophylla by comparing mapped bioclimatic variables with its observed distribution. First, for each individual bioclimatic variable, we assess its predictive capacity and determine the threshold that best predicts the T. heterophylla distribution. Spatial variation in the thresholds of bioclimatic variables may be due to adaptations to local climates (Rehfeldt et al., 1999), as shown for Tsuga canadensis (eastern hemlock: Kessell, 1979). We use a base map with a 2-km grid cell resolution to capture the species–climate relationship in mountainous areas. Second, we use bioclimatic envelope models to assess how accurately climate can predict the T. heterophylla distribution (Guisan & Zimmermann, 2000; Pearson & Dawson, 2003). We focus on areas of prediction outside of the observed distribution (overprediction) and assess whether such areas are consistent with potential non-climatic controls (Svenning & Skov, 2004). If bioclimatic models are poorly constructed, they may falsely suggest either weaker or stronger climatic control than actually exists (underfitting and overfitting, respectively). As our goal is to assess potential non-climatic controls operating in areas of overprediction, we use an approach that errs on the side of overfitting in order to minimize false overprediction. Specifically, we use a complex bioclimatic model with variable interactions to fit the species–climate relationship in addition to a simple threshold bioclimatic model. We also model only the core distribution of T. heterophylla in order not to include outlying populations that may occur in anomalous local conditions unrepresentative of the regional climate. Such populations, if included in constructing a bioclimatic model, may falsely increase the breadth of the climatic envelope and lead to false overprediction.

METHODS

Data sources

We defined the observed distribution of T. heterophylla using zonal ecosystems based on the Biogeoclimatic Zones of British
Columbia (Meidinger & Pojar, 1991). This classification was developed from extensive field surveys and aerial photo interpretation, and eventually mapped using elevation and aspect rules that vary among regions. Climatic data were not used in the mapping of zonal ecosystems, although it was assumed that the maps of zonal ecosystems reflect climatic gradients (Meidinger & Pojar, 1991). We did not use maps inclusive of all outlying populations because microclimates or local topographic effects rather than regional climate may maintain such populations. Zonal ecosystems are particularly suitable for describing the core distribution of late-successional species such as *Tsuga heterophylla* because zones are defined by the expected late-successional dominants, even if they are rare due to disturbance. Another advantage of zonal ecosystems is that they are defined at finer spatial resolutions than most species-distribution data (Little, 1971) and thus are suitable for areas with steep climatic gradients where coarse resolution data would not adequately describe the species–climate correlation.

We created the 2-km grid cell resolution map using a mosaic of previous classifications (Table 1); 98.9% of the zonal ecosystem was defined by two classification schemes (Broad Ecosystem Inventory and Ecoregions of Oregon) with a concordant overlap in northern Oregon. In all, the *Tsuga heterophylla* distribution was 302,008 km² (75,502 2-km grid cells, Albers equal-area projection), of which 18% occurs in the interior (Fig. 1). We defined the coastal and interior regions as the areas that are separated by a line equidistant between the two sub-distributions and that encompass up to 140 km from the distribution elsewhere. These regions extend the analysis out to areas with substantially different climatic regimes (e.g. east of the Rocky Mountains; Fig. 1).

Our analyses used four bioclimatic variables that are expected to have a strong functional relationship with the distribution of *Tsuga heterophylla*: actual evapotranspiration (AET); mean temperature of the coldest month (MTCO); water balance deficit (DEF); and growing degree-days on a 5°C base (GDD5). We calculated these four variables (see Table 2 for methods) from gridded climate normals (1961–90) obtained from the PRISM climate-mapping project (Daly et al., 1994). These variables were selected based on the results of earlier studies. Gavin & Hu (2005) used the same data with a regression tree analysis to select variables from a list of nine potential variables. That study found AET to be the most predictive of the *Tsuga heterophylla* distribution, followed by MTCO and GDD5. McKenzie et al. (2003) used general linear models to select variables to predict conifer species in Washington State, and found AET, DEF and soil degree-days, among others, to be the most predictive. Other studies have shown that AET and DEF are functionally significant predictors of vegetation in general (Stephenson, 1998).

### Univariate analyses

The degree of segregation of *Tsuga heterophylla* along each bioclimatic gradient was quantified with receiver-operating characteristic analysis (ROC: Swets, 1988). ROC analysis
involves (1) determining the fraction of \textit{T. heterophylla} grid points that exceed a threshold (the true positive fraction, TPF) and the fraction of non-\textit{T. heterophylla} grid points that exceed the same threshold (the false positive fraction, FPF), and (2) plotting TPF as a function of FPF over the full range of thresholds (an ROC curve). The area under the ROC curve (AUC) is a measure of the segregation of \textit{T. heterophylla} grid points towards one end of the range of values of a bioclimatic variable. AUC values near 0.5 indicate no segregation of \textit{T. heterophylla} grid points to either end of the gradient; values near 1 indicate strong segregation of \textit{T. heterophylla}; and values near 0 also indicate segregation, but in the opposite direction.

We assessed the significance of AUC in contiguous one-degree latitudinal bands by comparing it with a distribution of AUC values computed from a null model of the \textit{T. heterophylla} distribution. The null model consisted of: (1) determining the range of bioclimatic values occupied by \textit{T. heterophylla}, (2) selecting such a range at a random location within the entire range of combined \textit{T. heterophylla} and non-\textit{T. heterophylla} grid points, and (3) randomly selecting grid points from this segment equal to the number of grid points that occur within the observed \textit{T. heterophylla} distribution. We constructed 95% confidence envelopes of AUC from 2000 realizations of the null model.

We calculated thresholds of bioclimatic variables from the 5th or 95th percentiles of each variable within the \textit{T. heterophylla} distribution in contiguous 1° latitudinal bands. Thresholds were calculated only where the \textit{T. heterophylla} distribution is significant according to the null model.
distribution was segregated to one end of the gradient as determined by a statistically significant AUC. We also calculated overall thresholds for the coastal and interior regions. Latitudinal trends in AUC and thresholds were determined with linear regression, and inter-region differences in AUC and thresholds were assessed qualitatively.

**Multivariate bioclimatic models**

**Model forms and training regions**

We applied multivariate bioclimatic models to examine (1) how closely bioclimatic models can fit the observed *T. heterophylla* distribution using a resubstitution test (hereafter termed within-region prediction), and (2) whether models trained on one region (e.g. coastal region) can predict the other region (e.g. interior region) and thus how climatic controls differ between the regions (hereafter termed cross-region prediction). For within-region prediction we used the GM–SMAP model based on multivariate Gaussian mixture distributions (described below). This model is inappropriate for cross-region prediction because the climate space (vectors of the four variables) differs greatly between the coastal and interior regions (Fig. 1b). Any model that includes variable interactions will perform poorly across regions with different variable interactions due to unrealistic extrapolation (Beerling et al., 1995; Sykes, 2001). This issue is quite pronounced when using four variables and their interactions, as in this paper. Therefore, for cross-region predictions we used a simple rectilinear envelope (RE) model that predicts the *T. heterophylla* distribution using coastal and interior-region thresholds calculated from the univariate analysis (similar to the BIOC-LIM model of Beaumont & Hughes, 2002).

**Gaussian mixture–sequential maximum a posteriori (GM–SMAP) segmentation**

The GM–SMAP model was demonstrated to have a greater capacity to predict the *T. heterophylla* distribution than either regression trees or smoothed response surfaces in Gavin & Hu (2005). This model was developed to ‘segment’ multi-spectral satellite images into land-cover types (Bouman & Shapiro, 1994); here it is used to predict a species distribution using several bioclimatic variables. First, ‘climatic signatures’ were computed by fitting GM distributions to the climate in each of two ‘classes’: the area within and outside the *T. heterophylla* distribution. GM distributions were fitted using the program CLUSTER (Bouman, 1998), which breaks each class into subclasses using the expectation-maximization (EM) algorithm of Dempster et al. (1977) and computes a multivariate Gaussian distribution for each subclass. The closest two subclasses are then merged and the EM algorithm is run again, and this merging is repeated until the number of subclasses yields the smallest Rissanen minimum description length (Bouman, 1998), an information-theoretical index that penalizes the fit of a model with the number of parameters in the model. The result is a response surface that optimally describes the training data and adapts to steep or gradual gradients in environmental space.

The training data included the entire *T. heterophylla* distribution and every third grid point (on a diagonal pattern) in the area outside the *T. heterophylla* distribution. The latter reduced the number of samples to a more tractable size for fitting the model; given a high degree of spatial autocorrelation in the 2-km grid resolution, it should minimally affect the model’s accuracy. The GM model fitted the climatic variability using a larger number of subclasses in the coastal region (131 and 134 subclasses for the area within and outside of the *T. heterophylla* distribution, respectively) than in the interior (60 and 104, respectively).

The GM distributions were used to make predictions using the SMAP algorithm (i.smap function in GRASS GIS software), a method that uses predictions at coarse spatial scales to guide predictions at fine spatial scales (Bouman & Shapiro, 1994). This method results in more contiguous predictions and increases the overall accuracy of predictions compared with a non-spatial method using the same GM distributions (McCauley & Engel, 1995).

**Model evaluation**

Predicted vs. observed distributions of *T. heterophylla* were compared in each region using the kappa ($\kappa$) statistic. $\kappa$ ranges from −1 (the predicted distribution complements the observed distribution) to 1 (the predicted and observed distributions are identical), and values close to 0 indicate that the predicted distribution is not a better fit than that expected by chance alone. To show the locations of overpredicted areas, we summarized overprediction for each 1° latitudinal band as a function of distance to the range limit of *T. heterophylla* in 10-km distance classes.

We examined whether areas of model overprediction, using within-region predictions from the GM–SMAP model, coincided with outlier populations of *T. heterophylla* beyond its distribution as defined by the zonal ecosystems. We focused on three areas that varied in the amount of overprediction: the northern and southern ends of the interior region and the southern end of the coastal region. In the USA, known occurrences in vegetation plots were obtained from the VegBank database (VEGBANK Contributors, 2004) and the Forest Inventory and Analysis (FIA) database (USDA Forest Service, 2005). Landowner privacy regulations prevent the US Government from distributing FIA data associated with precise plot locations. We therefore obtained from the Forest Service a polygon map of *T. heterophylla* presence generalized to a hexagonal grid (27 km between polygon centres) that was not subject to the plot-location errors in the publicly available FIA data. In British Columbia, known occurrences were obtained from the Biogeoclimatic Ecosystem Classification database (http://www.for.gov.bc.ca/hre/becweb/index.htm). Overall, we searched > 46,000 plots for the presence of *T. heterophylla* in these three areas.
RESULTS

Univariate analyses

In the coastal region, AET is the strongest predictor of the *Tsuga heterophylla* distribution, with *T. heterophylla* preferring sites with higher AET than other locations at the same latitude (Figs 2 & 3). Similarly, DEF is also a strong predictor over most of the coastal region, with *T. heterophylla* preferring lower DEF than other locations at the same latitude (Figs 2 & 3). The predictive power of AET and DEF, as measured by the area under the ROC curve (AUC), is significantly greater than the null model (Fig. 3), but it decreases from a maximum in the south (AUC = 0.97 and 0.96 for AET and DEF, respectively) to a minimum in the north (AUC = 0.83 and 0.53 for AET and DEF, respectively). MTCO is also a strong predictor over most of the coastal region (AUC = 0.83–0.99), with *T. heterophylla* preferring higher MTCO at a given latitude, but this preference is only marginally significant from 45–50° N (AUC = 0.73–0.84). GDD5 is a moderate predictor (AUC = 0.61–0.94), with *T. heterophylla* preferring higher GDD5 at all latitudes except 45–48° N where *T. heterophylla* occurs closer to the low end of the GDD5 gradient (AUC = 0.26–0.30).

In the interior region, all four bioclimatic variables are weaker predictors than in the coastal region (Fig. 3). However, as in the coastal region, AET is the strongest predictor, although it is insignificant at two out of eight latitudinal bands (AUC = 0.59–0.84). MTCO is a weak predictor in the south (AUC = 0.50), and it increases to significant levels in the north (> 50° N; AUC = 0.68–0.78). DEF is a poor predictor (AUC = 0.70–0.36), significant only at 47–48° N. GDD5 is not predictive in the south (AUC = 0.38), but it increases in strength with latitude and reaches significant levels at the northern end (AUC = 0.72).

The thresholds of all four bioclimatic variables show that *T. heterophylla* has greater climatic tolerances in the interior region compared with the coastal region (Fig. 2a). The *T. heterophylla* distribution extends into lower AET, MTCO and GDD5 in the interior region than in the coastal region (380 vs. 419 mm; −10.9 vs. −9.0 °C; 579 vs. 650 degree-days). In addition, the distribution extends into higher DEF in the interior region (168 mm) than in the coastal region (145 mm).

Threshold values show latitudinal trends for several variables but not for others (Fig. 2b). In the coastal distribution, the threshold values have no latitudinal trend for AET, varying from 343 to 537 mm (linear regression, \( P = 0.149 \)), but they decrease with increasing latitude for MTCO (from 6.6 to

---

**Figure 2** (a) Distribution of bioclimatic variables within and outside the *Tsuga heterophylla* distribution (black and grey box plots, respectively) in the coastal and interior regions. Boxes show the middle quantile and vertical lines extend to the 5th and 95th percentiles. (b) Box plots of bioclimatic variable distributions for contiguous 1° latitudinal bands and separately for the coastal and interior regions. Open circles indicate the thresholds for *T. heterophylla* at latitudes where the *T. heterophylla* distribution is significantly segregated to one end of the bioclimatic gradient.
The coastal-trained RE model fits the coastal region much more accurately than the interior-trained model fits the interior region (difference in $\kappa = 0.534$; Table 3; Fig. 4a,b). The interior RE model overpredicts the area occupied by $T$. heterophylla by $> 80\%$ of the interior region, and the overpredictions occur at distances $> 100$ km from the species’ actual distribution at several latitude bands (Fig. 5). Overprediction by the interior model is relatively minor in the coastal region (e.g. at $55^\circ$ N). The coastal RE model, based on narrower thresholds than the interior RE model, results in a greater $\kappa$ in the interior region (due to less overprediction; Table 3). However, it incorrectly delimits (underpredicts) large areas at the north end of the interior (red areas on Fig. 4a).

Using GM–SMAP, the coastal-trained model also fits the coastal region more accurately than the interior-trained model fits the interior region (difference in $\kappa = 0.199$; Table 3; Fig. 4c,d). Areas of overprediction based on the coastal model rarely extend $> 30$ km from the coastal $T$. heterophylla distribution (Fig. 5). In contrast, the interior model overpredicts a large area (c. $30\%$ of the total predicted area) up to $140$ km from the interior $T$. heterophylla distribution. Because of climatic dissimilarity between the two regions (Fig. 1b), we do not interpret cross-region predictions from the GM–SMAP model.

A finer-resolution examination of the GM–SMAP model in three areas shows that the predicted $T$. heterophylla distribution is a close match to the observed distribution, with a few exceptions (Fig. 6). At the south end of the coastal region, the model accurately predicts the $T$. heterophylla distribution in most areas, and the few areas of overprediction coincide with outlier $T$. heterophylla populations detected by vegetation plots (Fig. 6c). At the north end of the interior region, the model is unable to predict narrow belts ($< 5$ km wide) of the $T$. heterophylla distribution, and two small areas of overprediction exist to the north (Fig. 6b). Small outlier populations of $T$. heterophylla occur up to $130$ km from its zonal distribution, but not near the areas of overprediction. At the south end of the interior region, the model predicts the $T$. heterophylla distribution accurately along its eastern and western borders, but it overpredicts extensively to the south and in disjunct patches to the east (Fig. 6a). Although outlier $T$. heterophylla populations coincide with some of these patches of overprediction to the east, few outlier populations occur to the south.

**DISCUSSION**

**Variation in bioclimatic thresholds**

The lack of latitudinal trend and high predictive capacity of AET suggest that it is a fundamental control on the distribution of $T$. heterophylla (Figs 2 & 3). This result is not surprising because $T$. heterophylla is known to occur in areas with mild and

---

**Table 3** Kappa ($\kappa$) statistics for predictions of the $Tsuga$ heterophylla distribution for two bioclimatic models each trained separately on the coastal and interior regions

<table>
<thead>
<tr>
<th>Model</th>
<th>Training region</th>
<th>Testing region</th>
<th>Coastal</th>
<th>Interior</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>Coastal</td>
<td>0.674</td>
<td>0.269</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interior</td>
<td>0.591</td>
<td>0.140</td>
<td></td>
</tr>
<tr>
<td>GM–SMAP</td>
<td>Coastal</td>
<td>0.731</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interior</td>
<td>–</td>
<td>0.532</td>
<td></td>
</tr>
</tbody>
</table>

The GM–SMAP model cannot be used for cross-region prediction because of dissimilar climate between regions.

---

**Figure 3** The degree to which $Tsuga$ heterophylla is segregated to one end of the bioclimatic gradient in $1^\circ$ latitudinal bands, as calculated by ROC analysis. Solid line, AUC for the observed $T$. heterophylla distribution; dashed lines, show the 95% confidence envelope for null models of randomly generated $T$. heterophylla distributions. *Slope of trend significant at $P < 0.05$; **$P < 0.01$.

---

**Multivariate bioclimatic models**

The coastal-trained RE model fits the coastal region much more accurately than the interior-trained model fits the interior region (difference in $\kappa = 0.534$; Table 3; Fig. 4a,b). The interior RE model overpredicts the area occupied by $T$. heterophylla by $> 80\%$ of the interior region, and the overpredictions occur at distances $> 100$ km from the species’ actual distribution at several latitude bands (Fig. 5). Overprediction by the interior model is relatively minor in the coastal region (e.g. at $55^\circ$ N). The coastal RE model, based on narrower thresholds than the interior RE model, results in a greater $\kappa$ in the interior region (due to less

---

$-16.6^\circ$ C, $P < 0.001$), DEF (210 to 0 mm, $P < 0.001$), and GDD5 (2456–333 degree-days, $P < 0.001$). Threshold variations in the interior are much smaller for AET (373–459 mm) and MTCO ($-10.0$ to $-11.8^\circ$ C), and they show no latitudinal trends.
humid climates. Other bioclimatic modelling studies of *Tsuga heterophylla* or similar species have also identified AET as the most predictive variable (Shao & Halpin, 1995; Stephenson, 1998; McKenzie et al., 2003). In contrast with these previous modelling results, however, our data reveal inter-regional variation in bioclimatic requirements. Specifically, *T. heterophylla* in the interior extends into areas with AET 39 mm lower than in the coastal region. In addition, the other three variables all show greater within-region (latitudinal) variation in thresholds than AET. For MTCO and GDD5, latitudinal variation in thresholds is quite large (> 10 °C and > 2000 degree-days, respectively).

Spatial variation in the thresholds of some bioclimatic variables probably reflects their correlation with a functionally significant variable (e.g. AET). For example, the high MTCO threshold in the southern portion of the coastal distribution probably results from the restriction of precipitation to the winter months in that area. Because of very dry summers, *T. heterophylla* must meet its AET requirements by growing at low elevations where temperatures are high enough for photosynthesis during a portion of the rainy season. Further north, precipitation is less restricted to the winter months, thus *T. heterophylla* may occur in areas with colder winters provided there is sufficient summer moisture (one exception is at 58°N where coastal *T. heterophylla* extends eastward into rainshadowed valleys with less summer moisture; Figs 1 & 2). This rationale also explains latitudinal variation in GDD5, because GDD5 is moderately correlated with MTCO in the coastal region (r = 0.85). Similarly, the latitudinal trend in DEF may be attributed to its relationship with AET. In the
north, areas with DEF > 20 mm also have very low precipitation (and low AET), which restricts T. heterophylla to very low DEF within these areas. In the south, T. heterophylla occurs in areas with very mild winters that are also very dry in the summer (relatively high DEF). While these co-varying trends suggest that non-AET variables are often redundant with AET, maps of areas delimited by individual thresholds show that there are indeed several areas where only DEF, GDD5 or MTCO are predictive. For example, DEF is the only limiting variable in the dry lowland valleys of western Oregon and Washington (see Fig. S1 in Supplementary Material).

Bioclimatic thresholds may also be affected by local adaptation causing clines in physiological thresholds (ecotypic variation). Common-garden experiments provided independ-
ent evidence of ecotypic variation in growth rate and frost hardness (Kuser & Ching, 1980, 1981). *Tsuga heterophylla* seeds originating from throughout its distribution, and germinated in an Oregon shadehouse at ambient temperature, showed that increasing provenance latitude resulted in increased survival following an early winter frost, decreased growth rates, and decreased shoot : root ratios (Kuser & Ching, 1980). Hannertz et al. (1999) found similar results in south-western British Columbia. Furthermore, compared with coastal provenances at the same latitudes, interior provenances have greater frost hardness and lower growth rates (Kuser & Ching, 1980, 1981). The finding of a frost-tolerance cline from these experiments matches our observed trend in MTCO thresholds (Fig. 2), but we suggest that the majority of this trend is attributable to the seasonality of precipitation mentioned above. One plausible scenario for clinal variation in MTCO thresholds is that AET provides a fundamental control on distribution limits (little local adaptation with respect to AET), but that within the AET threshold there are local adaptations to seasonal temperature variables. Unfortunately, these common-garden studies did not address moisture requirements and yielded inconsistent results regarding budburst dates, thus they do not offer direct evidence that clines in the other three variables result from local adaptation.

The occurrence of variation in bioclimatic thresholds presents difficulties for bioclimatic modelling (Loehle & LeBlanc, 1996). If a single AET threshold were sufficiently predictive throughout the observed distribution of *T. heterophylla*, predicting past and future distributions would be limited primarily by the accuracy of palaeo- and projected climatic data. However, multiple variables and their interactions (the GM–SMAP model) were required to fit the interior distribution (Fig. 4), suggesting a more complex species–climate relationship. Complexity such as spatial variation in thresholds may be modelled by the variable interactions of the GM–SMAP model. For example, a GDD5 × MTCO interaction may be more predictive than individual variables separately because the threshold of one variable appears correlated with the other (Fig. 2). The drawback of using multiple variable interactions is that models may become too specific to the environmental space of the training data, lowering accuracy when predicting in different climatic regimes. This problem is demonstrated by the inability of the GM–SMAP model, which includes variable interactions, to make cross-region predictions.

### Climatic and biotic controls of the *T. heterophylla* distribution

Both the RE and GM–SMAP models fit the coastal distribution rather well, suggesting strong climatic control of *T. heterophylla* occurrence. Their overprediction within 10 or 20 km of the *T. heterophylla* distribution (Fig. 5) probably reflects limitations in the spatial resolution of the climate data compared with that of the zonal ecosystem data. In valley bottoms, a one grid point–wide line often represents the *T. heterophylla* distribution, which is probably too fine to be captured by the climatic data. Furthermore, the SMAP algorithm makes predictions in the context of neighbouring grid points, effectively smoothing the predictions. We expect that finer-scale features such as narrow valleys could be accurately predicted by downscaling the temperature data using locally computed lapse rates and a fine-scale (e.g. 500-m) digital elevation model (Hamann & Wang, 2005). However, increased accuracy at fine spatial scales would not affect the major, coarse-scale patterns of overprediction detected in this study.

The interior *T. heterophylla* distribution is more poorly constrained by climate than the coastal distribution, regardless of the bioclimatic variables or form of model used. In fact, the complex GM–SMAP model (using all variable interactions) could not predict the interior distribution as accurately as a single AET threshold predicted the coastal distribution (Fig. 4d, \( \kappa = 0.532 \) vs. Fig. S1, \( \kappa = 0.553 \)). The weaker climatic control of the interior distribution is upheld with further exploratory analyses. First, using latitude-specific thresholds increases the fit of the interior RE model very little, reflecting the fact that these thresholds vary only a minor amount by latitude (Fig. 2). Second, the steeper climatic gradients in the coastal distribution may have resulted in a smaller area of overprediction. Indeed the sensitivity of the area of overprediction to changes in thresholds in the coastal distribution was c. 50% of that in the interior distribution. However, this sensitivity difference accounts for < 20% of the large discrepancy in the overpredicted areas using coastal vs. interior thresholds (Fig. 5a,b).

The overprediction of a few specific localities distal from the main interior distribution suggests that *T. heterophylla* has not fully expanded into its potential interior range. The most extensive overprediction occurs in Idaho and north-western Montana (Fig. 6). Many of these areas are patchy habitat islands on windward sides of mountain ranges (Fig. 6), and in Idaho these areas support dense forests with species that usually co-occur with *T. heterophylla*. Many of these species probably have broader climatic envelopes than *T. heterophylla*, such as the canopy-dominant trees *Thuja plicata* (western redcedar) and *Pinus monticola* (western white pine). Several other species have distributions that overlap widely with the coastal *T. heterophylla* distribution but are also common south of the interior *T. heterophylla* distribution, such as *Alnus rubra*, *Coriunus nutallii*, *Blechnum spicant* and *Viola sempervirens* (Daubenmire, 1952, 1975; Johnson, 1968; Habeck, 1978). Lorain (1988) identified 41 vascular plant species with this geographical pattern. Daubenmire (1952, 1975) speculated that this large number of disjunct coastal species in the unglaciated valleys of northern Idaho (the Clearwater and Lochsa valleys) indicated the location of a Pleistocene refugium, which is supported by the phylogeographical patterns of several plant and animal species (Brunsfeld et al., 2001). However, *T. heterophylla* is absent from this refugium and has not been observed more than 4 km south of its zonal distribution (c. 46°45’ N, Fig. 6; Habeck, 1978, 1987; Cooper et al., 1991). It is surprising that *T. heterophylla* has not colonized moist
drainages in these areas, as other species with broadly similar climatic requirements exist in abundance.

Areas of unoccupied potential range in the interior may be attributable to a combination of dispersal limitation and interspecific competition. Pollen records (Mehringer, 1996; Rosenberg et al., 2003) indicate that *T. heterophylla* arrived in northern Idaho relatively recently (c. 3000–2000 yr BP). Thus time may have been a limiting factor for *T. heterophylla* to disperse into patchy habitat islands. However, outlying *T. heterophylla* populations exist (Fig. 6), and *T. heterophylla* trees can produce very large seed crops (10^6 seeds ha^-1) that disperse effectively for at least hundreds of metres (Clark, 1970). These factors argue against dispersal as a sole constraint on the modern range of this species. We suggest that late arrival in the interior, in combination with interspecific competition, has limited colonization of potential range. The interior region supports diverse forest communities with a mixture of species typical of the moist (coastal) and dry (interior) regions (Cooper et al., 1991), creating a competitive seedling environment. In addition, fires are frequent at the southern end of the interior distribution (Rollins et al., 2001), and in the high light conditions following fire *T. heterophylla* is outcompeted by diverse shrub and early successional tree species (Mueggler, 1965). Thus frequent fire over the past several thousand years (Brunelle & Whitlock, 2003) could have prevented or slowed *T. heterophylla* colonization by limiting the availability of safe sites (Green, 1983).

The mismatch in potential and occupied distribution of *T. heterophylla* should be viewed in terms of climatic change and local adaptation, which would have resulted in a continually changing potential distribution over time. Thus the size of overprediction also changes over time. The onset of cooler conditions 3100–2500 yr BP (Luckman et al., 1993) may have expanded the potential distribution of *T. heterophylla* south and east in Idaho and Montana, not all of which has been colonized. Alternatively, local adaptation of interior populations may have greatly expanded the predicted distribution (*sensu* Case & Taper, 2000).

In terms of future climatic change, the significance of local adaptation is unclear (Harte et al., 2004). Climatic change will probably occur too rapidly for the emergence of new ecotypes, and will necessitate the reorganization of current ecotypes on the landscape. Given sufficient climatic change, loss of potential distribution may cause the extinction of marginal ecotypes (Davis & Shaw, 2001). This study showed that a large portion of the interior distribution of *T. heterophylla* might be vulnerable to warmer and drier conditions because much of the interior distribution occurs near its AET threshold (Fig. 2). It is possible, however, that increased atmospheric CO2 concentrations will increase water-use efficiency and lower the AET threshold, complicating our ability to predict distributions in future (and past) climates (Williams et al., 2001). Regardless, turnover of dominant forest species in cedar–hemlock forests is likely to be slow (Habeck, 1978). Such anticipated lags in vegetation change suggest that non-climatic controls on distribution limits may become more important in the future.

**ACKNOWLEDGEMENTS**

This study was supported by a National Science Foundation Grant and a Packard Fellowship in Science and Engineering. We thank Elizabeth LaPoint from the USDA Forest Service for the analysis of the Forest Inventory and Analysis plot data, Will MacKenzie from the British Columbia Ministry of Forests for the British Columbia plot data, and Steve Brunsfeld for discussions on forest zones in Idaho. We are grateful for comments from two anonymous referees.

**REFERENCES**


Ministry of Forests, Research Branch, Victoria, BC, Canada.


**SUPPLEMENTARY MATERIAL**

The following supplementary material is available for this article online from http://www.Blackwell-Synergy.com:

**Figure S1** Areas delimited by the thresholds of four bioclimatic variables with respect to the distribution of *Tsuga heterophylla*.

**BIOSKETCHES**

**Daniel Gavin** is an assistant professor in the Department of Geography at the University of Oregon. He is interested in Quaternary palaeoecology, biogeography, fire history and forest ecology.

**Feng Sheng Hu** is an associate professor in the Departments of Plant Biology and Geology at the University of Illinois. He is interested in climatic change and Quaternary ecology.

Editor: Glen MacDonald
Figure S1. Areas delimited by the thresholds of four bioclimatic variables with respect to the distribution of *Tsuga heterophylla*. AET=actual evapotranspiration, MTCO=mean temperature of the coldest month, DEF=climatic water deficit, and GDD5=growing degree-days on a 5°C base. Thresholds were calculated separately for the coastal and interior sub-distributions of *T. heterophylla*. Thresholds were set at the 95th percentile (AET, GDD5, and MTCO) or the 5th percentile (DEF) of the range of values within the *T. heterophylla* distribution.