



Climatic niche, ecological genetics, and impact of climate change on eastern white pine (*Pinus strobus* L.): Guidelines for land managers

Dennis G. Joyce^{a,*}, Gerald E. Rehfeldt^b

^a Ontario Ministry of Natural Resources, 70 Foster Drive, Suite 400, Sault Ste. Marie, Ontario, Canada P6A 6V5

^b Rocky Mountain Research Station, USDA Forest Service, Forestry Sciences Laboratory, 1221 S. Main, Moscow, ID 83843, USA

ARTICLE INFO

Article history:

Received 5 November 2012

Received in revised form 17 December 2012

Accepted 18 December 2012

Keywords:

Bioclimatic models

Ecological genetics

Climatic niche

Climate response models

Seed transfer guidelines

Seed zones

ABSTRACT

Approximately, 123,500 forest inventory and ecological ground plots representing eastern North America were used to predict the contemporary distribution of eastern white pine (*Pinus strobus* L.) from climate. The random forests classification tree procedures produced an 8-variable algorithm that had an 8% overall error rate. Erroneous predictions of presence, or errors of commission, were 13%, while falsely predicting absence, or errors of omission were 1%. Climate-based multiple regression models were developed to describe patterns of genetic variation among 112 populations representing the range of *P. strobus* in Ontario, Canada east of Lake Superior. Degree days >5 °C was the best predictor of variation in phenology and growth potential, with 5-year height providing the greatest resolution of inter-population variation ($R^2 = 0.68$). Cold hardiness in the fall was most closely associated with mean minimum temperature ($R^2 = 0.26$). Height growth data from four disparate provenance test series that together included a total of 354 provenances corroborated the range-wide applicability of the regional genetic models. Although variation in growth potential in the central Appalachian Mountains was most closely associated with mean minimum temperature, degree days >5 °C remained the best predictor of range-wide variation in growth potential ($R^2 = 0.41$). The contemporary distribution and inter-population genetic variation were projected into future climates predicted by three General Circulation Models, two scenarios, and three time steps. All projections indicate early and sustained deterioration in the contemporary habitat. Concurrence among projections regarding the redistribution of suitable habitat to the north of the contemporary distribution identifies geographic locations with the highest probability of supporting vigorous stands of *P. strobus*. Concurrences among genetic projections clarify the intraspecific redistribution required to conserve adaptive variation. The projections have direct relevance in developing management strategies for accommodating the changing climate.

© 2013 Published by Elsevier B.V.

1. Introduction

The Intergovernmental Panel on Climate Change has presented unequivocal evidence that the climate is changing (IPCC, 2007). Over the last 30 years, the mean global temperature has increased by 0.6 °C. During this same time period, isotherms marking the average annual temperature have moved pole-ward about 160 km, and the inertia of the global climate system ensures an additional warming of 0.4 °C over the next two decades (IPCC, 2007). If greenhouse gas emissions continue at the current business-as-usual rates, the speed of isotherm movement will at least double in the latter half of the current century (Hansen, 2009).

Documented ecological changes in the phenology and distribution of plants and animals that have already occurred continue to accrete (Parmesan, 2006). Studies of the contemporary distribution of forests and trees species have shown the overall pattern is one of range contraction (Allen et al., 2010; Breshears et al., 2005; Gitlin et al., 2006; Jump and Peñuelas, 2006; Michaelian et al., 2011; Rehfeldt et al., 2009; Shaw et al., 2005; Worrall et al., 2010). Evidence of expanding ranges in tree species at the leading northern edge have been both positive (Tchebakova et al., 2010; Woodall et al., 2009), and negative (Masek, 2001; Zhu et al., 2011).

Over the 21st century, the magnitude of climate shifts is expected to exceed the ecological amplitude of many tree species eventually leading to extirpation over substantial portions of contemporary ranges (Barns, 2009; Iverson et al., 2008; Kirschbaum and Fischlin, 1996). In situ adaptational rescue of populations in decline is difficult (Lavergne et al., 2010) because the multiple-generation adaptive capacity of local populations deteriorates rapidly under intensive selection pressure (Burger and Lynch, 1995; Lynch

* Corresponding author. Address: 449 Walls Road, Sault Ste. Marie, Ontario, Canada P6A 5K6. Tel.: +1 705 945 5894/779 2545; fax: +1 705 945 6667.

E-mail addresses: dennis.joyce@ontario.ca (D.G. Joyce), jrehfeldt@gmail.com (G.E. Rehfeldt).

and Lande, 1993) leading to the expectation that the speed of climate change will exceed the capacity of many populations to adapt in situ (e.g., Rehfeldt et al., 2001).

Especially at middle to northern latitudes, persistence of tree species is projected to be contingent, in large part, on colonizing emergent suitable habitat beyond the northern limits of contemporary distributions (Williamson et al., 2009). But, post-glacial migration rates of North American tree species (McLachlan and Clark, 2004; McLachlan et al., 2007) are a small fraction (less than 2%) of the rate climate gradients are shifting (Hansen, 2009). A seed-dispersal simulation model applied to five species of trees native to the eastern United States indicated that there is only a small probability of colonization at distances beyond 20 km (Iverson et al., 2008). The probability of repeated long distance colonization is further constrained by the expected small size of the initial colonizing population, the small size of the seed crops, and the level of inbreeding.

The conclusion drawn by many authors is that many North American tree species will experience range contraction at the trailing southern edges with little prospect of effective stochastic range expansion to the north (Barns, 2009; Iverson and Prasad, 1998; Iverson et al., 2008; Keenan et al., 2011; Ledig et al., 2010; Rehfeldt, 2004; Rehfeldt and Jaquish, 2010; Thomas, 2004; Thuiller, 2003). In the face of these conclusions, forest managers are challenged to choose the most appropriate management options for maintaining and enhancing productive forests. Preparing to mitigate adaptational and dispersal lags through assisted redistribution of native tree species is fundamental component of this challenge (Ledig et al., 2010). While the challenge is clear, the task is daunting. A low tolerance for failure; uncertainties associated with the timing of climate change; and the expectation that species vulnerabilities and responses will be unique (e.g., Lavergne et al., 2010; Barns, 2009) confound the development of efficient management response.

A useful concept for assessing potential impacts of climate on forests is the climatic niche. Hutchinson (1957) quantified the environmental niche concept as an n -dimensional hypervolume, every point in which corresponds to a state of the environment which would permit a species to persist indefinitely. At a regional scale, the hypervolume dimensions are largely defined by climatic variables. Consequently, the use of climatic conditions to describe a species' niche has become central to much ecological reasoning and theory (Pulliam, 2000).

An array of correlative niche modeling techniques (Elith and Leathwick, 2009; Hijmans and Elith, 2011; Pearson and Dawson, 2003) are now widely used to predict the contemporary realized niche of tree species and then to project the geographic redistribution of suitable habitat under future climate as a foundation for developing management strategies (Barns, 2009; Iverson and Prasad, 1998; Iverson et al., 2008; Keenan et al., 2011; Ledig et al., 2010; Rehfeldt, 2004; Rehfeldt and Jaquish, 2010; Thomas, 2004; Thuiller, 2003). The central premise of this approach is that the contemporary realized niche of a species provides the best available metric for projecting the future distribution of suitable habitat (Pearson, 2007). For simplicity the term 'climate profile' (sensu Rehfeldt et al., 2006) will be used here to refer to the realized climatic niche of a species.

Determining a species contemporary climate profile and projecting redistribution of suitable habitat is fundamental to considering ecological impacts of climate change, but it is imperative to include ecological genetics in the models (Matyas, 2006). Common garden studies of inter-population variation in wide-ranging tree species (provenance tests), have documented that habitat-correlated clinal genetic variation among populations is widespread (Morgenstern, 1996). Ecological races typically morph gradually across climatic gradients and adaptive differences in climatically

proximal populations are subtle so that distinct boundaries between ecological races are absent. Yet, populations originating from contrasting severe and mild climatic conditions also exhibit inherent contrasting balances in the inter-correlated traits associated with the annual development cycle (see Morgenstern, 1996 for general review). Such ecological genetic variation is invariably interpreted as adaptive.

Visualizing the spatial heterogeneity of clinal variation requires a method of classifying continuous variation into climatic ecotypes or climatotypes (Turesson, 1925). A robust approach to climatotype classification involves the use of the confidence interval surrounding a population mean (Rehfeldt, 2004; Rehfeldt et al., 2004; Rehfeldt and Jaquish, 2010). The expectation being that all populations within this statistical space are relatively homogeneous genetically. Depending on the disparity among multiple populations, climatotypes can be either overlapping or discrete.

Well-designed common garden (provenance) tests of multiple seed sources support the development of models that articulate the ecological genetic structure of a species. Historically, ecological genetic profiles have been based on geographic variables that are surrogates for climatic gradients (Campbell and Sorensen, 1978; Matyas and Yeatman, 1992; Rehfeldt, 1988, 1991; St. Clair et al., 2005). With the availability of high-precision climate models, ecological genetic profiles have begun to shift from the largely descriptive geographic models to climate-based predictive models that can be projected into future climate conditions (e.g., Bower and Aiken, 2008; Rehfeldt, 2004; Richardson et al., 2009).

Throughout this paper, the terms provenance and seed sources refer to a geographic location of seed collections; population refers to the trees native to the provenance and their wind-pollinated offspring. The term growth potential follows Rehfeldt (1990) as the inherent capacity for height growth when environmental conditions approach the optimum. Height of young trees acts as a surrogate for growth potential when provenance tests are located in mild, relatively stress-free environments (e.g., Leites et al., 2012; Rehfeldt et al., 1999).

Eastern white pine is one of the most commercially valuable trees in eastern North America. It is most abundant in a broad latitudinal band extending from northwest Ontario, Minnesota and Wisconsin eastward to the New England states and the Maritime Provinces (Iverson et al., 2008). The natural range also extends southward primarily along the Appalachian Mountains reaching its southern limits in Georgia. It occurs on nearly all soils within its range, but is most competitive on well drained sandy soils of low to medium site quality (Wendel and Smith, 1990).

A wealth of provenance tests of *Pinus strobus* have documented latitudinal clines in phenology (Lu et al., 2003; Santamour, 1960), timing of cold acclimation (Maronek and Flint, 1974; Mergen, 1963; Lu et al., 2002), and survival, height and diameter growth at ages up to 28 years, (Abubaker and Zsuffa, 1990; Beaulieu et al., 1996; Demeritt and Garrett, 1996; Demeritt and Kettlewood, 1976; Fowler and Heimbucher, 1969; Funk, 1971; Funk et al., 1975; Garrett et al., 1973; Genys, 1968, 1987, 1990; Joyce et al., 2002a; King and Nienstadt, 1968; Li et al., 1997; Sluder, 1963; Sluder and Dorman, 1971; Sprackling and Read, 1976; Wright, 1970; Wright et al., 1963; Wright et al., 1979). While insightful for understanding the contemporary pattern of adaptive variation, geophysical models are of little use in predicting adaptive responses to changing climate.

The present analyses draw on a high-precision climate model, a comprehensive inventory data set, and sophisticated analytical software to quantify the contemporary species climate profile and ecological genetic profile for the entire botanical range of *P. strobus*. Contemporary distribution and intraspecific genetic variation are projected into future climate scenarios and mapped in order to assess spatially explicit variation in vulnerability of the

contemporary eastern white pine resource, and to predict emergent suitable habitat as the climate changes. The analytical steps include: (1) assembling contemporary geo-referenced species occurrence data base and predicting climatic attributes of each location; (2) defining and mapping the species climate profile (sensu Rehfeldt et al., 2006); (3) assessing the impact of climate change projections on both the contemporary and future distribution of suitable habitat; (4) applying climate-based distribution modeling techniques to disparate provenance test data to elucidate the ecological genetic profile within the climate profile; and, (5) mapping the ecological genetic profile for both contemporary and selected future climate scenarios.

2. Methods

2.1. *P. strobus* presence–absence data

A *P. strobus* presence–absence dataset was derived from approximately 39,000 forest inventory and ecological ground-plot data obtained from provincial offices in Ontario, Quebec, Nova Scotia, Newfoundland and Labrador, and 74,000 US Forest Inventory and Analysis (FIA) Program permanent sample plots representing the United States east of the 100th meridian.

The US FIA plots are systematically located to sample the vegetation on forested and non-forested lands (see Alerich et al., 2004; Bechtold and Patterson, 2005). Legal restrictions prevent the US Forest Inventory and Analysis Program from revealing the precise geographic location of their plots. Consequently, the plot data available to the public is based on ‘fuzzy’ coordinates that approximate actual locations, but the degree of geographic imprecision is not available. Climatic estimates for US plots as well as illustrations based on climate variables inherit the imprecision of these fuzzy coordinates.

Operational records of Ontario Ministry of Natural Resources, as well as published geographic locations from *P. strobus* provenance tests (Beaulieu et al., 1996; Genys, 1968) added about 400 additional locations where *P. strobus* has been reported to occur. To assure that our observations included a full range of climates in which *P. strobus* does not occur, a random sample of 10,000 observations was drawn from a database (Rehfeldt et al., 2012) of Biotic Communities (Brown et al., 1998) that were geographical proximal to the distribution of *P. strobus* but from communities in which *P. strobus* does not occur: about 50 from Adirondack–Appalachian Subalpine Forests and Tundra; 500 from Gulf Coastal Grasslands; 1000 from the Northern Tundra; 1000 from Manitoba including both Canadian Taiga and Great Plains Grassland; 1500 from the Canadian Taiga north of 54.5°N and east of –105°; and 2000 from the Great Plains Grassland from 25°N to 49°N and east of –100° longitude. These procedures assembled a total of approximately 123,500 locations of which *P. strobus* was documented to occur in 9381.

2.2. Genetic data

Both published and unpublished measurements of numerous provenance tests were assembled from several disparate studies at differing ages. Results of analyses of most of these studies detected significant differences among provenances. Most of the results, moreover, demonstrated clines in genetic variability associated with the latitude of the seed source. The primary thrust of our analyses is to re-define the clines in terms of climate, using data and techniques not available when the previous studies were conducted.

Our approach to the analysis of these data is to treat thoroughly a comprehensive data set from tests conducted in Ontario,

combine the Ontario results with historical data from tests conducted in Quebec and Maryland to consider range-wide clines in genetic variation, and to both verify and supplement the range-wide cline with somewhat less comprehensive historical data from tests conducted by the US Forest Service (USFS) and those designed for the southern Appalachian region.

2.2.1. Ontario test series

The data employed to develop an ecological genetic model for Ontario was obtained from a study of 112 open-pollinated seed sources representing the geographic distribution of *P. strobus* in Ontario east of Lake Superior. Trees were grown for the first year in plastic containers (65 cm³) in a greenhouse in Sault Ste. Marie, Ontario. A randomized complete block design was used with 11 replications of 10-tree-row-plots.

After over wintering in a dark freezer maintained at –2 °C, three replications were placed in a greenhouse on April 23, 1996. Total tree height was recorded three times a week from April 24th until June 6th when all shoot elongation was complete. The three replications were then transferred to a shaded holding area next to the greenhouse and maintained until September. In September, one needle fascicle from the top of each seedling was removed and measured for length. On September 3, 5, and 7, 1996, additional needle fascicles were collected from each tree and placed in a plastic bag containing needles from all trees in each plot. The bags were stored temporarily at 2 °C until allocated to one of three freezing treatments (–7°, –9°, and –11 °C) conducted between September 6 and 16. Freezing treatments were applied using a freezing chamber that reduced temperatures at a rate of 5° per hour until the target temperature was reached. The target temperature was maintained for 2 h and then increased at a rate of 5° per hour. After removal from the freezing chamber, bagged needles were placed in a shaded greenhouse for approximately one week, at which time the proportion of needle fascicles exhibiting damage (flaccid and necrotic tissue) was recorded for each plot.

In May of the second year (1996), the remaining eight replications were assigned to one of two cleanly cultivated test sites and planted at a spacing of 1 m. within and between rows. The first test was located at the Ontario Forest Research Institute Arboretum near Sault Ste. Marie, Ontario (46°33′ latitude, –84°27′ longitude, 225 elevation). The second test was located near Turkey Point Provincial Park on the north shore of Lake Erie (42°42′ latitude, –80°37′ longitude, 220 m elevation). Both tests were maintained as farm-field tests where vegetation was managed to minimize competition. Total tree height, damage agent, and severity of damage were recorded annually through the fifth growing season (1999). Moderate and severely damaged trees were excluded from analyses involving height.

2.2.2. Quebec test series

Beaulieu et al. (1996) reported the mean 10-year height across three field performance tests in Quebec for 67 range-wide provenances. Two of the three tests contained seven randomized blocks while the third test contained six blocks. All three tests employed 4-tree-row plots. The average height across all three tests was 188.4 cm. Two pairs of provenances were excluded because they had identical geographic coordinates, disparate performance and no information on elevation.

Dr. Beaulieu also made available for our analyses the unpublished 4-year mean heights for a range-wide sample of 130 provenances that were grown in nursery beds at the Valcartier Forest Experiment Station Erie (46°56′ latitude, –71°30′ longitude, 183 m elevation). Only 27 of these provenances were included in the field performance tests described above. Mean population heights were based on 12 replications of 4-tree-row plots (Beaulieu – personal communication July 2011).

2.2.3. Maryland test series

Mean 2-year height (cm) for populations representing 119 range-wide provenances growing in a nursery bed at the State Forest Tree Nursery in Harmans, Maryland is available from Genys (1968). The reported mean heights were obtained by measuring the four largest trees, one per quarter plot of 50 cm nursery bed plots in each of four replications.

For our analyses, data from 21 populations were discarded because of either missing geographic references or survival so poor that seedling height could not reflect the growth potential of the province. In addition, geographic referencing of nearly one-half of the populations appears questionable; latitude and longitude were recorded to only the nearest 5 min. For example, the geographic coordinates for one of the Maryland sources was actually in Pennsylvania, and the coordinates for one of the Ontario provenances centers it in Lake Erie. Nonetheless, with the exception of such flagrant errors, we have assumed the remaining geographic references are accurate.

Genys (1987) also reported for the same test series, the 16-yr height expressed as a percentage of the mean of all trees in the plantation, for 108 of the same populations in two Maryland field tests. Each test had four replications of 4-tree square plots. Because the mean 16-year height across the two tests was strongly correlated with performance at each individual test ($r > 0.85$), we used only the mean height of populations across both tests. Doing so alleviates potential concerns associated with low survival of northern latitude provenances at one of the test sites.

2.2.4. USFS and southern appalachian test series

The US Forest Service initiated a range-wide provenance test series of 31 *P. strobus* provenances in 1955. Field tests were established in 13 Eastern Seaboard locations plus one test each in West Virginia, Kentucky and Ohio. The number of provenances represented at each site ranged from 16 to 28. An additional 9 tests, with a standard 16 provenances, were planted in the Lake States, and Kentucky. Twelve of the standard provenances were also tested in two locations in southern Ontario. The design varied among tests, but most commonly, 48 trees per population were planted at each test. Survival, height and various other traits were monitored over time (Garrett et al., 1973). Mean population height reported as a percentage of all trees planted at a site, is available from numerous sources (Demeritt and Garrett, 1996; Demeritt and Kettlewood, 1976; Fowler and Heimbuger, 1969; Funk, 1971; Funk et al., 1975; Garrett et al., 1973; King and Nienstadt, 1968; Sluder, 1963; Sluder and Dorman, 1971; Wright et al., 1963). For consistency, we used mean height at age 10.

An additional series of tests was initiated solely for the southern Appalachian region which previously had been treated summarily. The series included 38 provenances from the south and central Appalachian regions planted in 9 tests in the Lake States and one test site in Tennessee and West Virginia. The geographic coordinates of most provenances are reported to the nearest 5 min and elevations to the nearest 100 feet. Many provenances were represented by 10-tree bulked seed lots, but several seed collections were from a single tree. Most tests contained five blocks with 4-tree plots. The number of provenances represented in individual tests varied between 14 and 38. Mean population height was expressed as a percent of individual test means at ages 7–11 years (Wright et al., 1979).

2.3. Climate estimates

For estimates of current and future climate, these analyses used a North American-wide climate model that was developed by applying thin-plate smoothing spline procedures (ANUSPIN, Hutchinson, 2000) to fit monthly normals for the period of

1961–1990 to geographic surfaces (accessible at URL: <http://forest.moscowfsl.wsu.edu/climate>). These surfaces are the most recent extension of the surfaces originally developed for the western United States (Rehfeldt, 2006) and Mexico (Saenz-Romero et al., 2010). The North American surfaces were based on weather station data for about 15,000 locations for precipitation and 12,000 for temperature.

A total of 34 candidate climatic variables, representing annual and seasonal precipitation and temperatures; frost free period and heat sums; and precipitation–temperature interactions were considered in these analyses (Table 1). Climate estimates for each of these variables for each location in the presence–absence data set were obtained by submitting a file containing the point identification, latitude, longitude, and elevation to the web-site identified above.

North American-wide spline-based climatic grids, also available at the above web site, were developed using the digitized elevation model of GLOBE Task Team (1999) which are available on 0.00833 degree grids. This model was used also to fill in missing elevations for point data.

The six future climate projections considered in these analyses included: (1) Canadian Center for Climate Modeling and Analysis (CCCMA), using the CGCM3 (T63 resolution) model and SRES A2 and B1 scenarios; (2) Geophysical Fluid Dynamics Laboratory (GFDL), using the CM2.1 model and SRES A2 and B1 scenarios; and (3) Met Office, Hadley Centre (UKMO), using the HadCM3 model, SRES A2 and B2 scenarios.

The climatic grids for future climate scenarios were developed by using general circulation model (GCM) output to calculate the expected monthly change in climate between the normalization period (1961–1990) and the decades centered on 2030, 2060, and 2090 for each weather station in the database. The relatively coarse grid projections of the GCMs were downsized to weather station point locations by weighting the average monthly change in climate calculated for the GCM cell centers lying within 400 km of a station. The inverse of the square of the distance from the station to the cell center was used for weighting. Once expected climatic values were generated for each weather station, thin-plate smoothing spline procedures were repeated to estimate monthly climate surfaces for average, minimum, and maximum temperature and precipitation for each GCM and each scenario. Derived variables were then calculated using the same procedures described above. The procedures for developing these grids are described in detail by Rehfeldt (2006) and Saenz-Romero et al. (2010). Estimates are available at URL: <http://forest.moscowfsl.wsu.edu/climate>.

2.4. Statistical procedures

2.4.1. Climatic profile

The statistical models developed through these analyses are based on regression tree analysis (RTA); a nonparametric data mining procedure (Breiman, 2001). RTA starts by identifying the best single predictor variable for partitioning the data set into two categories (referred to as branches). The partitioning is repeated to add branches until no further improvement in the model can be achieved. The result is a classification tree that predicts the condition of the dependent variable, either qualitative or quantitative, at each terminal node. The resulting classification tree has two components; a binary 'tree' structure showing the data partitions, and a set of simple linear logistic models, fitted one to each partition.

The methodology for modeling the binary variable, species occurrence, used in these analyses essentially duplicated those of Rehfeldt et al. (2009).

P. strobus occurred in approximately 8.6% of the approximately 123,500 records in the occurrence data set. To address the

Table 1

Acronyms and definitions of climatic variables used in Random Forest and parametric regression procedures. Mean, minimum and maximum values are for the ground plots with *Pinus strobus* present. Temperature-related variables are defined in units of °C and precipitation values in mm. Numeric values under Rank refer to the climate profile. The alphabetic value refers to the genetic model.

Acronym	Definition	Mean	Minimum	Maximum	Rank
ADI ^a	Annual dryness index: $(DD5)^{0.5}/MAP$	0.0458	0.0211	0.0817	5
ADIMINDDO	Annual dryness and cold index: $((DD5)^{0.5}/MAP) * MINDDO$	79.7260	3.5906	266.5515	–
D100 ^a	Julian date the sum of degree-days > 5 °C = 100	120.9	24.0	171.0	–
DD0 ^a	Annual mean degree-days < 0 °C based on monthly mean temperatures	943.9	0.0	2485.0	–
DD5 ^a	Annual mean degree-days above 5 °C	1999.1	919.0	4006.0	A
DD5MTCM	$(DD5 * MTCM)/1000$	–15.53	–30.95	45.19	–
FDAY ^a	Julian date of first freezing temperature in autumn	269.6	245.0	321.0	–
FFP ^a	Length of the frost-free period	130.4	77.0	250.0	–
GSD5 ^a	Mean degree-days > 5 °C within FFP	1621.1	598.0	4364.0	–
GSP ^a	Precipitation, April through September	549.4	335.0	1105.0	–
GSPDD5	$(GSP * DD5)/1000$	1117.19	436.64	3666.62	–
GSPMTCM	$(GSP * MTCM)/1000$	–4.61	–10.55	5.70	–
GSPTD	$(GSP * TDIFF)/100$	155.27	97.29	223.49	4
MAP ^a	Mean annual precipitation	1001.5	482.0	2215.0	–
MAPDD5	$(MAP * DD5)/1000$	2047.7	632.4	6784.8	–
MAPMTCM	$(MAP * MTCM)/1000$	–8.09	–19.05	11.48	–
MAPTD	$(MAP * TDIFF)/100$	279.95	189.91	438.57	3
MAT ^a	Mean annual temperature	6.23	–0.70	18.90	–
MMAX ^a	Maximum temperature in the warmest month	26.16	16.90	33.40	1
MMIN ^a	Minimum temperature in the coldest month	–14.33	–27.10	5.01	–
MMINDDO ^a	Annual degree-days < 0 °C based on monthly minimum temperatures	1728.1	65.0	3547.0	–
MTCM ^a	Mean temperature in coldest month	–	–8.75	–21.80 8.90	–
MTCMGSP	MTCM/GSP	–0.0167	–0.0651	0.0139	–
MTCMMAP	MTCM/MAP	–0.0097	–0.0452	0.0069	–
MTWM ^a	Mean temperature in the warmest month	19.78	13.80	27.30	–
PRATIO	GSP/MAP	0.5598	0.3997	0.7664	2
PRDD5	$(GSP /MAP) * DD5$	1113.53	428.08	2611.56	8
PRMTCM	$(GSP/MAP) * MTCM$	–5.07	–15.15	4.42	–
SDAY ^a	Julian date of last freezing temperature in spring	137.3	72.0	171.0	–
SDI ^a	Summer dryness index: $(GSD5)^{0.5}/GSP$	0.0734	0.0361	0.1048	–
SDIMINSSO	SDI * MMINDDO	139.17	7.23	383.52	–
TDCSP	TDIFF/GSP	0.0532	0.0179	0.1176	–
TDIFF ^a	MTWM – MTCM	28.53	18.40	39.40	6
TDMAP	TDIFF/MAP	0.0302	0.0089	0.0817	7

^a Variables used in bootstrap sampling procedures described in text.

statistical concern that classification trees be based on approximately equal representation across classes (Breiman, 2001), five samples were taken from the parental database using the three-step sampling protocol of Rehfeldt et al. (2009). First, plots with *P. strobus* present were duplicated ($2N = 18,762$), and defined as 40% of the total sample. The doubling of presence observations permitted a higher proportion of the absence observations to be included in each sample, and concentrated errors of prediction into predicting presences when absent. Second, an additional 40% of the sample targeted absence plots with climatic attributes that made them difficult to separate from those with a positive occurrence of *P. strobus*. To do this, the presence plots were used to define an 18-variable hypervolume (sensu Hutchinson, 1957). (See Table 1 for list of variables.) A random sample approximately equal to 40% of the total (about 18,200) was then randomly selected from the portion of the plot data lacking *P. strobus* that was within the hypervolume expanded by ± 0.01 standard deviation in all dimensions. The last 20% of the sample focussed on data points without *P. strobus* that were outside the hypervolume. The procedure involved determining the values for the first two principle components for the entire data set for the same 18 derived variables used in step 2, and subtending each principle component into 10 uniform classes for a total of 20 classes. Approximately 1% (460) of the total sample was randomly selected from within each principle component class of the non-occurrence plots outside the hypervolume. In total, samples were approximately 46,000 records.

The Random Forests classification tree (Breiman, 2001), available in R (R Development Core Team, 2004; Liaw and Wiener,

2002) was used to construct a set of 100 classification trees from each of the five bootstrap samples. Each set of 100 'trees' is referred to as a 'forest'. Using five 'forests' increased the probability that each observation within the hypervolume would be used in at least one 'forest'.

The Random Forests RTA began by reserving one third of each sample, referred to as the out-of-bag (OOB) subsample, for the purpose of computing classification errors. Breiman's algorithm was used to build classification trees from the remaining two thirds of the sample. Each node of a 'tree' was created by taking a random sample of the climatic variables approximately equal to the square root of the number of independent variables still available, and identifying the variable that minimized the classification error. Additional nodes were created using the same procedure until no further improvement could be achieved. OOB errors were calculated as the proportion of misclassifications relative to the total number of observations in the 'forest'. The OOB error term contains two types of misclassifications; (a) errors of commission where presence is predicted for absence plots; and (b) errors of omission where absence is predicted for plots with *P. strobus* present.

For each forest, a Random Forest stepwise procedure using 100 'trees' was used to rank the importance of all 35 climatic variables as predictors of presence for each forest. At each step, the least important variable was identified on the basis of the mean-decrease-in-accuracy statistic. The final ranking of variables was based on the mean importance of each variable across all five 'forests'. The number of variables in the final model was determined on the basis of the mean OOB error from across all five 'forests'. The number of variables in the model was judged to be parsimoni-

ous at the point where culling an additional variable resulted in a sudden increase in the OOB error (Rehfeldt and Jaquish, 2010). The final model was based on 100 ‘trees’ in each of five ‘forests’.

Once the ‘forests’ were created, the terminal nodes on each of the 500 ‘trees’ function as ‘votes’ concerning the predicted occurrence *P. strobus*. To make a prediction, a data point is run through all the ‘trees’ of all the ‘forests’ and the resulting percentage of positive votes provides a likelihood estimate of the climate being suited to *P. strobus*. The ‘votes’, therefore, have direct relevance to the climatic niche; to the current limits of distribution; and to the likelihood that *P. strobus* will be present. We chose to view ‘voting’ summaries as an estimate of the likelihood of occurrence of *P. strobus* in the current climate or the likelihood that a future climate would be suitable, even though our sampling protocol upset the balance between presence and absence data points occurring in natural landscapes. The climate of an observation was considered suited to *P. strobus* when a majority of the votes was affirmative (Rehfeldt et al., 2009; Rehfeldt and Jaquish, 2010; Saenz-Romero et al., 2010). Collectively, all suitable climates are referred to as the climate profile of *P. strobus*.

2.4.2. Ecological genetic profile

The comprehensive Ontario test series forms the cornerstone to developing an understanding of the ecological genetics of *P. strobus*. The premise for this test series is that the inter-correlated traits associated with the annual development cycle in *P. strobus* are the means by which natural selection adapts populations to climate. Farm-field testing can generate high quality genetic information at young ages through extensive site preparation, close spacing, and nearly complete weed control. The best assay of growth cycle traits is achieved when damaged trees are removed from the data set before analyses (e.g., Joyce et al., 2002b).

Analysis of variance of the Ontario data using SAS 9.1.3, PROC MIXED (SAS® Institute Inc., 2002–2003) was conducted separately for each test site. Block effects were initially treated as random in order to evaluate the effectiveness of blocking in removing large-scale environmental effects. Blocks were then considered as fixed effects for the purposes of calculating the proportion of the phenotypic variance (i.e., intraclass correlation) attributable to population effects. The linear model was:

$$Y_{jk} = U + B_j + P_k + B_j P_k + e_{jk}$$

where Y_{jk} is the observed plot mean for the k th population in the j th block. U is the overall test mean, B_j is the block effect, P_k is the population effect, $B_j P_k$ is the interaction of the j th block and the k th population, and e_{jk} is the residual of the k th population in the j th block. The interaction mean square was used to test the significance of the population effect.

The analysis of variance for freezing damage to needles used a linear model that included the main effect of treatment (temperature) as well as a treatment-by-block interaction, which were considered as fixed effects.

Freezing damage data was transformed using the arcsine square root function for the purposes of the analysis of variance. The linear model was:

$$Y_{ijk} = U + T_i + B_j + TB_{ij} + P_k + B_j P_k + e_{ijk}$$

where Y_{ijk} is the observed plot mean of the k th population in the j th block, in the i th temperature. U is the overall mean, T_i is the temperature effect, B_j is the block effect, P_k is the population effect, TB_{ij} is the interaction of the i th temperature and the j th block, $B_j P_k$ is the interaction of the j th block and the k th population, and e_{ijk} is the residual of the k th population in the j th block of the i th temperature. The $B_j P_k$ interaction mean square was used to test the significance of the population effect.

Intraclass correlations (r_t) were estimated as follows:

$$r_t = \sigma_p^2 / (\sigma_p^2 + \sigma_{bp}^2 / n_t)$$

where n_t is the harmonic mean number of trees per plot and σ_p^2 and σ_{bp}^2 are the population and the block-by-population variance components, respectively.

The error mean square divided by the number of blocks provided an estimate of the mean within-population variance. The square root then provided a mean within-population standard deviation. Confidence intervals were calculated using a significance level of 0.2 in order to manage the type II error, the probability of declaring no differences when differences are indeed present, because it is considered to be of greater concern than the type I error, declaring significant differences when none exist.

$$C.I._{0.2} = t_{A/2} * \text{sqrt}((\text{EMS})/b)$$

where $C.I._{0.2}$ is the confidence interval, $t_{A/2}$ is the value of ‘ t ’ at the desired significance level; EMS (an estimate of the root within-population variance) is the error mean square and ‘ b ’ is the number of blocks. Performance of populations was considered as the observed mean plus or minus the confidence interval, and two populations were judged to be significantly different when the mean of one was outside the confidence interval of the other.

2.4.3. Geographic patterns of genetic variation

To assess geographic patterns of genetic variation in the Ontario test series we ran stepwise multiple regression analysis using 34 candidate climate variables (Table 1) describing the climate of each provenance and population mean responses in six variables representing phenology, growth potential, and cold hardiness. The stepwise program was used to cull superfluous climatic variables and examine intercorrelations among the influential variables retained for further analyses.

The development of a regression model to predict range-wide variation in growth potential requires a single synthetic variable, visualized as an index to growth potential, that is derived from height measurements in the Ontario, Quebec and Maryland test series ($n = 3, 2,$ and $2,$ respectively). To create the variable, population means for each of seven height variables were scaled to values between 0 and 1 by subtracting the minimum population value from each population mean and dividing the remainder by the range in population means. To assure that distributions of scaled height were centered on a common climate, we calculated for each test the mean scaled height at the midpoint of the distribution of provenance climates, using all provenances in all tests and concentrating on the climate variables with the highest simple correlations with scaled height. The deviation of these test-specific means from 0.5 was used to center test series’ distributions by adding the deviation to the scaled height. For instance, to center the distributions so that a scaled height of 0.5 corresponded to 1700 DD5, about 0.03 would be added to scaled heights for the Ontario data, 0.10 subtracted the Quebec data, and 0.01 added to the Maryland data.

Data from the USFS and southern Appalachian test series were reserved to verify and supplement range-wide patterns, particularly pertaining to the provenances in the Appalachian region largely because of the small number of populations represented at each test site. These series of tests, however, included southern populations tested at southern sites; a condition not available from the more northern test series.

2.5. Mapping

Mapping was based on the 0.00833° grids for contemporary and projected climates described above. Estimates of the pertinent climate variables for each grid cell were inserted into the climate profile and ecological genetic profile algorithms to map projections of

the respective statistical models. All mapping was done with the ArcGIS software.

3. Results

3.1. Bioclimatic profile

Errors of prediction averaged across the five forests for each step in the stepwise elimination of variables were a constant 7.7% through the 19th step after which the out-of-bag errors gradually increased to 8.3% for the 7-variable model; to 10.6% for the three-variable model, and peaked at 22.4% for the one-variable model. We selected the 8-variable model with an out-of-bag error of 8.1% as being a reasonably parsimonious balance between predictive strength and simplistic models (see Rehfeldt et al., 2006, 2009). The out-of-bag error was composed of a 13.0% error of commission (predicting occurrence for plots where *P. strobus* was absent), and only 1.1% for the more serious errors of omission (predicted absence for plots documented to contain *P. strobus*).

The eight predictors in the final model are ranked according to their relative importance, as determined by the mean decrease in accuracy, are identified in Table 1. The ranking of the predictors was judged from the stepwise elimination sequence; that is the most important variable was the last variable remaining. The final model included a summer warmth variable (MMAX), two precipitation and temperature seasonality variables (PRATIO & TDIFF), and five variables expressing interactions between precipitation and temperature.

Histograms of the two most important climate profile variables illustrate the frequency of *P. strobus* occurrence within the distribution of eastern North American forests (Fig. 1). The figure shows that white pine is prevalent where summers are warm but not hot and growing season precipitation accounts for one-fifth to one-third of the total precipitation.

We assume that the likelihood of *P. strobus* would be present at a site is reflected in the proportion of correct 'votes' cast by the 100 classification trees from each of the five 'forests' (i.e., 500 'votes'). For the plots containing *P. strobus*, 63.4% received at least 95% of the 'votes' cast (Fig. 2). An additional 33.9% received at least 85%. Only 2.6% of the plots containing *P. strobus* received less than 75% of the 'votes'. No plots containing *P. strobus* received <55% of the total 'votes'. In contrast, 76.2% of the plots not containing *P. strobus* received <25% of the 'votes', with only 7.2% of the plots without this species receiving >60%. These statistics illustrate the ability of the model to accurately predict the likelihood that *P. strobus* would be present across the range of climates in which it occurs naturally.

Previous studies (e.g. Rehfeldt and Jaquish, 2010; Rehfeldt et al., 2009, 2006) have shown a strong concurrence between Little's (1971) range maps and climate profiles defined as locations with the likelihood of occurrence >50%. The application of the same standard to the *P. strobus* model accounts for all plots documented to contain *P. strobus*. In addition, 11% of the absence plots without *P. strobus* are predicted to have suitable climate. An examination of the spatial distribution of the misclassified plots (not presented here) showed the errors of commission to be interspersed with plots containing *P. strobus* throughout its range-wide distribution. Consequently, these classification errors are more likely associated with other factors such as edaphic constraints, stage of ecological succession, logging history, or human development than imprecision in the description of the realized climatic niche of *P. strobus*.

Mapped predictions of the geographic distribution are presented in Fig. 3 in two strata; (a) a likelihood of occurrence greater than 85% (green), and (b) a likelihood of occurrence between 50% and 85% (yellow). Little's (1971) digitized range map is included

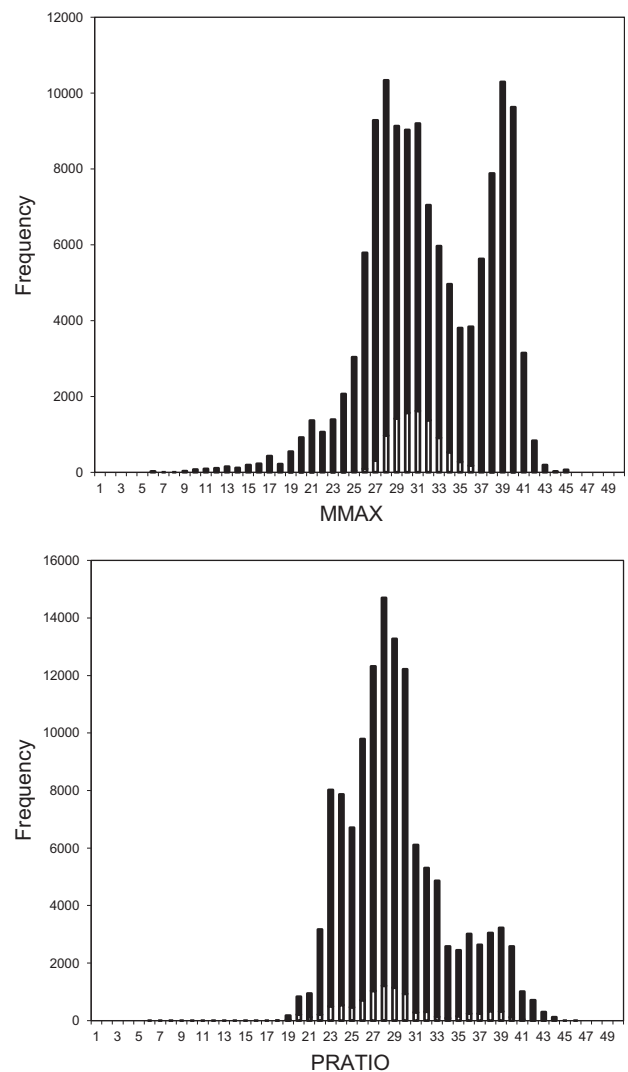


Fig. 1. Frequency of plots that contain *Pinus strobus* (white bars) relative to the frequency of the entire eastern North American plot database (black bars) plotted for 50 classes of the two most important predictors, mean maximum temperature (MMAX) and the ratio of growing season precipitation to annual precipitation (PRATIO).

for reference (USGS, 2005). As shown previously, the range map defines limits of distribution while the climate profile maps suitable habitat.

Fig. 4 illustrates at high resolution the precision of geographically referenced maps made from the climate profile in relation to the two-dimension mapping of range maps prepared decades previously. The distribution of observations in our database shows clearly that in northern Minnesota and adjacent Ontario, the older range map overstates the northern limits of the contemporary distribution (Fig. 4A). A similar situation occurs in Quebec north of the St. Lawrence River (Fig. 4B), but the localized northern limits *P. strobus* west of the Laurentian Mountains (red arrow) are underestimated by the range map. At the southern limits in Tennessee, Georgia and North Carolina (Fig. 4C), the occurrence of *P. strobus* in the Cumberland Mountains northwest of Knoxville is accurately represented by the range map, but our model also suggests suitable climates beyond the limits of the range map where *P. strobus* is largely absent today. In the Great Smokey Mountains to the south east of Knoxville (Fig. 4C), the highest elevations (dark gray) are outside the contemporary climate profile but within the limits of the range map. Similarly, in the central Appalachians (Fig. 4D), *P.*

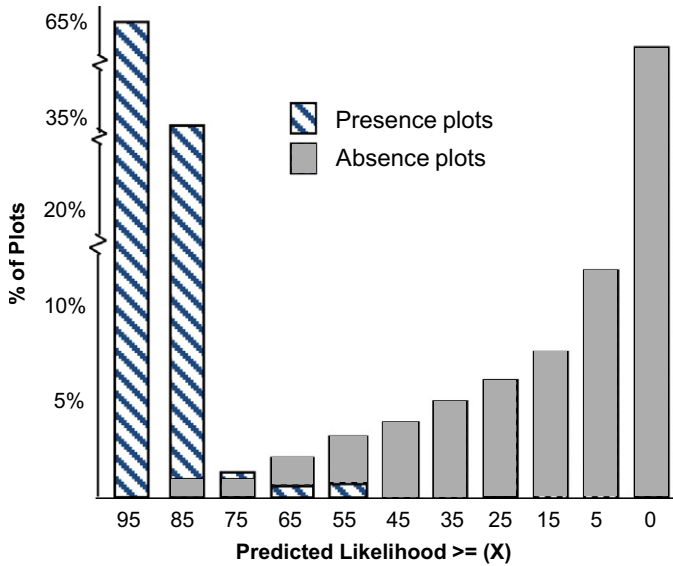


Fig. 2. Predicted likelihood of database ground plots either with (cross-hatched) or without (gray) *Pinus strobus* plotted as frequency distributions presented as the proportion of the total 'votes' cast by 100 classification trees in each of five Random Forests model.

strobus is largely restricted to the east-facing slope, a distinction not discernible in the range map. This abrupt edge of the climate profile along the Virginia – West Virginia border is likely a function of cold temperatures associated with the boundary of the 'Eastern Subalpine Forest and Tundra' biomes (see Brown et al., 1998). Our data also correctly illustrates that *P. strobus* occurs with equal frequency on either side of range limits purported by the range map along the Ohio – West Virginia border.

The likelihood of occurrence reported here also infers abundance as well as limits of distribution. Figs. 3 and 4 depict the same pattern of relative abundance reported by Iverson and Prasad (1998) for the United States. In the remainder of this paper, the mapped climate profile is used in place of the range map to represent the contemporary distribution of *P. strobus* and the foundation for projecting the future geographic distribution of suitable habitat.

Differences among the general circulation model projections of the future distribution of the contemporary climate profile are obvious yet the differences are largely one of timing (Fig. 5). All three GCM's project climates that should force severe early and sustained shifts in the contemporary climate profile. There is a consensus for the 2030 projections that the trailing edge of climate profile in the Appalachians and Lake States will recede toward the Canadian border. The 2060 projections continue the trend of rapid contraction at the southern limits with expansion of suitable

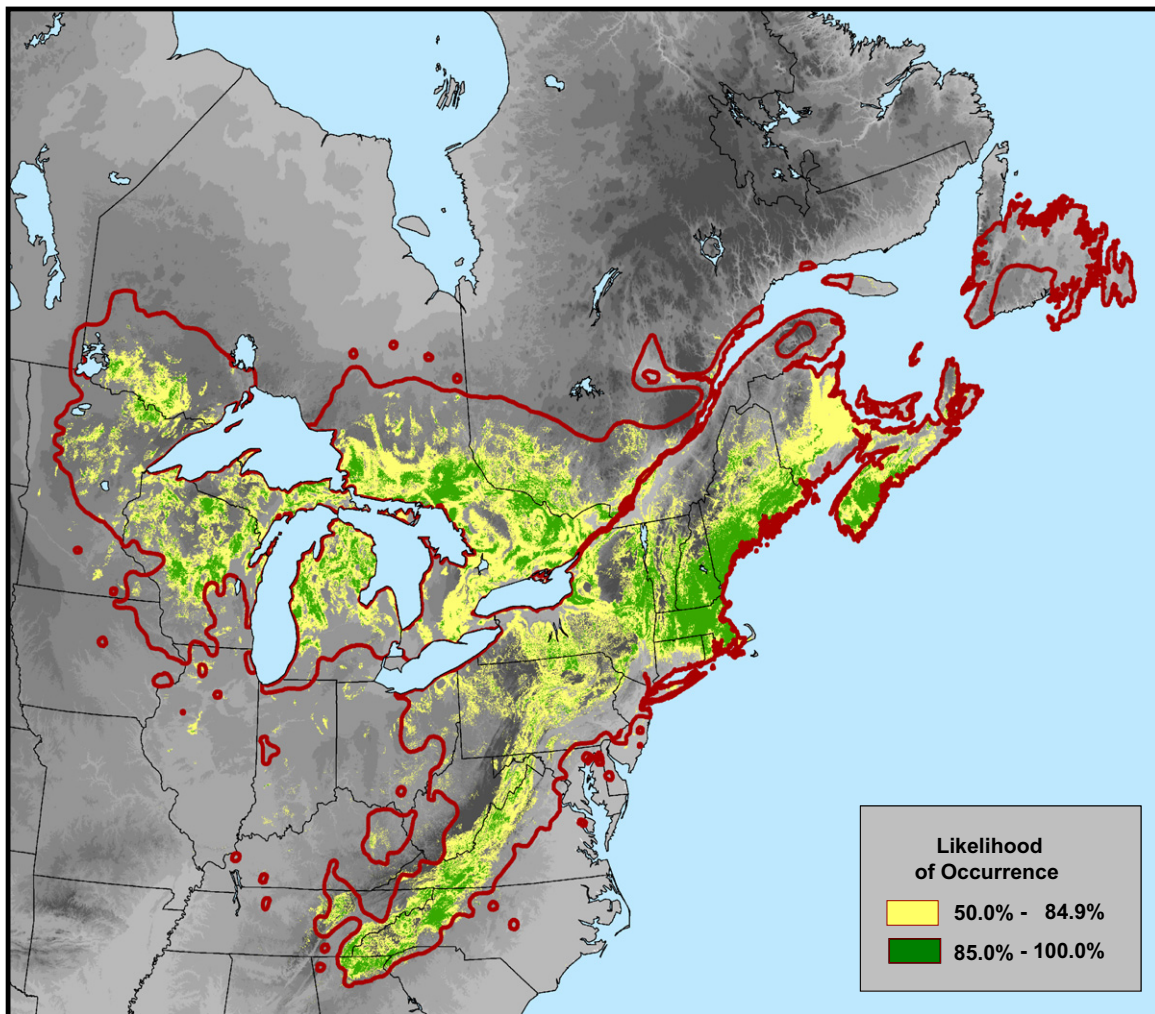


Fig. 3. Mapped climatic profile of *Pinus strobus* in relation to Little's (1971) digitized range map (red line). Likelihood of occurrence, the predicted number of votes, is superimposed on the digital elevation model of the Globe Task Team (1999) in 100 meter intervals (lightest, 0–100 m).

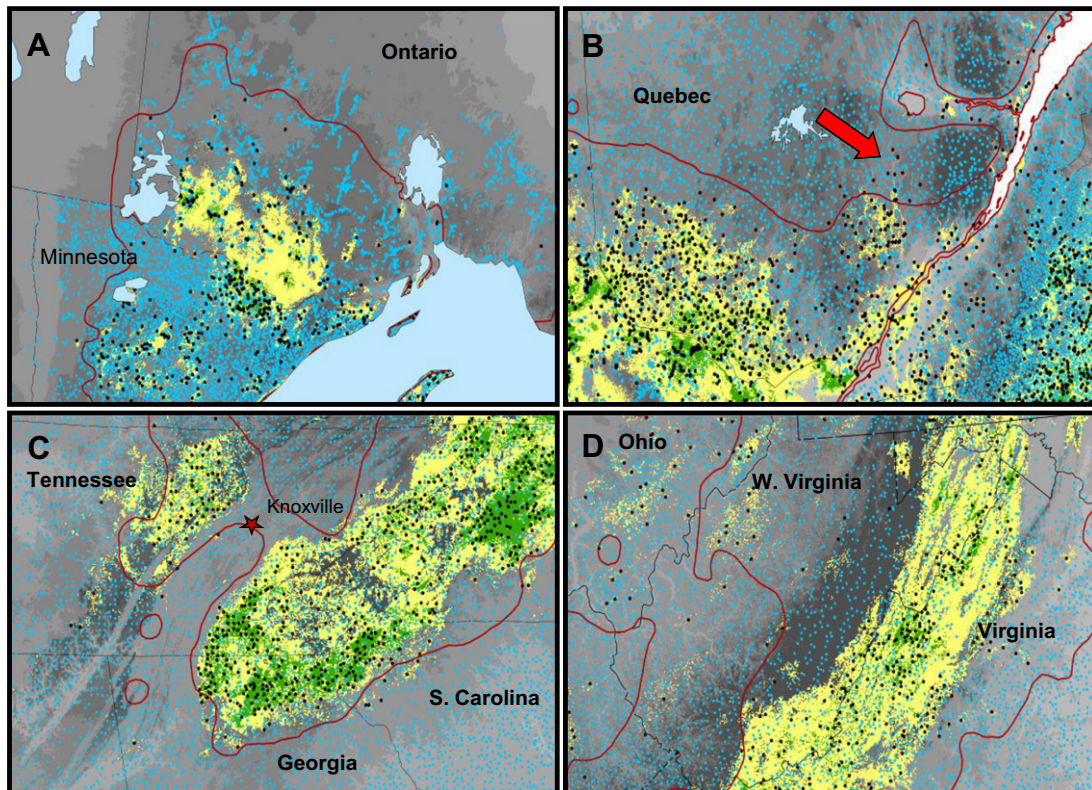


Fig. 4. Mapped climate profile of *Pinus strobus*. Yellow and green represent 50–85%, and 85–100% likelihood of occurrence, respectively. Black dots indicate observations where *P. strobus* is present, and blue dots represent locations without *P. strobus*. Red arrow in Fig. 4b indicates area west Laurentian Mountains referred to in the text.

habitat to the north. Also obvious in Fig. 5 are the tremendous differences in suitable future climate niche projected by the three GCM's. The UKMO projections suggest that *P. strobus* may become a minor component of future forests. But, the CCCMA and GFDL projections suggest that the species could flourish to the north of the contemporary distribution through 2060. By 2090, climate profile projections are centered in northern Quebec, Labrador, and the Maritime Provinces with relatively little overlap with the contemporary distribution. While the CCCMA projects suitable habitat in the Hudson Bay lowlands, soil moisture conditions may restrict suitable habitat to specific microsites. That the projections generally follow the same trajectory but differ in timing was also noted in analyses of impacts to North American biomes (Rehfeldt et al., 2012) for which the conclusion was reached that the result reflects the similarity among GCM's for temperature projections but the widely disparate projections for precipitation. This conclusion seems applicable for *P. strobus* also because interactions of temperature and precipitation contribute heavily to the climate profile (Table 1).

Also evident across all projections is that the spatial extent of high likelihood of occurrence (>85%) is substantially reduced relative to the contemporary distribution. Given that 97% of the plots with *P. strobus* have greater than 85% likelihood of occurrence, these projections suggest diminished competitive ability in emergent suitable habitat. The complexity of the climate profile model precludes assessing the causal agents of these projections, but undoubtedly includes shifts in the moisture balance.

Despite the disparate projected impacts of the GCM's (Fig. 5), overlaying the six mapped projections (3 GCM's and 2 IPCC scenarios) shows a high degree of consensus in the projected redistribution of suitable habitat (Fig. 6). By adopting a majority threshold (at least four projections agreeing) the uncertainty in mapped potential suitable habitat can be taken into consideration (Rehfeldt

et al., 2009; Rehfeldt and Jaquish, 2010). Fig. 6 shows that the highest likelihood of suitable habitat at 2030 is in a latitudinal band extending from the upper peninsula of Michigan across the north shore of Lake Huron to up-state New York and through New England to the Maritime Provinces (Fig. 6A). By 2060, the models concur that suitable habitat most likely will be centered in Ontario north and east of the Great Lakes, and in the Maritime Provinces (Fig. 6B). The figure also suggests that the highest elevations in western North Carolina appear to have some potential as a refugium for *P. strobus* through 2060.

3.2. Ecological genetics

3.2.1. Ontario test series

Summaries of analyses of variance for six variables (Table 2) show statistically significant ($p < .001$) population effects for all traits. These effects accounted for nearly one-half or more of the total variance in four of the traits. These effects, moreover, were intercorrelated such that populations with superior height growth also had longer needles, grew longer, and had greater freezing damage. But, population effects were strongest for height growth and late growth.

Stepwise multiple regression analysis for predicting genetic variation associated with climatic gradients were all statistically significant ($p > 0.0001$). Univariate models were optimal for all variables, and degree-days $> 5^{\circ}\text{C}$ (DD5) dominated the regression models for growth (Table 3a). The strongest model (Ht5) accounted for 68% of the variation in mean height of populations. In contrast to growth and phenology, freezing injury to needles was most strongly associated with mean minimum temperature (MMIN), but only accounted for 26% of the variation among populations.

The general geographic pattern of genetic variation for growth potential (not shown) is a cline of increasing potential from the

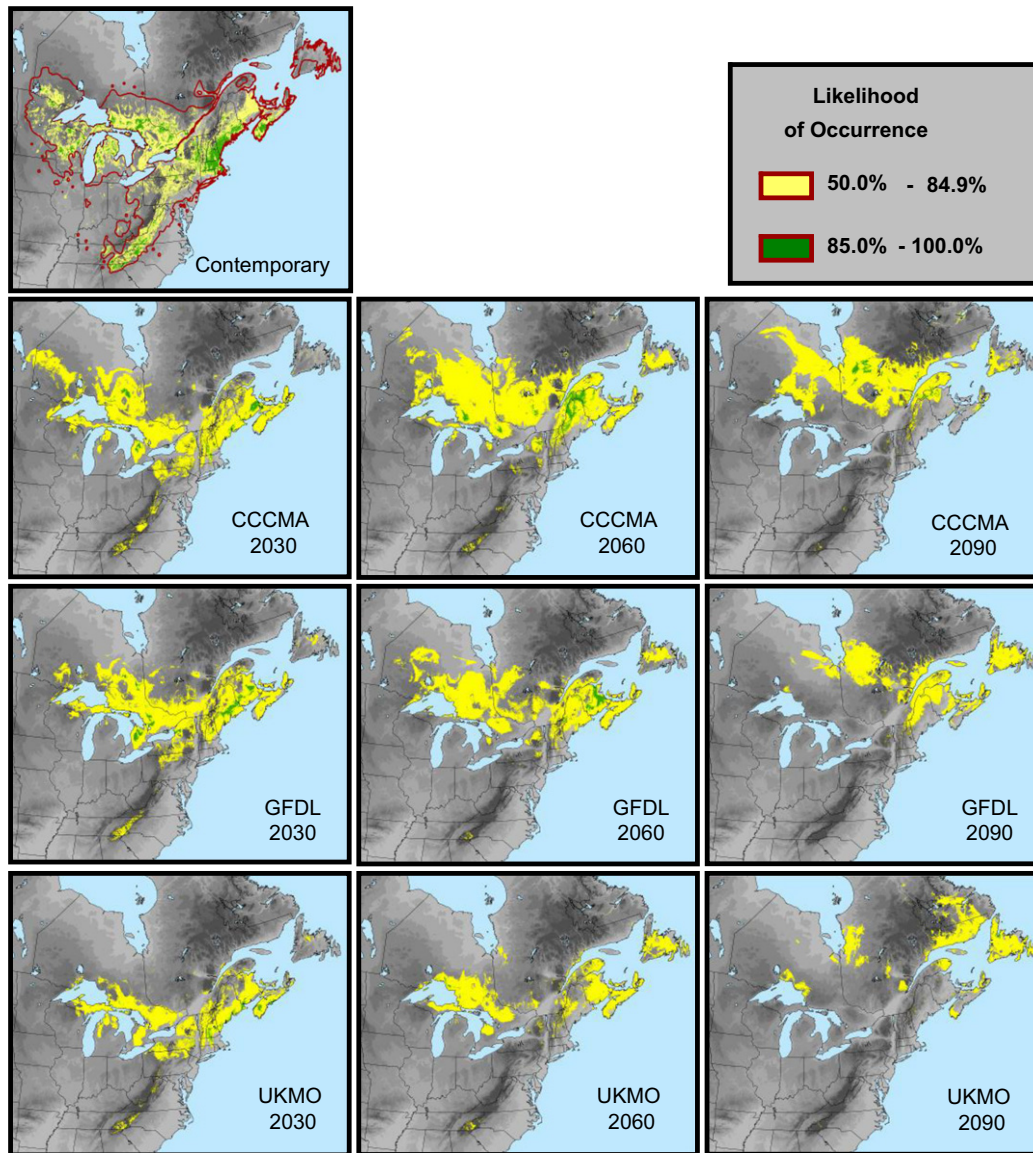


Fig. 5. Mapped projected distribution of the contemporary climate profile for *Pinus strobus* under the IPCC A2 scenario for the decades centered on 2030, 2060, and 2090 and three general circulation models: CCCMA, Canadian Center for Climate Modeling and Analysis; GFDL, Geophysical Fluid Dynamics Laboratory; UKMO, Hadley Centre.

northwest to the southeast of the study area. Genetic variation in needle damage from freezing studies is arranged along a north–south cline. This latter result corresponds favorably with that of Maronek and Flint (1974) who found significant effects among populations for freezing tolerance but resulting latitudinal clines nonetheless were relatively flat. However, in the Ontario test series, the relative flatness of the cold hardiness cline and the strong correlation between DD5 and MMIN ($r = 0.80$) for the sampled populations means that genetic variation in timing of cold acclimation is largely redundant from a practical perspective.

3.2.2. Range-wide clines: Synthesis of Ontario, Quebec and Maryland tests

Multiple regression analysis of mean height of populations on climate variables for both the Quebec and Maryland test series were remarkably consistent with the Ontario growth potential model (Table 3b and c, respectively). In the Quebec test series, DD5 provided the best fitting model for predicting 10-year mean height of 61 populations in three field performance tests. The strongest univariate models for the other three height variables

were not driven by DD5, but the intent of the analyses was to test the effectiveness of the DD5 rather than create models de novo. A comparison of the models based on DD5 against the best models showed that there were only minor differences in predictive power.

3.2.3. Synthesis of Ontario, Quebec and Maryland tests

Multiple regression analysis of the synthetic variable ‘Scaled Ht’ found DD5 to be the best climatic predictor (Table 3d). When scaled heights are plotted against DD5, the effectiveness of the scaling procedure becomes apparent, and the scatter demonstrates the utility of the procedure for combining disparate data sets (Fig. 7). However, in the field tests of both range-wide studies conducted in Maryland and Quebec, a few of the populations with the highest provenance DD5 values were shorter than predicted by a linear model. As a result, a univariate model based on the natural log of DD5 produced a slightly better fit to the data (Table 3d).

In interpreting Fig. 7, it is important to note that the climates of the Quebec tests and one of the Maryland tests were substantially cooler than those of the southern provenances, and the remaining

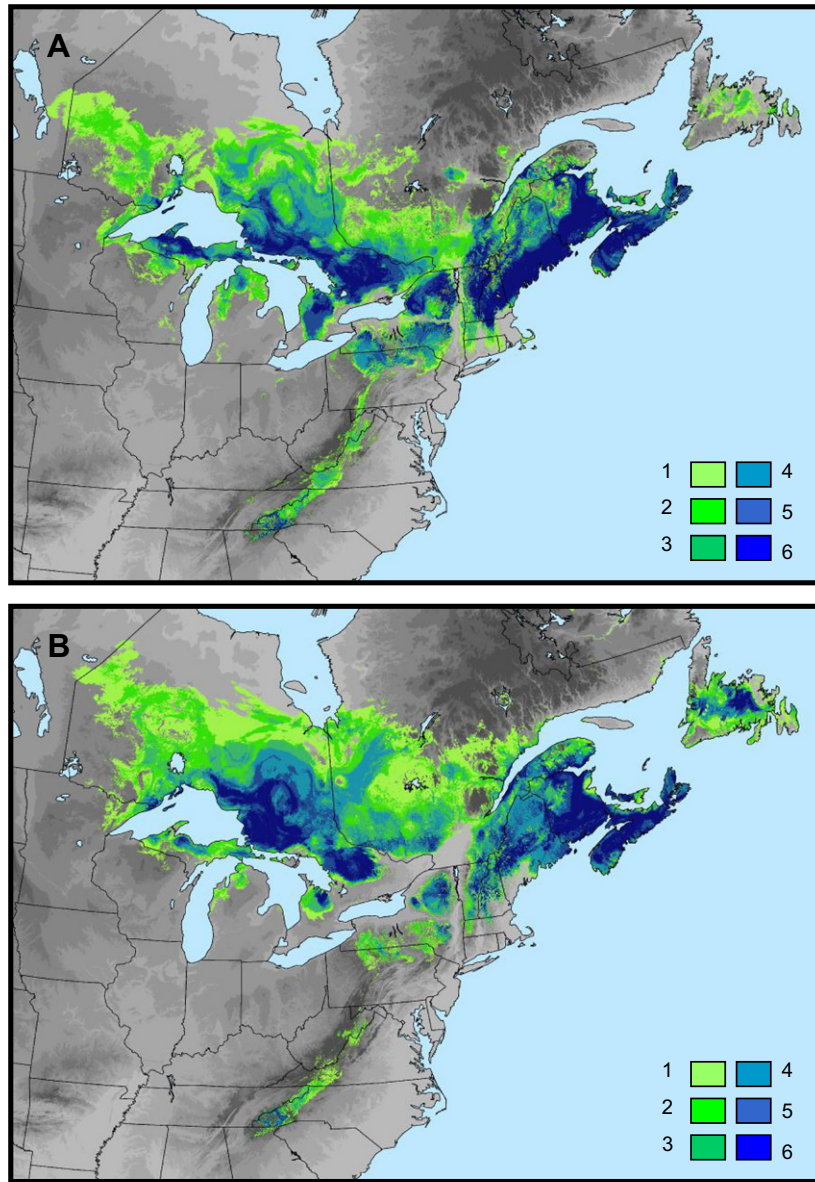


Fig. 6. Mapped projections of the climate profile for the decades centered on 2030 (A) and 2060 (B) superimposed for three GCMs (CCCMA, GFDL, and UKMO) and two greenhouse gas emissions scenarios. Shading indicates the number of projections that agree on habitat suitability for *Pinus strobus*: light to dark green represent 1–3 projections, light to dark blue represent 4–6 projections.

Table 2

Intraclass correlations (the ratio of the population variance component to the sum of variance components for random model effects) from analysis of variance for eight variables representing 112 eastern *Pinus strobus* populations in Ontario.

Variable	Mean	n_{ij} ^a	r^b	Root w/population variance	C.I. ^c
<i>Greenhouse</i>					
2-Year height (mm)	166.4	8.7	0.529***	6.6553	8.58
Late growth (mm)	17.2	8.7	0.485***	2.3192	2.99
Needle length (mm)	101.0	8.7	0.183***	2.4127	3.11
<i>Farm-field tests</i>					
3-Year height (cm)	30.7	9.2	0.555***	1.3555	1.75
5-Year height (cm)	159.1	8.8	0.532***	5.1258	6.61
Needle damage (%) ^d	0.811	9.4	0.274***	0.1353	0.17

^a Harmonic mean number of trees per plot.

^b Population intraclass correlation.

^c Confidence Interval at alpha = 0.2.

^d Analysis was based on the arcsine transformation of the proportion of needles damaged.

*** Statistical significance at $p < 0.001$.

Table 3
Results of multiple regression analyses to describe genetic variation in *Pinus strobus* in response to climate.

Dependent variable	Independent variable ^a	R^2	Model	
			Intercept	Slope
<i>(a) Ontario</i>				
5-Year height (Northern Ontario)	DD5 ^{***}	0.686	81.729	0.046
3-Year height (Southern Ontario)	SMI ^{b,***}	0.522	15.779	4.032
	DD5 ^{***}	0.499	10.944	0.012
2-Year height (greenhouse)	DD5 ^{***}	0.358	105.808	0.036
Late growth (greenhouse)	DD5 ^{***}	0.253	0.210	0.010
Needle length (greenhouse)	DD5 ^{***}	0.133	92.609	0.005
Cold damage (greenhouse)	MMIN ^{b,***}	0.263	0.879	0.020
	DD5 ^{***}	0.179	0.184	0.0002
<i>(b) Quebec</i>				
4-Year height	SMI ^{b,***}	0.366	32.881	6.263
	DD5 ^{***}	0.320	33.141	0.012
10-Year mean height ^c	DD5 ^{***}	0.441	127.844	0.033
<i>(c) Maryland</i>				
2-Year height	MAT ^{b,***}	0.477	11.683	0.555
	DD5 ^{***}	0.458	8.108	0.004
16-Year mean height ^d	MAPMTCM ^{b,***}	0.261	107.728	1.420
	DD5 ^{***}	0.204	72.978	0.012
<i>(d) Scaled and adjusted height</i> (all 7 height variables above)				
	DD5 ^{***}	0.367	0.035	0.0003
	ln(DD5) ^{***}	0.412	-4.674	0.699

^a Variables defined in Table 1.

^b Univariate model with the largest R^2 .

^c Mean height across three field performance tests.

^d Mean height across 2 field performance tests.

*** Statistically significant at $p < .001$.

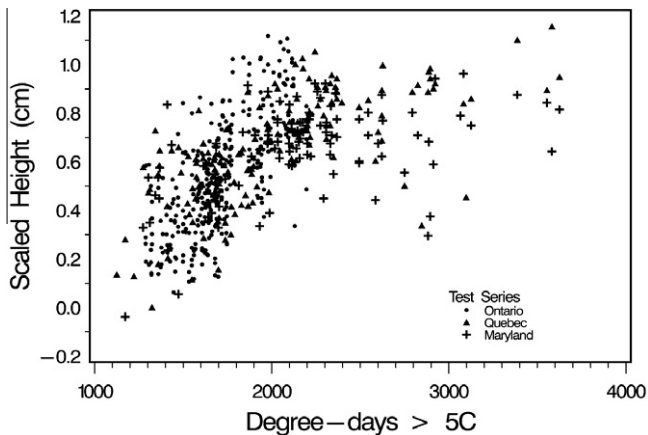


Fig. 7. Plot of the synthetic variable 'scaled height', derived from disparate height growth data from three test series of *Pinus strobus* populations, against degree-days $> 5^\circ\text{C}$ at provenance origin.

Maryland nursery and field tests were outside the bioclimate profile. Whether southern populations could express their innate growth potential in such test environments is problematic (Leites et al., 2012; Rehfeldt et al., 1999). This may account, therefore, for the apparent lack of close fit at high dd5 in Fig. 7, resulting, perhaps from variables related to freezing tolerance (see Table 5) contributing to population adaptedness in that region.

The 5-year height confidence interval (Table 2) was used as to classify and map the contemporary range-wide clinal pattern of ecological genetic variation predicted by the ln(DD5) model (Table 3, Fig. 8). Altitudinal gradients are apparent in the central Appalachian Mountains and the Adirondack Mountains of New York, but otherwise the pattern is closely associated with latitude. As documented by the Ontario test series, the region with the lowest growth potential is centered on the east end of Lake Superior. The yellow interval indicates those populations predicted to have average growth potential. With the exception of a small region in

West Virginia, the populations from the central Appalachian Mountains are expected to have superior growth potential with populations from the southern Appalachians (Georgia, Tennessee and North Carolina) predicted to have the greatest growth potential.

The effectiveness of a single climate variable in predicting variation in growth potential of three independent provenance test series is remarkable, but consideration of potential bias in the sample of the provenances used to build the model is a prerequisite to practical application (Draper and Smith, 1981). While the sample of provenances is range-wide, concern is centered on the robustness of the sample in terms of the predictive variable, DD5. The frequency distribution of DD5 for the 9831 observations containing *P. strobus* have a mean of 1999, and a range from 919 to 4009 degree-days $> 5^\circ\text{C}$. An overlay of the distribution of the provenances included in the Ontario, Quebec and Maryland test series shows that the sum of provenances included in those test series represent a robust sample of the overall distribution below 2500 DD5 (Fig. 9). The sampling of the warmer tail of the distribution (DD5 > 2800) only included 17 out of 349 provenances (5%). The distribution of observations with DD5 > 2800 is closely associated with the south and central Appalachian Mountains. Consequently, the applicability of the ecological genetic model for the Appalachian Mountains south of Pennsylvania merits consideration.

3.2.4. Verification: USFS and southern Appalachian test series

The USFS test series included a range-wide selection of provenances tested across a broad range of test environments but with a relatively small number of provenances in each test. In Table 4, the test sites in this series of tests are arranged and grouped according to DD5. The simple correlations of DD5 with population growth, expressed as a proportion of the plantation mean, show that DD5 is strongly correlated with performance at the warmest test sites where trees of the three southernmost provenances DD5 > 2800 (2913, 3029, 3305 for Tennessee, Georgia, and North Carolina provenances, respectively) can express their growth potential. As test sites become cooler, the relative performance of

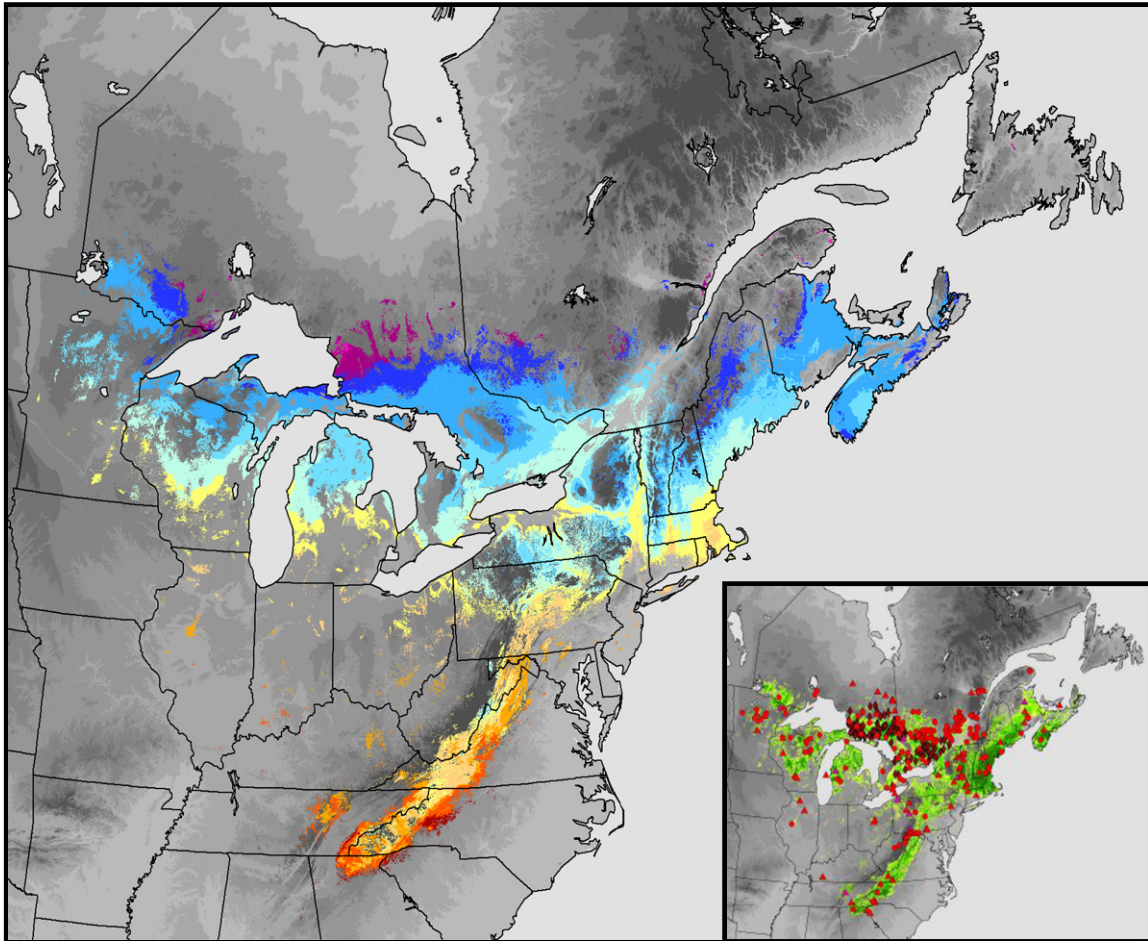


Fig. 8. Mapped pattern of genetic variation for growth potential (scaled height) of *Pinus strobus* predicted from a regression driven by $\ln(\text{DD5})$ using data from three disparate provenance tests (Table 3). Clinal variation has been classified by using confidence intervals for 5-year height reported in Table 2. Red, orange and yellow identify the distribution of populations predicted to have superior growth potential. Blue and purple indicate populations predicted to have the lowest growth potential. Inset locates provenances represented in scaled height data set.

these southernmost seed sources deteriorated with respect to the other provenances, causing the correlation of performance with DD5 of the provenance to be reduced to non-significance. We note, moreover, that a strong argument could be made that the observed patterns reflect the coldness of winters of the provenance as much as the warmth of summers; correlations of MMIN and DD0 with relative growth at the test site are essentially the same as with DD5. Nonetheless, while consistent with previous analyses (Garrett et al., 1973; Genys, 1990) these results suggest that the performance of southern Appalachian populations is strongly conditioned by genotype by environment interactions at these test sites. The result is also consistent with Fig. 7 which shows the southern Appalachian provenances to be similar to each other genetically but different from those to the north. These differences are expressed phenotypically by genotype by environment interactions dependent on climatic disparity of the test sites. In terms of clinal variation being dependent on the relative warmth of the climate of the seed source, these results provide a strong validation.

The population performances in the southern Appalachian test series are considered in Table 5 where, except for the last entry, test sites also are arranged according to their DD5. The Nebraska test site is listed last because the combination of relative high DD5 and low MMIN place it outside the climate profile of *P. strobus*. At this site, therefore, the Appalachian provenances could not express their growth potential. Interpretations of the results in Table 5 are consistent but not as clear as those in Table 4. Because

of the restricted geographic context of the study, simple correlations become more dependent on precisely which provenances were included in the test. Nonetheless, Table 5 shows that the southern Appalachian populations grew the best at the warmest test sites (Group A), while inter-regional differences were muted at the cooler sites of Group B. In the coldest tests (group C), the growth of populations from both regions were essentially the same. While DD5 remains a robust predictor of growth performance particularly on the warmest sites, that MMIN is a stronger predictor in the southern and central Appalachian Mountains reinforces the adaptive role of cold hardiness in ecological genetic variation. Overall, these analyses confirm the veracity of the predicted range-wide pattern of growth potential predicted from the Ontario test series, but includes the caveat that growth performance is a function of adaptation to winter temperatures as well.

4. Discussion

4.1. Bioclimate and genetic profiles

The Random Forests regression tree algorithm of Breiman (2001) applied to five bootstrap sample ‘forests’ consistently produced models that falsely predicted positive occurrence (errors of commission) for less than 13% of the observations. Such modest errors of commission are remarkable given the importance of

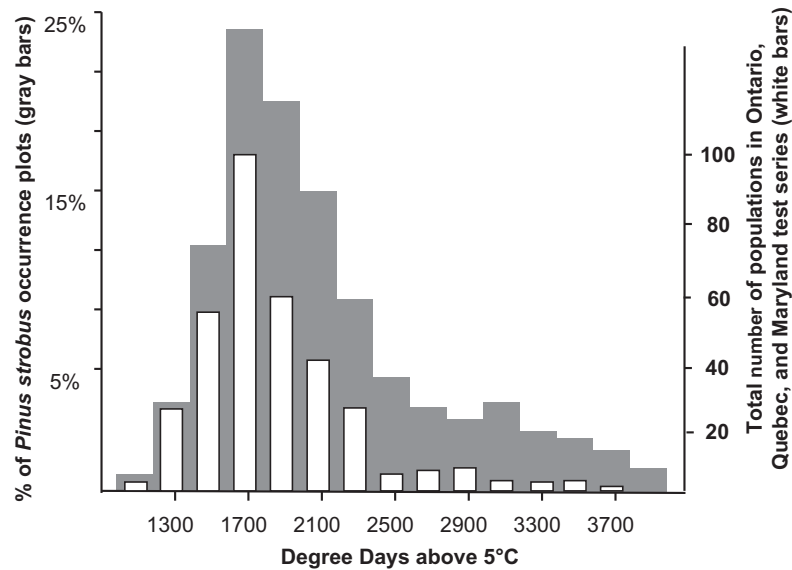


Fig. 9. Frequency distribution according to classes of degree-days > 5 °C for ground plots with *Pinus strobus* present (gray bars) compared to that for the provenances included in the Ontario, Quebec, and Maryland test series (white bars).

Table 4

Tests in the USFS range-wide *Pinus strobus* test series listed in descending order by predicted degree-day > 5 °C (DD5) and grouped according to the mean performance of the three southern Appalachian populations included in the study. CorrDD5 refers to the correlation between provenance DD5 and population mean height at age 10 expressed as a percentage of individual test means.

Test Location	N ^a	Latitude (decimal)	Longitude (decimal)	Elevation (m) ^b	DD5 ^c	CorrDD5	S. Appalachian Mean Ht ^d (%)
<i>Group A</i>							
Hardin County, Ill.	16	37.47	88.50	55	3499	0.634	123
Maryland (11) ^e	27	38.50	77.33	91	3422	0.779	126
Perry County, Ind.	16	38.14	86.61	62	3309	0.666	123
Laurel County, Kent.	16	36.95	64.36	91	3287	0.735	128
Kentucky (13)	16	36.98	84.28	335	3231	0.704	130
Pennsylvania (8)	28	39.87	75.68	122	2961	0.514	122
Ohio (12)	16	39.47	82.20	213	2893	0.582	135
Pennsylvania (7)	28	40.62	79.58	293	2555	0.681	125
West Virginia (10)	26	39.18	79.58	524	2499	0.613	115
<i>Group B</i>							
Pennsylvania (6)	22	41.83	79.25	360	2196	0.546	115
Vermont (3)	26	44.47	73.00	30	2116	ns	105
Maine (2)	27	43.53	70.67	91	2108	ns	98
Maryland (9)	24	39.67	79.25	823	2101	0.595	110
<i>Group C</i>							
Wexford Co., Mich. ^f	16	44.20	85.80	84	2045	0.703	115 (Ht9)
Massachusetts (5)	22	42.73	73.23	320	2004	ns	92
Maine (1)	26	44.88	68.65	30	1911	ns	80
Cass Co., Minn. ^f	16	47.40	94.50	119	1868	ns	75 (Ht7)
Forest Co., Wisc. ^f	16	45.40	88.50	125	1800	ns	76 (Ht9)
H. S. Nursery., Wisc. ^f	16	45.50	89.50	151	1788	ns	95 (Ht4)
Schoolcraft Co., Mich. ^f	16	46.00	86.40	61	1773	ns	93 (Ht9)
New York (4)	23	44.43	74.22	553	1619	ns	90

^a Number of populations represented in the test.

^b Elevations were estimated from a digital elevation model for the geographic coordinates reported in the literature.

^c Degree-day > 5 °C for individual tests.

^d Mean Ht refers to the mean height of the Tennessee, Georgia and North Carolina populations at age 10 expressed as a percentage of each test mean. The height measurements for the Midwest states ranged from ages 4 to 9.

^e For the Eastern Seaboard tests, the number in parentheses refers to test number as identified Garrett et al. (1973). The reference report (Funk et al., 1975) did not number the Midwestern state tests.

^f H. Sauer Nursery – North Carolina population was not included in these tests (King and Nienstadt, 1968).

edaphic conditions and anthropogenic and stochastic environmental disturbance at the site level. Erroneous predictions of absence (errors of omission) of 1% are indicative of the power of regression trees to properly classify data (Rehfeldt, 2006). But it also worth noting that the excellent fit achieved by this model required high

quality predictions of climate conditions for each location in a robust eastern North American ground plot network. The most important variable in the bioclimate profile of *P. strobus* is mean maximum temperature, but variables representing the complex interactions between precipitation and temperature dominate

Table 5

Tests in the Southern Appalachian *Pinus strobus* test series listed by degree-days > 5 °C (DD5) of the test site, and arbitrarily grouped according to the relative performance of the populations from the southern and central Appalachian provenances. CorrDD5 and CorrMMIN refer to the correlation between DD5 and the mean monthly minimum temperature in the coldest month (MMIN) with population mean height expressed as a percentage of the test site mean.

Test location	Height @age	N ^a	Latitude	Longitude	Elevation (m) ^b	DD5	MMIN (°C)	Corr DD5	Corr MMIN	Southern mean Ht ^c	Central mean Ht ^c (%)
<i>Group A</i>											
Tullahoma, Tenn.	7	16/13	35.40	86.30	300	3694	−3.3	0.559	0.712	114%	82
Pope Co., Ill.	11	10/4	37.30	88.70	120	3593	−5.2	0.490 ns	0.474 ns	111%	88
Jackson Co., Ill.	9	14/5	37.70	89.30	150	3480	−6.2	0.552	0.647	101%	79
Lafayette, Ind.	9	16/7	40.40	86.80	180	2851	−9.7	0.557	0.712	98%	78
<i>Group B</i>											
Cass Co., Mich.	9	19/11	42.00	86.00	280	2516	−9.1	0.537	0.592	105%	92
Lenawee Co., Mich.	10	14/9	42.00	84.20	200	2505	−10.2	0.409	0.244 ns	100%	97
Kalamazoo Co., Mich.	10	20/17	42.20	85.20	270	2495	−10.0	0.449	0.466	106%	98
Loudonville, Ohio	8	17/11	40.60	82.30	378	2403	−10.0	0.449	0.550	106%	84
Parsons, W. Virg.	8	17/4	39.10	79.70	700	2266	−9.0	0.465	0.524	106%	83
<i>Group C</i>											
Roscommon Co., Mich.	8	18/11	44.50	84.60	300	1961	−13.7	0.171 ns	0.157 ns	103%	97%
Plattsmouth, Neb.	9	21/10	41.00	95.90	335	2931	−12.0	−0.058 ns	0.015 ns	101%	98

^a Number of populations represented in the test. The split numbers refer to the provenances originating in southern (Tennessee, Georgia, and North Carolina) and central regions (West Virginia, Virginia and Maryland), respectively.

^b Elevations were estimated from a digital elevation model for the geographic coordinates reported in Wright et al. (1978).

^c Southern- and central-Mean Ht refers to the mean height (expressed as a percentage of individual test mean height) of all populations originating in the southern and central regions, respectively.

the overall model. Overall the model indicates that the distribution of *P. strobus* is a function of seasonal temperatures and moisture balance.

When the 100 classification trees of each of the five ‘forests’ were merged, and a standard of 50% of the ‘trees’ predicting occurrence was adopted as the threshold for defining a bioclimate profile, no observed occurrences of *P. strobus* were misclassified, and

only 11% of locations without *P. strobus* exceeded the threshold. The result is a robust understanding of the climatic conditions associated with the contemporary distribution of *P. strobus*.

The effectiveness of close observation of phenology, growth potential and cold hardiness in short term tests maintained under intensive culture to develop high quality ecological genetic models is well documented in the literature (e.g., Campbell and Sorensen,

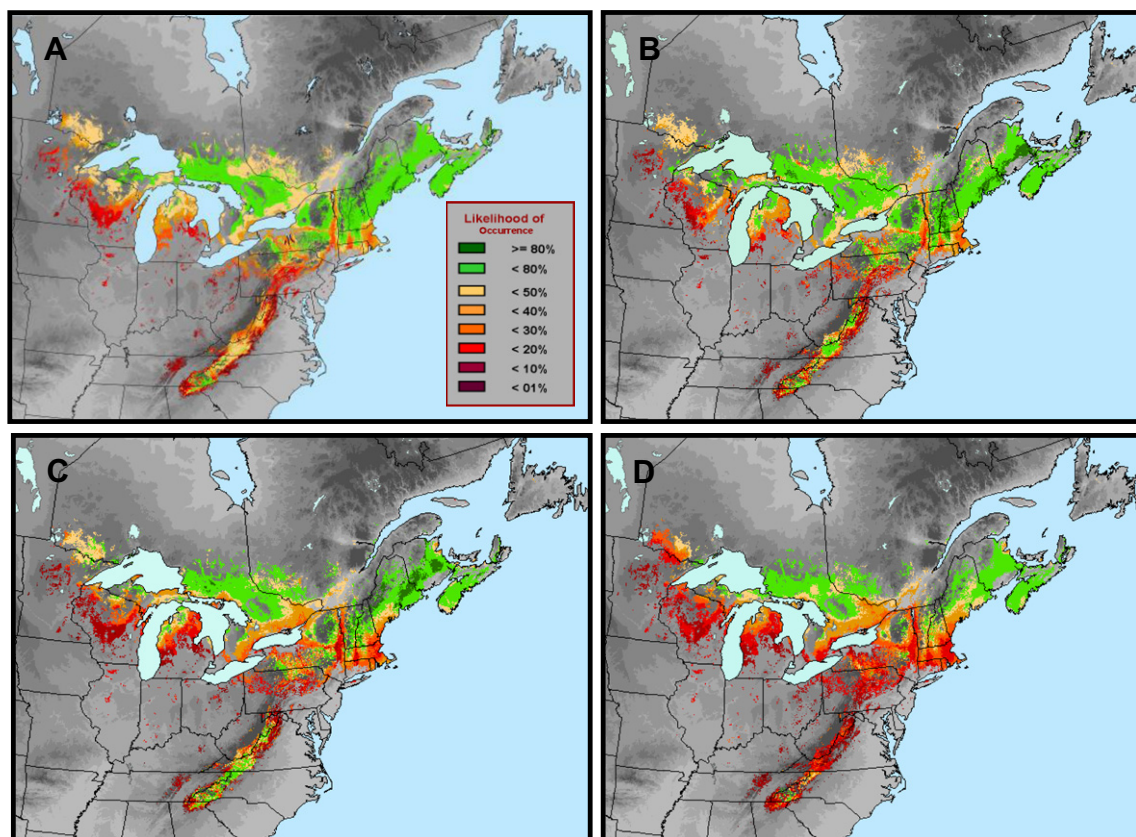


Fig. 10. Mapped projections of the mean likelihood of occurrence, for three GCM's (CCCMA, GFDL, and UKMO), of *Pinus strobus* within its contemporary distribution for the decades centered on 2030 and 2060. (A) 2030 mean IPCC scenario B1; (B) 2030 mean IPCC scenario A2; (C) 2060 mean IPCC scenario B1; (D) 2060 mean IPCC scenario A2.

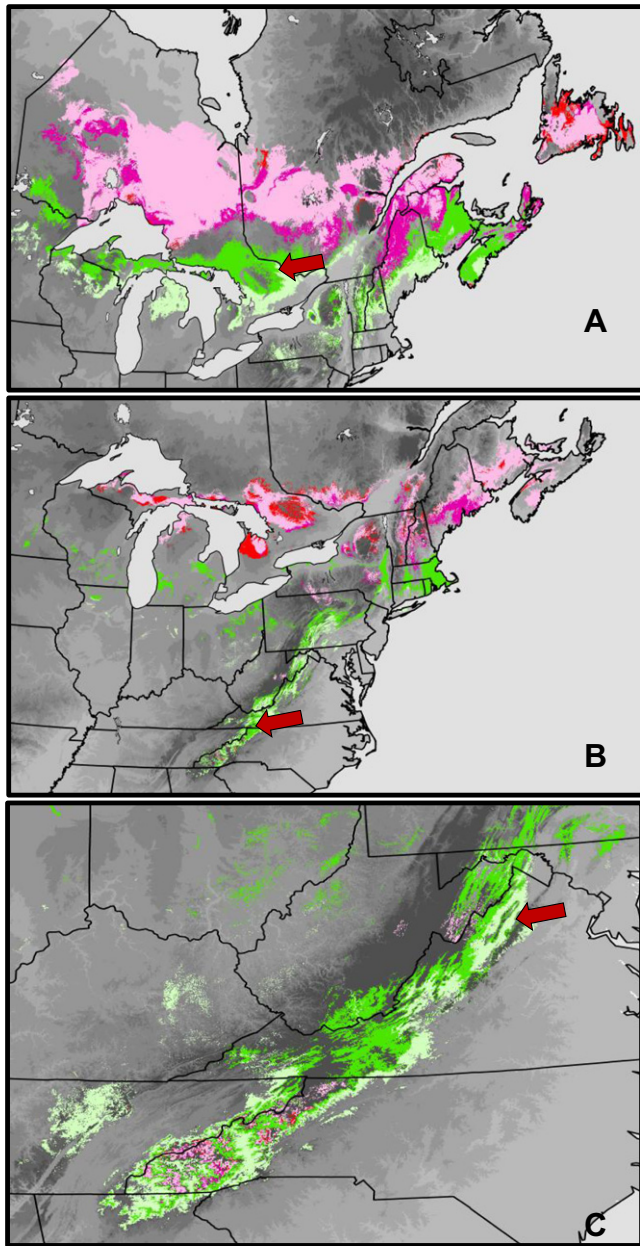


Fig. 11. Examples of contemporary seed transfer zones for three disparate *Pinus strobus* provenances (red arrows), and projected shifts in suitable habitat for the decade centered on 2060, IPCC Scenario A2. Provenance DD5 = 1765, 2587, and 2913 for figures A, B, and C, respectively. Contemporary seed transfer zones are presented as plus (light green) or minus (dark green) one 5-yr height confidence interval (Table 2). Suitable habitats in 2060 are presented as an overlay of the three GCM's (CCCM – pink, GFDL – purple, and UKMO – red).

1978; Rehfeldt, 1991, 1995; St. Clair et al., 2005). The Ontario test series reaffirms that, phenology, growth potential, and cold hardiness are intercorrelated adaptive traits. Variation in growth potential provided the highest resolution of genetic differentiation among populations, with univariate models based on degree-days $>$ above 5 °C accounting for between 36% and 68% of the variance among populations. Mean minimum temperature of the coldest month was a slightly stronger predictor of cold hardiness in the autumn, but added little to the pattern of adaptive variation in Ontario because of the strong correlation between the two predictor variables ($r = 0.80$).

Despite the disparity in study design, test locations, maintenance, and provenance representation among the Ontario, Quebec

and Maryland provenance test series considered in developing the range-wide ecological genetic profile, degree-days $>$ 5 °C was a remarkably consistent predictor of interpopulation variation in growth potential. The DD5 model accounted for 41% of the variation in standardized height growth synthesized from the studies with the three largest provenance samples (Ontario, Quebec and Maryland test series). Nonetheless, average minimum temperature in the coldest month was also a strong predictor.

While degree days $>$ 5 °C was consistently the strongest predictor of growth potential, summer and winter temperature variables are strongly intercorrelated throughout the contemporary range of *P. strobus*. Consequently, growth potential in *P. strobus* is associated with both summer and winter temperatures and the predictive power of the respective climatic variables are a function of their relative spatial heterogeneity and composition of the tests. The superior strength of mean minimum temperature as a predictor of growth potential in the central Appalachian region reaffirms that timing of cold acclimation is a fundamental component of adaptation. Ecologically, the ramifications of planting genotypes from warm climates that have high growth potential in climates with cold winters increases the probability of freezing damage; planting genotypes from cool climates that have low growth potential in warm climates will place them at a competitive disadvantage. Arguing the relative merits of degree-days or minimum temperatures in controlling adaptive clines in *P. strobus* is, therefore, moot. The important point is that adaptation of populations is a function of the temperature regime of the provenance. Both summer warmth and winter cold condition this response.

The bioclimate and ecological genetic profiles of *P. strobus* are powerful tools for considering the future analogs of the spatial distribution of climatic conditions under which contemporary populations will be competitive. The entire southern half of the contemporary spatial and genetic distribution of *P. strobus* is projected to be extirpated by mid-century as habitat deteriorates. Where habitat remains suitable, the cumulative stresses of increasing maladaptation of endemic climatypes elevates both acute and chronic population susceptibility to insect and disease attack, and will likely lead to ineffective seed production. Under these projections, multiple-generation persistence of populations in the contemporary distribution will be the exception. And, projecting the emergence of suitable habitat in the future does not imply occupancy. Indeed, projected geographic shifts for *P. strobus* are at least two orders of magnitude greater than estimates of post-glacial dispersal rates in trees (McLachlan and Clark, 2004; McLachlan et al., 2007). The effectiveness of stochastic colonization is dubious, but extensive population decline and extirpation in the contemporary distribution is inevitable.

The uncertainty is largely centered on the timing of population decline. The temporal limit of species-specific phenotypic plasticity as a buffer against climatic stress is unknown. Ultimately, the transition will be in response to environmental stochasticity (e.g., weather, disease, insect predation, fire). Documented transition in the distribution of trees species in North America since the last glaciation has been associated with wildfire (Pielou, 1992), but insect and disease outbreaks as well as abiotic stresses could have predisposed landscapes to a transition in forest composition. While there is uncertainty in timing of transition, designing and implementing strategies for maintaining forest growth and productivity must be proactive.

Under the paradigm of a rapidly changing climate, decisions regarding optimal seed transfer guidelines must balance risk across the entire rotation. Regardless of the current habitat suitability, there is little point in planting trees in locations that are expected to be outside the bioclimate profile by mid-rotation. Similarly, matching climatypes to contemporary habitat may maximize survival and productivity during the establishment phase of a

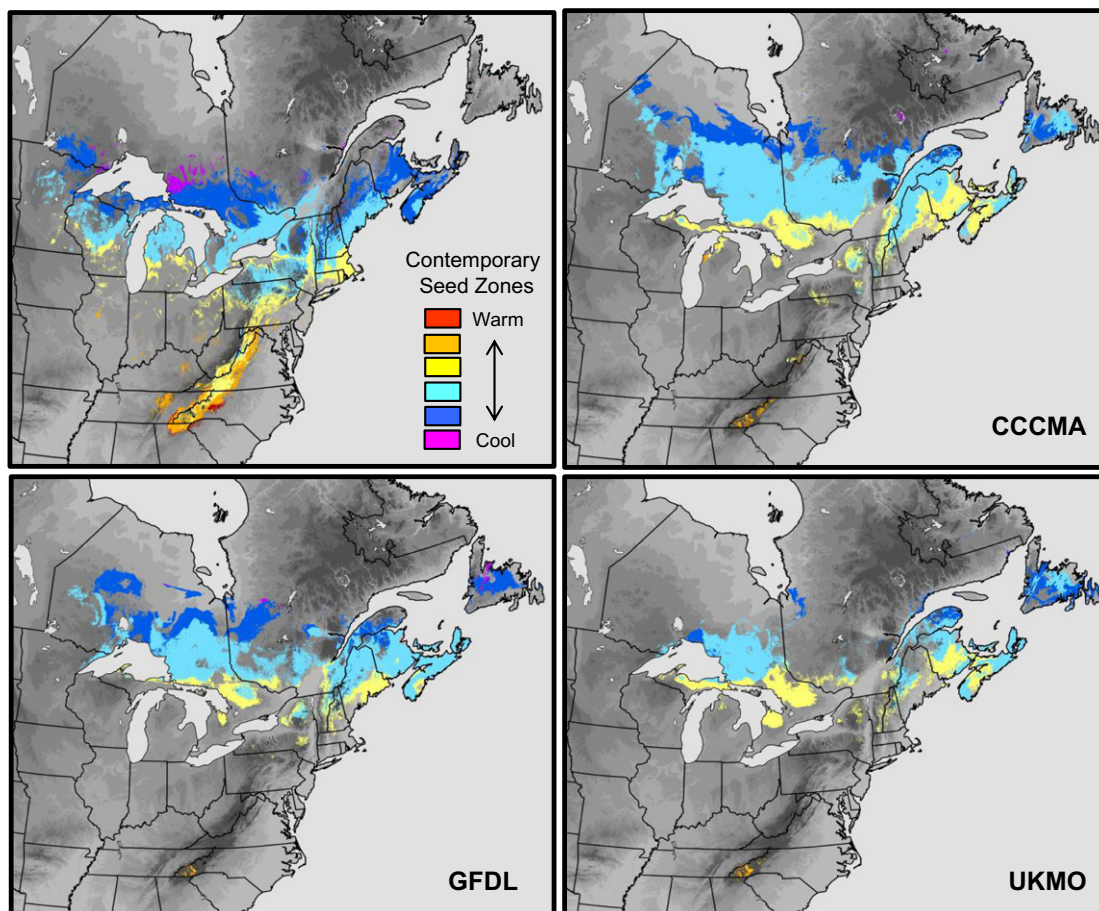


Fig. 12. Seed zones for *Pinus strobus* for the contemporary climate (A), and projected geographic shifts in these zones for the IPCC 2060 A2 scenario and the three general circulation models. (B) Canadian Center for Climate Modeling and Analysis (CCCMA); (C) Geophysical Fluid Dynamics Model (GFDL), and (D) Met Office, Hadley Centre (UKMO).

plantation, but carries substantial risk of plantation decline in mid rotation due to increasing maladaptation. Conversely, proactive transfer into inhospitable contemporary environments that are projected to be suitable habitat in the future is problematic. Over long rotations, there may be no effective management strategy for achieving either conservation or economic objectives. The optimal management strategy will likely include iterative transfers oriented toward accepting some short term risks and shortened rotations to reduce risks of decline and high mortality at later ages. While substantial work is needed to assess these risks, we assume in the section that follows that strategies will evolve toward shorter rotations.

4.2. Applications

The bioclimate and ecological genetic profiles provide the landscape scale context for designing and implementing a reasoned response to managing a forest land base on the verge of transition. At a site level, silvical characteristics (edaphic conditions moisture regime, light requirements) interact with biotic (competition, insects and disease) and abiotic factors (e.g., drought and fire frequency) to determine local abundance. Silvicultural practices are employed to influence the establishment, density, health, and species composition at the site level. Silviculture programs of the twenty first century are faced with the daunting tasks of conserving the existing resource in the short term by mitigating the elevated stress associated with climate warming and focusing reforestation on a redistribution of species and genotypes to the appropriate array of climates. (Rehfeldt et al., 2006).

4.2.1. Forest management

The dramatic and sustained geographic shifts in the climate profile projected over the twenty first century carries with it the expectation that as the quality of the contemporary habitat declines, the existing *P. strobus* resource will experience range-wide decline, dieback and eventually widespread extirpation. Projected changes in the likelihood of occurrence for the contemporary distribution of *P. strobus* provide some insight on the spatial variation in relative risk of the existing eastern white pine resource to forest decline. The mean likelihood of occurrence of the contemporary distribution for the three GCM's for 2030 under IPCC B scenario (Fig. 10A) indicates that the low elevations in the southern Appalachians are at relatively high risk, but there is also a broad east/west band of high risk extending from Massachusetts through Pennsylvania and across the southern limits in the Lake States. As expected the mean likelihood of occurrence for the 2030 IPCC A2 scenario projects more wide-spread high level risk (Fig. 10B). The 2060 IPCC B (Fig. 10C) and A2 (Fig. 10D) scenarios project the longer term progression of relative risk of forest decline. Maps such as these are basic to a climate change response monitoring program, and could be used to prioritize silvicultural efforts such as stand thinning to alleviate stresses associated with climate change.

4.2.2. Forest renewal

Operational planting programs rarely return seed and stock to the donor stand, so artificial reforestation is a program of assisted migration. Seed transfer guidelines and seed zones are typically used to manage the risks of maladaptation. Under the paradigm

of a rapidly changing climate, planting programs are now faced with developing standards that accept elevated risk of maladaptation in the short term in order to optimize adaptation over the life of the stand. A few jurisdictions have begun to modify seed transfer standards to address the expected ecological impacts of climate change (e.g., O'Neil et al., 2008).

In principle, regardless of the geographic distance involved, the best match of individual seed sources to planting site results from transfers along rather than across climatic gradients. For example, under a stable climate paradigm, biologically sound seed transfer guidelines for *P. strobus* can be derived for individual seed sources by defining the area enclosed by one 5-year height confidence interval (Table 2) on either side of the provenance DD5. Fig. 11A presents the contemporary range-wide seed transfer zone for a population originating in southern Ontario (red arrow) as the provenance DD5 (1765) plus (light green) and minus (dark green) one 5-year height confidence interval. Limits of seed transfer in southern Ontario are straight forward. But, long distance transfer to the Minnesota-Ontario border and substantial areas of Maine, New Brunswick and Nova Scotia are predicted to be equally viable.

The projected redistribution of this seed transfer band under IPCC scenario A for the decade centered on 2060 is also presented as an overlay of the projections from the three GCM's (CCCMA – pink, GFDL – purple, and UKMO – red). While the climate of the contemporary distribution is expected to become inhospitable by 2060, substantial new suitable habitat is projected to develop to the north. Fig. 11B is a similar treatment of a population from a Virginia provenance (DD5 = 2587). Biologically sound transfer opportunities include much of the south and central Appalachians as well as some locations in the Lake States and Massachusetts. Projected future suitable habitat for the Virginia population is similar to the contemporary distribution of the Ontario provenance considered in Fig. 11A. That is, as the climate changes, populations such as the Virginia provenance may be suitable for transfer into Ontario to address adaptation concerns. The projections for populations from provenances with DD5 > 2800 are grim (Fig. 11C). For example, a population from northern Virginia (DD5 = 2913) is well suited to contemporary environment in the mid elevations of the south and central Appalachians, but by 2060 the suitable range truncates and shifts to the highest elevations in western North Carolina and a small area in West Virginia. By 2090 (not shown), suitable habitat is projected to all but disappear.

Actively managing the redeployment of multiple localized populations when transfer guidelines will require repeated review and revision as the climate continues to change is operationally impractical on a broad scale. The alternative is to subdivide the observed clinal variation into geographically discrete seed zones based on discrete climatypes and manage redeployment on the basis of seed zone rather than individual local populations. Climatype classes with non-overlapping confidence intervals would be representative of discrete climatypes. On that basis, the ecological genetic profile of *P. strobus* (Fig. 8) can be converted into six discrete seed zones with the reasonable expectation that populations within one classification will differ adaptively on average from populations in another zone (Fig. 12A). The warmest zone (red) is peripheral and is restricted to the lowest slopes of the southern Appalachians. The orange zone is restricted to the south and central Appalachian Mountains. The yellow zone represents the highest elevations in the Appalachians and a latitudinal band from the southern limits in the Lake States, to Massachusetts. The light blue zone straddles the international boundary and the dark blue zone dominates the Canadian distribution. The coolest zone (purple) is limited in spatial extent with only minor areas beyond the eastern shore of Lake Superior.

The shifts in seed zones projected by the three GCM's, within the projected *P. strobus* climate profile under the 2060 IPCC A2

scenario are presented in Fig. 12B–D. Across the three GCM's the projected shifts in seed zones are consistently projected to be displaced from their contemporary distribution. The dark blue zone appears to be best suited for the northern fringe of the advancing front, while the yellow zone is expected to be best suited for the trailing edge of the future distribution. The light blue zone dominates the emergent suitable habitat. The orange and red zones shift to a small area in western North Carolina. Clearly, strict adherence to seed zones based on contemporary climate are almost assured of having disastrous consequences as the climate continues to warm. By implication, management concerns for conserving *P. strobus* are unique to each discrete climatype associated with the seed zones, but of the particular concern are the impacts in the southern Appalachian region. The projections consistently indicate that the northern half of the ecological genetic profile should have new suitable habitat develop to the north of the current distribution. In contrast, suitable climatic conditions for the central and southern Appalachians climatypes are projected to erode and shift toward the highest elevations in western North Carolina. By 2090, suitable climatic conditions in the south are all but eliminated from western North Carolina (not shown).

5. Conclusions

The integration of the bioclimate and ecological genetic profiles produced substantial refinement in the autecology of *P. strobus*. The contemporary distribution of *P. strobus* was most closely associated with mean maximum temperature, but five of the eight variables in the bioclimate profile represent interactions between precipitation and temperature. On a range-wide scale, adaptive variation was mostly closely associated with summer temperatures, but winter temperatures were the strongest predictors of genetic variation in the central Appalachian Mountains. Mapped projections of these models refine Little's (1971) range map, and provide the first illustration of the ecological genetic profile in *P. strobus*.

Mapped projections of the three GCMs and two IPCC scenarios demonstrate a high degree of concurrence regarding the redistribution of both suitable habitat and climatypes through the decade centered on 2060. While disparity among projections increases for the decade centered on 2090, all projections indicate a severe reduction in suitable habitat within the contemporary distribution. The consistency of the projected redistribution for both the climate and ecological genetic profiles places *P. strobus* on a growing list of temperate and boreal tree species that appear to be poorly equipped to persist in their contemporary distribution as the climate warms. The magnitude and speed of climate change is projected to lead to widespread range recession; preclude requisite adaptation of populations in situ; and exceed the capacity of marginal populations to colonize increasingly disjunct emergent suitable habitat. It seems clear that maintaining optimal growth and productivity of *P. strobus* as the climate warms will require substantial artificial reforestation efforts focussed on matching climatypes to suitable habitat.

Mapped projections provide a framework for developing management strategies for both conservation and economic goals. The identification of trailing edge of the contemporary distribution and its expansion over time is the prerequisite for targeting mitigation efforts to conserve the existing resource. The congruence of the projected redistribution of suitable habitat provide the foundation for identification of the land base with the highest probability of being suitable for *P. strobus* planting programs. Areas likely to be suitable as refugia, such as the area in western North Carolina are also evident. The ecological genetic profile of *P. strobus* provides the cornerstone for (a) identifying priority populations for

inclusion in ex situ conservation programs such as seed banks; (b) guiding seed transfer for conservation and operational reforestation; (c) identifying the geographic context for selecting genotypes for inclusion in operational seed orchards; and (d) fostering trans-jurisdictional synergy in seed procurement and deployment. Within this framework, adherence to silvicultural principles oriented toward matching species to site and managing stochastic environmental effects to favor *P. strobus* will be essential to operational success at the site level.

When considering the artificial regeneration of *P. strobus* in the coming decades, the limitations of supportive models should be kept in mind. Practical application of these models should be based on the paradigm that they provide a foundation for formulating alternative testable hypotheses. Management actions would then be treated as formal scientific investigations focussed on testing and refining models. Monitoring of both the actual climate change and biological responses are inherent components of this iterative adaptive management paradigm.

Acknowledgments

The provinces of Nova Scotia, Ontario, Quebec, and the Forest Inventory and Analysis, US Forest Service, graciously granted us access to their data files. Jean Beaulieu, Canadian Forest Service provided access to unpublished provenance test data from the Quebec test series. We appreciate the technical support of Nicholas Crookston, Rocky Mountain Research Station, Moscow, Idaho. Robert Sinclair, Don Higgs, and Terri Winters provided technical support during the course of data collection from the Ontario test series.

References

- Abubaker, H.I., Zsuffa, L., 1990. Provenance variation in eastern white pine (*Pinus strobus* L.): 28th-year results from two southern Ontario plantations. In: Garrett, P.W. (Ed.), Proceedings of a Symposium on White Pine Provenance and Breeding. USDA, For. Serv., Northeast. Exp. Sta., Gen Tech, Report, NE-155, pp. 69–76.
- Alerich, C.A., Klevgard, L., Liff, C., Miles, P.D., 2004. The Forest Inventory and Analysis. Database: Database Description and Users Guide, Version 1.7. <http://www.ncrs2.fs.fed.us/4800/fiadb_documentation/FIADB_v17_1221104.pdf>.
- Allen, C.D. et al., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manage.* 259, 660–684.
- Barns, B.V., 2009. Tree response to ecosystem change at the landscape level in eastern North America. *Forstarchiv* 80 (3), 76–89.
- Beaulieu, J., Plourde, A., Daoust, G., Lamontagne, L., 1996. Genetic variation in juvenile growth of *Pinus strobus* in replicated Quebec provenance–progeny tests. *For. Genet.* 3 (2), 103–112.
- Bechtold, W.A., Patterson, P.L., 2005. The Enhanced Forest Inventory and Analysis. Program—National Sampling Design and Estimation Procedures. USDA, For. Serv., South. Res. Sta., Gen Tech Report SRS-80, Asheville, NC, 85 pp.
- Bower, A.B., Aiken, S.A., 2008. Ecological genetics and seed transfer guidelines for *Pinus albicaulis* (Pinaceae). *Am. J. Bot.* 95 (1), 66–76.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Breshears, D.D. et al., 2005. Regional vegetation die-off in response to global-change-type drought. *Proc. Natl. Acad. Sci., USA* 102, 15144–15148.
- Brown, D.E., Reichenbacher, F., Franson, S.E., 1998. A Classification of North American Biotic Communities. University of Utah Press, Salt Lake City.
- Burger, R., Lynch, M., 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution* 49, 151–163.
- Campbell, R.K., Sorensen, F.C., 1978. Effect of test environment on the expression of clines and on delimitation of seed zones in Douglas fir. *Theor. Appl. Genet.* 51, 233–246.
- Clair, J.B., Mandel, N.L., Vance-Boland, K.W., 2005. Geneecology of Douglas-fir in western Oregon and Washington. *Ann. Bot.* 96, 1199–1214.
- Demeritt, M.E., Jr., Garrett, P.W., 1996. Adaptation of Eastern White Pine Provenances to Planting Sites USDA, For. Serv., Northeast. For. Exp. Sta., Res. Pap. NE-703, 7pp.
- Demeritt, M.E., Kettlewood, H.C., 1976. Eastern white pine seed source variation in the northeastern United States: 16-year results. In: Proc. Twelfth Lakes States Forest Tree Improv. Conf. St. Paul MN August, 1975. USDA, For. Serv., North Central For. Exp. Sta., Gen Tech, Report NC-26, pp. 80–87.
- Draper, N.R., Smith, H., 1981. Applied Regression Analysis, second ed. John Wiley and Sons, Inc., New York, NY.
- Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* 40, 677–697. <http://dx.doi.org/10.1146/annurev.ecolsys.110308.120159>.
- Fowler, D.P., Heimbuger, C., 1969. Geographic variation in eastern white pine, 7-year results in Ontario. *Silv. Genet.* 18, 123–129.
- Funk, D.T., Allen, R., Williams, R.D., 1975. Fifteen-year performance of eastern white pine seed source tests in the lower Ohio valley. In: Proceedings of the Central States Forest Tree Improvement Conference, vol. 9, pp. 153–158.
- Funk, D.T., 1971. Eastern White Pine Seed Source Trials: Ten-year Results from Three Midwestern Plantations. USDA, For. Serv., North Central For. Exp. Sta., Res. Note NC-113, 4p.
- Garrett, P.W., Schreiner, E.J., Kettlewood, H., 1973. Geographic Variation of Eastern White Pine in the Northeast. USDA, For. Serv., Northeast. For. Exp. Sta., Res. Pap. NE-274, 14p.
- Genys, J.B., 1968. Geographic variation in eastern white pine. Two-year results of testing range-wide collections in Maryland. *Silv. Genet.* 17 (1), 6–12.
- Genys, J.B., 1987. Provenance variation among different populations of *Pinus strobus* from Canada and the United States. *Can. J. For. Res.* 17, 228–235.
- Gitlin, A.R., Stultz, C.M., Bowker, M.A., Stumpf, S., Paxton, K.L., Kennedy, K., Munoz, A., Bailey, J.K., Whitham, T.G., 2006. Mortality gradients within and among dominant plant populations as barometers of ecosystem change during extreme drought. *Conserv. Biol.* 20, 1477–1486.
- GLOBE Task Team, 1999. The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, Boulder.
- Hansen, J., 2009. Storms of My Grandchildren: The Truth About the Coming Climate Catastrophe and Our Last Chance to Save Humanity. Bloomsbury USA, New York.
- Hijmans, R.J., Elith, J., 2011. Species Distribution Modeling with R. <cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf>.
- Hutchinson, G.E., 1957. Concluding remarks. *Cold Spring Harbor Symp. Quant. Biol.* 22, 415–457.
- Hutchinson, M.F., 2000. ANUSPLIN Version 4.1 User's Guide. Australian National University, Centre for Resource and Environmental Studies, Canberra.
- IPCC, 2007. Summary for policymakers. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1–18.
- Iverson, L.R., Prasad, A.M., 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecol. Monogr.* 68, 465–485.
- Iverson, L.R., Prasad, A.M., Mathews, S., 2008. Modelling potential climate change impacts on the trees of northeastern United States. *Mitig. Adapt. Strat. Glob. Change* 13, 487–516.
- Joyce, D.G., Lu, P., Sinclair, B., 2002a. Genetic variation in height growth among populations of eastern white pine (*Pinus strobus* L.) in Ontario. *Silv. Genet.* 51 (4), 136–142.
- Joyce, D.G., Ford, R., Fu, Y.B., 2002b. Spatial patterns of tree height variations in a black spruce a farm-field progeny test and neighbors-adjusted estimations of genetic parameters. *Silv. Genet.* 51, 13–18.
- Jump, A.S., Peñuelas, J., 2006. Rapid climate change-related growth decline at the southern edge of *Fagus sylvatica*. *Glob. Chang. Biol.* 12, 163–174.
- Keenan, T., Serra, J.M., Lloret, F., Ninyerola, M., Sabate, S., 2011. Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO₂ matters! *Glob. Change Biol.* 17, 565–579.
- King, J.P., Nienstadt, H., 1968. Variation in eastern white pine seed sources planted in the Lake States. *Silv. Genet.* 18, 83–86.
- Kirschbaum, M., Fischlin, A., 1996. Climate change impacts on forests. In: Watson, R., Zinyowera, M.C., Moss, R.H. (Eds.), *Climate Change 1995 – Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analysis. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)*. Cambridge University Press, Cambridge, pp. 95–129.
- Lavergne, S., Mouquet, N., Thuiller, W., Ronce, O., 2010. Biodiversity and climate change: evolutionary and ecological responses of species and communities. *Annu. Rev. Ecol. Evol. Syst.* 41, 321–350.
- Ledig, F.T., Rehfeldt, G.E., Sáenz-Romero, C., Flores-López, C., 2010. Projections of suitable habitat for rare species under global warming scenarios. *Am. J. Bot.* 97 (6), 970–987.
- Leites, L.P., Robinson, A.P., Rehfeldt, G.E., Marshall, J.D., 2012. Height-growth response to climate changes differs among populations of Douglas-fir: a novel analysis of historic data. *Ecol. Appl.* 22 (1), 199–206.
- Li, P., Beaulieu, J., Daoust, G., Plourde, A., 1997. Patterns of adaptive genetic variation in eastern white pine (*Pinus strobus*) from Quebec. *Can. J. For. Res.* 27, 199–206.
- Liaw, A., Wiener, M., 2002. Classification and regression by Random Forests. *R News* 2, 18–22.
- Little, E.L., Jr., 1971. Atlas of United States Trees. vol. 1. Conifers and Important Hardwoods. USDA, Misc Publ. 1146, Washington, D.C.
- Lu, P., Joyce, D.G., Sinclair, R.W., 2002. Geographic variation in cold hardiness among eastern white pine (*Pinus strobus* L.) provenances in Ontario. *For. Ecol. Manage.* 178, 329–340.
- Lu, P., Joyce, D.G., Sinclair, R.W., 2003. Effect of selection on shoot elongation rhythm of eastern white pine (*Pinus strobus* L.) and its implications to seed transfer in Ontario. *For. Ecol. Manage.* 182, 161–173.

- Lynch, M., Lande, R., 1993. Evolution and extinction in response to environmental change. In: Kareiva, P.M., Kingsolver, J.G., Huey R.B. (Eds.), *Biotic Interactions and Global Change*. Sinauer, Sunderland, MA, pp. 234–250.
- Maronek, D.M., Flint, H.L., 1974. Cold hardiness of needles of *Pinus strobus* L. as a function of geographic source. *For. Sci.* 20, 135–141.
- Masek, J.G., 2001. Stability of boreal forest stands during recent climate change: evidence from Landsat satellite imagery. *J. Biogeogr.* 28, 967–976.
- Matyas, C., Yeatman, C.W., 1992. Effect of geographic transfer on growth and survival of jack pine (*Pinus banksiana* Lamb.) populations. *Silv. Genet.* 41 (6), 370–376.
- Matyas, C., 2006. Migratory, genetic and phenetic response potential of forest tree populations facing climate change. *Acta Silvatica et Lignaria Hungarica* 2, 33–46.
- McLachlan, J.S., Clark, J.S., 2004. Reconstructing historical ranges with fossil data at continental scales. *For. Ecol. Manage.* 197, 139–147.
- McLachlan, J.S., Hellmann, J.J., Schwartz, M.W., 2007. A framework for debate of assisted migration in an era of climate change. *Conserv. Biol.* 21, 297–302.
- Mergen, F., 1963. Ecotypic variation in *Pinus strobus* L. *Ecology* 44, 716–726.
- Michaelian, M., Hogg, E.H., Hall, R., Arsenault, E., 2011. Massive mortality of aspen following severe drought along the southern edge of the Canadian boreal forest. *Glob. Change Biol.* 17, 2084–2094.
- Morgenstern, E.K., 1996. *Geographic Variation in Forest Trees: Genetic Basis and Application of Knowledge in Silviculture*. UBC Press, Vancouver.
- O'Neill, G.A., Ukrainetz, N.K., Carlson, M.R., Cartwright, C.V., Jaquish, B.C., King, J.N., Krakowski, J., Russell, J.H., Stoehr, M.U., Xie, C.-Y., Yanchuk, A.D., 2008. Assisted Migration to Address Climate Change in British Columbia: Recommendations for Interim Seed Transfer Standards. B.C. Min. For. Range, Res. Br., Victoria, B.C. Tech. Rep. 048.
- Parmesan, C., 2006. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* 37, 637–669.
- Pearson, R.G., Dawson, T.P., 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* 12, 361–371.
- Pearson, R.G., 2007. *Species' Distribution Modeling for Conservation Educators and Practitioners*. Synthesis. American Museum of Natural History. <<http://ncep.amnh.org>>.
- Pielou, E.C., 1992. *After the Ice: The Return of Life to Glaciated North America*. University of Chicago Press, Chicago.
- Pulliam, H.R., 2000. On the relationship between niche and distribution. *Ecol. Lett.* 3, 349–360.
- R Development Core Team, 2004. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. <<http://www.r-project.org/>>.
- Rehfeldt, G.E., Jaquish, B.C., 2010. Ecological impacts and management strategies for western larch in the face of climate-change. *Mitig. Adapt. Strat. Glob. Change* 15, 283–306.
- Rehfeldt, G.E., Ying, C.C., Spittlehouse, D.L., Hamilton, D.A., 1999. Genetic responses to climate in *Pinus contorta*: Niche breadth, climate change, and reforestation. *Ecol. Mongr.* 69 (3), 375–407.
- Rehfeldt, G.E., Wykoff, W.R., Ying, C.C., 2001. Physiologic plasticity, evolution, and impacts of a changing climate on *Pinus contorta*. *Clim. Change* 50, 355–376.
- Rehfeldt, G.E., Tchebakova, N.M., Parfenova, E.I., 2004. Genetic responses to climate and climate-change in conifers of the temperate and boreal forests. *Recent Res. Dev. Genet. Breeding* 1, 113–130.
- Rehfeldt, G.E., Crookston, N.L., Warwell, M.V., Evans, J.S., 2006. Empirical analyses of plant–climate relationships for the western United States. *Int. J. Plant Sci.* 167, 1123–1150.
- Rehfeldt, G.E., Ferguson, D.E., Crookston, N.L., 2009. Aspen, climate, and sudden decline in western USA. *For. Ecol. Manage.* 258, 2353–2364.
- Rehfeldt, G.E., Crookston, N.L., Sáenz-Romero, C., Campbell, E.M., 2012. North American vegetation model for land-use planning in a changing climate: a solution to large scale classification problems. *Ecol. Appl.* 22 (1), 119–141.
- Rehfeldt, G.E., 1988. Ecological genetics of *Pinus contorta* from the Rocky Mountains (USA): a synthesis. *Silv. Genet.* 37, 131–135.
- Rehfeldt, G.E., 1990. Genetic differentiation among populations of *Pinus ponderosa* from the upper Colorado River Basin. *Bot. Gaz.* 151, 125–137.
- Rehfeldt, G.E., 1991. A model of genetic variation for *Pinus ponderosa* in the Inland Northwest (USA): applications in gene resource management. *Can. J. For. Res.* 21, 1491–1500.
- Rehfeldt, G.E., 1995. Genetic-variation, climate models and the ecological genetics of *Larix occidentalis*. *For. Ecol. Manage.* 78, 21–37.
- Rehfeldt, G.E., 2004. Interspecific and Intraspecific Variation in *Picea engelmannii* and Its Congeneric Cohorts: Biosystematics, Geneecology, and Climate-Change. USDA, For. Serv., Rocky Mount. Exp. Sta., Gen Tech Report RMRS-134, Fort Collins, CO, 18pp.
- Rehfeldt, G.E., 2006. A Spline Climate Model for the Western United States. USDA, For. Serv., Rocky Mount. Exp. Sta., Gen Tech Report 165, Fort Collins, CO, 21pp.
- Richardson, B.A., Rehfeldt, G.E., Kim, M.S., 2009. Congruent climate-related geneecological responses from molecular markers and quantitative traits for western white pine (*Pinus monticola*). *Int. J. Plant Sci.* 170 (9), 1120–1131.
- Saenz-Romero, C., Rehfeldt, G.E., Crookston, N.L., Duval, P., Beaulieu, J., 2010. Contemporary and projected spline climate surfaces for Mexico and their use in understanding climate-plant relationships. *Clim. Change* 102, 595–623.
- Santamour, F.S., Jr., 1960. Seasonal Growth of White Pine Seedlings from Different Provenances. USDA, For. Serv., Northeast. For. Exp. Sta., For Res Note 105, 4pp.
- Shaw, J.D., Steed, B.E., DeBlander, L.T., 2005. Forest Inventory and Analysis (FIA) annual inventory answers the question: what is happening to pinyon-juniper woodlands? *J. For.* 103, 280–285.
- Sluder, E.R., Dorman, K.W., 1971. Performance in the Southern Appalachians of Eastern White Pine Seedlings from Different Provenances. USDA, For. Serv., Southeast. For. Exp. Sta., Res. Pap. SE-90, 11pp.
- Sluder, E.R., 1963. A White Pine Provenance Study in the Southern Appalachians. USDA, For. Serv., Southeast. For. Exp. Sta., Res. Pap. SE-2, 16p.
- Sprackling, J.A., Read, R.A., 1976. Eastern White Pine in Eastern Nebraska: A Provenance Study of Southern Appalachian Origins. USDA, For. Serv., Rocky Mount. Exp. Sta., Res. Pap. RM-179, 8pp.
- Tchebakova, N.M., Parfenova, E.I., Soja, A.J., 2010. Potential climate-induced vegetation change in Siberia in the 21st century. In: Baltzer, H. (Ed.), *Environmental Change in Siberia*. Springer, Dordrecht, The Netherlands, pp. 67–82.
- Thomas, C.D., 2004. Extinction risk from climate change. *Nature* 427, 145–148.
- Thuiller, W., 2003. BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* 10, 2020–2027.
- Turesson, G., 1925. The plant species in relation to habitat and climate. *Hereditas* 6, 147–236.
- USGS, 2005. Digital Representations of Tree Species Range Maps, from Little, E.L., Jr., 1971. *Atlas of United States Trees*, vol. 1. Conifers and Important Hardwoods. USDA, Misc. Publ. 1146, Washington, DC. <<http://climatechange.cr.usgs.gov/data/atlas/little>>.
- Wendel, G.W., Smith, H.C., 1990. *Pinus strobus* L: eastern white pine. In: Burns, R.M., Honkala, B.H. (Tech. Coords.), *Silvics of North America*. vol. 1. Conifers. USDA, Washington, DC, Agric Handbook. No. 654, pp. 476–488.
- Williamson, T., Colombo, S., Duinker, P., Gray, P., Hennessey, R., Houle, D., Johnston, M., Ogden, A., Spittlehouse, D., 2009. Climate change and Canada's Forests from Impacts to Adaptation. Co-published by: Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre. <<http://sfmnetwork.ca/>> or the Canadian Forest Service Bookstore <<http://bookstore.cfs.nrcan.gc.ca/>>.
- Woodall, C.W., Oswalt, C.M., Westfall, J.A., Perry, C.H., Nelson, M.D., Finley, A.O., 2009. An indicator of tree migration in forests of the eastern United States. *For. Ecol. Manage.* 257, 1434–1444.
- Worrall, J.J., Marchetti, S.B., Egeland, L., Mask, R.A., Eager, T., Howell, B., 2010. Effects and etiology of sudden aspen decline in southwestern Colorado, USA. *For. Ecol. Manage.* 260, 638–648.
- Wright, J.W., Lemmings, W.L., Bright, J., 1963. Geographic variation in eastern white pine: 6-year results. *Q. Bull. Mich. Agric. Exp. Sta.* 45, 691–697.
- Wright, J.W., Amiel, R.J., Cech, F.C., Kriebel, H.B., Jokela, J., Lemmings, W.A., Matheson, A.C., Merritt, C., Read, R.A., Roth, P., Thor, E., Thulin, I.J., 1979. Performance of eastern white pine from the southern Appalachians in eastern United States, New Zealand, and Australia. In: Proc. Twenty-sixth Northeastern Forest Tree Improvement Conference, pp. 203–217.
- Wright, J.W., 1970. Ten years of provenance research at Michigan State University and the next steps in tree improvement. In: Proceedings of the Ninth Lake States Forest Tree Improvement Conference, St. Paul, MN, USDA, For. Serv., North Central For. Exp. Sta., Res. Pap. NC-47, pp. 8–11.
- Zhu, K., Woodall, C., Clack, J., 2011. Failure to migrate: lack of tree range expansion in response to climate change. *Glob. Change Biol.* <http://dx.doi.org/10.1111/j.1365-2486.2011.02571.x>, 11 pp.