



# Modeling the occurrence of 15 coniferous tree species throughout the Pacific Northwest of North America using a hybrid approach of a generic process-based growth model and decision tree analysis

Nicholas C. Coops, Richard H. Waring, Clayton Beier, Raphael Roy-Jauvin & Tongli Wang

## Keywords

3-PG model; Climate analysis; Decision tree analysis; Species geographical distribution

## Abbreviations

DEM = Digital Elevation Model; FIA = Forest Inventory and Analysis; LAI<sub>max</sub> = maximum Leaf Area Index; SRTM = Shuttle Radar Topography Mission; VPD = Vapor Pressure Deficits

Received 15 November 2010

Accepted 21 January 2011

Co-ordinating Editor: Geoffrey Henebry

**Coops, N.C.** (corresponding author, nicholas.coops@ubc.ca): Department of Forest Resource Management, 2424 Main Mall, University of British Columbia, Vancouver, Canada V6T 1Z4

**Waring, R.H.** (richard.waring@oregonstate.edu): College of Forestry, Oregon State University, Corvallis, Oregon 97331, USA

**Beier, C.** (clay.beier@gmail.com) & **Roy-Jauvin, R.** (rroyj063@interchange.ubc.ca): Department of Forest Resource Management, 2424 Main Mall. University of British Columbia, Vancouver, Canada V6T 1Z4

**Wang, T.** (tlwang@mail.ubc.ca): Centre for Forest Conservation Genetics, Department of Forest Sciences, 2424 Main Mall. University of British Columbia, Vancouver, Canada V6T 1Z4

## Abstract

**Question:** Can we interpret how climatic variation limits photosynthesis and growth for one widely distributed species, and then relate these responses to model the geographic distributions of other species?

**Location:** The forested region of the Pacific Northwest, United States and Canada.

**Methods:** We first mapped monthly climatic data, averaged for the period 1950 to 1975 at 1 km resolution across the region. The recorded presence and absence of 15 native tree species were next mapped at 1 km resolution from data acquired on 22 771 field survey plots. To establish seasonal limits on photosynthesis and water use, a process-based growth model (3-PG, Physiological Processes to Predict Growth) was parameterized for Douglas-fir (*Pseudotsuga menziesii*), one of the most widely distributed species in the region. Automated decision tree analyses were used to predict the distribution of different species by creating a suite of rules associated with the relative constraints that soil drought, atmospheric humidity deficits, suboptimal and subfreezing temperatures would impose on the growth of Douglas-fir.

**Results:** The 3-PG process-based modeling approach, combined with automated decision tree analyses, predicted presence and absence of 15 conifers on field survey plots with an average accuracy of  $82 \pm 12\%$ . Predictive models of current distribution for each species differed in the number of, order in, and physiological thresholds selected. A deficit in the soil water balance, followed by departures from optimum temperatures in the summer were the two most important variables selected in predicting species distributions.

**Conclusions:** Although empirical models using different sampling techniques and statistical analyses may be more accurate in predicting current distribution of species, the hybrid approach presented in this paper provides a greater mechanistic understanding of the limits to growth and tree distributions. These attributes of process-based models make them particularly useful in designing mitigating strategies to projected changes in climate.

## Introduction

In an earlier paper (Coops et al. 2009a, b), we introduced a hybrid modeling approach that progresses in two steps. First, a process-based growth model (3-PG, Physiological

Principles Predicting Growth) developed by Landsberg & Waring (1997) was used to evaluate the extent that four climatically related variables limit photosynthesis for Douglas-fir (*Pseudotsuga menziesii*), one of the most widely

distributed tree species in western North America. Second, we related the relative limitations on Douglas-fir photosynthesis to field observations of the distribution of other tree species using an automated decision tree analysis (Coops et al. 2009a,b). These species-specific decision tree models predicted presence and absence for six tree species with a weighted accuracy averaging 87% using climatic data acquired over an 18-yr period (1980-1997).

By incorporating a process-based growth model in the first stage of the analyses, we gain several advantages over purely empirical correlations with climatic data. The most obvious advantage of process-based models is that they incorporate non-linear biological responses in a documented manner and provide an explanation for the cause of any reduction or increase in growth. For example, although sub-alpine tree species are well adapted to frequent frosts in the spring, a slight increase in temperature will increase photosynthesis and growth. At the same time, less frequent frost provides an opportunity for outbreaks of needle cast (Woods et al. 2005) and bark beetles (Bentz et al. 2010).

As a result of such analyses, we can appreciate why more favorable growth conditions for sub-alpine species might still lead to their eventual replacement by more temperate tree species. Similarly, because process-based models calculate a soil water balance that is linked to the canopy leaf area, they explain why an open ponderosa pine forest may experience less constraint on photosynthesis and transpiration during a summer drought than denser stands growing in areas receiving significantly more precipitation (Runyon et al. 1994; Law et al. 2001).

Physiological models such as 3-PG grow stands of trees at rates commensurate with both climate and soil conditions. By assuming uniform soil conditions, climatically induced changes in site growth potential can be 3-PG spatially projected estimates of maximum Leaf Area Index ( $LAI_{max}$ ) are generally in good agreement across the Pacific Northwest Region with values derived with satellite-borne remote sensing (Coops et al. 2001). The LAI imposes limits on light absorbed by photosynthetic pigments as well as on transpiration. Thus a canopy with an LAI of 1.0 absorbs only about 60% of the incident radiation (approximately 50% photosynthetically active) and transpires less than  $1 \text{ mm d}^{-1}$  (Angell & Miller 1994) whereas a forest canopy with an LAI above 6.0 absorbs > 95% of incident radiation and transpires up to  $4 \text{ mm d}^{-1}$  (Waring et al. 2008).

In this paper we increase the number of Pacific Northwest conifers analysed from six to fifteen, and expand the area from the USA to include the Canadian province of British Columbia. To accommodate the increased spatial and taxonomic scope of the analysis, we increased the reference base of permanent survey plots by sixfold.

Further, we refined the climatic analysis to permit a more detailed interpretation of the relative importance of seasonal variation in four climatically related variables that impose non-linear constraints on photosynthesis: (1) soil drought, (2) atmospheric humidity deficits, (3) suboptimal temperature and (4) the frequency of frost. By expanding the analysis, we hoped to extend the approach without making the analysis more complicated.

## Methods

### Description of study area

Within the Pacific Northwest Region of North America, the distribution of flora is largely correlated with spatial variations in temperature and precipitation. The maritime influence of the Pacific Ocean provides conditions that are generally favorable for tree growth and the region supports some of the most productive forest in North America (Whittaker 1960). Franklin & Dyrness (1973) provided detailed descriptions of plant communities and their associated climates. The extent that such historical descriptions offer a basis for predicting future floristic composition under a changing climate is unknown. It is unlikely that they can account for the differential responses expected by species that evolved in different geological epochs and are known to form new combinations as environmental conditions change (Raven & Axelrod 1974; Williams et al. 2007).

### Climatic data

Mean monthly climate spatial surfaces were generated using ClimateWNA, which downscals precipitation and temperature data generated at 2 to 4 km by PRISM (Parameter-elevation Regressions on Independent Slopes Model, Daly et al. 2002) to 1 km. The downscaling is achieved through a combination of bilinear interpolation and elevation adjustment (see Wang et al. 2006). To provide the required elevation data for ClimateWNA at 1 km a 90 m Digital Elevation Model (DEM) was resampled from the Shuttle Radar Topography Mission (SRTM). Mean monthly atmospheric vapor pressure deficits (VPD) for daylight periods were estimated by assuming that the water vapor concentration throughout the day was equivalent to that held at saturation for the average monthly minimum temperature (Kimball et al. 1997). The number of days per month with subfreezing temperatures (less than  $-2^{\circ}\text{C}$ ) was estimated from empirical equations with mean minimum temperature (Coops et al. 1998).

Monthly estimates of total incoming short-wave radiation were calculated following a modeling approach detailed by Coops et al. (2000) that first calculates the potential radiation at the top of the atmosphere then adjusts for slope, aspect and elevation (Garnier & Ohmura

1968; Swift 1976), and finally for variation in water vapor and the effects of clouds on the fraction of diffuse to direct beam incoming radiation (Running et al. 1987) based on a previously published relationship with the difference between mean daily maximum and minimum temperatures and latitude (Coops et al. 2000). The latter conversion takes advantage of a correlation between monthly mean temperature extremes and the transmissivity of the atmosphere (Bristow & Campbell 1984). The modeling approach, when compared with direct measurements, predicted both the direct and diffuse components of mean monthly incoming radiation with 93 to 99% accuracy on flat surfaces, and on sloping terrain accounted for > 87% of the observed variation with a mean error less than  $2 \text{ MJ} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$  (Coops et al. 2000).

### Species occurrence data

Across the Pacific Northwest, tree species presence/absence data were assembled from a number of sources. For British Columbia, tree species were taken from the centroids of stand-level polygons located in protected forested areas and from vegetation resource inventory plots collected across all forest lands using a three-phase, photo- and ground-based sampling design (Schroeder et al. 2010). The spatial accuracy of the plot coordinates was estimated at roughly  $\pm 500 \text{ m}$ .

In the USA, tree species data were acquired from US Forest Service, Forest Inventory and Analysis (FIA) surveys. Plot data are recorded on a permanent sampling grid established across the conterminous USA at a density of approximately one plot per 2400 ha (Bechtold & Patterson 2005). Only trees with DBH > 2.54 cm were considered in the sample. As actual FIA plot locations are confidential,

we used the publically available coordinates, which have similar spatial accuracy as the Canadian data (i.e.  $\pm 500 \text{ m}$ ). Additional information on the FIA data sets is available in Schroeder et al. (2009). The presence/absence data from both Canada and the USA were combined into one database with a total of 22 771 plots (Table 1).

### Description of the 3-PG model

The 3-PG model contains a number of simplifying assumptions that have emerged from studies conducted over a wide range of forest types and include the use of monthly climate data (rather than daily or annual) with little loss in the accuracy of model predictions. Each month, the most limiting climatic variable on photosynthesis is selected, based on departure from conditions that are defined as optimum (expressed as unity) or completely limited (expressed as zero) for a particular species or genotype. The ratio of actual/potential photosynthesis decreases in proportion to the reduction in the most limiting environmental factor. The fraction of production not allocated to roots is partitioned among foliage, stem and branches based on allometric relationships and knowledge of annual leaf turnover (Landsberg et al. 2003).

The basic model works as follows: absorbed photosynthetically active radiation (APAR) is estimated from global solar radiation and LAI, and the portion utilized is calculated by reducing APAR by an amount determined by a series of modifiers that take values between 0 (system 'shutdown') and one (no constraint) to limit gas exchange via canopy stomatal conductance (Landsberg & Waring 1997). The monthly modifiers include: averaged daytime vapor pressure deficits (VPD), the frequency of subfreezing conditions, soil drought and mean daily temperature.

**Table 1.** Scientific and common names, and proportion of plots with presence and absence of the species across the Pacific Northwest region defined as the province of British Columbia and the USA: Washington, Oregon, California, Idaho, Montana, Nevada and Utah.

Species	Common name	Presence	Absence	% Present	%Absent	Total (n)
<i>Pseudotsuga menziesii</i>	Douglas-fir	6643	16 128	29.2	70.8	22 771
<i>Thuja plicata</i>	Western redcedar	3107	19 664	13.6	86.4	22 771
<i>Tsuga heterophylla</i>	Western hemlock	2106	20 665	9.3	90.8	22 771
<i>Tsuga mertensiana</i>	Mountain hemlock	473	22 298	2.1	97.9	22 771
<i>Pinus contorta</i>	Lodgepole pine	4742	18 029	20.8	79.2	22 771
<i>Pinus ponderosa</i>	Ponderosa pine	2393	20 378	10.5	89.5	22 771
<i>Picea sitchensis</i>	Sitka spruce	364	22 407	1.6	98.4	22 771
<i>Pinus albicaulis</i>	Whitebark pine	644	22 127	2.8	97.2	22 771
<i>Chamaecyparis nootkatensis</i>	Alaska yellow cedar	887	21 884	3.9	96.1	22 771
<i>Abies procera</i>	noble fir	82	22 689	0.4	99.6	22 771
<i>Picea engelmannii</i>	Engelmann spruce	1365	21 406	6.0	94	22 771
<i>Larix occidentalis</i>	Western larch	670	22 101	2.9	97.1	22 771
<i>Abies lasiocarpa</i>	Subalpine fir	2369	20 402	10.4	89.6	22 771
<i>Abies amabilis</i>	Pacific silver fir	764	22 007	3.4	96.6	22 771
<i>Abies grandis</i>	Grand fir	598	22 173	2.6	97.4	22 771

Drought limitations are imposed as a function of soil properties and a simple water balance that calculates when soil water supply is less than transpiration, estimated with the Penman–Monteith equation. In the current format, we did not separate precipitation as rain from that as snow because such a separation is best achieved at daily rather than monthly time steps (Coughlan & Running 1997).

A major simplification in the 3-PG model is that it does not require detailed calculation of autotrophic respiration, assuming that it is a fixed fraction (0.47, SE  $\pm$  0.04) of gross photosynthesis (Landsberg & Waring 1997; Waring et al. 1998; Law et al. 2001).

We further simplified the approach by selecting Douglas-fir, the most widely distributed species in the region, to characterize the importance of climatic constraints on photosynthesis and growth across all forested environments, as we have done previously for other purposes (Swenson et al. 2005; Waring et al. 2005; Coops et al. 2007, 2010).

We make the implicit assumption that a species presence or absence at a given site is a function of integrated physiological responses to climatic variation that cause relative differences in the growth rates of competing species. Although we recognize that physical limitations, such as the accumulation of snow or those of chemical origin, such as salt-spray or pollutants, can also limit photosynthesis and growth, our modeling excludes these interactions in this paper. We do not assume that Douglas-fir physiological tolerances exactly match those of other species. Instead, once we characterize geographically the relative important of seasonal climatic constraints on Douglas-fir photosynthesis, we translate these limitations for other species in reference to how they depart from conditions favorable for Douglas-fir. The alternative would be to parameterize the 3-PG model for each species, a task for which the required data are generally lacking, except for a few widely distributed species such as ponderosa (Coops et al. 2005) and lodgepole pine (Coops & Waring 2010).

The parameter values for Douglas-fir matched those reported in Waring & McDowell (2002), with a few exceptions. To limit the analysis to climatic effects, we set the available water-holding capacity at 200 mm for a sandy loam soil throughout the region, which through sensitivity analysis is a value that will capture the effect of seasonal drought (Nightingale et al. 2007).

To limit the analysis to climatic variation, we assigned a constant soil fertility rank of 50% of maximum, which results in an even partitioning of growth above and below ground. Previous research indicates that the error of using constant values for soil fertility is mainly significant in areas where soils are extremely infertile or toxic (e.g. serpentine-derived soils) (Coops & Waring 2001). We set the maximum

quantum efficiency equal to  $0.04 \text{ mol} \cdot \text{C} \cdot \text{mol} \cdot \text{photon}^{-1}$ , equivalent to  $2.2 \text{ g} \cdot \text{C} \cdot \text{MJ}^{-1}$  absorbed photosynthetically active radiation (PAR), about mid-way between reported minimum and maximum values in the literature (Landsberg et al. 2003). To take into account seasonal adjustments in temperatures responses (Hember et al. 2010), we set the minimum, optimum, and maximum temperatures at  $-7^\circ\text{C}$ ,  $18^\circ\text{C}$  and  $40^\circ\text{C}$ , respectively. The photosynthetic response at temperatures less than  $-2^\circ\text{C}$  was truncated to zero, because below that threshold stomata are closed (Running et al. 1975; Hadley 2000). The extent that other species encounter environments that would impose different restrictions than those on Douglas-fir is incorporated through the second component of the modeling approach using an automated decision tree analysis, which is described in more detail below.

### Decision tree modeling for individual species

We applied the 3-PG model across the region using the spatial climate layers to predict coverages of stand growth and LAI, using the mean climate derived from ClimateW-NA from 1950 to 1975. Model simulations were run for 50 yr of stand development by which time stands have obtained maximum LAI and maximum canopy closure. The 3PG model was then stopped and at each of the 22 771 plots the monthly climatically-restricting modifiers to photosynthesis were extracted for each of the four climatic modifiers ( $12 \times 4 = 48$ ). We derived seasonal averages and an annual extreme from these monthly climatic modifiers for modeling.

To assess the extent that the 3-PG physiological modifiers might serve to predict presence or absence of each of 15 selected tree species, a decision tree analysis, similar to that developed by Coops et al. (2009a, b) was applied to predict presence and absence for each of the 15 species, based on the maximum constraints that each of the four climate modifiers imposes on photosynthesis throughout the four seasons, as well as the maximum constraint throughout the entire year.

Decision tree analysis is increasingly selected in ecological research because of a number of features, including an ability to deal with collinear datasets, to exclude insignificant variables, and to allow for asymmetrical distribution of samples (De'ath 2002; Melendez et al. 2006; Schwalm et al. 2006). The technique automatically separates the dependent variables (presence or absence of a tree species) into a series of choices that identifies the importance of each constraining variable.

Decision Tree Regression (DTREG, Sherrod 2010) software was used to develop a classification tree for each species using a tenfold cross validation technique where the total dataset is partitioned randomly into ten equally

sized groups, a model is developed on nine of the groups, and then tested against the remaining 10% of the data not used in model development. This process, also known as k-fold partitioning, is repeated 10 times with the results merged to produce a final classification tree with an overall accuracy accessed by averaging the independent results of the 10 simulations (Breiman et al. 1984). This type of K-fold partitioning provides a realistic test of model accuracy. It is not a true independent validation of the model, however, as an independent dataset of species occurrence was not available over the entire area (see Beale et al. 2008 for a discussion of potential biases).

Models were developed with a maximum of five (decision) branches, and a misclassification weighting of 1.5, which constrained the model development and ensured the decisions trees did not over-fit the dataset. Once each model was created a 'confusion matrix' was developed, which provides an indication of the positive and negative predictive power of the model as well as a number of other statistics (Fielding & Bell 1997). Forbes (1995) provides some basic criteria of what measures could be extracted from confusion matrices to establish the accuracy of distribution models.

The Kappa ( $\kappa$ ) statistic has been recommended as a metric to assess the accuracy of the developed models and is useful particularly in ecological research (see review by Monserud & Leemans 1992). This statistic calculates the proportion of specific agreement across categories (classes) and meets most of Forbes's criteria. If the prediction is in complete agreement with observed, then  $\kappa = 1$ ; if there is no agreement (other than what would be expected by chance) then  $\kappa = 0$ . In cases where one class size significantly exceeds the other (as is often the case in vegetation distribution modelling) the maximum  $\kappa$  achieved for perfect agreement will be less than one, when this happens, the Kappa statistic can be rescaled based on the observed marginal frequencies (Vach 2005; Ben-David 2008). We generated maps predicting each species' distribution and compared them with recorded presence on survey plots as well as with published range maps (Critchfield & Little 1966; Little 1971) (URL: <http://esp.cr.usgs.gov/data/atlas/little/>, last accessed 11.01.2011).

## Results

Spatial variation in the four climatic modifiers in the season that they most constrain photosynthesis of Douglas-fir is presented in Fig. 1. Late summer drought is typical throughout most of the interior of the region when the maximum available water supply is set at 200 mm (Fig. 1a). Severe water limitations on photosynthesis therefore result in some months, particularly for parts of Washington State and the interior of British Columbia. Along coastal British Columbia and in the

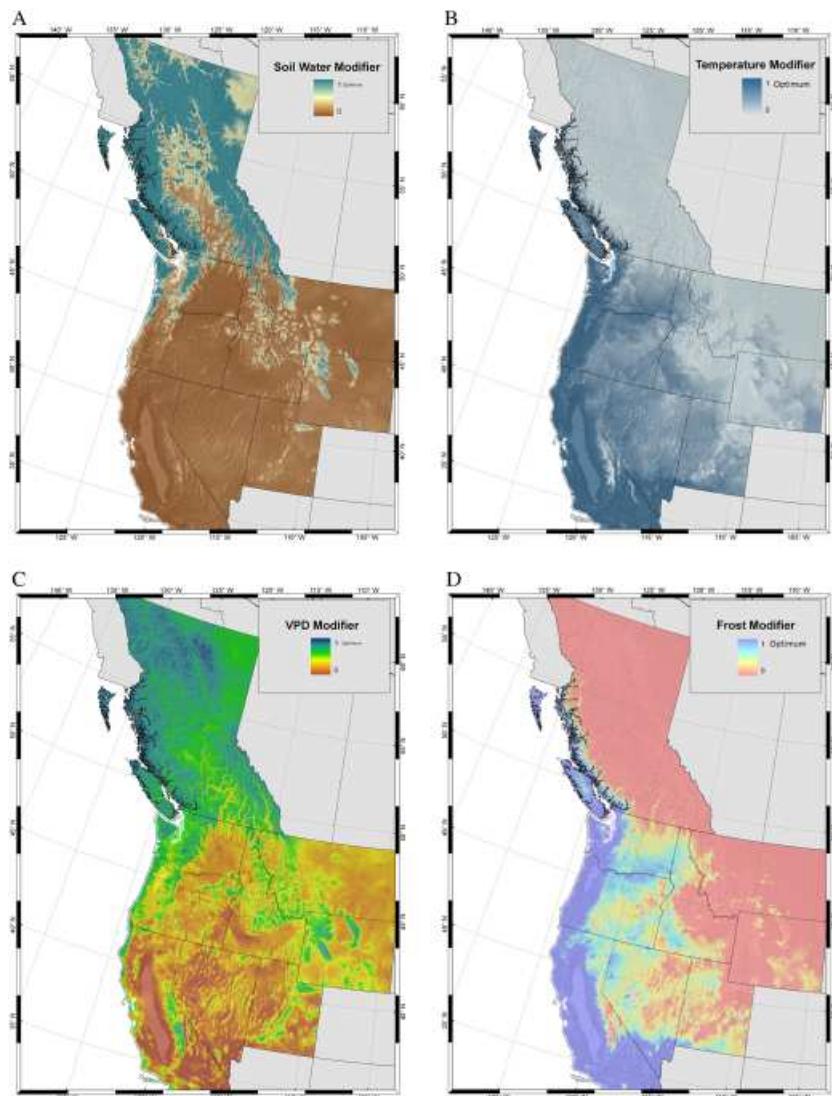
Canadian Rocky Mountains, there is little water stress because ample precipitation fully recharges the soil profile and transpiration does not exhaust this reservoir during the growing season based on simulated values of LAI.

High evaporative demand during the summer is typical throughout the central valley in California, and for much of the area on the eastern sides of the Cascade and Sierra Mountains. Mountainous areas toward the interior and northern portions of the region remain cool enough to minimize evaporative demand throughout the year. As Fig. 1b shows, the stress imposed by high evaporative demand never reduces seasonal (or monthly, not shown) photosynthesis to 0. Deviations from optimum temperature (Fig. 1c) and limitations imposed by frequent frost (Fig. 1d) show comparable patterns to one another, with the area most unfavorable for Douglas-fir located at high latitudes and to the east of the Cascades and Sierra mountain ranges. The coastal mountains of the Pacific Northwest are buffered from extremes in temperature, whereas diurnal variation increases with elevation and with movement inland.

Based on presence data recorded on plots, we selected a number of species to illustrate the range in seasonal variation in the climatic modifiers (Fig. 2). Sitka spruce (*Picea sitchensis*) (Fig. 2a), is currently restricted to a narrow band along the coast in areas where precipitation is sufficient to maintain a soil water balance within 10% of the Douglas-fir optimum throughout the year and where temperature (both frost frequency and minimum temperature) are moderate (with only a reduction of about 50% in reference to optimum conditions for Douglas-fir). Because of the maritime climate along the coast, Sitka spruce in its current range is less limited by evaporative demand than any of the other species reported on.

Lodgepole pine (*Pinus contorta*) (Fig. 2b) is well adapted to low temperatures in the winter, spring and autumn, which severely restrict photosynthesis in its range by > 80%. In summer, evaporative demand, and soil water stresses reduce growth by up to 30% with soil water limitations in fall imposing limits of up to 40% on potential photosynthesis. Western hemlock (*Tsuga heterophylla*) covers a large range through the Pacific Northwest, extending from the coast of Alaska to northern California and eastward into Idaho. Within its range, soil moisture is generally not limiting (Fig. 2c). Temperatures too are close to optimum for Douglas-fir, except during winter (Fig. 2c). Similarly, evaporative demand creates relatively little constraint throughout the year, remaining within 40% of optimum for Douglas-fir.

In contrast to western hemlock, western larch (*Larix occidentalis*), the only deciduous conifer modeled, is principally distributed inland across a range of elevations from 500 m in the north to 2400 m at the southern extremes of its range. Figure 2d indicates its general tolerance to



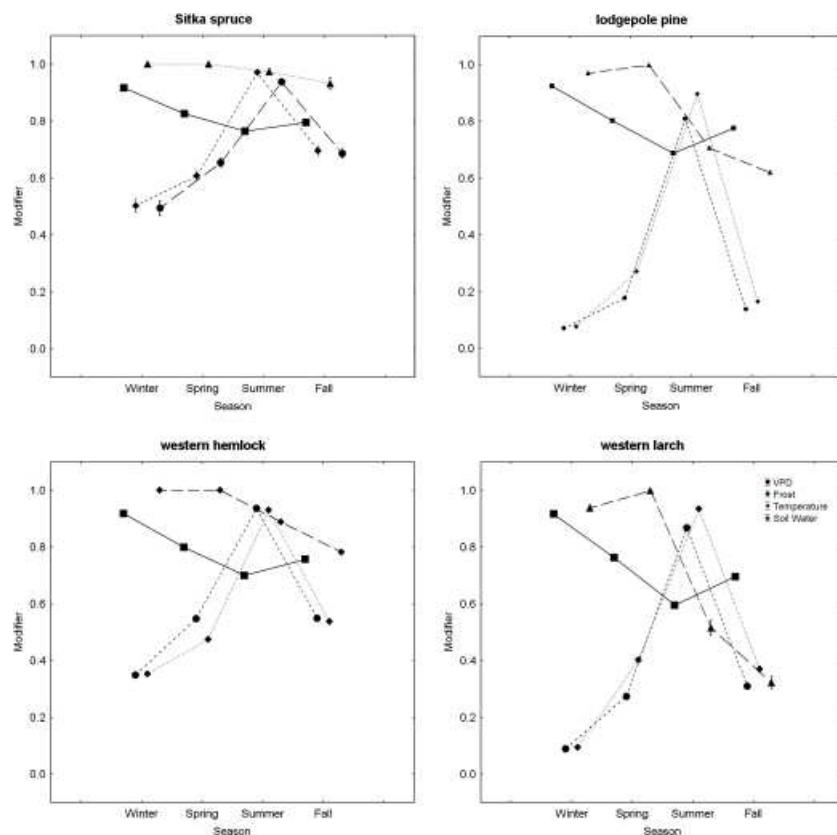
**Fig. 1. (a)-(d).** Geographic variation in the maximum limitation on photosynthesis of the four climatic modifiers throughout the year, referenced to responses by Douglas-fir stands at 50 yr of age. Modifiers are scaled between 0 (complete restriction in photosynthesis) to 1 (optimum).

moderate soil water stress, which decreases to less than 50% of that characterizing Douglas-fir distribution throughout summer and fall. In the winter, western larch in the winter is exposed to low temperatures that would greatly limit photosynthesis, even if leaves were present. With regard to evaporative demand, western larch rarely is subjected to < 60% of optimum for photosynthesis in reference to Douglas-fir.

Although all four modifiers were used in the decision tree analysis, their importance differs. For the species discussed above, the distribution of Sitka spruce was defined with the fewest number of modifiers, with the model principally responding to modest if any limits on soil water availability and lack of frost in the autumn. In

the case of lodgepole pine, five modifiers were required, including frost in the summer and spring, and soil water status in summer, winter and annually. For western hemlock, soil water in the autumn and winter, as well as frost in spring, defined its distribution and for western larch, all four of the physiological restrictions were represented with a near-equal four-way split between frost and temperature in autumn, VPD in summer and soil water status in spring.

Figure 3 summarizes how often, and to what extent, individual modifiers incorporated into 15 specific models accounted for variations in their distributions across the Pacific Northwest. In the winter, frost, temperature and VPD have little significance because there is little light



**Fig. 2. (a)-(d).** Seasonal mean variation in the four climatic modifiers for selected tree species (Sitka spruce, lodgepole pine, western hemlock and western larch).

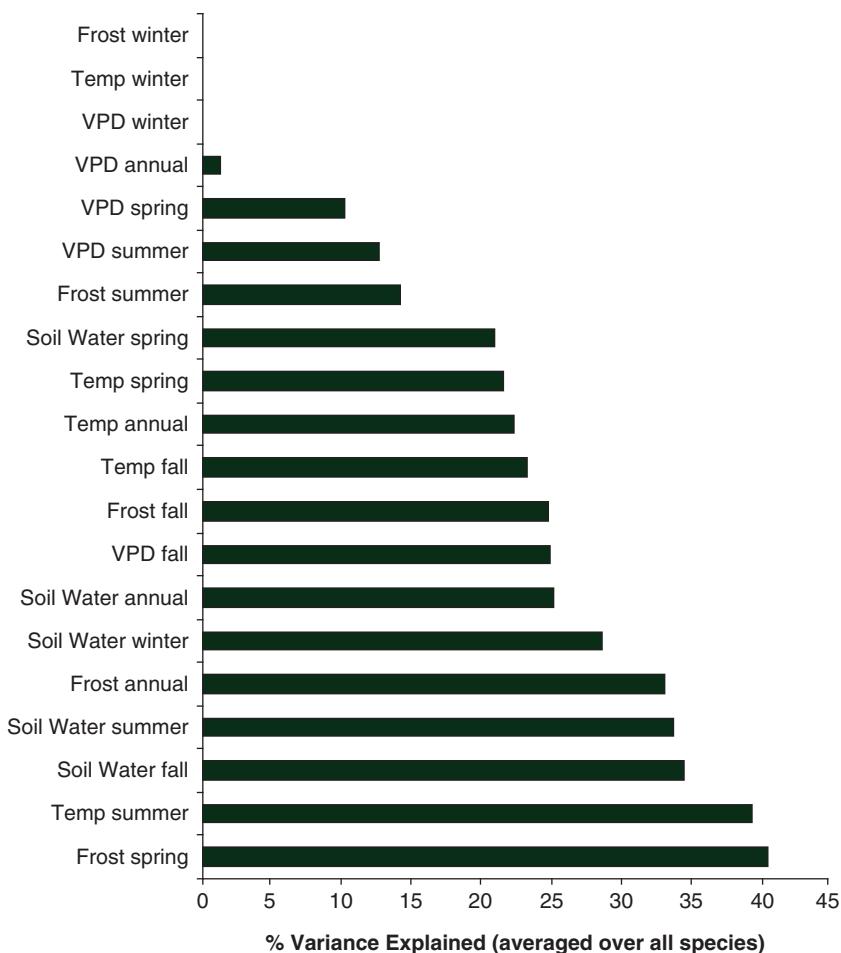
available for photosynthesis. In other seasons, sub-freezing temperatures in the spring and sub-optimal temperatures in the summer accounted, on average, for approximately 35% of the variance in species distributions across the region. The availability of soil water in the summer and autumn were the next most common variables included in the 15 specific-specific models. The influence of VPD on species distributions in the autumn ranked eighth in overall importance but still accounted, on average, for about 25% of the observed variance in predicting species distributions.

Accuracy assessments of the models are referenced to the percentage of plots on which a species was correctly assigned as being present or absent, and then combined into a weighted value, proportional to the number of plots associated with each of the two categories. The value for each model is also provided. The overall accuracy of the models averaged 82% (Table 2). The most accurate model with respect to modeling its presence was noble fir (*Abies procera*) (92%) with a value of 0.92 indicating excellent agreement with the plot observations. In contrast, the poorest prediction of species presence was for lodgepole pine (68%) with a of 0.50, indicating only moderate

agreement. The most accurate model with respect to modeling absences was again noble fir (93%), with the poorest predictions for Engelmann spruce (*Picea engelmannii*) (68%). Overall the presence of a species was predicted with slightly more accuracy than its absence (82% vs 81%, respectively). Overall was 0.74.

The location of the presence/absence survey data plots, and the Little (1971) range maps are shown in Fig. 4. In the case of the Little (1971) maps, the ranges are independent of current land-use, whereas the FIA and British Columbia plots in this study are located only in areas recognized as forest. As a result, general range maps often include areas where a species is absent, particularly with respect to extremes in elevation (Rehfeldt et al. 2006). Nonetheless, they are instructive for comparisons with field surveys and model predictions.

The species range predictions for Sitka spruce indicate that it is present along the Oregon, Washington and British Columbian coast as well as on Vancouver Island and Haida Gwai. These results are in close agreement with the range maps and the presence/absence data, even matching small isolated populations along the coast. Because of its adaption to low temperatures, lodgepole

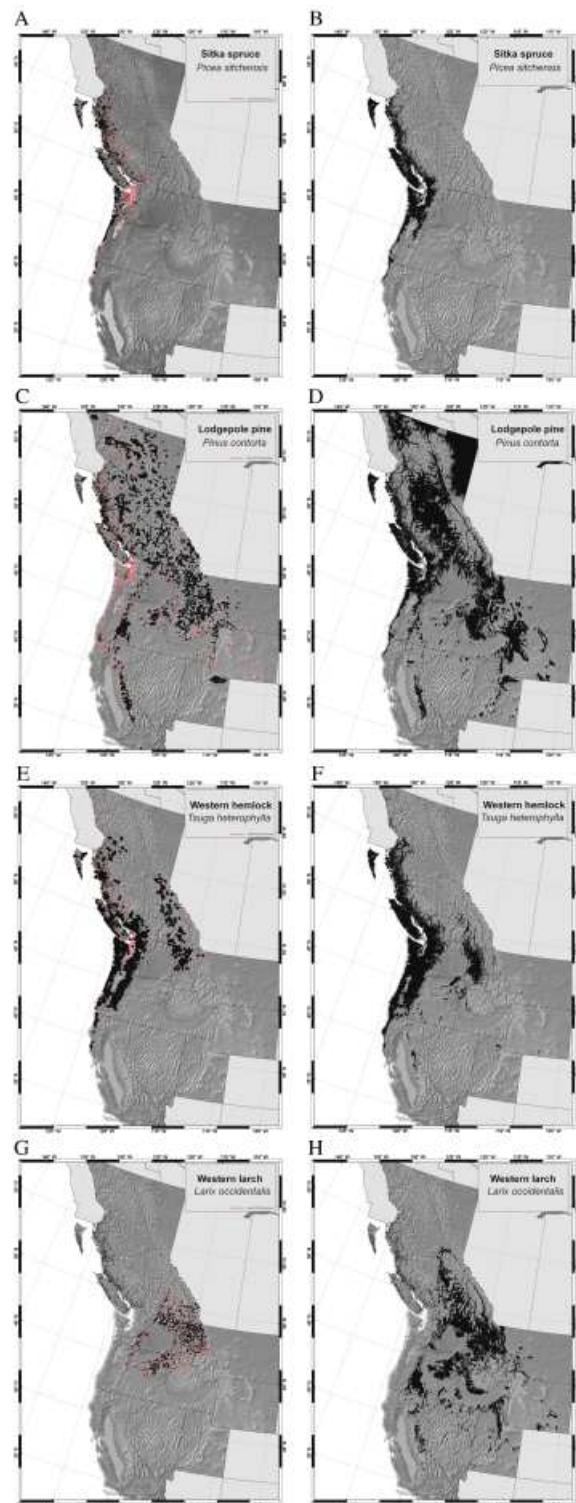


**Fig. 3.** Summary of the important modifiers for all 15 species modeled in the Pacific Northwest.

**Table 2.** Per cent correct classification for classification trees for 15 selected species. \*Landis & Koch (1977) have suggested the following ranges of agreement for the kappa ( $\kappa$ ) statistic: < 0 poor; 0–0.21 slight; 0.21–0.4 fair; 0.4–0.6 moderate; 0.61–0.8 substantial; 0.81–1.0 excellent.

Species	Presence Accuracy (%)	Absence Accuracy (%)	Overall average (%)	$\kappa$ (*)
Douglas-fir	76	70	73	0.584
Lodgepole pine	68	72	70	0.495
Ponderosa pine	84	82	83	0.785
Western red cedar	80	77	78	0.711
Mountain hemlock	78	78	78	0.711
Sitka spruce	90	88	89	0.884
Pacific silver fir	85	83	84	0.821
Whitebark pine	84	82	83	0.797
Englemann spruce	78	68	73	0.574
Western hemlock	81	78	80	0.740
Alaska yellow cedar	89	91	90	0.879
Noble fir	93	93	93	0.921
Western larch	81	81	81	0.754
Subalpine fir	76	75	76	0.654
grand fir	80	83	82	0.805

pine currently has an extensive range in northwestern North America from the Yukon, south into California. The predictions agree very well with both the plot locations and Little's (1971) range map for the species. The correspondence for western hemlock is also good with some obvious departures between the field survey plots, Little's range map, and the modeled distribution. The species is predicted to grow further south than is recorded by Little (1971) and to some extent as reported in the field records. The western hemlock predictions extend the species' range to the interior, and the coastal variants, joining in central British Columbia whereas both the plot locations and Little's (1971) map recognize two distinct groupings. Of the four species, predictions for western larch are the most divergent from other sources. The distribution of western larch from the Little's (1971) range map and the survey locations indicates the species occurs principally in the USA with some populations in interior valleys of British Columbia. The decision tree predictions extend the species' range to a larger area than where it has been



**Fig. 4. (a)-(h).** Maps of predicted species occurrence in reference to presence data recorded on survey plots (●), and to more general range distributions (in red) from Little (1971) (left), and distributions predicted using decision-tree rules (right).

recorded in the Klamath Mountains of southwestern Oregon and northwestern California. The species is also predicted to occur slightly further north in central British Columbia than recorded on the field plots.

## Discussion

The dependence on a widely distributed species, such as Douglas-fir, to provide a physiological benchmark for other coniferous species seems justified, based on the generally good accuracy of the decision tree models in predicting a species' presence or absence on field plots. We recognize that it is important to confirm consistency in the order that variables are selected, and in the thresholds defined. In a previous study, Coops et al. (2009a, b) used only annual extremes as physiological constraints. We found that the distributions predicted for three species in common with this paper's analysis were well matched, and that the same modifiers, in most cases, were selected in the decision tree models.

We recognize that better agreement between predicted and observed distributions of species can be obtained with an increase in the number branches, and model runs, as demonstrated by Rehfeldt et al. (2006). To apply these more statistically sophisticated techniques, however, compromises the ability to define physiological thresholds useful to ecologists and managers. For example, Rehfeldt et al. (2006), by applying Random Forest Regression Tree Analyses, started with a combination of 21 climatic indices to predict the current distributions of 25 plant communities. Their most commonly selected climatic index was the ratio of summer to annual precipitation, followed by the summer–winter temperature differential, the mean annual temperature and mean annual precipitation. Rehfeldt & Jaquish (2010) further refined the approach to increase the predictive power of where genotypes of western larch occur. This was accomplished by double-weighting the climatic information where larch was recorded (on 2.5% of the total plots) and by selecting 40% of such plots to characterize the environment in each sample taken to construct hundreds of decision trees.

There are clear benefits to obtaining a more balanced sample between plots with and without a species present. In the case of noble fir the proportion of plots with the species present was < 1% of the total dataset, reducing confidence in the analysis. To make more accurate estimate of climatically important physiological thresholds it would be desirable to employ methods particularly designed for predicting the occurrence of rare species (Marcot et al. 2001).

We believe that the hybrid approach presented in this paper offer several advantages over existing, more correlative approaches. By starting a climatic analysis with a model that creates leaves and roots, as well as stems, we

obtain mechanistic insights in regard to the seasonal importance of different variables in different geographic areas. This permits one to take into account latitudinal variation in the incident radiation, and its contribution to photosynthesis and water vapor transfer, dependent on the modeled LAI and stomatal conductance at the time of canopy closure (assumed by age 50 yr).

In addition, the physiological modifiers extracted from process model runs for a series of years provide insights as to the season for which they are important, and trends that might emerge from an analysis of interannual or longer-term variation in climatic conditions. Managers, challenged with perpetuating the presence of one species over another, might consider the implications of altering the energy balance and potential damage from spring frost by maintaining a denser rather than a more open canopy (Aussenac 2000). Alternatively, the effects of increasing drought might be reduced through intensive thinning to lower LAI well below its potential (Simonin et al. 2007).

It may be desirable to include additional variables to those now considered in most process-based models. We recognize that in areas where temperatures fall well below freezing that snow and ice can severely limit some species but not others (Waring 1969). The 3-PG model indirectly identifies areas with heavy snowpack as sites with many subfreezing days and with more precipitation each month than is lost through evaporation and transpiration. With the present configuration of the 3-PG model, sites with heavy snowpack rarely experience drought, high evaporative demand or winter temperatures favorable for growth of Douglas-fir. More subtle effects of changes in the duration of a snowpack are missed. Early disappearance of snow cover may expose soils to subfreezing temperatures in the spring that can kill small-diameter roots (Tierney et al. 2001). Similarly, salt spray typical of coastal areas could also be included in more comprehensive models along with ozone, which is particularly harmful when stomata are open under modest humidity deficits (Panek et al. 2002). Rising levels of atmospheric CO<sub>2</sub> concentrations are also important and have been included in many process-based models (Cramer et al. 2001).

By predicting a species' presence in a particular environment, we gain a measure of its competitive range under specified climatic conditions, but not the extent that it might dominate. To assess the latter, the frequency and kind of disturbance would be necessary to quantify. Also, to incorporate biological interaction with insects and pathogens would require more details regarding the age and spacing of individual trees (Coops et al. 2009a, b), along with the population dynamics and distributions of insects and pathogens (Seidl et al. 2007; Raffa et al. 2008). Nevertheless, if climatic conditions were to shift gradually to those less favorable for a native species, we would expect

the hybrid approach used in this paper to be able to predict where constrictions and expansions in a species' range should occur, which are attributes that can be measured.

Overall, our objectives were met. We gained more insight with a seasonal analysis than from annual extremes, and were able to predict the distributions of 15 species with similar accuracies to those attained previously for six, only three of which were in common. As a result, we conclude that the hybrid approach, which links physiological growth modifiers with a decision tree analysis, has many advantages over more purely empirical methods, particularly when predicting the distribution of species in the future. We can and should take advantage of advances in sampling design and continue to compare predictions made with a variety of approaches. Where similar trends are predicted, and the significant variables identified are closely related, our power of inference will be greatly increased over that provided by any single modeling approach.

## Acknowledgements

This study was supported by the National Aeronautics and Space Administration (NASA Grant NNX09AR59G) to Waring as part of the Biodiversity and Ecological Forecasting program and a Canadian NSERC Discovery grant to N.C.C. We are very grateful for the helpful and insightful comments of the two anonymous reviewers and the associate editor. We also thank Prof. Andreas Hamann and Dr Todd Schroeder (USFS) for access to species datasets in British Columbia and the United States. More detailed information on the model outputs for each species are available at: <http://www.pnwspecieschange.info/>.

## References

- Angell, R.F. & Miller, R.F. 1994. Simulation of leaf conductance and transpiration in *Juniper occidentalis*. *Forest Science* 40: 5–17.
- Aussenac, G. 2000. Interactions between forest stands and microclimate: ecophysiological aspects and consequences for silviculture. *Annales des Sciences Forestieres* 57: 287–301.
- Beale, C.M., Lennon, J.J. & Gimona, A. 2008. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences of the United States of America* 105: 14908–14912.
- Bechtold, W.A. & Patterson, P.L. 2005. *The enhanced forest inventory and analysis national sample design and estimation procedures*. Southern Research station, SRS-80. United States Department of Agriculture, Forest Service, Asheville, NC, US.
- Ben-David, A. 2008. Comparison of classification accuracy using Cohen's weighted kappa. *Expert Systems with Applications* 34: 825–832.
- Bentz, B.J., Régnière, J., Fettig, C.J., Hansen, E.M., Hayes, J.L., Hicke, J.A., Kelsey, R.G., Negrón, J.F. & Seybold, S.J. 2010. Climate change and bark beetles of the western United

- States and Canada: direct and indirect effects. *BioScience* 60: 602–613.
- Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.G. 1984. *Classification and regression trees*. Wadsworth International Group, Belmont, CA, US.
- Bristow, K.L. & Campbell, G.S. 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agricultural & Forest Meteorology* 31: 159–166.
- Coops, N.C. & Waring, R.H. 2001. Assessing forest growth across southwestern Oregon under a range of current and future global change scenarios using a process model, 3-PG. *Global Change Biology* 7: 15–29.
- Coops, N.C. & Waring, R.H. 2010. A process-based approach to estimate lodgepole pine (*Pinus contorta* Dougl.) distribution in the Pacific Northwest under climate change. *Climatic Change* doi: 10.1007/s10584-010-9861-2.
- Coops, N.C., Waring, R.H. & Landsberg, J.J. 1998. Assessing forest productivity in Australia and New Zealand using a physiologically-based model driven with averaged monthly weather data and satellite derived estimates of canopy photosynthetic capacity. *Forest Ecology and Management* 104: 113–127.
- Coops, N.C., Waring, R.H. & Moncrieth, J. 2000. Estimating mean monthly incident solar radiation on horizontal and inclined slopes from mean monthly temperatures extremes. *International Journal of Biometeorology* 44: 204–211.
- Coops, N.C., Waring, R.H. & Landsberg, J.J. 2001. Estimation of potential forest productivity across the Oregon transect using satellite data and monthly weather records. *International Journal of Remote Sensing* 22: 3797–3812.
- Coops, N.C., Waring, R.H. & Law, B.E. 2005. Predicting the past and future productivity of ponderosa pine in the Pacific Northwest using a process model, 3-PG. *Ecological Modelling* 183: 107–124.
- Coops, N.C., Coggins, S.B. & Kurz, W.A. 2007. Mapping the environmental limitations to growth of coastal Douglas-fir stands on Vancouver Island, British Columbia. *Tree Physiology* 27: 805–815.
- Coops, N.C., Waring, R.H. & Schroeder, T.A. 2009a. Combining a generic process-productivity model and a statistical classification method to predict presence and absence of tree species in the Pacific Northwest, USA. *Ecological Modelling* 220: 1787–1796.
- Coops, N.C., Waring, R.H., Wulder, M.A. & White, J.C. 2009b. Prediction and assessment of bark beetle-induced mortality of lodgepole pine using estimates of stand vigor derived from remotely sensed data. *Remote Sensing of Environment* 12: 1058–1066.
- Coops, N.C., Hember, R.A. & Waring, R.H. 2010. Assessing the impact of current and projected climates on Douglas-fir productivity in British Columbia, Canada. *Canadian Journal of Forest Research* 40: 511–524.
- Coops, N.C., Gaulton, R. & Waring, R.H. 2011. Mapping site indices for five Pacific Northwest conifers using a physiologically based model. *Applied Vegetation Science* DOI: 10.1111/j.1654-109X.2010.01109.x.
- Coughlan, J.C. & Running, S.W. 1997. Regional ecosystem simulations: a general model for simulating snow accumulation and melt in mountainous terrain. *Landscape Ecology* 12: 119–136.
- Cramer, W., Bondeau, A., Woodward, F.I., Prentice, I.C., Betts, R.A., Brovkin, V., Cox, P.M., Fisher, V., Jonathan, A., Foley, J.A., Friend, A.D., Kucharik, C., Lomas, M.R., Ramankutty, N., Sitch, S., Smith, B., White, A. & Young-Molling, C. 2001. Global response of terrestrial ecosystem structure and function to CO<sub>2</sub> and climate change: results from six dynamic global vegetation models. *Global Change Biology* 7: 357–373.
- Critchfield, W.B. & Little, E.L. Jr. 1966. Geographic distribution of the pines of the world. *United States Department of Agriculture Miscellaneous Publication* 991: 1–97.
- Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L. & Pasteris, P. 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research* 22: 99–113.
- De'ath, G. 2002. Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology* 83: 1105–1117.
- FIA. 2007. The forest inventory and analysis database: Database description and users guide. Version 3. Forest inventory and analysis program. Available at [http://www.fia.fs.fed.us/tools-data/docs/pdfs/FIADB\\_user%20guide\\_v3-0\\_P2\\_06\\_01\\_07.pdf](http://www.fia.fs.fed.us/tools-data/docs/pdfs/FIADB_user%20guide_v3-0_P2_06_01_07.pdf) [United States Department of Agriculture, Forest service]. (accessed 11 January 2011)
- Fielding, A.H. & Bell, J.F. 1997. A review of methods for the measurement of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49.
- Forbes, A.D. 1995. Classification algorithm evaluation: five performance measures based on confusion matrices. *Journal of Clinical Monitoring and Computing* 11: 189–206.
- Franklin, J.F. & Dyrness, C.T. 1973. *Natural vegetation of Oregon and Washington*. United States Forest Service, General Technical Report PNW-8 Portland, OR, US.
- Garnier, B.J. & Ohmura, A. 1968. A method of calculating direct solar radiation on slopes. *Journal of Applied Meteorology* 7: 796–800.
- Hadley, J.L. 2000. Effect of daily minimum temperature on photosynthesis in eastern hemlock *Tsuga canadensis* L. in autumn and winter. *Arctic, Antarctic, & Alpine Research* 32: 368–374.
- Hember, R.A., Coops, N.C., Black, T.A. & Guy, R.D. 2010. Simulating gross primary production across a chronosequence of coastal Douglas-fir stands with a production efficiency model. *Agricultural & Forest Meteorology* 150: 238–253.
- Kimball, J.S., Running, S.W. & Nemani, R. 1997. An improved method for estimating surface humidity from daily minimum temperature. *Agricultural & Forest Meteorology* 85: 87–98.
- Landsberg, J.J. & Waring, R.H. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. *Forest Ecology and Management* 95: 209–228.

- Landsberg, J.J., Waring, R.H. & Coops, N.C. 2003. Performance of the forest productivity model 3-PG applied to a wide range of forest types. *Forest Ecology and Management* 172: 199–214.
- Law, B.E., Goldstein, A.H., Anthoni, P.M., Unsworth, M.H., Panek, J.A., Bauer, M.R., Francheboud, J.M. & Hultman, N. 2001. Carbon dioxide and water vapor exchange by young and old ponderosa pine ecosystems during a dry summer. *Tree Physiology* 21: 299–308.
- Little, E.L. Jr. 1971. Atlas of United States trees, conifers and important hardwoods.: Volume 1. United States Department of Agriculture [miscellaneous publication 1146].
- Marcot, B.G., Holthausen, R.S., Raphael, M.G., Rowland, M.M. & Wisdom, M.J. 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *Forest Ecology and Management* 153: 29–42.
- Melendez, K.V., Jones, D.L. & Feng, A.S. 2006. Classification of communication signals of the little brown bat. *Journal of the Acoustical Society of America* 120: 1095–1102.
- Monserud, R.A. & Leemans, R. 1992. Comparing global vegetation maps with the kappa statistic. *Ecological Modeling* 62: 275–293.
- Nightingale, J.M., Coops, N.C., Waring, R.H. & Hargrove, W.W. 2007. Comparison of MODIS gross primary production estimates for forests across the USA with those generated by a simple process model, 3-PG. *Remote Sensing of Environment* 109: 500–509.
- Panek, J.A., Kurpius, M.R. & Goldstein, A.H. 2002. An evaluation of ozone exposure metrics for a seasonally drought-stressed ponderosa pine ecosystems. *Environmental Pollution* 117: 93–100.
- Raffa, K.F., Aukema, B.H., Bentz, B.J., Carroll, A.L., Hicke, J.A., Turner, M.G. & Romme, W.H. 2008. Cross-scale drivers of natural disturbance prone to anthropogenic amplification: the dynamics of bark beetle eruptions. *BioScience* 58: 501–517.
- Raven, P.H. & Axelrod, D.I. 1974. Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* 61: 539–673.
- Rehfeldt, G.E. & Jaquish, B.C. 2010. Ecological impacts and management strategies for western larch in the face of climate-change. *Mitigation* 15: 283–306.
- Rehfeldt, G.E., Crookston, N.L., Warwell, M.V. & Evans, J.S. 2006. Empirical analysis of plant–climate relationships for the western United States. *International Journal of Plant Science* 167: 1123–1150.
- Running, S.W., Waring, R.H. & Rydell, R.A. 1975. Physiological control of water flux in conifers: a computer simulation model. *Oecologia* 18: 1–16.
- Running, S.W., Nemani, R.R. & Hungerford, R.D. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapo-transpiration and photosynthesis. *Canadian Journal of Forest Research* 17: 472–483.
- Runyon, J., Waring, R.H., Goward, S.N. & Welles, J.W. 1994. Environmental limits on net primary production and light-use efficiency across the Oregon transect. *Ecological Applications* 4: 226–237.
- Schroeder, T.A., Hember, R.A., Coops, N.C. & Liang, S. 2009. Validation of solar radiation surfaces from MODIS and reanalysis data over topographically complex terrain. *Journal of Applied Meteorology and Climatology* 48: 2441–2458.
- Schroeder, T.A., Hamann, A., Coops, N.C. & Wang, T. 2010. Occurrence and dominance of six Pacific Northwest conifer species. *Journal of Vegetation Science* 23: 586–596.
- Schwalm, C.R., Black, T.A., Amiro, B.D., Arain, M.A., Barr, A.G., Bourque, C.P.-A., Dunn, A.L., Flanagan, L.B., Giasson, M.-A., Lafleur, P.M., Margolis, H.A., McCaughey, J.H., Orchansky, A.L. & Wofsy, S.C. 2006. Photosynthetic light use efficiency of three biomes across and east–west continental-scale transect in Canada. *Agricultural and Forest Meteorology* 140: 260–286.
- Seidl, R., Baier, P., Rammer, W., Schopf, A. & Lexer, M.J. 2007. Modelling tree mortality by bark beetle infestation in Norway spruce forests. *Ecological Modelling* 206: 383–399.
- Sherrod, P.H. 2010. DTREG manual Available at <http://www.dtreg.com/DTREG.pdf> (accessed 11 January 2011)
- Simonin, K., Kolb, T.E., Montes-Helu, M. & Koch, G.W. 2007. The influence of thinning on components of stand water balance in a ponderosa pine forest stand during and after extreme drought. *Agricultural and Forest Meteorology* 143: 266–276.
- Swenson, J.J., Waring, R.H., Fan, W. & Coops, N.C. 2005. Predicting site index with a physiologically based growth model across Oregon, USA. *Canadian Journal of Forest Research* 35: 1697–1707.
- Swift, L.W. 1976. Algorithm for solar radiation on mountain slopes. *Water Resources Research* 12: 108–112.
- Tierney, G.L., Fahey, T.J., Goffman, P.M., Hardy, J.P., Fitzhugh, R.D. & Driscoll, C.T. 2001. Soil freezing alters fine root dynamics in a northern hardwood forest. *Biogeochemistry* 56: 175–190.
- Vach, W. 2005. The dependence of Cohen's kappa on the prevalence does not matter. *Journal of Clinical Epidemiology* 58: 655–661.
- Wang, T., Hamann, A., Spittlehouse, D.L. & Aitken, S.N. 2006. Development of scale-free climate data for western Canada for use in resource management. *International Journal of Climatology* 26: 383–397.
- Waring, R.H. 1969. Forest plants of the eastern Siskiyous: their environmental and vegetational distribution. *Northwest Science* 43: 1–17.
- Waring, R.H. & McDowell, N. 2002. Use of a physiological process model with forestry yield tables to set limits on annual carbon balances. *Tree Physiology* 22: 179–188.
- Waring, R.H., Landsberg, J.J. & Williams, M. 1998. Net primary production of forests: a constant fraction of gross primary production? *Tree Physiology* 18: 129–134.

- Waring, R.H., Milner, K.S., Jolly, W.M., Phillips, L. & McWethy, D. 2005. A basis for predicting site index and maximum growth potential across the Pacific and Inland Northwest USA with a MODIS satellite-derived vegetation index. *Forest Ecology & Management* 228: 285–291.
- Waring, R., Nordmeyer, A., Whitehead, D., Hunt, J., Newton, M., Thomas, C. & Irvine, J. 2008. Why is the productivity of Douglas-fir higher in New Zealand than in the Pacific Northwest, USA? *Forest Ecology and Management* 255: 4040–4046.
- Whittaker, R.H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30: 279–338.
- Williams, J.W., Jackson, S.T. & Kutzbach, J.E. 2007. Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Science USA* 104: 5738–5742.
- Woods, A., Coates, K.D. & Hamann, A. 2005. Is an unprecedented *Dothistroma* needle blight epidemic related to climate change? *BioScience* 55: 761–769.