

Modeling the distribution of white spruce (*Picea glauca*) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas

Bettina Ohse · Falk Huettmann ·
Stefanie M. Ickert-Bond · Glenn P. Juday

Received: 3 March 2009 / Revised: 20 May 2009 / Accepted: 11 June 2009
© Springer-Verlag 2009

Abstract Most wilderness areas still lack accurate distribution information on tree species. We met this need with a predictive GIS modeling approach, using freely available digital data and computer programs to efficiently obtain high-quality species distribution maps. Here we present a digital map with the predicted distribution of white spruce (*Picea glauca*) in Alaska (4 km resolution, accuracy over 90%). Our presented concept represents a role-model for predicting tree species distribution for remote areas world-wide. Although this model intends to be accurate in making predictions rather than to give detailed biological mechanistic explanations, it can also be used as a baseline for further research and testable hypothesis on the importance of the environmental

variables used to build a generalizable model. Further, we emphasize that work like presented here is a pre-condition for assessing human impacts and impacts of climate change on species distribution in a quantitative and transparent fashion, allowing for improved sustainable decision-making world-wide.

Keywords Species distribution models (SDM) · Alaska · Tree species · Predictive modeling · Open access (OA) · White spruce *Picea glauca*

Abbreviations

AGDC	Alaska Geospatial Data Clearinghouse
CODATA	Committee on Data for Science and Technology
ESRI	Environmental Systems Research Institute
FGDC	Federal Geographic Data Committee
FIA	Forest inventory and analysis
GIS	Geographic information system
ICSU	International Council for Science
IDW	Inverse distance weighting
IPY	International Polar Year
NAD83	North American Datum of 1983
NBII	National Biological Information Infrastructure
NDVI	Normalized difference vegetation index
NSF	National Science Foundation
OA	Open access
OECD	Organisation for Economic Collaboration and Development
PRISM	Parameter-elevation regressions on independent slopes model
ROC	Receiver operating characteristic
SDM	Species distribution models

Electronic supplementary material The online version of this article (doi:10.1007/s00300-009-0671-9) contains supplementary material, which is available to authorized users.

B. Ohse (✉)
Institute of Botany and Landscape Ecology,
University of Greifswald, Grimmer Straße 88,
17489 Greifswald, Germany
e-mail: bettina.ohse@uni-greifswald.de

B. Ohse · G. P. Juday
Forest Science Department, University of Alaska Fairbanks,
Fairbanks, AK 99775, USA

F. Huettmann
EWHALE Lab, Biology and Wildlife Department,
Institute of Arctic Biology, University of Alaska Fairbanks,
Fairbanks, AK 99775-7000, USA

S. M. Ickert-Bond
Herbarium and Department of Biology and Wildlife,
Institute of Arctic Biology, University of Alaska Museum
of the North, 907 Yukon Dr., PO Box 756960, Fairbanks,
AK 99775-6960, USA

Introduction

The distribution of tree species in remote areas is currently not well known. Most areas in mountainous regions and remote islands for instance still suffer from a lack of detailed plant species distribution mapping. This is also true for Alaska. To date, literature on Alaska's tree species is only available as atlases with coarse range maps (Hultén 1968; Viereck and Little 2007). However, reliable information even about simple presence or absence within this broad range of a single species cannot be sufficiently detected from such maps. Furthermore, species distribution information is given in eco-classifications (Viereck et al. 1992) and in forestry related articles (e.g. Farr and Harris 1979; LaBau and Alden 2000). Several maps show ecosystem or vegetation types of Alaska (Fleming 1997; Gallant et al. 1995), however, they are plant communities and do not show single species distributions as such. Articles concerned with single tree species either report on a certain region within Alaska (Hennon and Trummer 2001; Murray 1980) or focus on species genetics (Viereck and Foote 1970). An exhaustive report on single species' ranges could not be found. Alaska is the home of over 15 tree species, and such species are of great interest for the assessment of wildlife habitat, vegetation type classification, and adaptive resource management. As large areas of Alaska are very difficult to access, there is a need for advanced approaches to mapping tree species. Here we investigate a predictive modeling approach that uses publicly available tools, data and environmental variables to predict tree species distribution promising a high accuracy.

Need for a species distribution model (SDM) of trees in Alaska

Plant–climate-relationships and the importance of various other environmental factors for the geographical distribution of plant species have been recognized early (Whittaker 1967) and are widely used to explain biogeographical patterns (e.g. Ellenberg 1988; Walter 1985). SDMs use these concepts to determine the ecological niche of a species based on several environmental variables. The ecological niche can be projected into geographical space, resulting in a predictive map of the species' distribution (Franklin 1995; Tsoar et al. 2007). SDMs are widely applied for the study of plant species distribution (e.g. Engler et al. 2004; Franklin 1998; Guisan et al. 1998). They are a crucial tool for obtaining better maps, which are needed to facilitate further research on the species themselves (Parviainen et al. 2008), for developing informed hypotheses on wildlife and habitat (Guisan and Zimmermann 2000), for classifying plant communities and for assessing their change in composition or distribution (Ferrier and Guisan 2006; Zimmermann and

Kienast 1999), and also for improving ecological theory and knowledge (Dunning et al. 1995). Furthermore, maps derived through predictive modeling are used to improve floristic and faunistic atlases (Araújo et al. 2005; Prasad et al. 2007-ongoing), to assess the impact of land-use change (Dunning et al. 1995) or to help decide on conservation priorities (Margules and Austin 1994). They are an inherent tool in modern Adaptive Management (Huettmann 2007; Walters 1986). Developing SDMs using publicly available data is easier, faster, and less expensive than mapping in the field. A more detailed literature overview on the use of SDMs and what it entails can also be found in Guisan and Thuiller (2005).

Concept of open access in predictive modeling

Open access (OA) offers an improved principle of sharing high-quality scientific information among scientists as well as with the global public. Also, it makes scientific methods transparent and repeatable to everybody, which should add to its credibility and increased trustworthiness. This concept becomes available due to recent advances in computing, databases and online data delivery. OA is a recent movement that is virtually promoted globally by ICSU, OECD, CODATA, NSF, the European Union as well as by global policies such as the Rio Convention, and mega-science programs such as the IPY. Latest publicly funded science in the US and Canada is based on such OA principles and becomes a requirement for publication and funding (National Research Council 2003; Interagency Working Group on Digital Data 2009). This paper provides a further example of applied OA principles. A list of free available datasets and tools used in this study can be found in Table 1.

White spruce

Our modeled species, white spruce (*Picea glauca* [Moench] Voss), is one of the most common tree species in Alaska, and is of ecological and commercial importance, occupying app. 25% (121,000 km²) of Alaska's boreal forest (Labau and van Hees 1990). It has good overall data available, but suffers critical data gaps throughout its Alaska and wilderness distribution. White spruce occurs both in floodplains and uplands (Viereck et al. 1986; Walker et al. 1986). It is the dominant treeline species in the two main mountain ranges of Alaska, and forms large stands in the highlands of Interior Alaska (Juday et al. 1999), occurring from 100 m to treeline (300–1,600 m) (Viereck and Little 2007). White spruce appears to grow best on south-facing slopes and well-drained, sandy soils along the edges of lakes and rivers, but not in areas with

Table 1 Open access datasets and tools used in this study

Dataset or software tool	Source	Service
Species data	http://arctos.database.museum/home.cfm	Online information on species of herbarium collection
Climate data	http://www.prism.oregonstate.edu/	Spatial climate research, education, analysis and mapping data services
Data on elevation, soil, permafrost, and surface geology	http://agdc.usgs.gov/data/projects/fhm/	Spatial resource data base, data meet FGDC meta-data standards
Hawth tools ^a	http://www.spatial ecology.com/htools/	Free extension for ESRI's ArcGIS (specifically ArcMap)
TreeNet ^b	http://salford-systems.com/treenet.php	Statistical analysis tool (90 days trial version)

^a Instead of ArcGIS, free software (e.g. GRASS GIS) could be used

^b For free long-term use (without a Graphical User Interface) also available as R-package (Generalized Boosted Models), see <http://cran.r-project.org/web/packages/gbm/index.html>

continuous permafrost (Viereck et al. 1992). White spruce is known to be an important habitat for moose (MacCracken and Viereck 1990; Risenhoover 1989), red squirrel (Brink and Dean 1966; Smith 1968), marten (Buskirk 1984; Slough 1989), and hare (Sinclair et al. 1988; Wolff 1978). White spruce also plays an important role in local timber production and fuel supply (Holsten et al. 1991; Viereck and Little 2007). However, maps on single tree species distribution can hardly be found, and even the FIA database does not contain information beyond south-east and south-central Alaska.

Methods

Datasets

Our training dataset consisted of 108 confirmed white spruce presence datapoints, available for this species as biogeoreferenced records (online at Arctos Multi-Institution, Multi-Collection Museum Database, University of Alaska Museum Herbarium). The points represent samples dating from 1900 to 2000 and 85% have a location uncertainty of c. 3,615 m (horizontal and vertical datums unknown). Thus, we used a buffer of 3,615 m radius for each point (equivalent pixel size: 6,407 m × 6,407 m). As the Arctos Database does not include absence data, we created 600 pseudo-absence points (Engler et al. 2004; Tsoar et al. 2007) using the publicly available Hawth's tools random sample tool in ArcGIS 9.2 (see Table 1). More specifically, half the points were randomly distributed all over Alaska, the other half only within non-forest vegetation types according to a digital vegetation cover map (Fleming 1997, online available from the AGDC), in order to obtain more absence points in areas where absence is more likely. In a multi-hypothesis fashion (sensu Burnham and Anderson 1998), we tested 24 environmental

variables and latitude and longitude as potential predictors for the distribution of white spruce (see Table 2 for complete list of predictors). For climate data we used the datasets on Alaska average monthly precipitation/monthly mean temperature, 1961–1990, by C. Daly (2 km × 2 km raster data, provided by PRISM). Elevation, aspect, and slope came from AGDC as 1 km × 1 km raster, aspect was used as a continuous variable, which ensures more transparency and accuracy than the traditional use of categories. We also used permafrost, soil, and surface geology (polygon data) from AGDC. In order to extrapolate the model results evenly to a large area, a regular grid was created in Hawth's Tools for the entire state of Alaska, carrying an even point spacing of 4 km.

Model

A model approach described in Fig. 1 was used to predict the distribution of white spruce, but is intended to represent a role-model for predicting any tree species distribution for remote areas anywhere in the world. The datasets were overlaid in ArcGIS 9.2 (step 1) and transformed to a consistent projection (Alaska Albers, geographic datum: NAD-83 Alaska). The values of each layer were extracted to the buffered presence points (mean values from raster datasets, prevailing class from polygon datasets) as well as to the 600 pseudo-absence points, resulting in a table with presence/absence as a response and climatic and bioclimatic variables as predictors. The environmental parameter values were also extracted to the regular grid (step 2).

For modeling the associations between the tree species and its environmental predictors, we favoured non-parsimonious (with 'parsimonious' referring to approaches based on few preselected determinant values) and non-linear modeling. Hence, we applied machine learning concepts, such as classification trees (Breiman et al. 1984; Breiman 2001) to obtain best possible predictions. These

Table 2 Environmental variables used for developing the models

Variable	Description
tmean_01	Mean temperature January ($^{\circ}\text{C} \times 10$)
tmean_04	Mean temperature April
tmean_05	Mean temperature May
tmean_06	Mean temperature June
tmean_07	Mean temperature July
tmean_08	Mean temperature August
tmean_09	Mean temperature September
tmean_g	Mean temperature growing season (May through September)
t_juljan	Temp difference July–January (means)
tmean_ann	Mean annual temperature
ppt_04	Precipitation sum April (mm)
ppt_05	Precipitation sum May
ppt_06	Precipitation sum June
ppt_07	Precipitation sum July
ppt_08	Precipitation sum August
ppt_09	Precipitation sum September
ppt_grow	Precipitation sum growing season (May through September)
ppt_ann	Annual precipitation sum
Elevation	Elevation (m)
Aspect	Aspect ($^{\circ}$)
Slope	Slope ($^{\circ}$)
Permafrost	Permafrost (classes)
Soil	Soil (classes)
Sufgeol	Surface geology (classes)

methods account for complex ecological and environmental interactions between variables (Guisan et al. 2006; Lawler et al. 2006), and even when using noisy data (Craig and Huettmann 2008) they show high performance with fine and coarse resolution datasets (Guisan et al. 2007). We used the boosting classification and regression tree software TreeNet (SalfordSystems, San Diego, CA, USA) to analyze the data and to build the model (Hastie et al. 2001; Friedman et al. 2000).

After initial testing, and for obtaining best results, we used TreeNet with the following settings: three nodes per tree, minimum number of six observations per terminal node, 100-fold cross validation (to ensure high model stability), and the option ‘balanced’ for equal weight of number of presence and absence points (Maggini et al. 2006). Here we followed the concept of using informed default settings, as promoted in Blackbox modeling for ease and convenience. It is known that this approach with its settings in most cases helps to achieve good modeling results in a fast and reliable manner (Craig and Huettmann 2008). Obtaining good but time-critical results is usually crucial for management-related applications as provided

here. First, we ran nine basic models (Table 3, models 1–9) to compare effects of temperature/precipitation with those of soil characteristics. We then compared ROC values and percent of correctly predicted presences (misclassification threshold 0.5, hereafter referred to as %corr).

From these exploratory runs, the models with $\text{ROC} > 0.75$ and $\% \text{corr} > 0.65$ (models 3, 5, 6, 7) were kept and slightly modified by dropping climate variables that consistently fell into the lower half of the variable ranking lists of most of the models (models 3, 5, 7), or were considered not important by TreeNet variable ranking (models 8, 9). Thus, we derived nine more models to further improve ROC values and %corr, and finally, the four best-performing models were chosen (models 6, 12, 14, 15, step 3). These four models were applied within TreeNet in order to predict presence/absence to the regular grid (step 4). The predicted value of relative occurrence for each gridpoint was mapped and points were interpolated using the IDW (Inverse Distance Weighting) tool. Thus, four maps with the statewide index of relative occurrence of white spruce were obtained (step 5). The concept followed principles described by Huettmann and Linke (2002), batch files of the TreeNet runs are available on request.

Accuracy assessment

Assessing the accuracy of a spatially explicit model means assessing prediction errors and spatial uncertainties. For a first comparison of models and their accuracies we interpreted ROC curves, derived from cost matrices (Bradley 1997; Fielding and Bell 1997). For a more detailed assessment, predicted map values (predicted index of relative occurrence) were compared to evaluation data points taken from four independent datasets with ‘presence only’ data (Fig. 1, step 6). As a measure of model performance, we found the Boyce index to be most suitable (Boyce et al. 2002), as it is independent of the prediction’s threshold between presence and absence and it is based on evaluation data using presence-only (Hirzel et al. 2006). The Boyce index F_i is an area adjusted frequency index. Lower habitat suitability classes should have F_i values < 1 (less evaluation points than with a random distribution) and high habitat suitability classes should have F_i values > 1 (more evaluation points than with a random distribution). F_i was then plotted against the mean index of relative occurrence for each class, resulting in a curve that is monotonically increasing for a model with high accuracy, and monotonically decreasing for models with low accuracy. Performing a (non-parametric) Spearman’s rank correlation of the Boyce indices of all classes versus the mean index of relative occurrence of each class provides an estimate about how stable the prediction of the specific class is, compared to the overall prediction accuracy of the model (step 7),

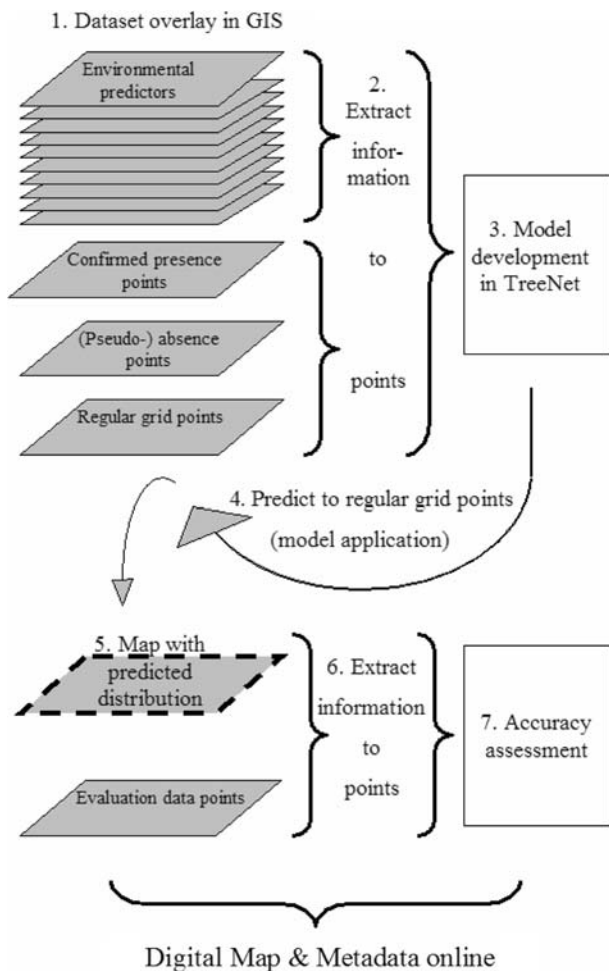


Fig. 1 Flowchart illustrating the concept of model development used in this study and proposed as a role model for future models

while the overall prediction accuracy of the model can be assessed by the Spearman's rank correlation coefficient r_s .

Data management

For this study, we used OA tools and data as this has many advantages and implications for studies and project goals like ours. Most of these data proved to be of sufficient and reliable quality and carried high-quality metadata (all climate data and elevation dataset). Only some data came with very basic descriptions (species data, soil, permafrost, surface geology) and details had to be requested by email. We operated these data on a PC within Excel and ArcGIS 9.2 and with the help of additional free tools (Hawth's tools). GIS data are presented in grids and shapefiles. Metadata were created within ArcCatalog and the freely available Metavist XML editor, and made globally available at the National Biological Information Infrastructure website (NBII, <http://mercddev3.ornl.gov/nbii/>). All data formats we used are supported by OpenGIS and OpenOffice.

Results

Model ranking with TreeNet

Model ranking was done by comparing ROC values and %correctly predicted presences (Table 3). Models 1–9 (exploratory runs) obtained relatively low ROC and %corr values (<0.8 and <0.7, respectively). Two exceptions were models 6 and 7, with model 6 (only elevation, aspect, and slope) reaching a slightly higher ROC value (0.806) and the highest value for %corr compared to all other 17 models (84.47). Model 7 (same as model 6 plus all temperature and precipitation variables) reached an even higher ROC value (0.869), but a lower %corr value (78.64). Model 10 showed only slightly improved values. The three models 11, 12, and 13 ranked highest according to the ROC values (all 0.875), with model 12 having the highest value for %corr compared to all other 17 models (79.61). Models 14 (improvement of model 11 by adding lat and long) and model 15 (improvement of model 11 by adding permafrost) scored with relatively high ROC (both 0.871) and %corr values (77.67 and 75.73, respectively). Adding soil (model 16), surfgeol (model 17), or permafrost + surfgeol (model 18) did not improve ROC or %corr values. However, it is worthwhile to point out that the best predictions were not achieved by the most parsimonious model, i.e. the one with the fewest predictors, giving further support for non-parsimonious non-linear model algorithms that can deal with highly complex data. This approach allowed us to identify interactions among variables and to determine systematically the variable combinations with the highest impact.

We chose (1) the model with the highest ROC value (model 12), (2) the one with the highest %corr value (model 6), (3) the best model including lat + long (model 14), and (4) the best one including at least one of the variables permafrost, soil or surface geology (model 15) as models for further consideration. For comparison of the ROC curves for the four most relevant models (6, 12, 14, and 15) see Supplemental Fig. 5a–d.

Variable importance

As an example, the variable importance, as obtained from TreeNet for model 12 (best-performing model) is shown in Table 4. Values represent absolute and relative importance, and thus aid in ranking the variables. The variable importance ranking shows a high contribution of aspect as a predictor variable, as well as total precipitation in August, followed by mean temperature in April. Precipitation sum of April, May and total precipitation sum between May and September are of minor importance, as are mean temperatures in May, September, and June, as well as elevation. The temperature differences between the warmest and

Table 3 ROC values and %correctly predicted presences (misclassification threshold 0.5)

Model no.	Variables included	ROC	%corr
1	All temperature variables	0.726	67.96
2	All precipitation variables	0.722	58.25
3	All temperature and precipitation variables	0.775	66.02
4	Surfgeol, permafrost, soil	0.704	46.30
5	Temperature, precipitation, elevation	0.777	66.99
6	Elevation, aspect, slope	0.806	84.47
7	Temperature, precipitation, elevation, aspect, slope	0.869	78.64
8	All 24 environmental variables	0.718	53.40
9	All 24 environmental variables, lat, long	0.716	53.40
10	tmean_01, 04, 05, 06, 09; ppt_04, 05, 06, 07, 08, grow; elevation	0.777	64.08
11	Same as 10, +aspect	0.875	78.64
12	tmean_04, 05, 06, 09; ppt_04, 05, 08, grow; elevation, aspect	0.875	79.61
13	Same as 12, +slope	0.875	77.67
14	Same as 12, +lat, long	0.871	77.67
15	Same as 12, +permafrost	0.871	75.73
16	Same as 12, +soil	0.772	54.37
17	Same as 12, +surfgeol	0.860	75.73
18	Same as 12, +permafrost, surfgeol	0.863	73.79

All values obtained by TreeNet, internal model test

Table 4 Variable importance (ranking) for model 12 according to TreeNet

Variable	Absolute importance	Relative importance
Aspect	48.70	100.00
ppt_08	28.30	58.11
tmean_04	22.46	46.12
ppt_04	20.12	41.33
ppt_05	19.91	40.88
ppt_grow	18.11	37.20
tmean_05	17.40	35.73
tmean_09	17.12	35.16
Elevation	16.83	34.56
tmean_06	15.75	32.34
tjuljan	14.66	30.10

coldest months are least important. The partial contribution of the variable values to the model can be seen in the single variable plots in Fig. 2a–d. The occurrence of white spruce appears to be favored by warm aspects of 150°–250° (SSE to SWW), whereas cooler aspects of 300°–50° (NWW to NE) appear to inhibit its presence (Fig. 2a). White spruce is also more likely to occur in areas with a total sum of precipitation in August below 75 mm and April mean temperatures above -4°C (Fig. 2b, c). The results from evaluation of the influence of elevation on the distribution of white spruce in Alaska are less clear. While an elevation below 1,000 m has a positive influence on the presence of white spruce (Fig. 2d), an elevation above 1,000 m has a negative influence.

Mapping the predicted distribution

Maps showing the predicted distribution of white spruce in Alaska, as obtained by the four chosen models (6, 12, 14, and 15), showed broad scale consistencies (for comparison see Supplemental Fig. 6). Visual comparison identified constantly low predicted values of relative occurrence for coastal regions, especially in the north and south, and higher values for Interior Alaska. Only some noise occurred on the small scale in the midlatitudes of Alaska, resulting in a ‘salt-pepper’ like pattern, which might indicate true mid-range values overall in the wider region. However, the north-east part of the Interior revealed consistently high values of relative occurrence for all models. The map in Fig. 3 shows the index of relative occurrence of white spruce as predicted by the best-performing model (model 12). Fully in agreement with IPY Metadata & Data Policy, all maps as well as the according metadata are made available, e.g. the IPY data repository of the Global Change Master Directory (<http://gcmd.nasa.gov>).

Model performance

The Boyce index (Fig. 4a) revealed similar patterns for models 12, 14, and 6 for low and middle classes, with all $F_{1-6} < 2$, but differing patterns for higher classes (F_{7-10}), with model 6 entirely omitting classes 9 and 10. In contrast to the broad pattern, only model 15 showed a more fluctuating curve for the Boyce indices. The spearman’s rank correlation for models 12, 14, and 6 showed that predictions for classes of high and low relative occurrence were

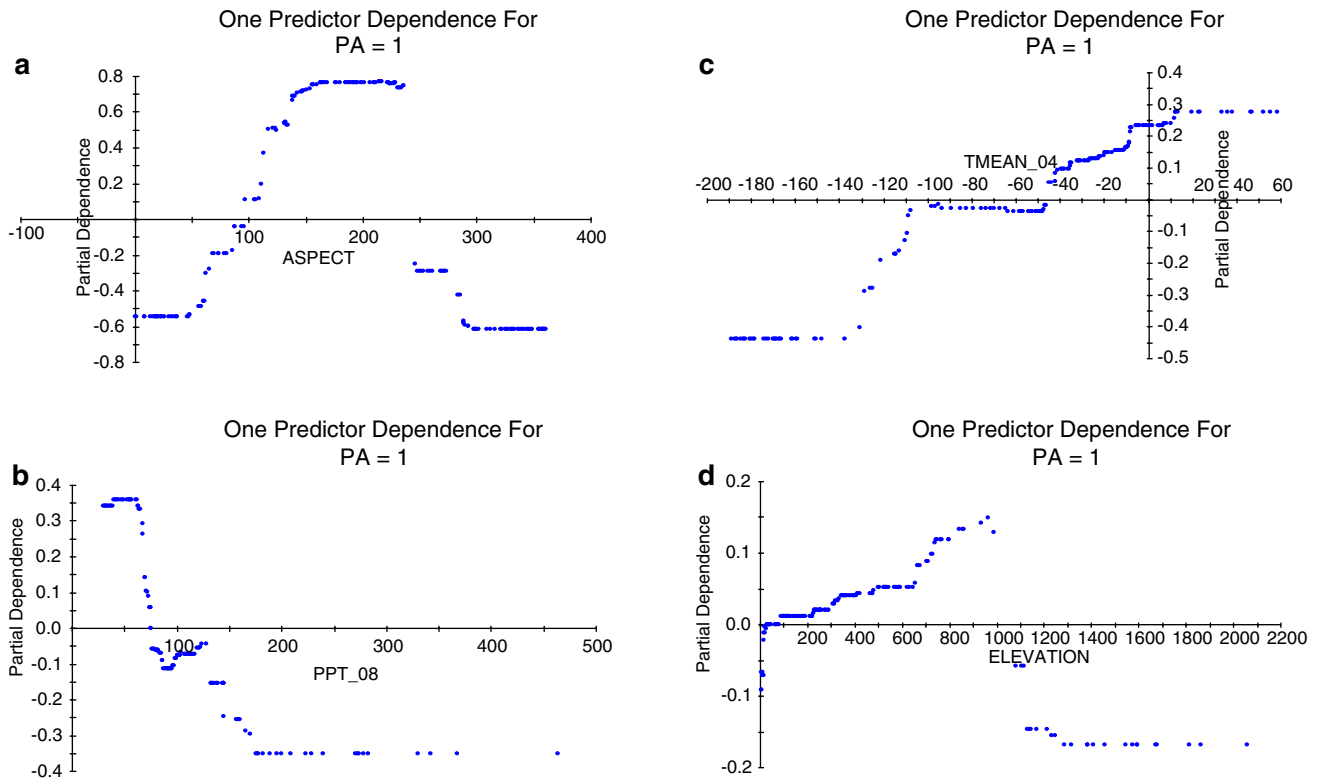
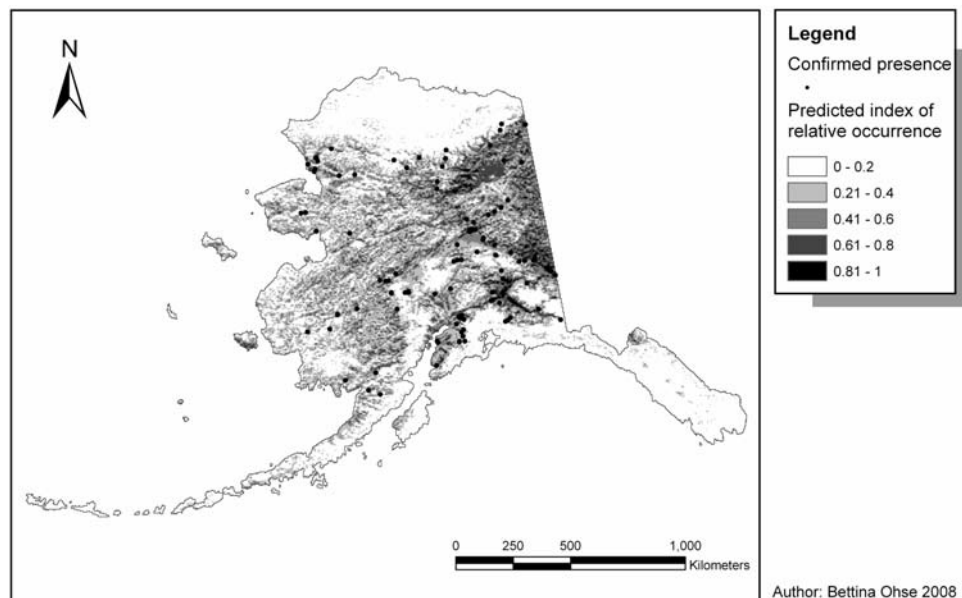


Fig. 2 Single variable plots as obtained by model 12 (best model, as selected by highest accuracy metric): thresholds of the most important variables used for predicting the habitat of white spruce (relative index; positive partial dependence indicates preference, negative

partial dependence indicates avoidance); **a** influence of aspect (degrees); **b** influence of August precipitation (mm); **c** influence of mean April temperature ($^{\circ}\text{C} \times 10$); **d** influence of elevation (m)

Fig. 3 Mapped distribution of White spruce in Alaska as predicted by model 12 (best model, as selected by highest accuracy metric)



more stable than were classes of mid-range relative occurrence (see trendline, Fig. 4b). This indicates that the models' ability to predict low and high relative occurrence was better than the ability to precisely predict relative

occurrence of mid-range, overlapping gray zones, on the pixel scale. However, models 12, 14, and 6 reached an r_s of over 0.9 (0.952, 0.905, and 0.907, respectively), whereas model 15 showed a large instability in predictive ability for

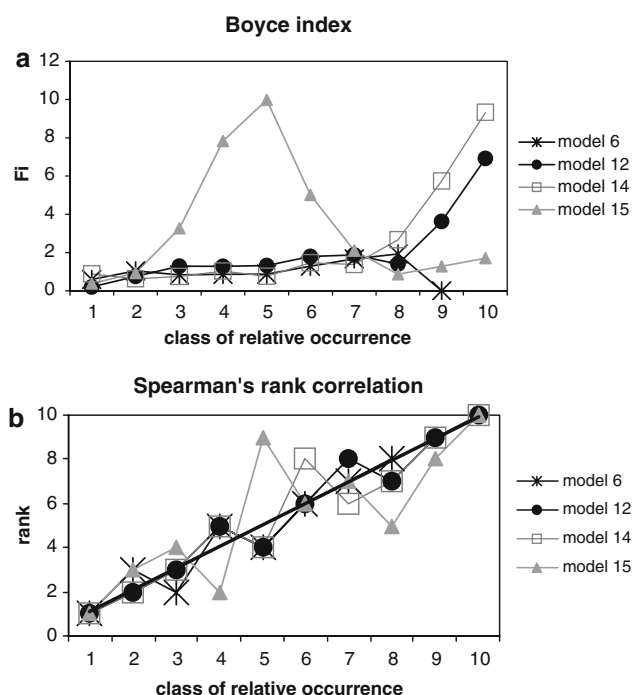


Fig. 4 Model prediction accuracy. Assessment of the four best-performing models with Boyce index and Spearman's rank correlation; **a** Boyce index (area adjusted frequency index) F_i for each class of each model; **b** Spearman's rank correlation between rank of F_i value and class of relative occurrence i for each class of each model (classes along the trendline indicate high model reliability; classes deviating from the trendline indicate reduced model reliability); Spearman's rank correlation coefficient $r_s = 0.907$ (model 6), 0.952 (model 12), 0.905 (model 14), and 0.649 (model 15), respectively

the mid-range classes, resulting in an r_s of 0.649. Thus, the best correlation, i.e. the most consistent prediction was achieved by model 12, having the least departures from the trendline.

Discussion

This study quantitatively models, predicts and maps for the first time the distribution of a tree species in a large wilderness area, with a high accuracy, using free online tools and data. As we focus on high prediction accuracy, we will discuss our methods and results in the context of which factors might potentially influence accuracy.

Freely available species data
(confirmed presence/absence)

Museum data generally prove to be very useful for SDM where other data on species locations are sparse (Stockwell and Peterson 2002, Graham et al. 2004). The authors argued that often the limitation on high resolution

comes with the environmental variables used as predictors, rather than with the species data (Fig. 1, step 1 and 2). However, as typical for wider parts of Alaska, 85% of the museum data we used came with an inherent location error of c. 3,615 km, whereas most of our predictors (elevation, aspect, slope with 1 km, climate data with 2 km cell size) were much more accurate. Effect of location error can be reduced by choosing an appropriate modeling technique (Fig. 1, step 3 and 4), such as Tree-Net, as predictions with boosted regression trees are only slightly influenced by location errors (Graham et al. 2008). Thus, we suggest that using museum data with location errors is still an option for broadscale SDMs and statewide predictions.

Often, museum data tend to be unevenly distributed in space and time and lacking a relevant research design due to opportunistic sampling, also referred to as sampling bias (Stockwell and Peterson 2002; Graham et al. 2004). In our study, data were more abundant along the road system, and few data existed elsewhere in the Interior, where white spruce is assumed to have the center of its range (see also Kadmon et al. 2004). Sampling bias might have the largest influence on prediction accuracy that cannot be accounted for, yet. However, additional information gained by using a multiparameter ecological approach and by considering interactions as presented here, should help mitigate sampling bias.

Resolution and choice of grain size

Cell size (also referred to as grain size) influences the accuracy of a prediction. If the cell size is too small, a slightly wrong geographic species location will result in an association with an environmental variable value of the neighbouring cell, i.e. with a different habitat (Fig. 1, step 2). If the cell size is too coarse, environmental conditions might be averaged, that do not provide an ecological meaning (Guisan et al. 2007). For making predictions for the entire state of Alaska (Fig. 1, step 4) the cell size used here (4 km \times 4 km) is fine enough to keep as much information as possible, but coarse enough for not introducing a much higher accuracy than the original data (with location error) provided. It was also found that differences between species are often higher than between techniques (Elith et al. 2006), suggesting that grain size might need to be adjusted to average patch sizes and/or overall range of a species. This would pose the need of further research on patch sizes and spatial autocorrelation, which could be done using remotely sensed data of vegetation cover, or average or monthly NDVI. However, as well as remote sensing has proved to be capable of revealing information on patch sizes of vegetation types, it cannot do so for single species, yet.

Predictor variables

Although this model is not meant to be a mechanistic biological model, some inferences about the ecological niche can be drawn from the single variable plots (Fig. 2) and variable ranking (Table 4). Often, climate parameters are chosen *a priori* and with a focus on a low number of variables, including only annual values (e.g. Thompson et al. 2006) or only values for the growing season (Calef et al. 2005). As a result climate parameters rarely get tested against each other for their performance. We found it important to not exclude any predictors from the beginning, starting unbiased and virtually uninformed, and therefore tested first all of the 18 climate parameters, and in various combinations. This approach is easily possible, as TreeNet handles large numbers of variables and interactions conveniently. All following steps of dropping variables, that led to model 12 as the best-performing model (see also Table 3), indicate that the excluded variables be of minor importance for the distribution of white spruce.

Our results show, that the most important variables are not necessarily those, which are usually given higher priority by other investigators and in the literature, such as mean temperature of the growing season. We found that taking aspect into account surprisingly increases ROC and %corr values. The importance of aspect for the type of microenvironment and thus vegetation distribution was already stated elsewhere (Van Cleve et al. 1983; Calef et al. 2005; Huettmann and Diamond 2001 for wildlife applications). In contrast, using slope as a predictor lowers %corr values, although topographic slope is often regarded as important (Van Cleve et al. 1983; Calef et al. 2005). Latitude and longitude do not cause significant changes in ROC and %corr values, but help cluster predicted occurrences spatially (see also Supplemental Fig. 6 for comparison of maps). However, using latitude and longitude as predictors reinforces sampling bias, because it gives more weight to areas that were sampled thoroughly and lower weight to areas with lower sampling effort. Table 3 shows that permafrost, soil, and surface geology do not help increase model performance values and thus indicate, either not to contribute to explaining the distribution of white spruce, or that these three datasets are not very suitable for the applied modeling approach. For example, the soil variable reduced the %corr value by more than 20%, which might be due to a mismatch of data, as this dataset includes 268 classes and is thus too specific for a species dataset with 108 presence points. This might result in a loss of generalization ability. The importance of permafrost, contrarily to our findings, is supported by Van Cleve et al. (1983), and indirectly by Calef et al. (2005), who use drainage type as a predictor variable, which is highly correlated with the persistence of permafrost.

However, both permafrost and drainage might be a function of elevation, aspect and slope, and thus are already included in the model.

Predictors found to be important (Table 4), can be used to “learn from the data”, because single variable plots (Fig. 2a–d) show the quantitative influence of each of the parameters on the distribution of white spruce (i.e. the partial dependence of white spruce on the specific parameter). Model 12 indicates the preference of white spruce for aspects ranging from 100° to 250° (Fig. 2a), which is consistent with (but more detailed than) the general idea of the typical white spruce habitat on south-facing slopes (Viereck and Little 2007). Little rainfall in August (total sum <80 mm) appear to favor the occurrence of white spruce (Fig. 2b), but we suggest that this variable is rather an indicator for distance to coast, than an actual climatic variable. Figure 2c shows the importance of time of snowmelt, as mean April temperatures above −4°C (it might be several degrees above zero within the days) help melt snow and thus provide moisture right at the start of the growing season. In contrast, mean April temperatures below −10°C (probably only around zero during the days) inhibit snow melting and moisture supply, and thus delay the start of the growing season, making these sites an unsuitable habitat for white spruce. According to Fig. 2d, white spruce favors elevations from slightly above 0 m (mainly along the rivers, where flowing water prevents the soil from permafrost) to about 1,000 m (highest occurrence of treeline, e.g. in the Alaska Range), which is, in a broad sense, consistent with the literature (Viereck and Little 2007).

Model performance

Model performance strongly depends on the choice of variables and the settings used. The model presented here should be seen as a first, conservative underestimate of model performance and accuracy, as there are many other settings we have not explored in concert, and thus, we could have missed the very best setting in TreeNet improving the model generalization and prediction accuracy further.

Comparing model results, maps, and accuracy assessments for the best four models, similarities and differences become evident. Most striking is that model 6 entirely omits to predict classes 9 and 10, and model 15 shows high instability in Boyce indices and a very fluctuating curve for the spearman’s rank (Fig. 4a, b). Models 12 and 14 obviously show the highest model stability (Fig. 4a, b). They only differ in patterns of distribution for different parts of the state, with model 14 tending to cluster indices of relative occurrence within the landscape (Supplemental Fig. 6), which is likely to be due to including latitude and

longitude as predictors. These variables stress on the locations of the confirmed presence points in such a way, that spatial sample bias is reinforced. Thus, we would propose the results of model 12 as the most reliable prediction, which is supported by a very high r_s .

The slight deviation of model stability for the mid-range classes would affect about 30% of the state-wide area, which might be due to the small patch size of lots of spruce stands. However, the most stable predictions are for the classes of high relative occurrence, proposing 138,192 km² of the state-wide area being covered with white spruce, which is in the same range with values proposed by Labau and van Hees (1990), who suggest about 121,000 km².

Quantitative comparisons to other white spruce maps are difficult, because few maps are published on this topic and most of them do not provide information on the methods used (e.g. Pojar and Mackinnon 1994). There are some areas in Alaska that were predicted by several of our models to have high potential for white spruce occurrence, although these areas were not recorded as such yet (e.g. some regions on the west coast, including offshore islands). They might have simply been undersampled, or, equally likely, range limitations, such as competition, disturbance, local extinction, or barriers to dispersal prevents white spruce from occurring there (Graham et al. 2004; Barry and Elith 2006). Only on a small scale, within the predicted range, e.g. in the westpart of Interior Alaska, values of relative occurrence are highly variable causing a salt-and-pepper like distribution of values on the maps. One explanation we found for this pattern was, that in the boreal and arctic, high variation in slope, aspect, drainage, postfire succession stage and vegetation cover results in larger changes in microenvironment over small distances than in humid midlatitudes (Van Cleve et al. 1983). It was found elsewhere that prediction errors might vary across landscape, and a call for an advanced model with spatial weighting was expressed (Fielding and Bell 1997; Fielding 2002), but is technically not available, yet.

Furthermore, we could not consider climatic trends yet, as our occurrence data span a time period of c. 100 years, and both temperature and precipitation data are averaging the period 1961–1990. Same applies for soil and permafrost characteristics, as those data were compiled once only in 1979 and 1965, respectively. However, according to Masek (2001), field investigations of tree stands at forest-tundra boundaries showed little indication of stand response to warming, yet. The boundaries were clearly mapped from satellite data, but no obvious change was apparent during the duration of the image time series (1970–1990th), constraining recent geographical expansion rates to <200–300 m per century. This might indicate time lags between forest response and climate change, or it reflects competition between trees and their surrounding

vegetation (Masek 2001). The relevance of climate variability in time might also depend on the magnitude and spatial distribution of climate change. Given these facts, we decided to start our role-model with a long-term stable condition, until data with higher temporal resolution become available.

So far, we have captured the white spruce distribution as one single, transparent and repeatable formula in a quantitative fashion and small binary software code, ready for digital use, and open for public assessment. As we were working with data publicly available, it is foreseeable that more and better data will help us improve our proposed and publicly available model even further. We would welcome such efforts.

Suggestions for further research

This study can be used as a baseline for decisions about where more sampling efforts are needed in the future, as we have recognized undersampling to have the most severe impact on our model predictions. Model performance is furthermore dependent on variables such as fire history (fire intensity, extent and frequency; Rupp et al. 2001; Calef et al. 2005), which will be important for delineating deciduous versus coniferous forest, and to define white spruce versus black spruce (*P. mariana*) habitats (Calef et al. 2005). Knowledge about dynamic dispersal, e.g. life history, seed production and seed release applied in a spatially explicit manner (as already used by Rupp et al. (2001) on a smaller scale), as well as information on tree pests, such as the spruce budworm (*Choristoneura occidentalis*), to account for the probability of local extinction will surely improve model performance. Still, it will be the potential niche that is modeled by using this algorithm, rather than the actual niche, unless competition is included as a variable, e.g. by applying a plant community model that accounts for interaction between different species (Ferrier et al. 2002; Ferrier and Guisan 2006; Zimmermann and Kienast 1999).

Seeking for a balance between habitat protection, conservation and recreation, and potential timber and fuel supply can only be successful with detailed knowledge about the potential niche of a species and its spatial distribution within the landscape, as provided here. This model also offers itself as a baseline for assessing land-use change or changes in species ranges due to climate change (Leathwick et al. 1996; Graham et al. 2004; Prasad et al. 2007-ongoing, Huettmann et al. unpublished), as it could potentially be modeled backward and forward in time. It would also be valuable to apply this model to and obtain maps for other tree species. Furthermore, for forestry and timber volume prediction purposes, this model could be adjusted by using tree volume data instead of presence/

absence for model calibration, or by linking the index of relative occurrence with timber volume. Definitely, this will affect forest management decisions, especially when pursuing sustainable forest management.

In this study, we operated data primarily in ArcGIS. However, there are options of exclusively using free software, such as GRASS GIS, which is applicable to geospatial data management and analysis, image processing, graphics/maps production, spatial modeling, etc. Further exploration of these options for modeling would help provide important tools for a broader research community. Without availability of high-quality data (Open access) accurate predictive modeling as presented here would not have been possible. Using these data and applying non-invasive methods helps preserve wilderness areas without disturbing them. Making model results publicly available helps connect scientists, resource managers, policy makers, and communities and shall enhance collaborative planning and management.

Acknowledgments We want to thank everybody who helped: Data for model evaluation were kindly supplied by C. Roland (Central Alaska Network Vegetation Monitoring Program, National Park Service), T. Loomis (ABRinc), S. Winslow (University of Alaska, Fairbanks), and K. Winterberger (Pacific Northwest Experiment Station). We furthermore want to thank V. Steen and T. McMillan for fruitful discussions. Of course, we highly appreciate the contributions of the two reviewers, who significantly helped improve this article. Thanks also to the UAF Department of Natural Resources Management, as well as the Department of Biology and Wildlife for providing technical support. B. Ohse wants to thank the Ev. Studienwerk Villigst and the Fulbright Commission for help fund her studies at UAF. This is EWHALE lab publication # 47.

References

- Araújo MB, Thuiller W, Williams PH, Reginster I (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Glob Ecol Biogeogr* 14:1–17
- Barry S, Elith J (2006) Error and uncertainty in habitat models. *J Appl Ecol* 43:413–423
- Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA (2002) Evaluating resource selection functions. *Ecol Modell* 157:281–300
- Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit* 30:1145–1159
- Breiman L (2001) Statistical modelling: the two cultures. *Statistical Sci* 16:199–215
- Breiman L, Friedman J, Olshen R, Stone C (1984) *Classification and regression trees*. Wadsworth, Belmont, CA
- Brink CH, Dean FC (1966) Spruce seed as a food of red squirrels and flying squirrels in interior Alaska. *J Wildl Manag* 30:503–512
- Burnham KP, Anderson DR (1998) *Model selection and inference: a practical information-theoretic approach*. Springer-Verlag, New York, USA
- Buskirk SW (1984) Seasonal use of resting sites by marten in south-central Alaska. *J Wildl Manag* 48:950–953
- Calef MP, McGuire AD, Epstein HE, Rupp TS, Shugart HH (2005) Analysis of vegetation distribution in Interior Alaska and sensitivity to climate change using a logistic regression approach. *J Biogeogr* 32:863–878
- Craig E, Huettmann F (2008) Using “blackbox” algorithms such as TreeNet and Random Forests for data-mining and for finding meaningful patterns, relationships and outliers in complex ecological data: an overview, an example using golden eagle satellite data and an outlook for a promising future. In: Hsiao-fan Wang (ed) *Intelligent data analysis: developing new methodologies through pattern discovery and recovery*. IGI Global, Hershey, PA, USA
- Dunning JB Jr, Stewart DJ, Danielson BJ, Noon BR, Root TL, Lamberson RH, Stevens EE (1995) Spatially explicit population models: current forms and future uses. *Ecol Appl* 5:3–11
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, JMcC Overton, Peterson AT, Phillips SJ, Richardson KS, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species’ distributions from occurrence data. *Ecogr* 29:129–151
- Ellenberg H (1988) *Vegetation ecology of Central Europe*, 4th edn. Cambridge University Press, Cambridge
- Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J Appl Ecol* 41:263–274
- Farr A, Harris AS (1979) Site index of Sitka Spruce along the Pacific Coast related to latitude and temperatures. *For Sci* 25:145–153
- Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. *J Appl Ecol* 43:393–404
- Ferrier S, Drielsma M, Manion G, Watson G (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New SouthWales. II. Community-level modelling. *Biodivers Conserv* 11:2309–2338
- Fielding AH (2002) What are the appropriate characteristics of an accuracy measure? In: Scott JM, Heglund PJ, Morrison ML, Hauffer JB, Raphael MG, Wall WA, Samson FB (eds) *Predicting species occurrences: issues of accuracy and scale*. Island Press, Covelo, CA, pp 271–280
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Fleming M (1997) A statewide vegetation map of Alaska using phenological classification of AVHRR data. In: Walker DA, Lillie AC (eds) *The second circumpolar arctic vegetation mapping workshop*, Arendal, Norway, 18–24 May 1996 and the CAVM-North American Workshop, Anchorage, Alaska, US, 14–16 January 1997. Institute of Arctic and Alpine Research, Boulder, CO, pp 25–26
- Franklin J (1995) Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients. *Proc Phys Geogr* 19:474–499
- Franklin J (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *J Veg Sci* 9:733–748
- Friedman JH, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting. *Ann Stat* 28:337–407
- Gallant AL, Binnian EF, Omernik JM, Shasby MB (1995) *Ecoregions of Alaska US Geological Survey Professional Paper 1567*. US Government Printing Office, Washington, DC
- Graham CH, Ferrier S, Huettmann F, Moritz C, Peterson AT (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol Evol* 19:497–503

- Graham CH, Elith J, Hijmans RJ, Guisan A, Peterson AT, Loiseleur BA, The Nceas Predicting Species Distributions Working Group (2008) The influence of spatial errors in species occurrence data used in distribution models. *J Appl Ecol* 45:239–247
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8:993–1009
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Modell* 135:147–186
- Guisan A, Theurillat J-P, Kienast F (1998) Predicting the potential distribution of plant species in an alpine environment. *J Veg Sci* 9:65–74
- Guisan A, Lehmann A, Ferrier S, Austin M, Overton JMcC, Aspinall R, Hastie T (2006) Making better biogeographical predictions of species' distributions. *J Appl Ecol* 43:386–392
- Guisan A, Graham C, Elith J, Huettmann F, NCEAS modelling Group (2007) Sensitivity of predictive species distribution models to change in grain size: insights from an international experiment across five continents. *Divers Distrib* 13:332–340
- Hastie T, Tibshirani R, Friedman JH (2001) The elements of statistical learning: data mining, inference, and prediction. Springer, New York
- Hennon PE, Trummer LM (2001) Yellow Cedar (*Chamaecyparis nootkatensis*) at the Northwest Limits of its Natural Range in Prince Williams Sound, Alaska. *Northwest Sci* 75:61–71
- Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecol Modell* 199:142–152
- Holsten EH, Thier RW, Schmid JM (1991) The sprucebeetle. USDA Forest Service Forest Insect and Disease Leaflet 127
- Huettmann F (2007) Modern adaptive management: adding digital opportunities towards a sustainable world with new values. *Forum Public Policy* 3:337–342
- Huettmann F, Diamond AW (2001) Seabird colony locations and environmental determination of seabird distribution: a spatially explicit seabird breeding model in the Northwest Atlantic. *Ecol Modell* 141:261–298
- Hultén E (1968) Flora of Alaska and neighbouring territories: a manual of vascular plants. University Press Stanford, Stanford
- Interagency Working Group on Digital Data (2009) Harnessing the power of digital data for science and society. Report to the Committee on Science of the National Science and Technology Council. Washington DC
- Juday GP, Barber V, Berg E, Valentine D (1999) Recent dynamics of white spruce treeline forests across Alaska in relation to climate. In: Kankaanpää S, Tasanen T, Sutinen M-L (eds) Sustainable development in Northern Timberline Forests. Proceedings of the Timberline Workshop, May 10–11, 1998 in Whitehorse, Canada. Finnish Forest Research Institute. Research Papers 734, pp 165–187
- Kadmon R, Farber O, Danin A (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecol Appl* 14:401–413
- LaBau VJ, Alden JN (2000) Unalaska, Alaska: Revisiting North America's Oldest Afforestation Effort. *J For* 98:24–29
- LaBau VJ, van Hees WS (1990) An inventory of Alaska's boreal forests: their extent, condition, and potential use. In: Condition, dynamics, anthropological influences. Proceedings of the International Symposium, 16–26 July, 1990. Archangelsk, Russia. State Forest Committee of the USSR. Part 6, pp 30–39
- Lawler JJ, White D, Neilson RP, Blaustein AR (2006) Predicting climate-induced range shifts: model differences and model reliability. *Glob Chang Biol* 12:1568–1584
- Leathwick JR, Whitehead D, McLeod M (1996) Predicting changes in the composition of New Zealand's indigenous forests in response to global warming: a modelling approach. *Environ Softw* 11:81–90
- MacCracken JG, Viereck LA (1990) Browse regrowth and use by moose after fire in interior Alaska. *Northwest Sci* 64:11–18
- Maggini R, Lehmann A, Zimmermann NE, Guisan A (2006) Improving generalized regression analysis for the spatial prediction of forest communities. *J Biogeogr* 33:1729–1749
- Margules CR, Austin MP (1994) Biological models for monitoring species decline: the construction and use of data bases. *Phil Trans Roy Soc B* 344:69–75
- Masek JG (2001) Stability of Boreal forest stands during recent climate change: evidence from Landsat Satellite Imagery. *J Biogeogr* 28:967–976
- Murray DF (1980) Balsam Poplar in Northern Alaska. *Can J Anthropol* 1:29–32
- National Research Council of the National Academies (2003) Sharing Publication-related Data and Materials: Responsibilities of Authorship in the Life Sciences. The National Academic Press, Washington DC, www.nap.edu
- Parviainen M, Luoto M, Rytätari T, Heikkinen RK (2008) Modelling the occurrence of threatened plant species in taiga landscapes: methodological and ecological perspectives. *J Biogeogr*. doi: 10.1111/j.1365-2699.2008.01922.x
- Pojar J, Mackinnon A (1994) Plants of the Pacific Northwest Coast: Washington, Oregon, British Columbia, and Alaska. Lone Pine Publishing, Redmond, WA
- Prasad AM, Iverson LR, Matthews S, Peters M (2007-ongoing) A climate change Atlas for 134 Forest Tree Species of the Eastern United States [database]. Northern Research Station, USDA Forest Service, Delaware, OH. <http://www.nrs.fs.fed.us/atlas/tree>
- Risenhoover KL (1989) Composition and quality of moose winter diets in interior Alaska. *J Wildl Manag* 53:568–577
- Rupp TS, Chapin FS, Starfield AM (2001) Modelling the influence of topographic barriers on treeline advance at the forest-tundra ecotone in northwestern Alaska. *Clim Chang* 48:399–416
- Sinclair ARE, Jogia MK, Andersen RJ (1988) Camphor from juvenile white spruce as an antifeedant for snowshoe hares. *J Chem Ecol* 14:1505–1514
- Slough BD (1989) Movement and habitat use by transplanted marten in the Yukon Territory. *J Wildl Manag* 53:991–997
- Smith M (1968) Red squirrel responses to spruce cone failure in interior Alaska. *J Wildl Manag* 32:305–317
- Stockwell DRB, Peterson AT (2002) Controlling bias in biodiversity data. In: Scott JM, Heglund PJ, Morrison ML, Hauffer JB, Raphael MG, Wall WA, Samson FB (eds) Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, pp 537–546
- Thompson RS, Anderson KH, Strickland LE, Shafer SL, Pelletier RT, Bartlein PJ (2006) Atlas of relations between climatic parameters and distributions of important trees and shrubs in North America—Alaskan Species and Ecoregions. USGS Professional Paper 1650-D
- Tsoar A, Allouche O, Steinitz O, Rotem D, Kadmon R (2007) A comparative evaluation of presence-only methods for modelling species distribution. *Divers Distrib* 13:397–405
- Van Cleve K, Dyrness CT, Viereck LA, Fox J, Chapin FS III (1983) Taiga ecosystems in interior Alaska. *Biosci* 33:39–44
- Viereck LA, Foote JM (1970) The Status of *Populus balsamifera* and *P. trichocarpa* in Alaska. *Can Field Nat* 84:169–173
- Viereck LA, Little EL (2007) Alaska trees and shrubs. Snowy Owl Books, Fairbanks
- Viereck LA, Van Cleve K, Dyrness CT (1986) Forest ecosystem distribution in the taiga environment. In: Van Cleve K, Chapin FS III, Flanagan PW, Viereck LA, Dyrness CT (eds) Forest ecosystems in the Alaskan taiga. Ecological studies 57. Springer-Verlag, New York, NY, pp 22–43
- Viereck LA, Dyrness CT, Batten AR, Wenzlick KJ (1992) The Alaska vegetation classification. General Technical Report No.

286. US Forest Service, Pacific Northwest Research Station, Portland, OR
- Walker LR, Zasada JC, Chapin FS III (1986) The role of life history processes in primary succession on an Alaskan floodplain. *Ecol* 67:1243–1253
- Walter H (1985) *Vegetation of the earth and ecological systems of geobiosphere*, 3rd edn. Springer, Heidelberg
- Walters CJ (1986) *Adaptive management of renewable resources*. McGraw Hill, New York
- Whittaker RH (1967) Gradient analysis of vegetation. *Biol Rev Camb Philos Soc* 42:207–264
- Wolff JO (1978) Food habits of snowshoe hares in interior Alaska. *J Wildl Manag* 42:148–153
- Zimmermann NE, Kienast F (1999) Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *J Veg Sci* 10:469–482