Linking climate, gross primary productivity, and site index across forests of the western United States

Aaron R. Weiskittel, Nicholas L. Crookston, and Philip J. Radtke

Abstract: Assessing forest productivity is important for developing effective management regimes and predicting future growth. Despite some important limitations, the most common means for quantifying forest stand-level potential productivity is site index (SI). Another measure of productivity is gross primary production (GPP). In this paper, SI is compared with GPP estimates obtained from 3-PG and NASA's MODIS satellite. Models were constructed that predict SI and both measures of GPP from climate variables. Results indicated that a nonparametric model with two climate-related predictor variables explained over 68% and 76% of the variation in SI and GPP, respectively. The relationship between GPP and SI was limited (R^2 of 36%–56%), while the relationship between GPP and climate (R^2 of 76%–91%) was stronger than the one between SI and climate (R^2 of 68%–78%). The developed SI model was used to predict SI under varying expected climate change scenarios. The predominant trend was an increase of 0–5 m in SI, with some sites experiencing reductions of up to 10 m. The developed model can predict SI across a broad geographic scale and into the future, which statistical growth models can use to represent the expected effects of climate change more effectively.

Résumé : Il est important d'évaluer la productivité forestière pour élaborer des régimes d'aménagement efficace et prédire la croissance future. Malgré certaines limitations importantes, le moyen le plus fréquemment utilisé pour quantifier la productivité potentielle à l'échelle du peuplement forestier est l'indice de qualité de station (IQS). La production primaire brute (PPB) est une autre mesure de la productivité. Dans cet article, IQS est comparé aux estimations de PPB obtenues à l'aide du modèle 3-PG et de l'instrument satellitaire MODIS de la NASA. Nous avons bâti des modèles qui prédisent IQS et les deux mesures de PPB à partir des variables climatiques. Les résultats ont montré qu'un modèle non paramétrique avec deux variables explicatives reliées au climat expliquait respectivement 68 % et 76 % de la variation de IQS et de PPB. La relation entre PPB et IQS était faible (R^2 de 36 % à 56 %) tandis que la relation entre PPB et le climat (R^2 de 76 % à 91 %) était plus forte que la relation entre IQS et le climat (R^2 de 68 % à 78 %). Le modèle qui a été développé pour IQS a été utilisé pour prédire IQS à partir de différents scénarios prévisibles de changement climatique. Une augmentation de IQS de 0 à 5 m était la tendance prédominante alors que certaines stations subissaient des réductions allant jusqu'à 10 m. Le modèle qui a été développé peut prédire IQS pour une large gamme d'échelles géographiques ainsi que dans le futur et la valeur de IQS ainsi obtenue peut être utilisée par les modèles statistiques de croissance pour illustrer plus efficacement les effets anticipés du changement climatique.

[Traduit par la Rédaction]

Introduction

Measures of productivity are critical in efforts to understand and maintain forest ecosystems sustainably for a range of products and services, while managing risks related to wildfire, pests, and diseases and, increasingly, climate change (Latta et al. 2010). Gross primary productivity (GPP) is a measure that characterizes the carbon-fixing ability of terrestrial ecosystems. When interest lies in the potential for forests to produce wood over time, site index (SI) is a preferred measure of productivity, particularly in growth and yield modeling applications involving even-aged forests (Skovsgaard and Vanclay 2008). Among the reasons for its widespread adoption are the relative ease with which SI can be estimated from field observations, its proven efficacy in predicting volume growth and yield, and the strength of the relationship between tree height and age in even-aged forests (McLeod and Running 1988).

Despite its utility in management-centered applications, SI has some limitations as a productivity measure, including a requirement that trees of known or measurable age and a particular species must be growing in even-aged cohorts. Soil SI has been one approach for addressing this shortcoming, but has not been widely used due to the limited availability and resolution of soils data (Carmean 1975). Measurement error and variability due to site-tree selection must be kept to a minimum to make accurate SI determinations. Further, species-specific height–age relationships must be known a priori

Received 9 December 2010. Accepted 1 June 2011. Published at www.nrcresearchpress.com/cjfr on 16 August 2011.

A.R. Weiskittel. School of Forest Resources, University of Maine, Orono, ME 04469, USA.
N.L. Crookston. Rocky Mountain Research Station, US Forest Service, Moscow, ID 83843, USA.
P.J. Radtke. College of Natural Resources and Environment, Virginia Tech, Blacksburg, VA 24060, USA.

Corresponding author: Aaron R. Weiskittel (e-mail: aaron.weiskittel@maine.edu).

to estimate SI (Monserud 1988; Nigh and Love 1999). Increasingly, managers and modelers are taking notice of the fact that SI may be affected over time by factors such as changing management regimes, genetic improvements, or climate change (Monserud and Rehfeldt 1990; Monserud et al. 2008).

Efforts to quantify terrestrial ecosystem productivity typically rely on GPP rather than SI, especially in applications involving climate change. Unlike SI, GPP can be measured in nonforested ecosystems and in forests where age, structure, or species composition do not meet the basic assumptions for SI. Further, GPP can be used to measure productivity at scales ranging from a single leaf to the entire globe (Reich and Bolstad 2001). In contrast, SI is primarily used when operating at the forest stand level. Direct measures or estimates of GPP may involve one of several approaches: (*i*) leaf-level or canopy-level carbon flux measurements coupled with canopy-scale models, (*ii*) canopy-scale accounting based on measurements of fluxes, storage, and respiration components, or (*iii*) remotely sensed observations of the fraction of absorbed photosynthetically active radiation coupled with ecosystem models (Gower et al. 1999; Running et al. 1999). Using GPP in management-oriented models has its own set of challenges, including the need to parameterize models for appropriate spatial, structural, and temporal scales and to calibrate predictions so that they are useful in predicting for populations of interest that span a broad range of management regimes and ecological conditions.

Models are increasingly being developed that are capable of accounting for the effects of human-caused changes in forest productivity, particularly those involving climate change (Waring et al. 2006). Both SI and GPP can serve to inform predictions generated by such models, even those designed to make management-oriented growth and yield predictions. One way to implement such models involves the updating of SI to account for climate or other factors that affect forest growth (Baldwin et al. 2001; Swenson et al. 2005). Other methods link statistical and physiological process-based models without directly modifying SI (e.g., Henning and Burk 2004), while some rely more completely on generalized process-based model predictions along with algorithms that allocate biomass using allometric relationships (e.g., Landsberg and Waring 1997). Regardless of the mechanism used, the ability of statistical growth and yield models to properly project future conditions will be limited unless such models can be adapted to account for climate effects on site productivity and other factors that influence growth and yield (Crookston et al. 2010).

The objective here is to address questions about whether SI and GPP are consistent measures of forest ecosystem productivity, especially in how they vary with climate-derived variables across the western United States and over time, given climate change. To approach this objective, several research questions were investigated: (*i*) what impact does ignoring tree species have on determining SI, since doing so would be advantageous in projecting changes in productivity that may result from climate change in coming decades; (*ii*) does the relationship between SI and GPP differ depending on how GPP is derived; (*iii*) does the relationship between SI and climate depend on the assumptions and resolution of the climate data; and (*iv*) how do predicted changes in SI differ depending on the assumptions made about climate change?

Methods and materials

Geographic scope

The geographic scope was chosen to match recent studies that have examined potential climate change impacts on species distributions and forest productivity across portions of the western United States (Rehfeldt et al. 2006; Latta et al. 2009). In addition, a climate-sensitive implementation of the Forest Vegetation Simulator model has recently been developed for forests of the western United States (Crookston et al. 2010). In particular, data from 11 western states were used, namely Arizona, California, Colorado, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, and Wyoming. It seemed likely that direct relationships between climate and productivity would be observable in this region, since it is known for having steep clines of productivity and a wide range of climate conditions.

SI

Individual tree height and age data used in this analysis were obtained from the USDA Forest Service Forest Inventory and Analysis (FIA) program, which comprises a network of field plots measured periodically across the forests of the United States. The FIA program uses a multiphase sampling design in which a systematic grid of hexagons is superimposed over the country to estimate forest and nonforest conditions. One FIA field plot is established at random within each hexagon where forest lands are present. The sampling intensity translates to about one field plot every 2500 ha in forested regions. Each field plot consists of three 7.32 m (24 ft) radius subplots that are arranged in a triangular pattern around a central subplot of the same size, with the full complement of subplots covering 672.5 m^2 (about one sixth of an acre).

At each FIA plot, one or more trees are measured for total height and age at breast height (1.37 m or 4.5 ft) as determined from increment core extraction and field examination of tree rings. SI is then computed as an attribute for the field plot from species-specific height–age equations and base ages typically 50 or 100 years that vary by region (e.g., Hanson et al. 2002). We denote the field-observed SI obtained from FIA measurements as SI_{FIA} (metres). In this analysis, only SI derived from a base age 50 year equation were used.

A variation on SI_{FIA} formulated for this study assigns a calculated 50 year base age height from the Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) height age model of Monserud (1984), which is denoted SI_{MON} (metres). The Monserud (1984) equation was selected for use because it represented the most common species in our database and corresponded to a broad geographic range. Rather than typical applications where SI is used in site-specific management applications spanning relatively short time frames, SI_{MON} was proposed here as a standardized measure of site productivity for long-term regional analyses. The species of dominant trees in FIA data was ignored in calculating SI_{MON} so that it could be computed without knowing the species composition of a site. A 50 year base age was applied uniformly across the study area, in contrast with SI_{FIA}, which adopts different base ages for various species and regions. Strength of correlation, mean bias ($SI_{FIA} - SI_{MON}$), and equivalence test of dissimilarity (Robinson and Froese 2004) were computed by species to determine the consistency of SI_{MON} as a productivity measure. Given the known problems associated with SI such as high measurement and prediction error (e.g., Goelz and Burk 1996), a critical threshold of $\bar{y} \pm 25\%$ was selected for testing equivalence (Parkhurst 2001). For each site tree in the FIA database, breast height age and total height were used to compute a SI_{MON} using the Monserud (1984) model. When more than one site tree was observed on an FIA plot, SI values were averaged. Furthermore, preliminary analysis indicated that similar results of predicting SI from climate and GPP were obtained when either SI_{MON} or SI_{FIA} was used.

Tree height and age observations were available for 21 554 FIA field plots measured between 1984 and 2007 in the western United States. Douglas-fir was the most common (37.5%) of the 61 species in the site-tree database, with ponderosa pine (*Pinus ponderosa* P. & C. Lawson) and lodgepole pine (*Pinus contorta* Dougl. ex Loud.) being the next two most common site-tree species at 15.3% and 8.3% of the observations, respectively.

Climate

Thirty-year (1961–1990) averaged monthly values for maximum, mean, and minimum daily temperatures along with monthly total precipitation were obtained from a spline-surface model developed from some 3000 climate stations throughout the western United States (Rehfeldt 2006). Climate values for each FIA plot location were generated from the spline surface based on longitude, latitude, and elevation (http://forest.moscowfsl.wsu.edu/climate; Rehfeldt 2006). In the publicly available FIA database used here, plot locations were offset by unspecified distances, generally ≤1 km, to protect landowner privacy and plot integrity (Coulston et al. 2006). Because elevations are highly spatially correlated with climate for the spatial extent of the study and unadjusted plot elevation is given in the FIA database, it was assumed that the plot location offsets would have little effect on climate data accuracy in this application (Prisley et al. 2009).

A suite of derived variables were computed from the monthly climate data (Table 1) based on the expectation that some of the variables would be relevant for predicting or explaining relationships between climate and GPP or SI (Rehfeldt et al. 2006; Waring et al. 2006). For comparison, climate variables were also obtained for FIA plots using the DAYMET database of 1 km^2 gridded daily temperature and precipitation estimates (http://www.daymet.org; Thorton et al. 1997). The spline-surface estimates from the DAYMET model overlap with those of the Moscow Forestry Sciences Laboratory in geographic scope and suitably overlap with the period of climate observations (1980–2003 for DAYMET versus 1961–1990 for the Moscow climate model).

GPP

Annual MODIS-derived GPP (megagrams of carbon per hectare per year) data corresponding to FIA plot locations were obtained for 2000–2004 from the repository of 1 km^2 GPP estimates at http://www.ntsg.umt.edu/modis/, which were generated using the MOD17 algorithm (Heinsch et al. 2003; Zhao et al. 2005). In addition to satellite-based estimates, GPP was calculated for each FIA plot using the physiological-based stand growth model 3-PG (Landsberg and Waring 1997). Although 3-PG is relatively simple compared with some process-based models, its performance has been demonstrated across a wide geographic range and different forest species (Landsberg et al. 2003). Annual 1 km^2 grid layers of GPP for 2000–2004 created using 3-PG were obtained from Nightingale et al. (2007). Mean annual GPP was estimated from these grids and the predicted value was extracted for each FIA plot based on its location on the grid. As in Nightingale et al. (2007), these two measures of GPP were highly correlated ($r = 0.88$).

Analyses

All computations were done using R (R Development Core Team 2009), including several freely available packages that extend the software's basic capabilities. The R packages used were "randomForest" and "maps". Relationships between SI and the MODIS and 3-PG measures of GPP were explored using scatterplots and spline curve fits, displayed graphically along with relevant values of R^2 and root mean square error (RMSE). Next, the ability to predict SI and GPP from climate was explored using the nonparametric regression tool random forest (RF; Breiman 2001), as implemented in R by Liaw and Wiener (2002). RF was used to explore regression relationships between climate variables and the three measures of productivity that we considered. The RF algorithm provides an assessment of how much influence each variable has on productivity. To determine the number of variables to use, we repeated the RF regression procedure sequentially, dropping the least influential variable at each step until there were only two predictors left. The fit statistics (e.g., R^2, RMSE) were then plotted over the number of variables in the sequentially fitted regression analyses and the optimal number of variables was identified by finding the point where the curve's rate of increase slowed to near zero. This approach was assumed to balance model parsimony with overfitting. The two-parameter model and the model with the optimal number of parameters were both presented to highlight trade-offs in predictive capacity.

Given the relative lack of use of RF in forestry, a brief explanation is warranted. The RF algorithm builds a set of regression trees (200 in our case), each based on a separate bootstrap sample of the data. The observations that are left out of the sample are called "out-of-bag" samples. Once the set of regression trees are created, a value of mean square error (MSE) is computed for each tree using the out-of-bag samples that correspond to the tree. The MSE from all of the trees is averaged and reported as overall MSE (the same observations are used to compute an estimate of R^2). We took the square root of this value to report RMSE. The importance scores for each variable are computed using the following logic. First, the observations of the variable are permuted breaking the relationship of the observed value of the variable with the observed values of all of the other variables. Second, predictions are made for the out-of-bag samples using the permuted data and MSE is computed as done for the nonpermuted data. The percent increase in MSE is then computed. The idea is relatively simple: if a variable is an important predictor, then permuting it should cause a relatively

Table 1. Climate variables from the US Forest Service Moscow Laboratory climate model.

Variable name	Definition	Mean	Minimum	Maximum
adi	sqrt(*dd5*)/*map*	0.0528	0.0072	0.3705
adimindd0	*adi* × *mmindd0*	85.63	0.94	502.67
d100	Julian date the sum of degree-days >5 °C reaches 100 °C	120.3	16.0	196.0
dd0	Annual degree-days <0 °C (based on monthly mean temperatures)	601.8	0.0	2187.0
dd5	Annual degree-days >5 °C (based on monthly mean temperatures)	1532	289	5959
dd5mtcm	(*dd5* × *mtcm*)/1000	–1.6	–25.2	67.3
fday	Julian date of the first freezing date of autumn	258.7	197.0	353.0
ffp	Length of the frost-free period	98.6	0.0	331.0
gsdd5	Degree-days >5 °C accumulating within the frost-free period	1011	6	5423
gsp	Growing season precipitation, April–September	285.2	48.0	847.0
gspdd5	(*gsp* × *dd5*)/1000	424.1	60.0	1410.6
gspmtcm	(*gsp* × *mtcm*)/1000	–0.8	–5.5	3.4
gsptd	(*gsp* × *tdiff*)/100	53.6	5.0	111.0
map	Mean annual precipitation	949.5	165.0	3375.0
mapdd5	(*map* × *dd5*)/1000	1572.2	172.5	6636.3
mapmtcm	(*map* × *mtcm*)/1000	–0.9	–13.4	17.3
maptd	(*map* × *tdiff*)/100	164.3	40.9	448.0
mat	Mean annual temperature	6.1	–3.1	21.3
mmax	Maximum temperature in the warmest month	24.8	14.1	40.5
mmin	Minimum temperature in the coldest month	–8.5	–23.0	5.0
mmindd0	Annual degree-days <0 °C based on monthly minimum temperatures	1528	24	4214
mtcm	Mean temperature in the coldest month	–2.9	–14.4	11.3
mtcmgsp	*mtcm*/*gsp*	–0.01	–0.09	0.17
mtcmmap	*mtcm*/*map*	–0.01	–0.05	0.04
mtwm	Mean temperature in the warmest month	16.2	8.6	32.5
pratio	*gsp*/*map*	0.36	0.09	0.80
prdd5	*pratio* × *dd5*	503.7	81.6	2589.7
prmtcm	*pratio* × *mtcm*	–1.7	–9.5	4.5
sday	Julian date of the last freezing date of spring	160	21	221
sdi	Summer dryness index = sqrt(*gsdd5*)/*gsp*	0.13	0.01	1.18
sdimindd0	*sdi* × *mmindd0*	151.2	6.2	790.7
tdgsp	*tdiff*/*gsp*	0.08	0.01	0.39
tdiff	*mtwm* – *mtcm*	19.2	5.5	34.3
tdmap	*tdiff*/*map*	0.03	0.01	0.18

Note: Temperature-related variables are defined in units of °C and precipitation values in mm.

large increase in error, and if it is not important, permuting it should cause relatively low increases in error.

RF models having the strongest ability to predict SI and GPP were subsequently used to forecast changes in productivity in western US forest lands under forecast climate change scenarios for 2060. Monthly downscaled predictions from three general circulation models (GCM) were used to demonstrate potential climate change effects on changes on SI and GPP, one from the Canadian Center for Climate Modeling and Analysis, another from the US Geophysical Fluid Dynamics Laboratory, and a third from the British Met Office's Hadley Center. Climate predictions were based on the A2 greenhouse gas emissions scenario as defined by Nakicenovic and Swart (2000). We used downscaled GCM data from the Moscow Forestry Sciences Laboratory (http://forest.moscowfsl.wsu.edu/climate), which were produced by adapting spline surfaces from contemporary climate data to future GCM outputs (Rehfeldt 2006). Projections from the RF model using inputs from the three GCMs were averaged to minimize any effects of downscaling or discrepancies between GCM predictions that may have unduly influenced future SI_{MON} projections.

Results

Relationship between SI_{FIA} and SI_{MON}

For species with at least five observations (45 of the 61 species), the average Pearson correlation between the SI_{FIA} and SI_{MON} was 0.84 ± 0.16 (mean ± SD). All of the correlations were statistically significant ($p < 0.05$). The weakest correlation was for *Pinus arizonica* ($r = 0.57$). A simple linear regression between SI_{FIA} and SI_{MON} using all observations produced a R^2 and RMSE of 0.76 and 4.36 m, respectively.

The overall mean difference between SI_{FIA} and SIA_{MON} was 0.63 ± 4.54 m (mean ± SD), which varied between species (Table 2). For 16 species, which comprised nearly 90% of the observations, equivalence tests with a null hypothesis of dissimilarity and a critical value of $\bar{y} \pm 25\%$ were rejected. In contrast, the hypothesis of dissimilarity was not rejected for only four of the 20 most common species in the database.

Relationship between GPP and SI

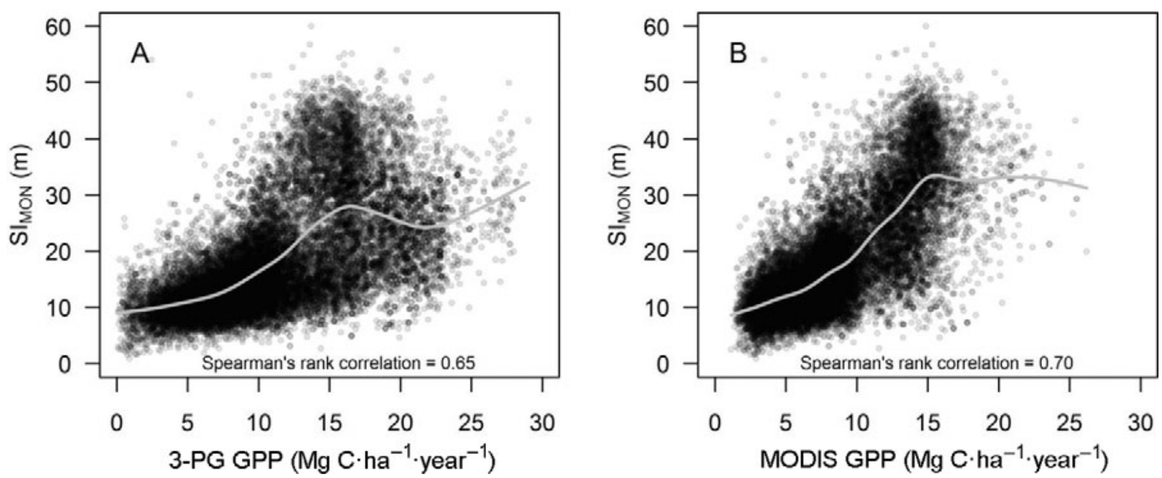
Scatterplots indicate that both measures of GPP and SI_{MON} were moderately related, with Spearman rank correlation co-

Table 2. Comparison between US Forest Service Forest Inventory and Analysis observed site index (SI_{FIA}) (base age 50 years) and site index estimated using the Monserud (1984) equation (SI_{MON}) for the top 20 species by frequency.

Species	*N*	Correlation between SI_{FIA} (m) and SI_{MON} (m)	Mean difference ± SE (m)	Result of equivalence test of dissimilarity
Pinus ponderosa	12290	0.83	–4.65±0.09	Rejected
Picea engelmannii	5659	0.90	0.51±0.08	Rejected
Pinus contorta	5415	0.90	–0.03±0.07	Rejected
Abies lasiocarpa	5225	0.90	1.09±0.07	Rejected
Abies concolor	2976	0.92	0.74±0.14	Rejected
Tsuga heterophylla	2761	0.96	0.55±0.23	Rejected
Abies grandis	1836	0.93	6.24±0.22	Not rejected
Pinus jeffreyi	1326	0.87	–8.61±0.19	Not rejected
Larix occidentalis	983	0.92	0.07±0.24	Rejected
Abies magnifica	783	0.97	4.32±0.22	Not rejected
Pinus flexilis	570	0.76	0.33±0.11	Rejected
Pinus albicaulis	547	0.73	–0.15±0.11	Rejected
Thuja plicata	538	0.92	4.45±0.30	Not rejected
Tsuga mertensiana	510	0.83	0.55±0.23	Rejected
Sequoia sempervirens	349	0.93	–0.38±0.54	Rejected
Pinus monticola	305	0.95	2.25±0.73	Rejected
Picea sitchensis	298	0.82	0.05±0.52	Rejected
Abies lasiocarpa var. *arizonica*	286	0.86	1.21±0.24	Rejected
Pinus lambertiana	275	0.98	1.98±0.43	Rejected

Note: The attributes include the Pearson's correlation coefficient, mean difference (SI_{MON} – SI_{FIA}), and the result of a an equivalence test of dissimilarity (Robinson and Froese 2004). The critical threshold for the test was $\bar{y}$ ± 25%.

Fig. 1. Relationships between predicted site index estimated using the Monserud (1984) equation (SI_{MON}) and gross primary production (GPP) as predicted by (A) 3-PG and (B) MODIS. The gray line is a lowess regression line.



efficient values of 0.65 for 3-PG and 0.70 for MODIS (Fig. 1). Similar patterns were noted when plotting the GPP versus SI_{FIA} (not shown). The relationship between SI and GPP was generally weaker at higher values of SI_{MON}, particularly in comparison with GPP estimated using 3-PG (Fig. 1).

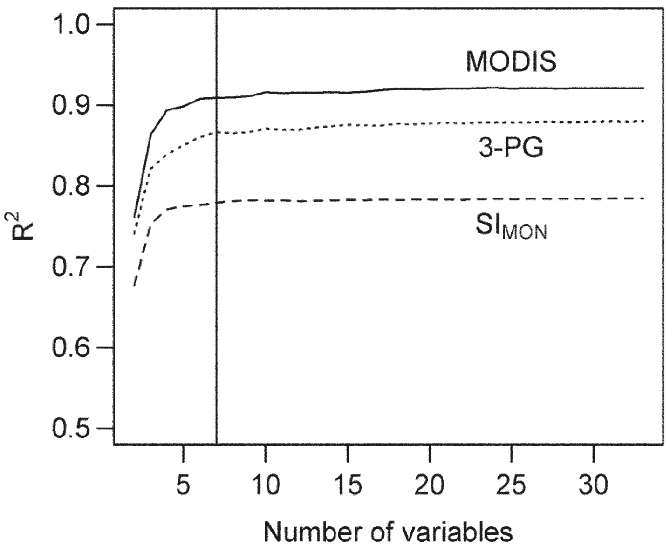
Relationship between climate and GPP and between climate and SI

RF results showed a steep increase in model goodness-of-fit for all response variables when the number of climate-related predictor variables increased from two to four. A relatively gradual increase was observed as the number of predictors was increased from four to seven (Fig. 2). Little additional increase in R^2 was realized by including more than seven predictors. Subsequent RF models were developed using seven predictors, with results for best models explaining 91%, 91%, and 78% of the total variation in MODIS GPP, 3-PG GPP, and SI_{MON}, respectively (Table 3).

Predicted SI given climate change

The distribution of SI_{MON} computed using FIA-observed tree height and age generally showed the greatest forest productivity occurring in coastal areas and along windward slopes of major mountain ranges (Fig. 3A). Observed SI_{MON} showed areas of relatively low productivity in large portions of the Rocky Mountains of Utah, central Colorado, and northwestern Wyoming (Fig. 3A). Spatial distributions of GPP across the region generally showed the same patterns of high and low productivity (Fig. 3).

Fig. 2. Changes in R^2 as the number of variables is increased in random forest models that predict MODIS gross primary production (Mg C·ha^{-1}·year^{-1}), 3-PG gross primary production, and site index estimated using the Monserud (1984) equation (SI_{MON} (m)).



The distribution of SI_{MON} values predicted from observed contemporary climate variables was somewhat narrower than that of the observed SI_{MON} data; however, both results showed similar patterns of areas having relatively high and low forest productivity across the region (cf. Figs. 3A and 4A).The 2060 spatial distribution of projected SI_{MON} based on an A2 greenhouse gas emissions scenario showed a decrease in the value in the coastal areas of Oregon and Washington, while increases were observed in the Idaho Central Rockies and western Montana. The distribution appeared to narrow considerably from the contemporary distribution, which is in part due to averaging the projections from the three GCMs (Fig. 5). In general, the British Met Office's Hadley Center GCM predicted the most severe reductions in SI_{MON}.

The average trend in SI_{MON} change from 2000 to 2090 over all FIA plots evaluated was a slight increase of 1.6 ± 4.9 m; however, mapping revealed that SI_{MON} was projected to decrease on many of the most productive sites while increasing in areas having low contemporary productivities (Fig. 4). Maps of future SI_{MON} produced from individual GCM climate inputs (not shown) exhibited greater variability in projected values and their spatial distributions but, in general, agreed with the three-GCM averages shown in Fig. 4.

Discussion

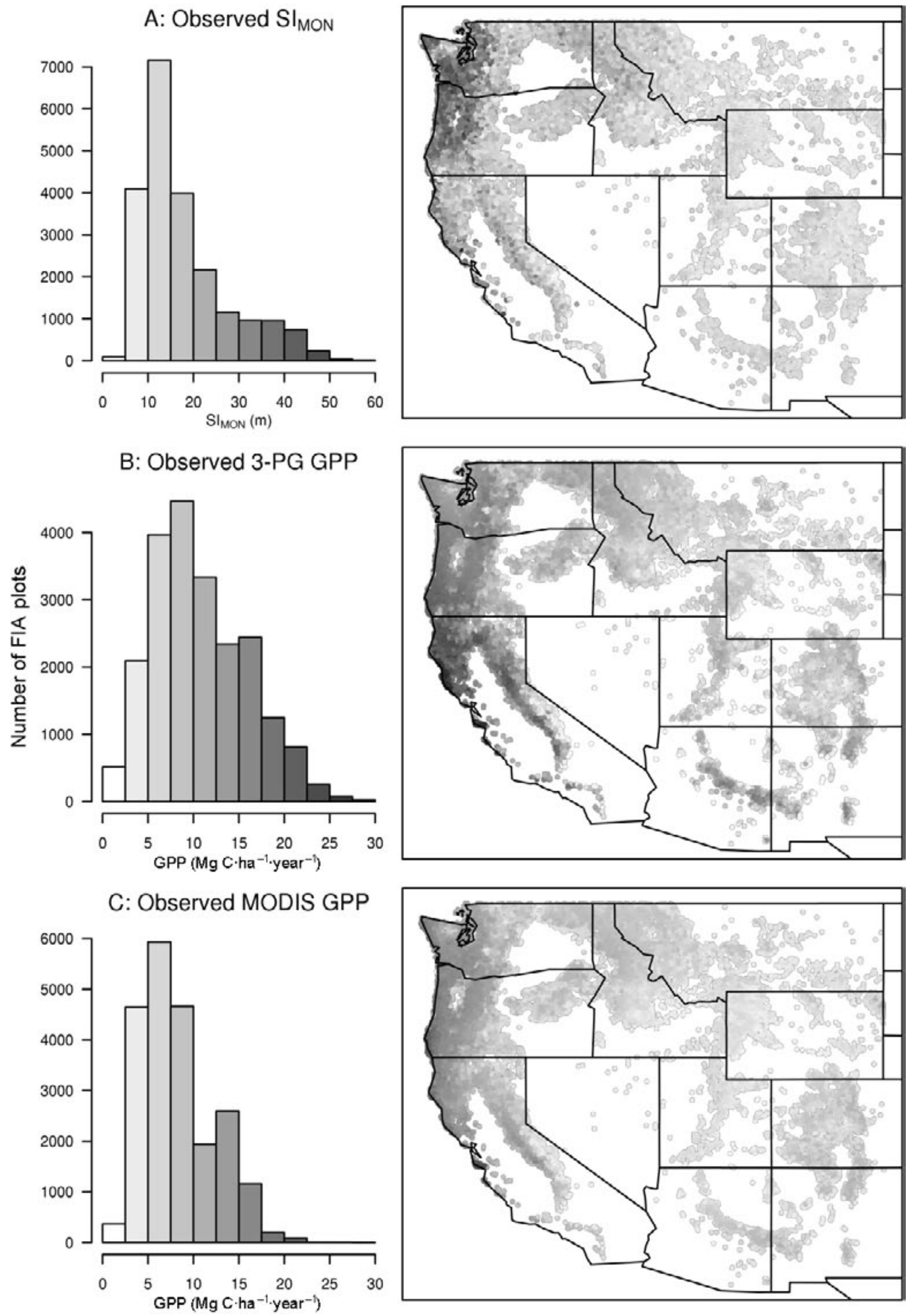
Multiple measures of potential forest productivity have been used in the past including SI (Skovsgaard and Vanclay 2008), maximum basal area (Skovsgaard and Vanclay 2008), yield-based measures (Schmoldt et al. 1985), maximum mean annual increment (Hanson et al. 2002), maximum leaf area index (Waring et al. 1978), and outputs obtained from a process-based model (Swenson et al. 2005; Rodriguez et al. 2009). Although imperfect, SI is currently the most common means for quantitatively assessing forest stand potential productivity and widely used in several statistical growth and

Table 3. Summary of the random forest variables used and their importance scores (%IncMSE) in the two- and seven-variable models for predicting MODIS gross primary production (GPP), 3-PG GPP, and SI estimated using the Monserud (1984) equation (SI_{MON}).

MODIS GPP (Mg C·ha^{-1}·year^{-1})				SI_{MON} (m)				3-PG GPP (Mg C·ha^{-1}·year^{-1})			
7-variable model		2-variable model		7-variable model		2-variable model		7-variable model		2-variable model	
Variable	%IncMSE	Variable	%IncMSE	Variable	%IncMSE	Variable	%IncMSE	Variable	%IncMSE	Variable	%IncMSE
mmax	51.6	*tdiff*	523.3	*prdd5*	49.0	*mmindd0*	446.8	*mmax*	51.6	*tdiff*	523.4
gsp	46.4	*gsp*	158.1	*mmax*	48.3	*gsptd*	155.2	*gsp*	46.4	*gsp*	157.5
prdd5	41.4			*mmindd0*	36.0			*prdd5*	41.4		
sdi	40.4			*maptd*	33.7			*sdi*	40.4		
tdiff	34.3			*sdi*	33.0			*tdiff*	34.3		
maptd	30.9			*gsptd*	32.1			*maptd*	30.9		
mapdd5	29.2			*tdgsp*	25.7			*mapdd5*	29.2		
R^2	0.91		0.76	R^2	0.78		0.68	R^2	0.91		0.76
RMSE	1.19		1.93	RMSE	4.59		5.54	RMSE	1.89		1.93

Note: The overall fit statics are also given, R^2 and root mean square error (RMSE). Variables are maximum temperature in the warmest month (*mmax*), growing season precipitation (*gsp*), interaction between the ratio of summer to total precipitation (*pratio*) and growing degree-days >5 °C (*dd5*) (*prdd5*), summer dryness index (*sdi*), summer–winter temperature differential (*tdiff*), interaction between mean annual precipitation (*map*) and *tdiff* (*maptd*), interaction between *map* and *dd5* (*mapdd5*), minimum growing degree-days <0 °C (*mmindd0*), interaction between *gsp* and *td* (*gsptd*), and *tdiff* divided by *gsp* (*tdgsp*).

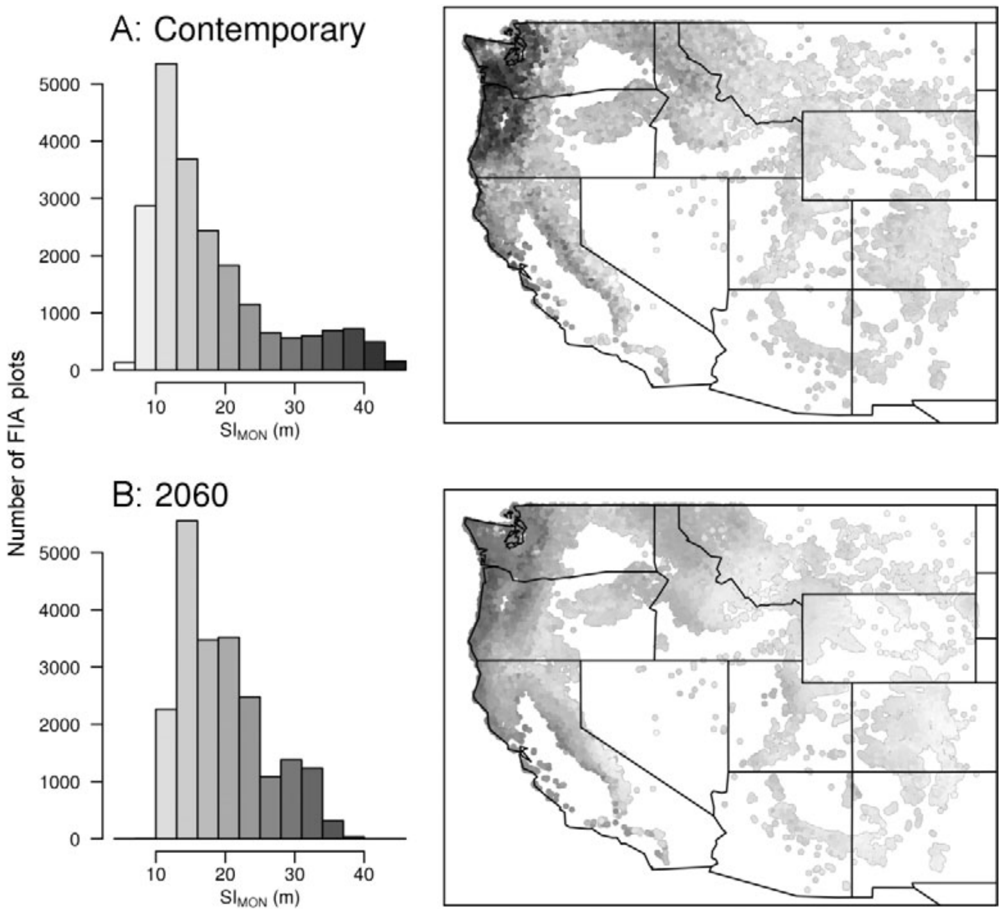
Fig. 3. Histograms and spatial map of (A) site index estimated using the Monserud (1984) equation (SI_{MON}), (B) 3-PG gross primary production (GPP), and (C) MODIS GPP. The darker shading indicates higher values.



yield models (Skovsgaard and Vanclay 2008). Previous work that has attempted to relate SI to climate, physiographic, and soil factors has generally been able to explain a limited amount of the observed variation (<30%; Carmean 1975). Multiple explanations for this have been offered including (*i*) SI is a complex interaction between environmental factors and genetics (Monserud and Rehfeldt 1990), (*ii*) sample size is too low to detect meaningful relationships (Monserud et al. 1990), (*iii*) past natural disturbances limit the strength of SI as a measure of productivity (Nigh and Love 1999), (*iv*) the true cause of forest productivity is not measured (Monserud et al. 1990), and (*v*) traditional parametric techniques like multiple linear regression have a limited capacity to detect complex relationships (Aertsen et al. 2010).

In this analysis, nonparametric modeling tools were used to assess large-scale variation in SI_{MON}, a field-based esti-

Fig. 4. Histogram and spatial map of predicted site index estimated using the Monserud (1984) equation (SI_{MON}) using the seven-parameter random forest model for the (A) contemporary and (B) future climate in 2060. The future SI_{MON} was the average from three general circulation models.

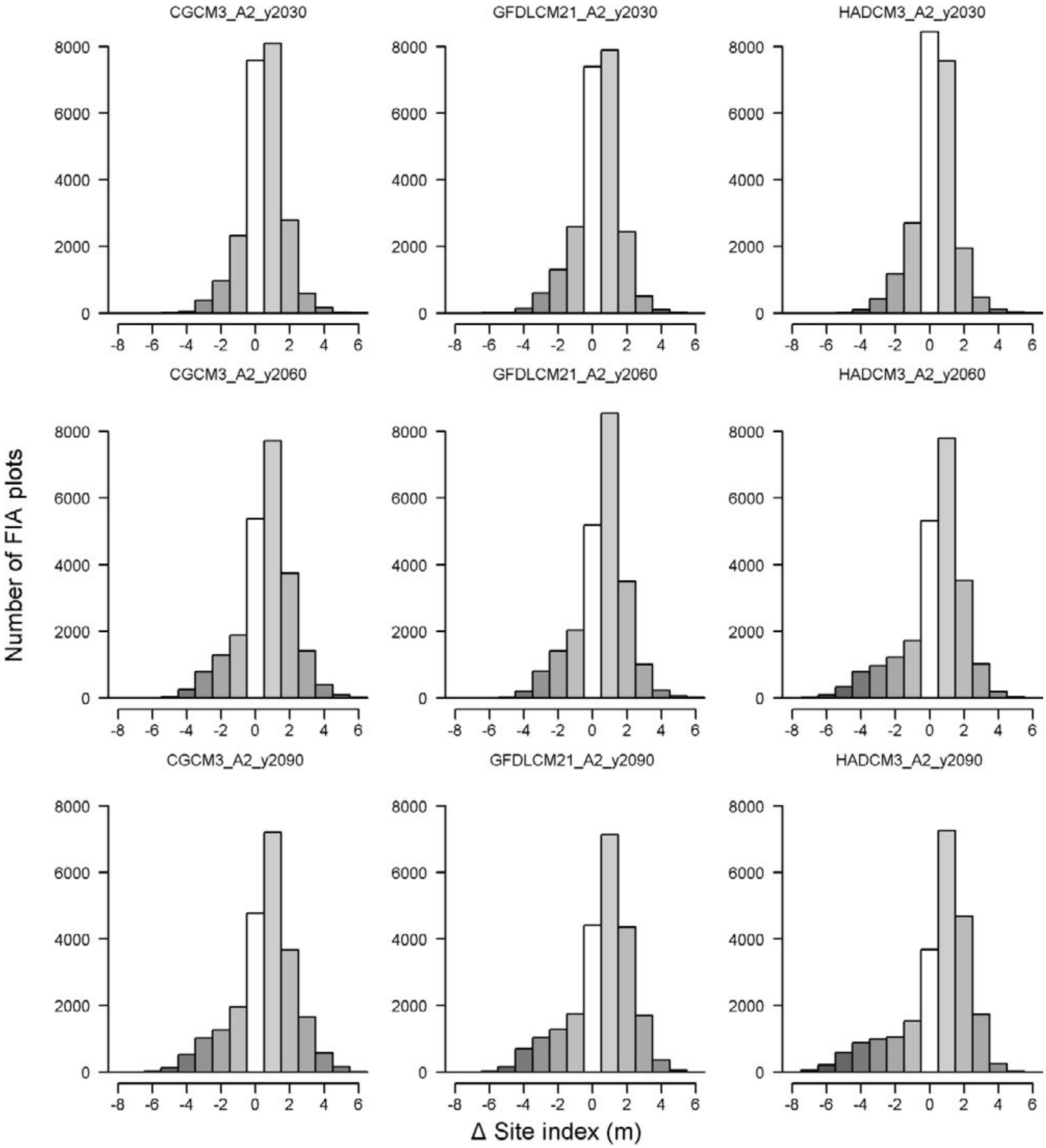


mate of SI. SI_{MON} was used in this analysis rather than SI_{FIA} because it was well correlated with SI_{FIA}, a consistent estimator of SI (i.e., multiple equations are used to estimate SI_{FIA} even for one species), and it can be applied across the landscape without knowing what species are growing there, which is particularly important for future climate scenarios. Overall, these results indicated that SI_{MON} provided a reasonable and consistent estimate of SI for the vast majority of species in this analysis. Predictions from the resulting nonparametric model explained approximately 78% of the variation in SI_{MON} with only a small number (seven) of climate-based predictors.

Process-based models have been proposed as more effective predictors of site productivity because they integrate both climate and soil factors, can mechanistically account for complex interactions between components, and can provide high-resolution productivity maps even for areas without trees currently present (Swenson et al. 2005). The drawbacks to process-based models are that they are often complex, difficult to parameterize, and rely on information that is not easily obtained (e.g., leaf area index, hourly or daily weather observations, soil maps, etc.) (Mäkelä et al. 2000). The 3-PG model (Landsberg and Waring 1997) has been widely used in a variety of forest types because it is relatively easy to parameterize, is robust in its predictions, and can run on information that is generally available for large regions (Landsberg et al. 2003). Consequently, 3-PG has been used to map productivity in the US Pacific Northwest (Swenson et al. 2005; Coops et al. 2011), portions of Chile (Rodriguez et al. 2009), British Columbia (Coops et al. 2010), and the entire United States (Nightingale et al. 2007). When compared with 3-PG, Waring et al. (2006) found that the MODIS midsummer enhanced vegetation index was an equally effective predictor of SI in the US Pacific and Inland Northwest and had the added benefit of not requiring information on soils or climate.

In this present analysis, GPPs derived from 3-PG and MODIS were used in predicting SI. However, both 3-PG and MODIS GPP had a more limited relationship with SI in terms of R^2 and RMSE when compared with the use of just climatic variables. Of the two models, MODIS was a more effective predictor of SI likely because it relies on remotely sensed estimates of leaf area index and daily surface climate (Running et al. 2004) rather than simulated leaf area index and interpolated monthly climate that 3-PG uses. However, both 3-PG and MODIS estimates of GPP are highly dependent on climate, as two-parameter RFs were able to explain over 75% of their original variation. Given this high sensitivity to climate, it appears that predicting SI directly from climatic variables may be a more parsimonious approach and avoids the complications and implicit assumptions inherent

Fig. 5. Histograms of predicted change in site index estimated using the Monserud (1984) equation for three general circulation models (GCM) by year (2030, 2060, and 2090) using the developed seven-parameter random forest model. The GCMs included one from the Canadian Center for Climate Modeling and Analysis (CGCM), another from the US Geophysical Fluid Dynamics Laboratory (GFDLCM), and a third from the British Met Office's Hadley Center (HADCM). Climate predictions were based on the A2 greenhouse gas emissions scenario as defined by Nakicenovic and Swart (2000).



to any process-based model. However, detailed process-based models will be necessary for predicting changes in productivity due to increases in CO_2 (Keenan et al. 2011), but the best approach for doing this is still unclear (Schwalm and Ek 2001).

In this present analysis, climate information was obtained from two distinct sources, namely a spline surface of monthly values from Rehfeldt (2006) and Gaussian weighting of daily observations (DAYMET; Thorton et al. 1997). Both sources integrated many observations from multiple climate stations across the study region into smoothed surfaces, one preserving data on a continuous spatial scale and the other gridded to a 1 km^2 resolution. Relatively little difference in predictive power was found between the various sources of climate data in this analysis. This suggests that multiple techniques used to derive climate data and the temporal resolution of it, the climate variables contain information that is effective for predicting SI despite being derived from only latitude, longitude, and elevation.

The two most influential climate variables identified in this analysis were *mmindd0* and *gsptd,* which were key predictors for some species in the analysis of Rehfeldt et al. (2006). The degree of winter coldness is reflected in *mmindd0*, while the availability of moisture during the warm season and range of temperature fluctuations are represented by *gsptd*. In their analysis, Latta et al. (2009) found that the interaction between annual temperature, precipitation, and a climate moisture index was the most effective for predicting productivity in the US Pacific Northwest. In Alberta, Monserud et al. (2006) found that the strongest linear predictors of lodgepole pine SI were *dd5*, the Julian date when *dd5* reached 100, and July mean temperature. The results of this analysis coincide with the results of Latta et al. (2009), as they indicated that both temperature and moisture are needed to predict productivity, whereas just temperature was needed in Monserud et al. (2006). Also, like Latta et al. (2009), models in this analysis were able to explain around 75% of the original variation in SI, while Monserud et al. (2006) could only explain 25%. However, the Monserud et al. (2006) SI data were obtained from stem-analyzed plots and not predicted from an equation based on observed tree height and age. Regardless of the variation explained, climate is clearly an important component of site productivity, as suggested by both Latta et al. (2009) and Monserud et al. (2006).

Climate change is expected to have a significant influence on species distribution patterns (Rehfeldt et al. 2006), tree growth (Way and Oren 2010), and potential site productivity (Monserud et al. 2008; Coops et al. 2010; Latta et al. 2010) and therefore on forest dynamics (Crookston et al. 2010). Depending on the climate change scenario and geographic area, the results for changes in potential site productivity have varied. Like this study, Bravo-Oviedo et al. (2010) found that Spanish *Pinus pinaster* (Aiton) SI either significantly increased or significantly decreased between –30% and 12.5% depending on geographic location and GCMs used. Monserud et al. (2008) found that SI generally increased by 3 m for lodgepole pine in Alberta, which represented a 26%–35% increase, but the results also were highly dependent on the GCM used. Latta et al. (2010) found a 2%–23% increase in their site productivity measure for the US Pacific Northwest. In British Columbia, Coops et al. (2010) found that SI may decrease up to 10% on average in the interior but increase between 10% and 30% along the coast. For Douglas-fir along the coast, Coops et al. (2010) indicated reductions up to 30% with an average reduction near 17%. Similar changes to those of Coops et al. (2010) were found in this present analysis for the western United States by 2060, as the coastal regions experienced reductions of 10%–30%, while SI for interior sites generally increased by 25%. Future studies are needed to further validate these findings. As noted by these other studies, several factors not accounted for in this analysis may ultimately control these potential changes, particularly genetics, CO_2 concentration, and alterations in disturbance rate and severity.

Summary

Measurements of forest site productivity are necessary but difficult to obtain. In this analysis, a correlation between multiple measures of site productivity existed, with each of them being strongly driven by climatic factors. Although process-based models offer a theoretical advantage of integrating climate and soils information in a mechanistic manner, they do not improve predictions of current SI when compared with purely climatic variables. Across a broad range of stand conditions, two climate variables (*mmindd0* and *gsptd*) representing changes in temperature and moisture availability during the growing season could explain a significant portion of the observed variation in SI_{MON}. Under a range of expected climate change scenarios, SI_{MON} was found to generally increase between 10% and 30% by 2060, but significant reductions up to 30% were also observed in some areas. Utilizing these expected changes in SI can help statistical growth and yield models to better represent future potential productivity without drastic changes to the existing equations, as demonstrated by Crookston et al. (2010).

Acknowledgements

This work was funded by the US Forest Service Global Change Research Program. Discussions with Gary Dixon, Robert Monserud, Peter Gould, and Greg Johnson were helpful in shaping the project. Thanks to the US Forest Service Forest and Inventory Analysis staff for compiling and providing the data. Special thanks to Pierre Bernier, the Associate Editor, and three anonymous reviewers for their help in improving a previous version of this manuscript.

References

Aertsen, W., Kint, V., van Orshoven, J., Özkan, K., and Muys, B. 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. Ecol. Model. **221**(8): 1119–1130. doi:10.1016/j.ecolmodel.2010.01.007.

Baldwin, V.C., Burkhart, H.E., Westfall, J.A., and Peterson, K.D. 2001. Linking growth and yield and process models to estimate impact of environmental changes on growth of loblolly pine. For. Sci. **47**: 77–82.

Bravo-Oviedo, A., Gallardo-Andrés, C., del Río, M., and Montero, G. 2010. Regional changes of *Pinus pinaster* site index in Spain using a climate-based dominant height model. Can. J. For. Res. **40**(10): 2036–2048. doi:10.1139/X10-143.

Breiman, L. 2001. Random forests. Mach. Learn. **45**(1): 5–32. doi:10.1023/A:1010933404324.

Carmean, W.H. 1975. Forest site quality evaluation in the United States. Adv. Agron. **27**: 209–269. doi:10.1016/S0065-2113(08)70011-7.

Coops, N.C., Hember, R.A., and Waring, R.H. 2010. Assessing the impact of current and projected climates on Douglas-fir productivity in British Columbia, Canada, using a process-based model (3-PG). Can. J. For. Res. **40**(3): 511–524. doi:10.1139/X09-201.

Coops, N.C., Gaulton, R., and Waring, R.H. 2011. Mapping site indices for five Pacific Northwest conifers using a physiologically based model. Appl. Veg. Sci. **14**(2): 268–276. doi:10.1111/j.1654-109X.2010.01109.x.

Coulston, J.W., Riitters, K.H., McRoberts, R.E., Reams, G.A., and Smith, W.D. 2006. True versus perturbed forest inventory plot locations for modeling: a simulation study. Can. J. For. Res. **36**(3): 801–807. doi:10.1139/x05-265.

Crookston, N.L., Rehfeldt, G.E., Dixon, G.E., and Weiskittel, A.R. 2010. Addressing climate change in the Forest Vegetation Simulator to assess impacts on landscape forest dynamics. For. Ecol. Manage. **260**(7): 1198–1211. doi:10.1016/j.foreco.2010.07.013.

Goelz, J.C.G., and Burk, T.E. 1996. Measurement error causes bias in site index equations. Can. J. For. Res. **26**(9): 1585–1593. doi:10.1139/x26-178.

Gower, S.T., Kucharik, C.J., and Norman, J.M. 1999. Direct and indirect estimation of leaf area index, f_{APAR}, and net primary production of terrestrial ecosystems. Remote Sens. Environ. **70**(1): 29–51. doi:10.1016/S0034-4257(99)00056-5.

Hanson, E.J., Azuma, D.L., and Hiserote, B.A. 2002. Site index equations and mean annual increment equations for Pacific Northwest Research Station Forest Inventory and Analysis Inventories, 1985–2001. U.S. For. Serv. Res. Note PNW-RN-533.

Heinsch, F.A., Reeves, M., Votava, P., Kang, S., Milesi, C., Zhao, M., Glassy, J., Jolly, W.M., Loehman, R., Bowker, C.F., Kimball, J.S., Nemani, R.R., and Running, S.W. 2003. User's guide GPP and NPP (MOD17A2/A3) products NASA MODIS land algorithm, version 2.0. Numerical Terradynamic Simulation Group, University of Montana, Missoula, Mt.

Henning, J.G., and Burk, T.E. 2004. Improving growth and yield estimtes with a process model derived growth index. Can. J. For. Res. **34**(6): 1274–1282. doi:10.1139/x04-021.

Keenan, T., Serra, J.M., Lloret, F., Ninyerola, M., and Sabate, S. 2011. Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO_2 matters! Glob. Change Biol. **17**(1): 565–579. doi:10.1111/j.1365-2486.2010.02254.x.

Landsberg, J.J., and Waring, R.H. 1997. A generalized model of forest productivity using simplified concepts of radiation use efficiency, carbon balance and partitioning. For. Ecol. Manage. **95**(3): 209–228. doi:10.1016/S0378-1127(97)00026-1.

Landsberg, J.J., Waring, R.H., and Coops, N.C. 2003. Performance of forest productivity model 3-PG applied to a wide range of forest types. For. Ecol. Manage. **172**(2-3): 199–214. doi:10.1016/S0378-1127(01)00804-0.

Latta, G., Temesgen, H., and Barrett, T.M. 2009. Mapping and imputing potential productivity of Pacific Northwest forests using climate variables. Can. J. For. Res. **39**(6): 1197–1207. doi:10.1139/X09-046.

Latta, G., Temesgen, H., Adams, D., and Barrett, T. 2010. Analysis of potential impacts of climate change on forests of the United States Pacific Northwest. For. Ecol. Manage. **259**(4): 720–729. doi:10.1016/j.foreco.2009.09.003.

Liaw, A., and Wiener, M. 2002. Classification and regression by randomForest. R News, **2**: 18–22.

Mäkelä, A., Landsberg, J., Ek, A.R., Burk, T.E., Ter-Mikaelian, M., Agren, G.I., Oliver, C.D., and Puttonen, P. 2000. Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol. **20**(5–6): 289–298. PMID:12651445.

McLeod, S.D., and Running, S.W. 1988. Comparing site quality indices and productivity in ponderosa pine stands of western Montana. Can. J. For. Res. **18**(3): 346–352. doi:10.1139/x88-052.

Monserud, R.A. 1984. Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type. For. Sci.**30**(4): 943–965.

Monserud, R.A. 1988. Variations on a theme of site index. *In* Forest Growth Modelling and Prediction. Proceedings of IUFRO Conference. No. 1, 23–27 August 1987, Minneapolis, Minn. *Edited by* A.R. Ek, S.R. Shifley, and T.E. Burk. U.S. For. Serv. Gen. Tech. Rep. NC-120. pp. 419–427.

Monserud, R.A., and Rehfeldt, G.E. 1990. Genetic and environmental components of variation of site index in inland Douglas-fir. For. Sci. **36**: 1–9.

Monserud, R.A., Moody, U., and Breuer, D.W. 1990. A soil-site study for inland Douglas-fir. Can. J. For. Res. **20**(6): 686–695. doi:10.1139/x90-092.

Monserud, R.A., Huang, S.M., and Yang, Y.Q. 2006. Predicting lodgepole pine site index from climatic parameters in Alberta. For. Chron. **82**(4): 562–571.

Monserud, R.A., Yang, Y., Huang, S., and Tchebakova, N. 2008. Potential change in lodgepole pine site index and distribution under climatic change in Alberta. Can. J. For. Res. **38**(2): 343–352. doi:10.1139/X07-166.

Nakicenovic, N., and Swart, R.E. 2000. Special report on emissions scenarios: a special report of working group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, U.K.

Nigh, G.D., and Love, B.A. 1999. How well can we select undamaged site trees for estimating site index? Can. J. For. Res. **29**(12): 1989–1992. doi:10.1139/x99-163.

Nightingale, J.M., Coops, N.C., Waring, R.H., and Hargove, W.W. 2007. Comparison of MODIS gross primary production estimates for forests across the U.S.A. with those generated by a simple process model, 3-PGS. Remote Sens. Environ. **109**(4): 500–509. doi:10.1016/j.rse.2007.02.004.

Parkhurst, D.F. 2001. Statistical significance tests: equivalence and reverse tests should reduce misinterpretation. Bioscience, **51**(12): 1051–1057. doi:10.1641/0006-3568(2001)051[1051:SSTEAR]2.0.CO;2.

Prisley, S.P., Wang, H.J., and Radtke, P.J. 2009. Combining FIA plot data with topographic variables: are precise locations needed? *In* Proceedings of the 2008 Forest Inventory and Analysis (FIA) Symposium. *Edited by* W. McWilliams, G. Moisen, and R. Czaplewski. USDA Forest Service, Rocky Mountain Research Station, Park City, Utah.

R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rehfeldt, G.E. 2006. A spline model of climate for the Western United States. U.S. For. Serv. Gen. Tech. Rep. GTR-RMRS-165.

Rehfeldt, G.E., Crookston, N.L., Warwell, M.V., and Evans, J.S. 2006. Empirical analysis of plant-climate relationships for the western United States. Int. J. Plant Sci. **167**(6): 1123–1150. doi:10.1086/507711.

Reich, P.B., and Bolstad, P. 2001. Productivity of evergreen and deciduous temperate forests. *In* Terrestrial global productivity. *Edited by* J. Roy, B. Saugier, and H.A. Mooney. Academic Press, San Diego, Calif. pp. 245–283.

Robinson, A.P., and Froese, R.E. 2004. Model validation using equivalence tests. Ecol. Model. **176**(3-4): 349–358. doi:10.1016/j.ecolmodel.2004.01.013.

Rodriguez, R., Real, P., Espinosa, M., and Perry, D.A. 2009. A process-based model to evaluate site quality for *Eucalyptus nitens* in the Bio-Bio Region of Chile. Forestry, **82**(2): 149–162. doi:10.1093/forestry/cpn045.

Running, S.W., Baldocchi, D.D., Turner, D.P., Gower, S.T., Bakwin, P.S., and Hibbard, K.A. 1999. A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sens. Environ. **70**(1): 108–127. doi:10.1016/S0034-4257(99)00061-9.

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M.S., Reeves, M., and Hashimoto, H. 2004. A continuous satellite-derived measure of global terrestrial primary production. Bioscience, **54**(6): 547–560. doi:10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2.

Schmoldt, D.L., Martin, G.L., and Bockheim, J.G. 1985. Yield-based measures of northern hardwood site quality and their correlation with soil-site factors. For. Sci. **31**: 209–219.

Schwalm, C.R., and Ek, A.R. 2001. Climate change and site: relevant mechanisms and modeling techniques. For. Ecol. Manage. **150**(3): 241–257. doi:10.1016/S0378-1127(00)00568-5.

Skovsgaard, J.P., and Vanclay, J.K. 2008. Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. Forestry, **81**(1): 13–31. doi:10.1093/forestry/cpm041.

Swenson, J.J., Waring, R.H., Fan, W.H., and Coops, N. 2005. Predicting site index with a physiologically based growth model across Oregon, USA. Can. J. For. Res. **35**(7): 1697–1707. doi:10.1139/x05-089.

Thorton, P.E., Running, S.W., and White, M.A. 1997. Generating surfaces of daily meteorology variables over large regions of complex terrain. J. Hydrol. (Amst.), **190**(3–4): 214–251. doi:10.1016/S0022-1694(96)03128-9.

Waring, R.H., Gholz, H.L., Grier, C.C., and Plummer, M.L. 1978. Variation in maximum leaf area of coniferous forests in Oregon and its ecological significance. For. Sci. **24**: 131–140.

Waring, R.H., Milner, K.S., Jolly, W.M., Phillips, L., and McWethy, D. 2006. Assessment of site index and forest growth capacity across the Pacific and Inland Northwest U.S.A. with a MODIS satellite-derived vegetation index. For. Ecol. Manage. **228**(1–3): 285–291. doi:10.1016/j.foreco.2006.03.019.

Way, D.A., and Oren, R. 2010. Differential responses to changes in growth temperature between trees from different functional groups and biomes: a review and synthesis of data. Tree Physiol. **30**(6): 669–688. doi:10.1093/treephys/tpq015. PMID:20368338.

Zhao, M.S., Heinsch, F.A., Nemani, R.R., and Running, S.W. 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. Remote Sens. Environ. **95**(2): 164–176. doi:10.1016/j.rse.2004.12.011.