Mapping large-scale forest dynamics: A geospatial approach

Article *in* Landscape Ecology · October 2012

DOI: 10.1007/s10980-012-9767-7

CITATIONS
16

READS
122

1 author:

Jingjing Liang
Purdue University
66 PUBLICATIONS **884** CITATIONS

SEE PROFILE

Landscape Ecol (2012) 27:1091–1108
DOI 10.1007/s10980-012-9767-7

RESEARCH ARTICLE

Mapping large-scale forest dynamics: a geospatial approach

Jingjing Liang

Received: 2 September 2011 / Accepted: 28 May 2012 / Published online: 22 June 2012

Abstract Digital map of forest dynamics is emerging as a useful research and management tool. As a key issue to address in developing digital maps of forest dynamics, spatial autocorrelation has been distinguished into "true" and "false" gradients. Previous ecological models are mostly focused on either "true" or "false" gradient, and little has been studied to simultaneously account for both gradients in a single model. The main objective of this study was to incorporate both gradients of spatial autocorrelation in a deterministic geospatial model to provide improved accuracy and reliability in future digital maps of forest dynamics. The mapping was based on two underlying assumptions—*unit homogeneity* and *intrinsic stationarity*. This study shows that when the factors causing the spatial non-stationarity have been accounted for, forest states could become a stationary process. A prototype geospatial model was developed for the Alaska boreal forest to study current and future stockings across the region. With areas of the highest basal area increment rate projected to cluster along the major rivers and the lowest near the four major urban developments in Alaska, it was hypothesized that moisture limitation and inappropriate human interference were the main factors affecting the stocking rates. These results could be of unprecedented value, especially for the majority of Alaska boreal region where little information is available.

Keywords Growth and yield · Matrix model · Universal kriging · Controlled trend surface · Alaska boreal forest

J. Liang (✉)
Davis College of Agriculture, Natural Resources, and Design, West Virginia University, PO Box 6125, Morgantown, WV 26506-6125, USA
e-mail: jingjing.liang@mail.wvu.edu

Introduction

Digital map of forest dynamics (DMFD), which has become feasible with the recent developments in geographic information system (GIS) and computer technology, is emerging as a useful research and management tool. DMFD is primarily composed of spatially and temporally integrated layers of data from which current and projected future stand attributes, such as stocking, species composition, and diversity at any user-specified location across a large-scale area could be instantly retrieved (Liang and Zhou 2010). When incorporated with a harvest or artificial regeneration component, DMFD could also be used to develop spatial-specific forest management regimes that will greatly reduce the time and cost in decision-making.

In mapping forest dynamics, spatial autocorrelation—a general property of most ecological attributes due to physical or community processes (*c.f.* Legendre

1993)—is a key issue to address (see Liang and Zhou 2010, p. 2351). Spatial autocorrelation represents the dependency among observations that is associated with their geographic locations and distances. As a basic statistical property of ecological variables and a fairly new ecological paradigm, spatial autocorrelation is "*functional*" (Legendre 1993, p. 1659) in ecosystem functioning, enabling estimation of random ecological variables in a geographic space beyond sampled areas. On the other hand, spatial autocorrelation, if not appropriately addressed, may affect the statistics properties of map data due to its violation of the assumption of independence of most standard statistical procedures (Legendre 1993), rendering the output maps less reliable.

Spatial autocorrelation has been distinguished into "true" and "false" gradients (Legendre et al. 1990). Representing a spatial trend across the region is a "true" gradient—global trend, which can be expressed as a function of geographic coordinates. A "false" gradient—autocorrelation represents a phenomenon in which observations at neighboring points are correlated with one another (Legendre et al. 1990). Different approaches have been used to address these two distinctive gradients. The "true" gradient has been accounted for mainly through trend surface (e.g. Gittins 1968; Kuuluvainen and Sprugel 1996) or controlled trend surface (CTS) models (Liang and Zhou 2010). The "false" gradient has often been addressed with geostatistical interpolation models, autoregressive models, and geographically weighted regression models (*c.f.* Miller et al. 2007). Kriging (Cressie 1993; Bailey and Gatrell 1995) is the most popular geostatistical interpolation approach since the estimators are unbiased (e.g. Bolstad et al. 1998; Pfeffer et al. 2003; Miller 2005). Previous ecological models are mostly focused on either "true" or "false" gradient, and little has been studied to simultaneously account for both gradients in a single model.

The main objective of this study was to incorporate both gradients of spatial autocorrelation in a deterministic geospatial model to provide improved accuracy and reliability in future digital maps of forest dynamics. A prototype geospatial model was developed for the Alaska boreal forest, the largest forest component of the Alaska landscape that occupies a vast area of roughly 500,000 km^2 (Olson et al. 2001). Growing under the most severe climate conditions in the world, including an air temperature as low as −70 °C and an annual precipitation that rarely exceeds 50 cm (Mitchell et al. 2004; Liang et al. 2011), this type of forest is characterized by large commercial stands of white spruce and cold-resistant hardwoods (Phelps 2005). Despite its high potential to be managed for timber and biomass (Liang 2010), little is known on which major factors could affect forest productivity and what the future stocking and species composition would be across the region, especially in remote areas with no forest inventory.

Materials and methods

Data

Two types of data were used in this study for model calibration and simulation: remote sensing data and ground-measured forest inventory data. Remote sensing data, introduced in §Projection and spatial inference, were only used for mapping and estimating initial stand basal area distribution. Forest inventory data consisted of remeasured stand population on 446 permanent sample plots established across interior and south-central Alaska since 1994 (see Cooperative Alaska Forest Inventory (CAFI), Malone et al. 2009). The sample area, stretching over 500 km from the Fairbanks area in the north to the Kenai Peninsula in the south and the International Boundary in the east, represents a wide range of species composition and physiographic conditions (Fig. 1). The fixed-size permanent sample plots, each circular in shape and 0.04 ha (0.10 ac) in size, were remeasured with a 5-year interval to monitor natural forest succession.

Periodic measurements on the permanent sample plots include physical site attributes, such as geographic coordinates (Λ, Φ), slope (L), aspect (α), and elevation (z); and tree data such as species, diameter at breast height (D), and status (recruitment, live, or dead) (Table 1). With the wide geographic distribution, sample plots are distanced from 0.3 to 540 km, with a mean of 282 km (Table 2). This large variation of plot distance made it possible to investigate large-scale spatial autocorrelation.

The Alaska boreal forest is largely dominated by only four species—*Betula neoalaskana* Sarg. (Alaska birch), *Populus tremuloides* Michx. (quaking aspen), *Picea glauca* (Moench) Voss (white spruce), and *Picea mariana* (Mill.) B.S.P. (black spruce). Due to

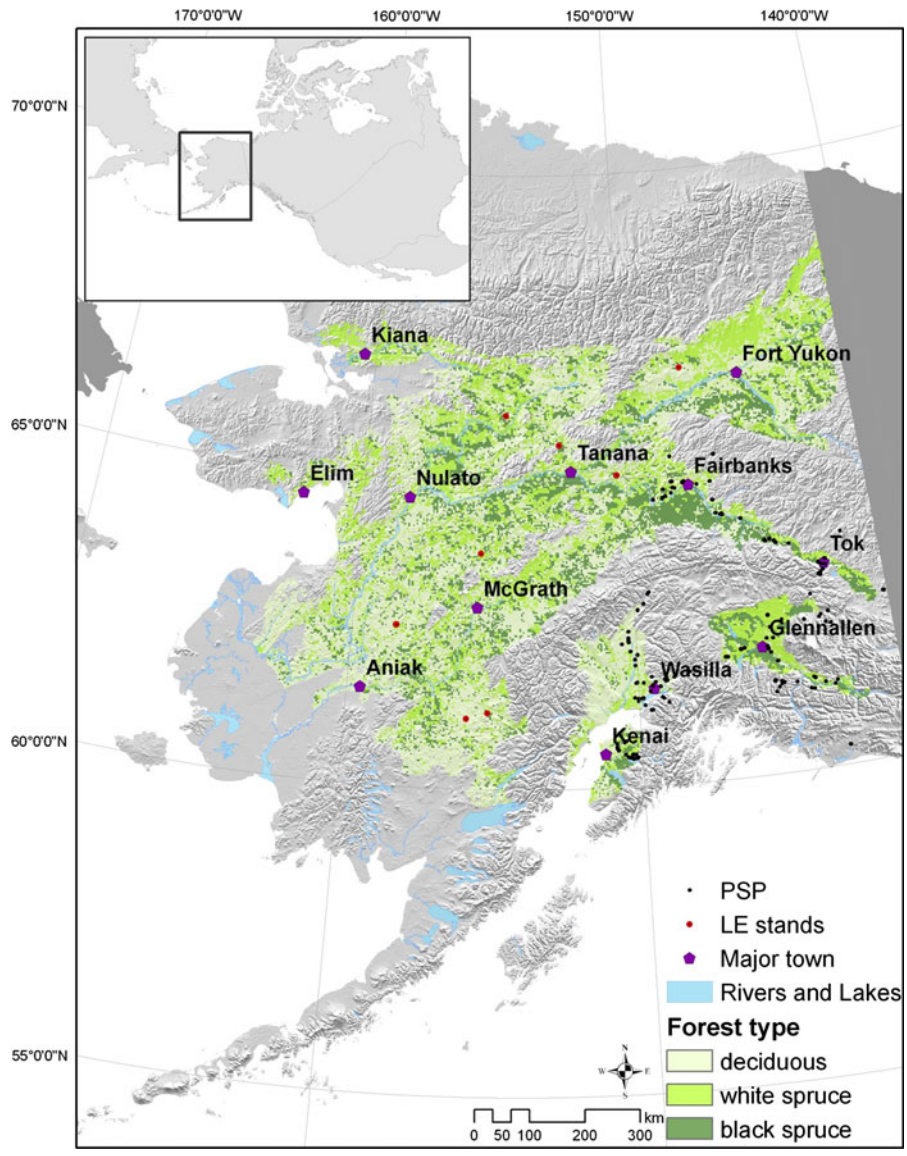

Fig. 1 Geographic distribution of the permanent sample plots (PSP), long-term evaluation (LE) stands, and their relative location within the North America Continent (*inset*), with forest types across the Alaska boreal region from the 2001 NLCD Landsat remote sensing data. Albers equal area conic projection with standard parallels

the similar growth pattern (Liang 2010) and spectral reflectance performance (Cipar et al. 2004), Alaska birch and quaking aspen were grouped here as deciduous species, accounting for 48 % of the total basal area, and white spruce and black spruce were grouped as coniferous species, covering 42 % of the total basal area. Other species in the region, including *Populus trichocarpa* Torr. & Gray (black cottonwood), *P. balsamifera* L. (balsam poplar), *Larix laricina* (DuRoi) K.Koch (Tamarack), and *Betula kenaica* W.H. Evans (Kenai birch), accounted for less than 10 % of the total basal area.

Model description

Assuming $B(\mathbf{s})$—stand basal area (m^2 ha^{-1}) at point locations $\mathbf{s}$—was a random process which satisfied the decomposition:

$$B(\mathbf{s}) = W(\mathbf{s}) + \delta(\mathbf{s}) + e, \qquad \mathbf{s} \in D \subset \Re^2 \tag{1}$$

where $W(\cdot) \equiv E(B(\cdot))$ was the deterministic mean structure representing large-scale variation (true gradient); $\delta(\cdot)$ was a zero-mean, intrinsically stationary process with known variogram; and e was a zero-mean white noise process independent of μ and W.

Table 1 Definition and units of GSM model variables

Variable	Definition	Unit
α	Plot aspect showing the direction to which the plot slope faces. 0 means no slope, 180 and 360 represented south- and north-facing slopes, respectively	Degree
B	Stand basal area	$m^2\ ha^{-1}$
F_1	Deciduous species indicator that takes the value 1 if a stand is dominated by deciduous species, and 0 if otherwise	
F_2	White spruce indicator that takes the value 1 if a stand is dominated by white spruce, and 0 if otherwise	
L	Plot slope	Percent
w	Euclidean distance between two locations	10^3 m
Λ	Easting of Albers equal-area conic projection coordinates	10^6 m
Φ	Northing of Albers equal-area conic projection coordinates	10^6 m
z	Plot elevation	m

Table 2 Summary statistics of GSM model variables. Sample was obtained from the first inventory of the 446 permanent sample plots

	B			Λ	Φ	w
	Deciduous	Coniferous	All species			
Mean	10.87	9.52	20.39	0.34	1.40	281.54
SD	11.06	9.27	10.47	0.13	0.15	159.98
Max	41.26	48.78	48.78	0.65	1.72	539.55
Min	0.00	0.00	0.00	0.17	1.17	0.30
n	446	446	446	446	446	446
	$\cos(\alpha)$	$\sin(\alpha)$	L	z	F_1	F_2
Mean	−0.01	0.00	10.17	356.24	0.50	0.44
SD	0.79	0.61	12.67	228.90	0.50	0.50
Max	1.00	1.00	77.00	957.19	1.00	1.00
Min	−1.00	−1.00	0.00	17.62	0.00	0.00
n	446	446	446	446	446	446

SD standard deviation of the mean, n total number of observations

The geospatial model was developed to estimate future forest stocking and species distribution throughout the Alaska boreal forest region. The model was composed of a dynamics module and a geospatial module. The dynamics module predicted point estimates of the future forest population structures. The geospatial module integrated these point estimates with large-scale forest stocking, species distribution, and topographic information to predict future forest dynamics across the region (see Fig. 2). The following sections describe how the two modules interacted in accounting for both the true and false spatial gradients in forest dynamics across the region.

True gradient control

The dynamics module of the geospatial model provided point estimates of forest dynamics for the inventoried areas. The module could be any type of forest growth model calibrated for the region, but in order to address the large-scale variation of forest dynamics, the CTS Matrix model developed for the Alaska boreal forest (Liang and Zhou 2010) was adopted for this study. Please refer to Liang and Picard (2012) for a comprehensive review of Matrix models.

This model predicted future population structure of a stand or a pixel on the map which represents the distribution of trees per unit area by species and size

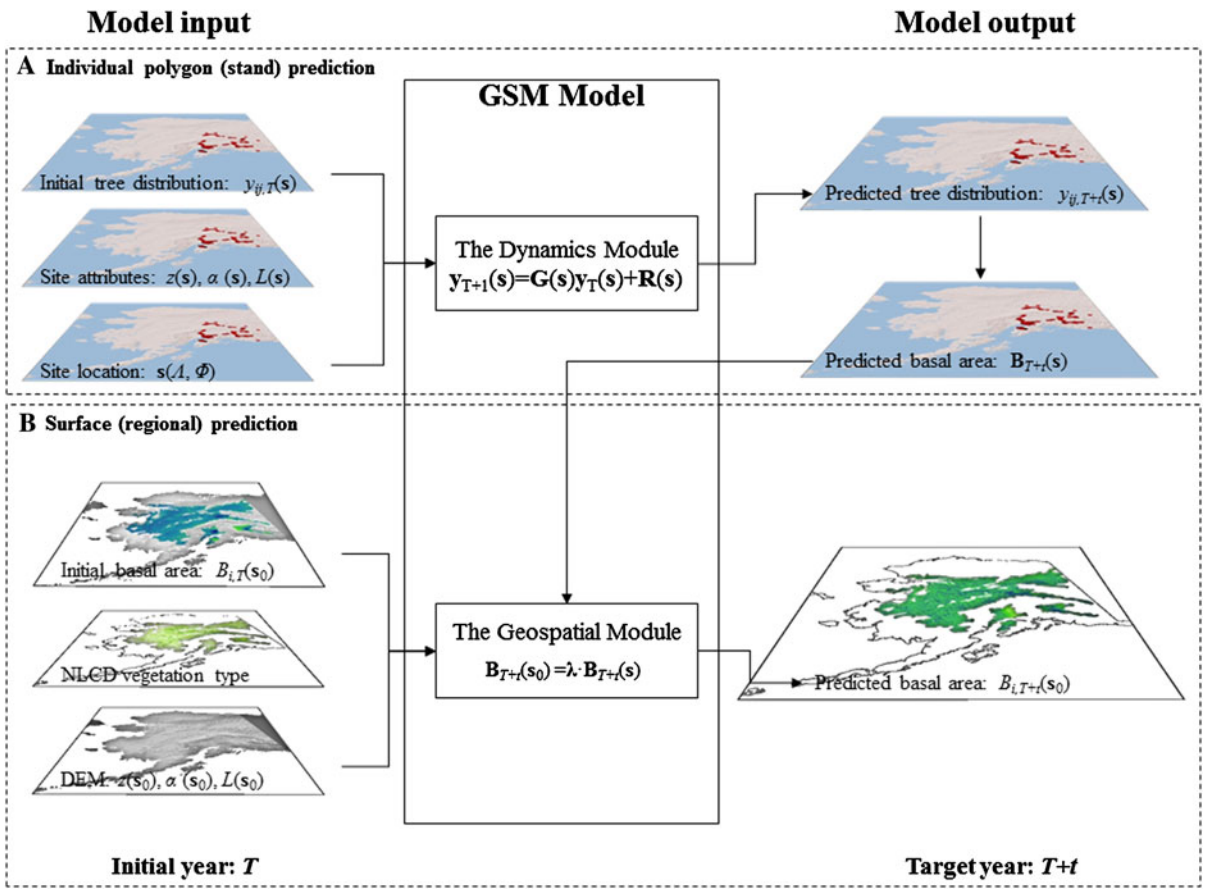

Fig. 2 Conceptual design of mapping with the GSM Model

(age) class. Future population structures are predicted from the initial structure with a regeneration vector, and a growth matrix that accounts for tree upgrowth and mortality. The CTS Matrix model further accounts for large-scale spatial autocorrelation in regeneration, upgrowth, and mortality:

$$\mathbf{y}_{T+1}(\mathbf{s}) = \mathbf{G}(\mathbf{s})\mathbf{y}_T(\mathbf{s}) + \mathbf{R}(\mathbf{s}) \quad (2)$$

where $\mathbf{G}(\mathbf{s})$ and $\mathbf{R}(\mathbf{s})$ represent a growth matrix and recruitment vector respectively which are state- and spatial-dependent. $\mathbf{y}_T(\mathbf{s})$ is a column vector of forest population structure:

$$\mathbf{y}_T(\mathbf{s}) = \begin{bmatrix} \mathbf{y}_{1j}(\mathbf{s})_T \\ \mathbf{y}_{2j}(\mathbf{s})_T \\ \mathbf{y}_{3j}(\mathbf{s})_T \\ \mathbf{y}_{4j}(\mathbf{s})_T \end{bmatrix}, \text{ and } \mathbf{y}_{ij}(\mathbf{s})_T = \begin{bmatrix} y_{i1}(\mathbf{s})_T \\ y_{i2}(\mathbf{s})_T \\ \vdots \\ y_{i19}(\mathbf{s})_T \end{bmatrix} \quad (3)$$

where $y_{ij}(\mathbf{s})_T$ represents the number of live trees per unit area of species i ($i = 1, 2, 3, 4$) and diameter class j ($j = 1,2,\ldots,19$) at time T of location $\mathbf{s}$. Equation 2 could be applied iteratively over t years to obtain $\mathbf{y}_{T+t}(\mathbf{s})$, the estimated population structure at time $T + t$ (see Fig. 2).

The predicted point basal area $\mathbf{B}_{T+t}(\mathbf{s})$ was then calculated with:

$$\mathbf{B}_{T+t}(\mathbf{s}) = \mathrm{b}\mathrm{y}_{T+t}(\mathbf{s}) \quad (4)$$

where $\mathrm{b} = [3.14 \times D_j^2/400]$ was a row vector of mean circular cross-sectional area (m^2) and D_j (cm) was the mid-point of diameter class j (see Fig. 2).

False gradient control

The geospatial module was developed here to obtain surface estimates of basal area $B_{T+t}(\mathbf{s})$ across the region. The control variables were selected from initial basal area $B_T(\mathbf{s})$, forest type $F_{1T}(\mathbf{s})$ and $F_{2T}(\mathbf{s})$ at year T, and time-independent physiographic variables

$P(\mathbf{s})$ which control for site radiation, nutrient availability, and moisture (Stage and Salas 2007):

$$\mathbf{B}_{T+t}(\mathbf{s}) = \mathbf{X}\alpha + \boldsymbol{\delta}(\mathbf{s}) + \mathbf{e}, \qquad \mathbf{s} \in D \subset \Re^2 \tag{5}$$

where $\boldsymbol{\alpha}$ was a column vector of parameters. The error vectors $\boldsymbol{\delta}(\mathbf{s})$ and $\mathbf{e}$ were zero-mean intrinsic stationary random process with known variograms, and zero-mean white noise process, respectively. $\mathbf{X}$ represented a row vector of trend at known location $\mathbf{s}$:

$$\mathbf{X} = [f_h(\mathbf{s})] = [\,1 \quad B_T(\mathbf{s}) \quad F_{1T}(\mathbf{s}) \quad F_{2T}(\mathbf{s}) \quad P(\mathbf{s})\,], \quad \mathbf{s} \in D \subset \Re^2 \tag{6}$$

where $P(\mathbf{s}) = [z(\mathbf{s}) \;\; \cos(\alpha(\mathbf{s})) \;\; \sin(\alpha(\mathbf{s})) \;\; L(\mathbf{s})]$, and $f_h(\mathbf{s}) \equiv \int_{\mathbf{s}} f_h(\mathbf{s})d\mathbf{s} \Big/ |\mathbf{s}|$ represented the mean value of the hth trend variable over the area of pixel $\mathbf{s}$ ($h = 1,2,\ldots,8$).

To account for the false gradient, a universal kriging (UK, *c.f.* Cressie 1993; Matheron 1969) model was adopted to estimate basal area $B_{T+t}(\mathbf{s}_0)$ at year $T + t$ of any unknown location $\mathbf{s}_0$ within the Alaska boreal region D:

$$\mathbf{B}_{T+t}(\mathbf{s}_0) = \mathbf{x}'\alpha + \delta(\mathbf{s}_0), \qquad \mathbf{s}_0 \in D \subset \Re^2 \tag{7}$$

where $\mathbf{x} = [f_h(\mathbf{s}_0)]'$, and the error term $\delta(\mathbf{s}_0)$ was an intrinsic stationary random process with known variograms $2\zeta(w)$. Basic assumptions behind the universal kriging model were intrinsic stationarity across the region, isotropic covariance, and a spherical variogram $2\zeta(\boldsymbol{w})$ (Cressie 1993).

According to the universal kriging theory (Cressie 1993), for any random location $\mathbf{s}_0$ within the Alaska boreal region D, the estimated values $\hat{\mathbf{B}}_{T+t}(\mathbf{s}_0)$ can be expressed as a linear combination of the observed values $\mathbf{B}_{T+t}(\mathbf{s})$ from the 446 permanent sample plots:

$$\hat{\mathbf{B}}_{T+t}(\mathbf{s}_0) = \lambda\mathbf{B}_{T+t}(\mathbf{s}), \qquad \mathbf{s}_0, \mathbf{s} \in D \subset \Re^2 \tag{8}$$

where $\boldsymbol{\lambda}$ was a row vector of linear coefficients estimated under the condition for uniformly unbiased predictor:

$$\lambda'\mathbf{X} = \mathbf{x}'(\mathbf{s}_0) \tag{9}$$

and under the minimization of kriging variance at location $\mathbf{s}_0$:

$$\boldsymbol{\sigma}^2(\mathbf{s}_0) = E\big(\hat{\mathbf{B}}_{T+t}(\mathbf{s}_0) - \mathbf{B}_{T+t}(\mathbf{s}_0)\big)^2 \tag{10}$$

Model estimation

In the geospatial module, the best linear unbiased estimator (BLUE) of $\mathbf{B}_{T+t}(\mathbf{s})$ and its predicted variances were obtained with universal kriging (UK, Cressie 1993; Matheron 1969), based on observed data from the 446 permanent sample plots. Sample variograms of $\hat{\mathbf{B}}_{T+t}(\mathbf{s})$ were computed using Cressie's robust variogram estimate (Cressie 1993) with non-spatial control factors from Eq. 6 as covariates. The computation was performed with geoR (Ribeiro and Diggle 2001)—a geostatistical analysis package of R (R Development Core Team 2010). A parametric empirical variograms model was then fitted by weighted least squares with a spherical covariance model (Cressie 1993).

To ensure goodness of fit and simplicity, the geospatial module was obtained as a reduced form of the full model (Eqs. 5, 6) based on two information criteria: Akaike information criterion (AIC, Akaike 1974) and Bayesian information criterion (BIC, Schwarz 1978). In this study, BIC was given a higher weight as BIC penalizes increased model parameters more than AIC and is more likely to result in a parsimonious model (Schwarz 1978). The candidate control factors were selected from the initial total stand basal area (B_T), forest type indicators (F_1, F_2), and physiographic variables (L, $\cos(\alpha)$, $\sin(\alpha)$, z), due to their obvious impacts on the dynamics of Alaska boreal forest (*c.f.* Liang 2010; Liang and Zhou 2010; Stage and Salas 2007). A parametric empirical variogram model with the lowest AIC and BIC scores was selected as the final model for each species group.

Projection and spatial inference

Underlying assumptions

The future maps of Alaska boreal forests were created in this study based on two underlying assumptions—*unit homogeneity* and *intrinsic stationarity*.

Unit homogeneity assumes that the mapping unit, i.e. *pixel* in this study, has homogeneous species composition and stocking. Technically this assumption can be met by partitioning an area of study to pixels of a size small enough to be homogeneous. In this study, the Alaska boreal region was divided into a grid of 2-km pixels because they represent the finest

resolution of the Alaska land-cover map (ALM, Rupp 2008) available for identification of initial vegetation type in the region.

Intrinsic stationarity assumes forest stocking is a spatial process with a spatially-dependent semivariogram and a constant mean that is subject to exogenous factors. In other words, by assuming *intrinsic stationarity*, we subscribe to the view that stocking of a forest stand could be determined to a certain confidence level from the exogenous factors such as physiographic attributes, and the conditions of its neighboring stands, including their stocking levels and geographic locations. This assumption could be violated by the spatio-temporal shifts of environmental conditions such as climate change, and this was thoughtfully discussed in §Model advantages and limitations.

Initial attributes

The initial attributes of all the 2-km pixels—representing the starting conditions of the Alaska boreal forest for the simulation—were composed of the following six variables. Forest type indicators (F_1, F_2) define the dominant species of each pixel. Initial stand basal area B_{2001} represents forest stocking, and physiographic variables L, α, and z stand for the slope, aspect, and elevation of each pixel, respectively.

Forest type indicators (F_1, F_2) came from the ALM raster data with a 2-km resolution (Rupp 2008) that is re-categorized using the mean growing season temperature and terrain information from the National Land Cover Database (NLCD) 2001 Landsat remote sensing data (http://www.mrlc.gov). There are five land cover classes in ALM: 0 = non-vegetated (e.g. water, rock, or ice), 1 = tundra, 2 = black spruce, 3 = white spruce, and 4 = deciduous trees (birch and aspen) (Fig. 1). For each 2-km pixel, $F_1 = 1$ if its land-cover class is 4 and $F_1 = 0$ otherwise; $F_2 = 1$ if its land-cover class is 2 or 3, and $F_2 = 0$ otherwise.

The initial basal area throughout the region was obtained with universal kriging from $\mathbf{B}_{2001}(\mathbf{s})$—the observed basal area of the 446 sample plots in 2001 and control variables selected (see §Model estimation) from forest type indicators, and physiographic variables L, $\cos(\alpha)$, $\sin(\alpha)$, and z. Physiographic variables were directly obtained from the Digital Elevation Model (DEM) of the U.S. Geological Survey EROS Alaska Field Office (http://agdc.usgs.gov/).

Mapping of future stocking and diversity

Region-wide maps of forest stocking in 2100 were developed with universal kriging (see §Model description) based on $\mathbf{B}_{2100}(\mathbf{s})$—the basal area of the 446 sample plots in 2100 projected by the dynamics module, and control variables selected (see §Model estimation) from forest type indicators and physiographic variables. Predicted basal area of all the pixels throughout the region, $\mathbf{B}_{2100}(\mathbf{s}_0)$, was then aggregated to map future forest stocking (Fig. 2, 8). A 100-year projection horizon was used in this study because with the slow growth rates of all the major species, a period of 100 years is a common cutting cycle for the management of Alaska boreal forest (Liang 2010).

Diversity of location $\mathbf{s}$, a joint entropy of species richness and evenness was measured with Shannon's index (Shannon 1948), which is generally regarded as a premier measure of ecological diversity (Gorelick 2006; Liang et al. 2007):

$$H(\mathbf{s}) = -100 \times \sum_{i=1}^{2} \frac{B_i(\mathbf{s})}{B(\mathbf{s})} \ln\left(\frac{B_i(\mathbf{s})}{B(\mathbf{s})}\right) \quad (12)$$

where $B_i(\mathbf{s})$ and $B(\mathbf{s})$ were the basal area of species i, and total stand basal area at location $\mathbf{s}$, respectively. Diversity would reach its maximum $H_{\max} = 100 * \ln(2) = 69.3$ for a mixed species stand where coniferous and deciduous species are of the same total basal area. Diversity would reach its minimum $H_{\min} = 0$ when a stand is pure coniferous or deciduous.

Model validation

The most straightforward way to validate the model would be to compare the predicted species distribution and forest stocking of year 2100 with observed ones. Since this is impossible, two alternatives were provided to validate this geospatial model in terms of climax states and vegetation, respectively.

Climax states validation assessed if the steady states reached through long-term simulations could be ecologically sound. For this validation, 10 stands were randomly selected from the Alaska boreal forest region D with similar physiographic conditions (Fig. 1), and their predicted population structures over 300 years were compared to the expected forest succession patterns. The stands, approximately 240 m in elevation, 2 % in slope, and around 30° in

aspect, differed greatly in initial stand states. Since climax states are sensitive to physiographic conditions but independent from initial states (e.g. Liang et al. 2005), similar predicted climax states should be expected over these 10 validation stands if the model is reliable.

Projected vegetation was descriptively examined by comparing the projected 2007 map of dominant species to the observed 2007 map of vegetation of Bonanza Creek, Alaska (64.76 N 148.28W, WGS84). The observed vegetation map, which comes from the reconstructed Color-infrared QuickBird imagery acquired on August 9, 2007 and June 18, 2007 and field measurements (Baird 2011), is done at the 1:12,500 scale using ArcGIS. The projected map was made by running the geospatial model on each 30-m pixel to determine the total basal area by species in 2007, and the species with the highest projected basal area was considered dominant for that pixel. An error matrix (Congalton 1991) was made to evaluate the accuracy of the predicted species distribution over 93,564 pixels of the entire area. The error matrix summarizes the number of pixels predicted to be in a particular species category relative to the actual observed number of pixels in that species category. Three measures of accuracy were presented here—*Producer's accuracy* represents the probability of a reference pixel being correctly classified by the model; *User's accuracy* indicates the probability that a pixel classified on the map actually matches that category on the ground (Congalton 1991); and *overall accuracy* was computed by dividing the total correct pixels (the sum of the major diagonal) by the total number of pixels.

Results

Model estimation and validation

The empirical variance of initial stocking displayed an obvious pattern of spatial autocorrelation (Fig. 3, dots and envelopes), for which a spherical variogram model was calibrated with Cressie's robust method (Fig. 3, solid line). All the parameters, including nugget, sill, and range were highly significant. Superior to the non-spatial estimates based on the same control variables, this map of initial stocking estimated by the geospatial model marked lower AIC and BIC scores (Fig. 3).

In order to find appropriate control variables to map future basal area, four candidate models, from the full model with all the potential control factors to the reduced model with no control factor (Table 3, Models #1 through 4) were investigated. According to Table 3, the model with B, F_1, F_2 had the least BIC scores consistently over all, deciduous, and coniferous species. Therefore, B, F_1, F_2 were selected as the control variables of the geospatial model.

The empirical variance of stand basal area of year 2100 predicted by the dynamics module displayed significant spatial autocorrelation over all, deciduous, and coniferous species (Fig. 4a–c, dots and envelopes). A spherical variogram model calibrated for

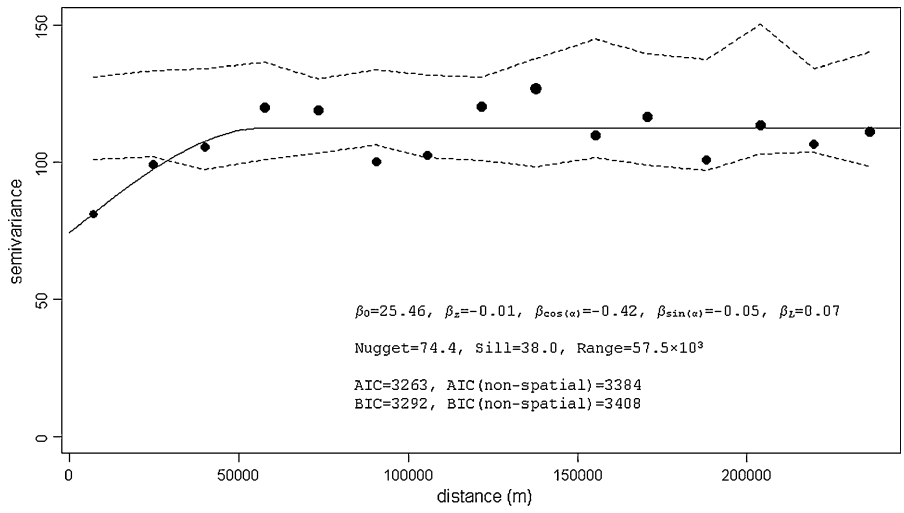

Fig. 3 Empirical semivariogram plot for the model to simulate initial basal area across the Alaska boreal region. Dot size was in proportion to the sample size in that range. Empirical semivariograms were calibrated with initial stand basal area from the 446 permanent sample plots using a spherical model and Cressie's robust method

Table 3 Performance of candidate GSM models

Model (species)		Spatial			Non spatial		
Covariates		*LL*	*AIC*	*BIC*	*LL*	*AIC*	*BIC*
Deciduous species							
1	B, z, cos(α), sin(α), L, F_1, F_2	−1608	3236	3277	−1653	3323	3360
2*	B, F_1, F_2	−1612	3237	3261	−1664	3339	3359
3	B	−1746	3499	3515	−1732	3470	3483
4	–	−1821	3649	3661	−1777	3557	3565
Coniferous species							
1	B, z, cos(α), sin(α), L, F_1, F_2	−1521	3062	3103	−1562	3142	3179
2*	B, F_1, F_2	−1531	3075	3099	−1576	3161	3182
3	B	−1763	3535	3551	−1701	3409	3421
4	–	−1809	3624	3636	−1720	3443	3452
All species							
1	B, z, cos(α), sin(α), L, F_1, F_2	−1628	3276	3317	−1694	3407	3444
2.*	B, F_1, F_2	−1636	3284	3309	−1720	3451	3471
3	B	−1705	3418	3435	−1751	3508	3520
4	–	−1711	3427	3440	−1759	3521	3530

LL log-likelihood

* Indicates the selected model based on *AIC* and *BIC* values

each species (Fig. 4a–c, solid line) featured lower AIC and BIC scores than non-spatial models (Tables 3, 4), which indicated the geospatial model was more accurate than the non-spatial models.

The geospatial model passed the climax state validation. Despite the wide range of initial stand states, the predicted basal areas of the 10 validation stands were able to finally converge to a typical late-seral state, in which the climax coniferous species accounted for approximately 90% of the total stand basal area (Fig. 5). As expected, the standard deviation of the 10 stands gradually declined with time as various stands converged to a single steady state. The predicted total stand basal area of the steady-state (14 m^2 ha^{-1}) resembled the typical stocking of a late-seral boreal forest in Alaska, suggesting the geospatial model was reliable for long-term applications.

The projected vegetation of Bonanza Creek in 2007 largely matched the observed map (Fig. 6). The error matrix (Table 5) shows that for a total of 93,564 30-m pixels in the area, 41,221 out of 50,869 deciduous pixels and 39,454 out of 42,695 coniferous pixels were correctly categorized by the model, respectively. Both *Producer's* and *User's accuracy* were greater than 80 % for all the species, and the *overall accuracy* was 86 %. As a potential reason for the difference between the predicted and observed species distribution, a pixel was defined deciduous or coniferous in this study if the basal area of that species category exceeds 50 percent of total basal area, whereas Baird (2011) determines the vegetation type of a pixel with field data and remotely sensed normalized difference vegetation index (NDVI) instead of basal area.

Maps of current and future stocking

Based on the foregoing model and observed stand basal area of the 446 permanent sample plots, initial basal area, i.e. total stand basal area of 2001, was mapped throughout the entire Alaska boreal region. The mean basal area was ranged from 6.07 to 32.23 m^2 ha^{-1}, with the highest basal area concentrated in three areas in the Tanana Flats near Fairbanks, the Copper River Valley surrounding Glennallen, and the Matanuska-Susitna Valley to the west of Wasilla, respectively. Some of the lowest basal area was estimated to be located at west Copper River Valley and west Kenai Peninsula (Fig. 7a). The standard deviation was expectedly at its lowest (8.97 m^2 ha^{-1}) on and near the 446 permanent sample plots. As the location became further apart from the

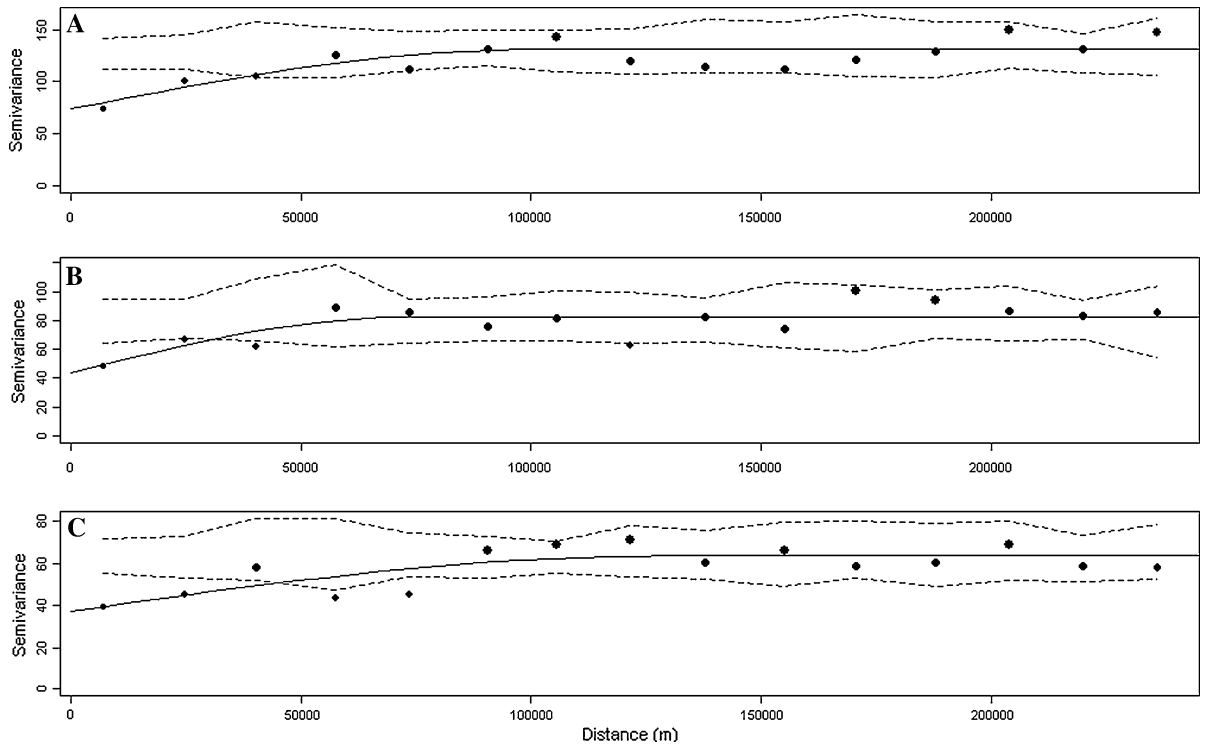

Fig. 4 GSM empirical semivariogram of basal area with the dot size in proportion to the sample size in that range. Empirical semivariograms were calibrated with data from the 446 permanent sample plots using a spherical model and Cressie's robust method for all (**a**), deciduous (**b**), and coniferous species (**c**)

Table 4 Estimated GSM parameters by species group with selected covariates

Model	Covariates				Nugget	Sill	Range (10^3 m)	WSS (10^6)
	Const.	B	F_1	F_2				
Deciduous	0.23	0.41	9.14	−5.21	43.78	38.71	73.75	3.23
(AIC, BIC)			(3237, 3261)					
($AIC_{non\text{-}spatial}$, $BIC_{non\text{-}spatial}$)			(3339, 3359)					
Coniferous	37.71	−0.20	−24.67	−15.12	37.29	26.26	127.93	1.68
(AIC, BIC)			(3075, 3099)					
($AIC_{non\text{-}spatial}$, $BIC_{non\text{-}spatial}$)			(3161, 3182)					
All species	37.84	0.21	−15.26	−20.58	73.42	57.22	98.61	5.90
(AIC, BIC)			(3284, 3309)					
($AIC_{non\text{-}spatial}$, $BIC_{non\text{-}spatial}$)			(3451, 3471)					

WSS weighted sum of squares value

sample plots, the standard deviation increased and peaked at 11.29 in the remote western Alaska (Fig. 7b).

According to the geospatial model projection, most of the basal area in 2100 was comprised of coniferous species, with higher basal area in the Tanana Flats south of Fairbanks, the Porcupine River Basin south of Fort Yukon, the Yukon Flats from Tanana to Nulato, and sporadic areas in the Copper River Basin near Glennallen. Low basal area was predicted for eastern Matanuska-Susitna Valley near Wasilla (Fig. 8a, c, e). More confidence could be placed on or near the

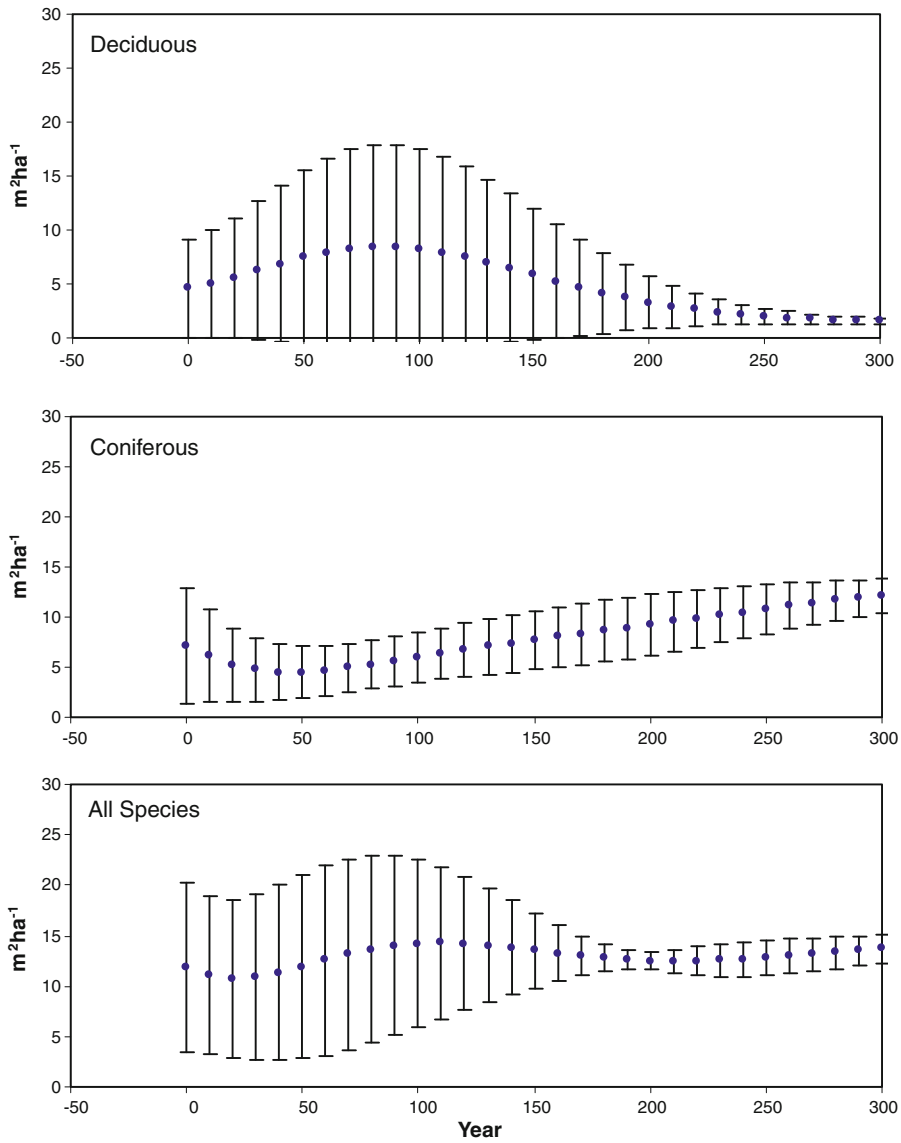

Fig. 5 GSM-projected long-term stand dynamics of 10 randomly selected stands in the Alaska boreal forest region, with vertical bars representing the standard deviation of the mean of the 10 stands. Predicted basal area converged to a steady state despite the huge variation in the initial states of the 10 stands

permanent sample plots where the standard deviation of predicted mean was lower than other areas (Fig. 8b, d, f).

Discussion

Forest states are not necessarily stationary or even continuous. However, this study shows that when the factors causing the spatial non-stationarity have been accounted for, forest states could become a stationary process. *Intrinsic stationarity* as a major modeling and mapping assumption implies that future forest states over the region may be a conceivable and spatially stationary process when exogenous factors have been controlled for. To this end, the essential questions are: are future forest states over the region conceivable, i.e. predictable, and is this process spatially stationary?

The first question was addressed by model validation (§Model validation). Future forest states were shown to be predictable as the post-sample projection of Bonanza Creek largely matched the observed map (Fig. 6) with an 86 % *overall accuracy*. The second question was addressed in §Model estimation and validation. The empirical variance of forest states displayed an obvious pattern of spatial autocorrelation (Figs. 3, 4), for which

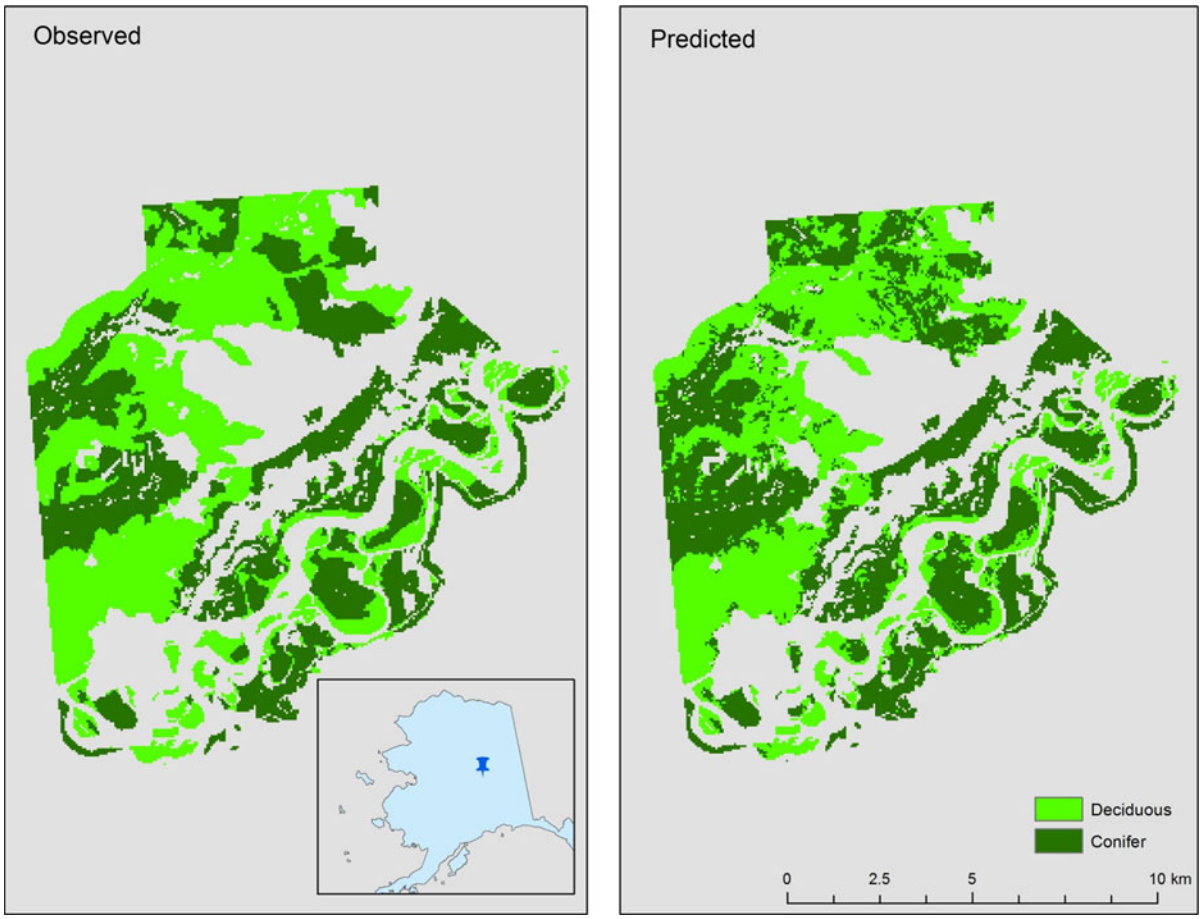

Fig. 6 Comparison between observed (*left*) and GSM predicted (*right*) vegetation map of Bonanza Creek in 2007. Observed species distribution comes from the reconstructed Color-infrared QuickBird imagery (Baird 2011). Inset pinpoints the relative location of Bonanza Creek in Alaska

Table 5 Error matrix for the comparison between observed and predicted species distribution maps (Fig. 6)

	Deciduous obs.	Conifer obs.	Predicted total	User's accuracy
Deciduous pred.	41,221	3,241	44,462	**93 %**
Conifer pred.	9,648	39,454	49,102	**80 %**
Observed total	50,869	42,695	93,564	
Producer's accuracy	**81 %**	**92 %**		

This table summarizes the number of pixels predicted to be in a particular species category (rows) relative to the actual observed number of pixels in that species category (columns)

spherical variogram models (Figs. 3, 4) were highly significant with better goodness-of-fit. When exogenous factors and the semivariogram are simultaneously taken into account, the projected maps of future forest states (Figs. 8, 9) are clearly patchy and discontinuous, resembling expected natural patterns.

Ecological implications

The geographic distribution (Fig. 9a) of basal area change from 2001 to 2100 indicated that moisture limitation may have a major effect on the stocking of the Alaska boreal forest. Most of the areas with positive basal area change were estimated to locate along the three major regional rivers—the Copper River, the Tanana River, and the Yukon River passing through Glennallen, Fairbanks, and Fort Yukon, respectively. The higher basal area increment rate of these areas, presumably due to a combination of natural through fall and capillary rise of soil water

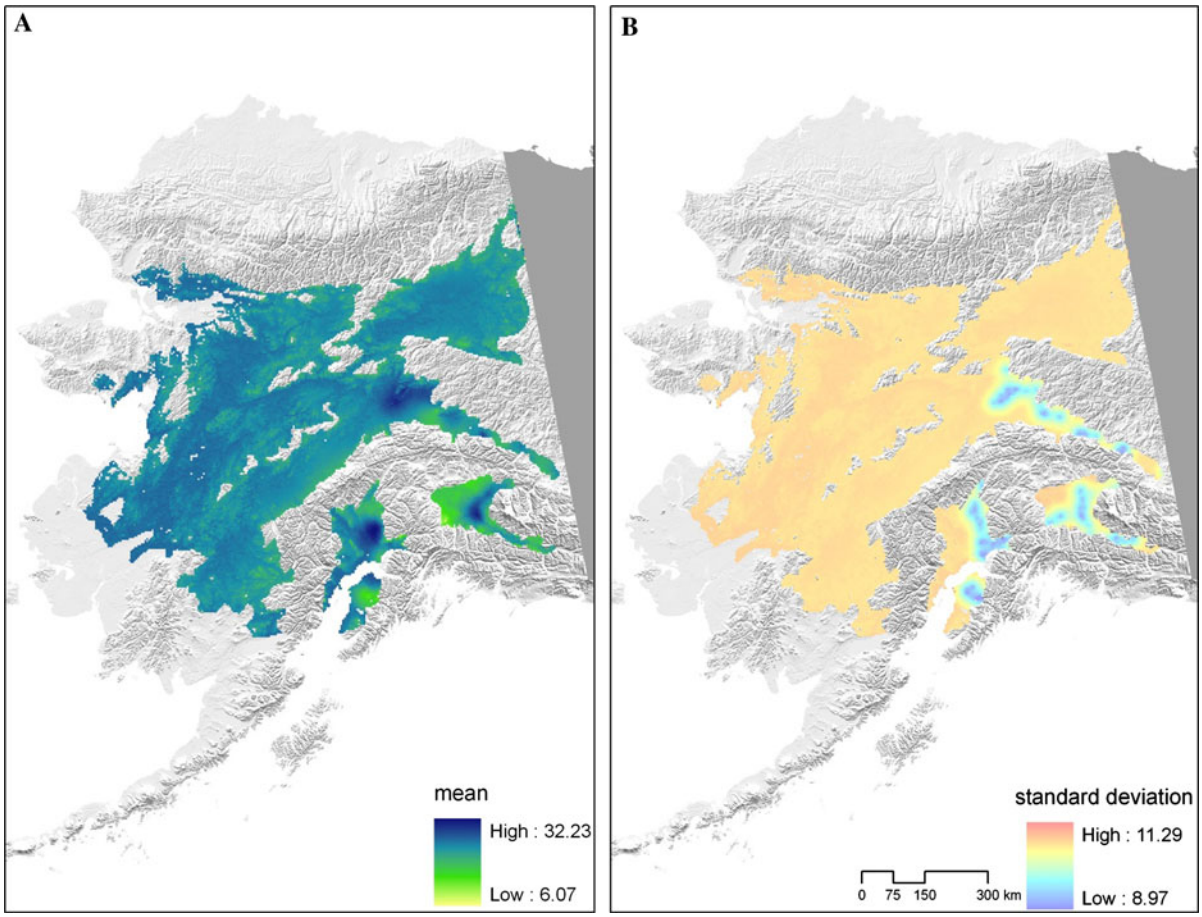

Fig. 7 Estimated basal area (m^2 ha^{-1}) distribution at year 2001 across the Alaska boreal forest region. Spatial estimation was performed using universal kriging (UK) with the sample basal area and DEM topographic (i.e. elevation, slope, and aspect) data

associated with the proximity of rivers, agreed with a recent finding that both deciduous and coniferous species displayed higher growth rates on floodplain than in upland control sites (Yarie 2008). Additionally, moisture-abundant lowland forests could have a higher stocking increment rate than dryer areas where drought-related disturbance agents, such as spruce bark beetle (Wermelinger 2004) and wildland fires (Johnstone et al. 2010) may further reduce basal area increment rates.

Another interesting observation was that the four major areas of negative stocking increment were located near major urban developments (Fig. 9a). I hypothesized that inappropriate human interference could impose a negative impact on stocking increment rate of the Alaska boreal forest. More specifically, the largest area of negative basal area change located in the east Matanuska-Susitna Valley near Wasilla and Anchorage—the two largest cities of Alaska—was likely caused by extensive historical removal of pole-sized coniferous trees in the area (Department of Community Development 2008). With this inappropriate high-grading practice, the residual stands, dominated by over-mature deciduous species, were destined to decline in basal area until coniferous species become predominant. The other areas of negative stocking increment, also located near major urban developments of Fairbanks, Delta Junction, and Glennallen, could also be related to inappropriate harvesting.

Management implications

The geospatial model digital map database was currently composed of estimated stand basal area by species at year 2100 throughout the Alaska boreal

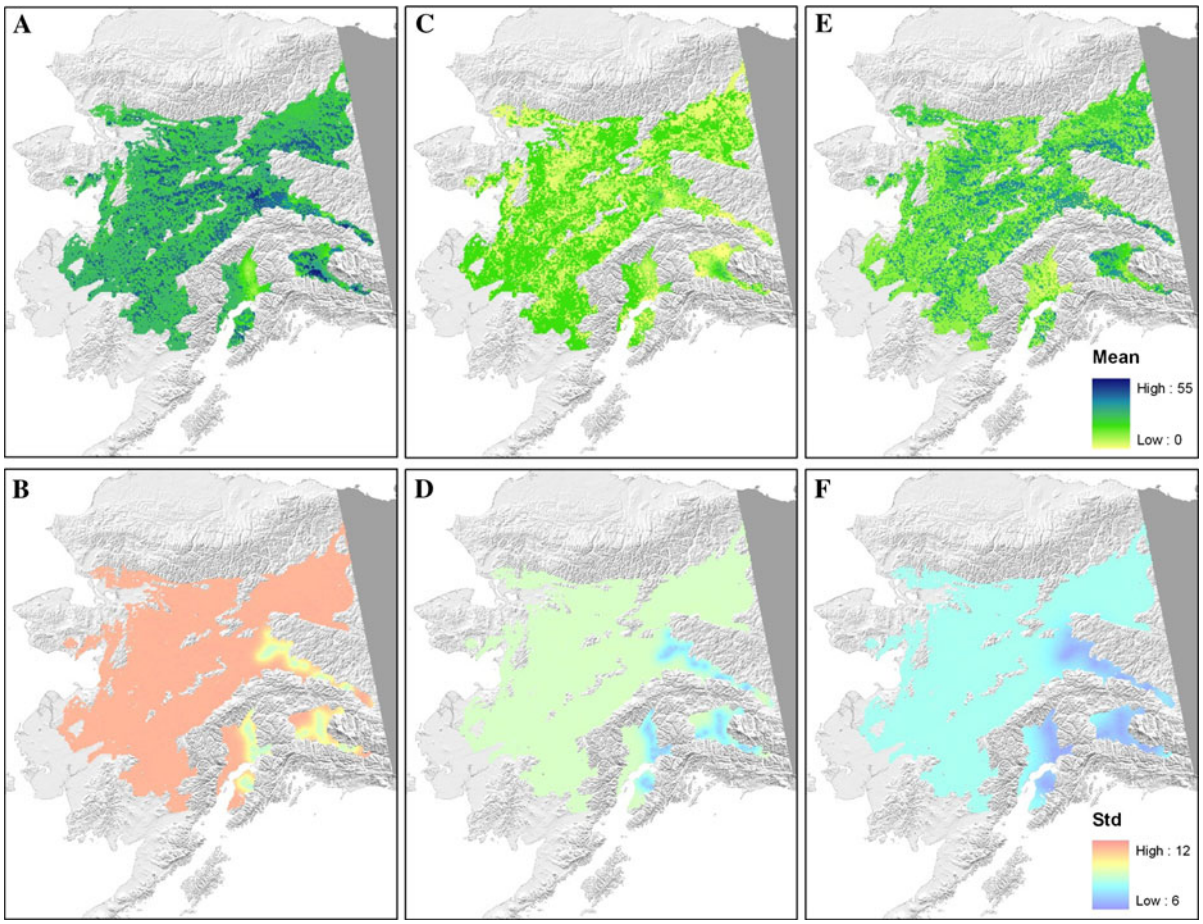

Fig. 8 GSM projected mean basal area (m^2 ha^{-1}) of all (**a**), deciduous (**c**), and coniferous species (**e**), with estimated standard deviation (m^2 ha^{-1}, **b**, **d**, **f**) across the Alaska boreal forest region at year 2100

forest region. The region was divided into 322,255 2-km pixels, and the estimated attribute values of these pixels at 2100 were stored in a single matrix (Table 6) and were further converted into a GIS layer representing forest stocking at 2100. The size of this matrix could be easily extended to accommodate the expansion of geographic coverage or additional attributes. Future temporal extension of the database, however, needs to be placed into separate matrices with each matrix representing a different future point in time, which would be more costly in terms of processing time and storage space than a geographic extension.

The digital map database created from this model has a high potential in facilitating forest management, especially at a large-scale. Stand-specific data of current and future species distribution and stocking, ready to be retrieved with stand coordinates, could assist the development of spatial-specific management regimes and greatly reduce the time and cost in decision-making. For instance, stands at 64.17°N 145.85°W (WGS84) were expected with increasing stocking under natural succession, but stands at 59.73°N 155.10°W may need a considerable thinning, as they were predominately coniferous, highly flammable, and were expected to lose stocking over time (Table 6).

Caution is advised for applying the present model to forest management. The issues are twofold. First, the present model was only able to project future stocking, species, and diversity, but the detailed stand attributes, such as diameter distribution and timber quality have being overlooked. Secondly, with the low confidence level for a majority of the region (Fig. 8b, d, f), the present model may only provide a vague guideline for most remote areas until future inventories will be expanded to cover the entire region.

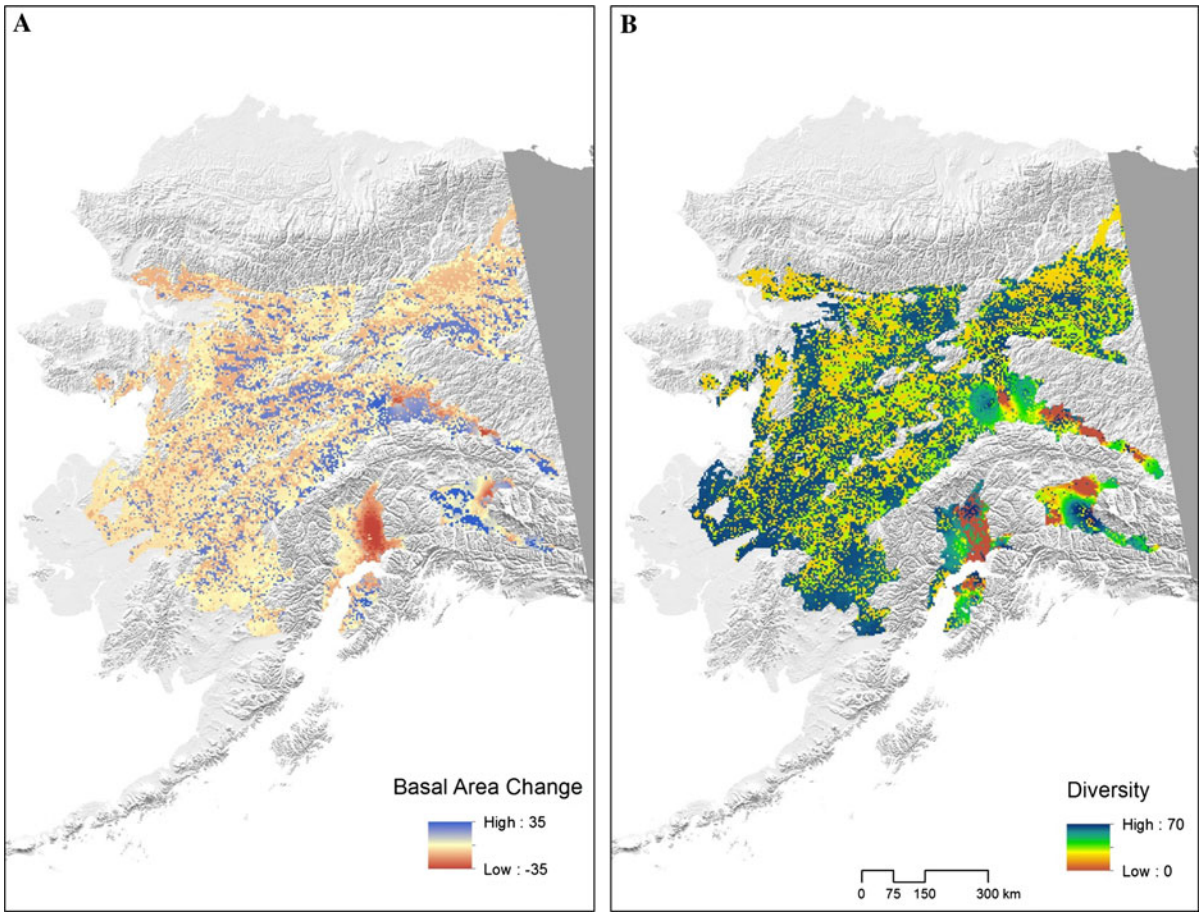

Fig. 9 GSM projected 100-year basal area change (m^2 ha^{-1}, **a**) and species diversity at year 2100 (**b**) across the Alaska boreal forest region

Model advantages and limitations

The geospatial model was of two major advantages over previous studies. First, the geospatial model enabled spatially explicit prediction of forest dynamics, greatly facilitating forest management projects at various scales, as spatial-specific forest population structures at stand, landscape, and regional levels could be instantly retrieved from the geospatial model digital map database.

As another major advantage, the geospatial model accounted for both regional trend and spatial autocorrelation. Previous spatial studies typically focus on either trend surface (e.g. Gittins 1968; Kuuluvainen and Sprugel 1996; Liang and Zhou 2010) or autocorrelation (e.g. Bolstad et al. 1998; Pfeffer et al. 2003; Miller 2005). Trend surface studies overlook spatial autocorrelation, whereas autocorrelation studies often have to assume constant trend (geostatistical interpolation models, autoregressive models, or geographically weighted regression models) or a trend local to each pixel (universal kriging). The geospatial model features an approach to incorporate the controlled trend surface into the geostatistical model of universal kriging. This approach maintained the accurate regional trend surface estimates (Liang and Zhou 2010), and mitigated the seemingly artificial patterns from the trend surface estimation (see Liang and Zhou 2010, pp. 2351).

The assumption of *intrinsic stationarity* could be violated by the shifts of environmental conditions such as climate change. This issue is admittedly difficult to address, as environmental conditions vary across both space and time, and it is a challenging task to account for temporal and spatial changes of climatic conditions across the region simultaneously. Although climate-

Table 6 Sample data sections of the GSM digital map database

Pixel	Coordinates	$B_{deciduous}$		$B_{coniferous}$	
		Mean	Std	Mean	Std
80	(−141.18, 68.60)	18.0275	9.0905	8.8456	8.2413
81	(−141.17, 68.60)	18.0275	9.0905	8.9509	8.2413
120	(−141.05, 68.56)	2.8105	9.0920	19.7199	8.2422
121	(−141.04, 68.56)	2.8105	9.0920	20.0484	8.2422
…	…	…		…	
178334	(−145.86, 64.17)	2.6158	7.4235	38.9369	6.6862
178335	(−145.85, 64.17)	2.6518	7.4097	38.6419	6.6842
178336	(−145.83, 64.17)	2.6786	7.4125	38.7153	6.6864
…	…	…		…	
211310	(−143.24, 63.35)	6.2072	7.4062	24.3193	6.5826
211311	(−143.22, 63.35)	6.0426	7.3772	24.6387	6.5695
211312	(−143.20, 63.35)	5.8415	7.3540	24.8751	6.5584
…	…	…		…	
322253	(−155.13, 59.73)	18.0275	9.0905	11.4574	8.2413
322254	(−155.12, 59.73)	18.0275	9.0905	11.4574	8.2413
322255	(−155.10, 59.73)	2.8105	9.0920	22.6265	8.2422

The basic database consisted of simulated future stocking of 322,255 2-km pixels of Alaska at year 2100, and can be further extended in space and time. Coordinates were in WGS84 Geodetic System

sensitive models of forest dynamics (e.g. Yan and Shugart 2005; Scheiter and Higgins 2009; Liang et al. 2011) could project future states of the sample points under climate change, it would be too risky to assume spatial stationarity across the region in future years, as geographic distribution of climatic conditions is known to change over time (*c.f.* IPCC 2001). Therefore, the present model should only be used as a baseline tool for the estimation of future forest dynamics. To this end, future studies could benefit from a spatio-temporal modeling framework that could simultaneously account for the temporal and spatial gradients in forest dynamics under the influence of climate change.

The present geospatial model was developed only to help approximate forest stocking for the locations with very limited data, and it is by no means meant to replace forest dynamics models in providing accurate estimation of forest stocking. The results were of less confidence as the study area distanced from the sampled area. This major limitation was reflected in Figs. 7 and 8 which show increased standard deviation of predicted means in remote areas. With the confidence level being quite low for a majority of areas in Alaska boreal region (Fig. 8b, d, f), the results from the current study should be used cautiously outside the sample areas. Nevertheless, the present model should be useful because not only does it provide projections to cover a region where most areas are of no data, it also presents the geographic distribution of confidence levels for the projection.

Conclusion

This paper presents a geospatial model, which accounted for both regional trend and spatial autocorrelation, to estimate dynamics of the Alaska boreal forest from existing forest inventory data in limited areas across the region. The spatially explicit estimates were aggregated into a digital map database to facilitate forest management at both micro and macro levels. The geospatial model accounted for both regional trend and spatial autocorrelation by incorporating the controlled trend surface into the geostatistical model of universal kriging. Not only did this approach maintain the accurate regional trend surface estimates, it also mitigated the seemingly artificial patterns of the trend surfaces by controlling for spatial autocorrelation. The accuracy and reliability of this

model were verified with two information criteria and an evaluation of long-term predictions and projected vegetation map. Precision of the projections (in terms of standard deviation) is admittedly low for the never inventoried areas in comparison to those with inventory data (Figs. 7, 8). However, for the majority of the region where little information is available, the results could be of unprecedented value.

With areas of the best basal area increment rate clustered along the major rivers and the worst ones near the four major urban developments in Alaska, it was hypothesized that moisture limitation and inappropriate human interference were the main factors affecting the stocking rates of the Alaska boreal forest. Forest management in Alaska was suggested to employ spatial-specific regimes to accommodate the geographic variation of forest dynamics. Despite limitations commonly associated with empirical geospatial models, the present model provided a useful baseline tool for the study of forest dynamics and management.

Acknowledgments The author is obligated to Dr. Dave Verbyla and Randy Peterson for their assistance with spatial analysis and mapping. Bonanza Creek species distribution map is kindly provided by Dr. Dave Verbyla. The author also wants to thank Dr. Mo Zhou, Dr. John Yarie, and Dr. Dave Valentine for their helpful comments. This research is supported in part by the Division of Forestry and Natural Resources, West Virginia University, USDA McIntire-Stennis Act Fund ALK-03-12, and the School of Natural Resources and Agricultural Sciences, University of Alaska Fairbanks.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman, Essex

Baird B (2011) Spatial and temporal trends in vegetation index in the Bonanza Creek experimental forest. University of Alaska Fairbanks

Bolstad PV, Swank W, Vose J (1998) Predicting Southern Appalachian overstory vegetation with digital terrain data. Landscape Ecol 13:271–283

Cipar J, Cooley T, Lockwood R, Grigsby P Distinguishing between coniferous and deciduous forests using hyperspectral imagery. In: Geoscience and remote sensing symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE international, 2004, vol 4, pp 2348–2351

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ 37:35–46

Cressie NAC (1993) Statistics for spatial data. J. Wiley, New York

Department of Community Development M (2008) Forest management plan. Matanuska-Susitna Borough, Palmer, AK, p 14

Gittins R (1968) Trend-surface analysis of ecological data. J Ecol 56(3):845–869

Gorelick R (2006) Combining richness and abundance into a single diversity index using matrix analogues of Shannon's and Simpson's indices. Ecography 29(4):525–530

IPCC (2001) Climate change 2001: the scientific basis. In: Houghton JT, Ding Y, Griggs DJ et al (eds) Contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, p 881

Johnstone JF, Hollingsworth TN, Chapin FS, Mack MC (2010) Changes in fire regime break the legacy lock on successional trajectories in Alaskan boreal forest. Glob Change Biol 16:1281–1295

Kuuluvainen T, Sprugel DG (1996) Examining age- and altitude-related variation in tree architecture and needle efficiency in Norway spruce using trend surface analysis. For Ecol Manage 88:237–247

Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? Ecology 74(6):1659–1673

Legendre P, Oden NL, Sokal RR, Vaudor A, Kim J (1990) Approximate analysis of variance of spatially autocorrelated regional data. J Classif 7:53–75

Liang J (2010) Dynamics and management of Alaska boreal forest: an all-aged multi-species Matrix growth model. For Ecol Manage 260(4):491–501

Liang J, Zhou M (2010) A geospatial model of forest dynamics with controlled trend surface. Ecol Model 221(19):2339–2352

Liang J, Picard N (2012) Matrix model of forest dynamics: an overview and outlook. For Sci. doi:10.5849/forsci.11-123

Liang J, Buongiorno J, Monserud RA (2005) Growth and yield of all-aged Douglas-fir/western hemlock stands: a Matrix model with stand diversity effects. Can J For Res 35(10):2369–2382

Liang J, Buongiorno J, Monserud RA, Kruger EL, Zhou M (2007) Effects of diversity of tree species and size on forest basal area growth, recruitment, and mortality. For Ecol Manage 243:116–127

Liang J, Zhou M, Verbyla D, Zhang L, Springsteen AL, Malone T (2011) Mapping forest dynamics under climate change: a matrix model. For Ecol Manage 262:2250–2262

Malone T, Liang J, Packee EC (2009) Cooperative Alaska Forest Inventory. Gen. Tech. Rep. PNW-GTR-785, USDA Forest Service. Pacific Northwest Research Station, Portland, p 42

Matheron G (1969) Le krigeage universel. Cahiers du Centre de Morphologie Mathematique. Ecole des Mines de Paris, Fontainebleau

Miller J (2005) Incorporating spatial dependence in predictive vegetation models: residual interpolation methods. Prof Geogr 57(2):169–184

Miller J, Franklin J, Aspinall R (2007) Incorporating spatial dependence in predictive vegetation models. Ecol Model 202:225–242

Mitchell TD, Carter TR, Jones PD, Hulme M, New MG (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record

(1901–2000) and 16 scenarios (2001–2100) Tyndall Working Paper. Tyndall Centre, Norwich, p 30

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on earth. Bioscience 51(11):933–938

Pfeffer K, Pebesma E, Burrough P (2003) Mapping alpine vegetation using vegetation observations and topographic attributes. Landscape Ecol 18(8):759–776

Phelps JE (2005) Emerging industry in Alaska's boreal forests: a hurting timber industry takes on a new edge. Alaska Business Monthly Oct-2005

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Ribeiro PJ Jr, Diggle PJ (2001) geoR: a package for geostatistical analysis. R-NEWS 1(2):ISSN 1609-3631

Rupp TS (2008) Projected vegetation and fire regime response to future climate change in Alaska. Preliminary Report prepared for U.S. Fish and Wildlife Service National Wildlife Refuge System. University of Alaska Fairbanks, Fairbanks, AK

Scheiter S, Higgins SI (2009) Impacts of climate change on the vegetation of Africa: an adaptive dynamic vegetation modelling approach. Glob Change Biol 15(9):2224–2246

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Stage AR, Salas C (2007) Interaction of elevation, aspect, and slope in models of forest species composition and productivity. For Sci 53(4):486–492

Wermelinger B (2004) Ecology and management of the spruce bark beetle Ips typographus-a review of recent research. For Ecol Manage 202(1–3):67–82

Yan X, Shugart HH (2005) FAREAST: a forest gap model to simulate dynamics and patterns of eastern Eurasian forests. J Biogeogr 32:1641–1658

Yarie J (2008) Effects of moisture limitation on tree growth in upland and floodplain forest ecosystems in interior Alaska. For Ecol Manage 256(5):1055–1063