Should Species Distribution Models Account for Spatial Autocorrelation? A Test of Model Projections Across Eight Millennia of Climate Change

Citation

Published Version

Permanent link

Terms of Use

Share Your Story

1   Running title: Projecting spatial species distribution models

2

3

4   Should species distribution models account for spatial autocorrelation? A test of model

5   projections across eight millennia of climate change

6

7

8   Sydne Record[1], Matthew C. Fitzpatrick[2], Andrew O. Finley[3], Sam Veloz[4], and Aaron M. Ellison[1]

9

10   [1]Harvard Forest, Harvard University, Petersham, MA

11   [2]Appalachian Lab, University of Maryland Center for Environmental Science, Frostburg, MD

12   [3]Departments of Forestry and Geography, Michigan State University, East Lansing, MI

13   [4]PRBO Conservation Science, Petaluma, CA

14

15   Key words: Bayesian, historical validation, spatial random effect, paleoecology

16

17   *Corresponding author's email: srecord@fas.harvard.edu

18   Article type: Research Paper

19 **ABSTRACT**

20 **Aim** The distributions of many organisms are spatially autocorrelated, but it is unclear whether

21 including spatial terms in species distribution models (SDMs) improves projections of species

22 distributions under climate change. We provide one of the first comparative evaluations of the

23 ability of a purely spatial SDM, a purely non-spatial SDM, and a SDM that combines spatial and

24 environmental information to project species distributions across eight millennia of climate

25 change.

26 **Location** Eastern North America.

27 **Methods** To distinguish between the importance of climatic versus spatial explanatory variables,

28 we fit three Bayesian SDMs to modern occurrence data for *Fagus* and *Tsuga*, two tree genera

29 whose distributions can be reliably inferred from fossil pollen: a spatially-varying intercept

30 model, a non-spatial model with climatic variables, and a spatially varying intercept plus climate

31 model. Using high temporal resolution paleoclimate data, we hindcasted the SDMs in 1,000 year

32 time steps for 8000 years, and compared model projections with palynological data for the same

33 periods.

34 **Results** For both genera, spatial SDMs provided better fits to the calibration data, more accurate

35 predictions of a hold-out validation dataset of modern trees, and higher variance in current

36 predictions and hindcasted projections than non-spatial SDMs. Performance of non-spatial and

37 spatial SDMs according to the Area Under the Curve of the Receive Operating Curve varied by

38 genus. For both genera, false negative rates between non-spatial and spatial models were similar,

39 but spatial models had lower false positive rates than non-spatial models.

40   **Main conclusions** The inclusion of computationally demanding spatial random effects in SDMs

41   may be warranted when ecological or evolutionary processes prevent taxa from shifting their

42   distributions or when the cost of false positives is high.

INTRODUCTION

The last decade has witnessed a marked increase in the application of models that project the potential geographic distributions of species by linking observations of species occurrences to environmental predictor variables. These models, commonly called bioclimatic envelope, ecological niche, or species distribution models (hereafter SDMs), are important tools for forecasting impacts of climatic change on biological diversity and for generating conservation plans and climate-change policy (Guisan & Thuiller, 2005). To project future distributions under different, plausible scenarios of climatic change, SDMs use statistical relationships between present-day distributions of species and climate (Elith *et al.*, 2010). Although generally successful at explaining and predicting current distributions of species (Franklin & Miller, 2009), impact assessments derived from SDMs have been criticized for their reliance on a number of largely untested ecological assumptions, methodological issues, and statistical concerns (e.g.*,* Pearson & Dawson, 2003; Dormann, 2007).

Chief among these issues is the failure of most SDMs to account for spatial dependence of occurrence data (Gelfand *et al.*, 2006; Bahn and McGill, 2007; Dormann, 2007; Elith *et al.*, 2010). Spatial autocorrelation arises in ecological data because nearby points tend to be more similar, in physical characteristics and/or species occurrences or abundances, than are pairs of locations that are farther apart (Legendre, 1993). When model assumptions about independent and identically distributed residuals are violated, there could be a bias in the regression parameter estimates, potentially leading to poor inference. Studies illustrate that failure to account for spatial autocorrelation can lead to misidentification of important driving variables and overly optimistic error rates (e.g.*,* Lichstein, *et al.*, 2002; Segurado *et al.*, 2006; Diez & Pulliam, 2007; Dormann, 2007), especially when small-scale patterns of explanatory variables

66 create instability in broad-scale regression parameter estimates (Hawkins *et al.*, 2007). Further,

67 models based solely on spatial interpolation can provide better fits to species range data than

68 models based on explanatory environmental variables (Bahn & McGill, 2007), suggesting that

69 spatial autocorrelation in unmeasured factors (e.g., population processes such as dispersal or

70 underlying resources such as soil moisture) may account for most of the observed distributional

71 patterns.

72   Analysis of spatial SDMs primarily has focused on predicting current or simulated

73 species' distributions using a hold-out dataset for model validation (Gelfand *et al.*, 2006; Wilson

74 *et al.*, 2010), but projections of spatial SDMs in changing climates over long time scales remain

75 largely untested. Observed changes in species distributions as a result of past climatic dynamics

76 provide a unique opportunity to compare projections of spatial and non-spatial SDMs

77 parameterized with current conditions (Pearman, *et al.*, 2008a; Nogués-Bravo, 2009; Dobrowski

78 *et al.*, 2011, Veloz *et al.*, 2012).

79   Projections to environmental conditions different from those used to calibrate SDMs are

80 subject to error (Heikkinen *et al.*, 2006) and may not be ecologically meaningful or statistically

81 valid if there are changes in correlations between variables across time and space (Elith *et al.*,

82 2010) or if species-environment relationships are not conserved (e.g., Fitzpatrick et. al., 2007,

83 Veloz *et al.*, 2012). It also is not known whether it is desirable to project models with spatial

84 random effects based on the partially observed spatial distribution of a species at one time point

85 into a new temporal domain.

86   In this study, we developed non-spatial and spatial SDMs for two genera of trees in

87 eastern North America. We calibrated the models with current climate data and Forest Inventory

88 and Analysis (FIA) data collected by the United States Forest Service. We then projected the

89   models back in time using paleoclimate simulations and extensive pollen records as independent

90   validation data. Our approach is similar to that of Pearman *et al.* (2008a), who used fossil pollen

91   to validate SDMs of European trees projected back to a single time in the mid-Holocene (6,000

92   years before present). However, the availability of new paleoclimate reconstructions, which

93   provide millennial snapshots of historic climate for the last 21,000 years before present, allowed

94   us to validate the models at a much finer temporal resolution.

95   To assess the usefulness of adding a spatial term to SDMs we consider the following: 1) a

96   spatially-varying intercept model with no climate variables; 2) a non-spatial model with climate

97   variables; and 3) a spatially-varying intercept model with climate variables. As detailed in the

98   Methods Section and Appendix S3, the spatially-varying intercept was introduced via spatial

99   random effects. The rationale for choosing these candidate models is a follows. If climate

100 variables describe a significant portion of the variability in the observed distribution and if these

101 variables change over time, then projections from models with climatic variables will show a

102 conservative shift away from the observed distribution. For the spatially varying intercept model

103 with climate variables, any projected shifts in distributions are tempered by the spatial random

104 effects. Depending on the amount of spatial autocorrelation, spatial random effects act to draw

105 the projected distribution back toward the observed distribution used to calibrate the model. If

106 climate variables do not describe a significant portion of the variability in the observed

107 distribution, then the spatial random effects will keep projected distributions close to the

108 observed distribution, i.e., the only learning for prediction will come from the observed

109 distribution and hence projected probability of species occurrence will be similar to the observed

110 probability of occurrence. With these three candidate models, we were able to tease apart

111 differences due to the spatial random effects alone, the climate variables alone, and their additive

112 effects. We parameterized and estimated model parameters following a Bayesian framework,

113 which provided full posterior distributions for model parameters and allowed us to estimate the

114 uncertainty in our statistical inferences. We focus on two tree genera, *Fagus* and *Tsuga*, whose

115 distributions can be readily inferred from fossil pollen and which possess contrasting life

116 histories.

117   We address three questions. (1) Do non-spatial SDMs of current distributions of *Fagus*

118 and *Tsuga* based on climate variables exhibit residual spatial autocorrelation? (2) Do SDMs with

119 spatial random effects that include or exclude climate variables provide better fits to the observed

120 distributions than non-spatial SDMs with climate variables only? (3) Do hindcasted spatial

121 SDMs better predict historic distributions than non-spatial SDMs?

122 **Methods**

123 ***Study genera***

124 We studied two tree genera, *Fagus* and *Tsuga*. In eastern North America, *Fagus* is represented by

125 only one species, *F. grandifolia* (Ehrh.) (American Beech), and *Tsuga* by only two, the

126 widespread *T. canadensis* (L.) Carr. (Eastern Hemlock), and the narrow endemic *T. caroliniana*

127 Engelm.) (Carolina Hemlock). For both *Fagus* and *Tsuga,* the relationship between local

128 abundance of trees and relative abundance of pollen in sediment cores has already been derived

129 (Davis, 1981). *Tsuga* is a conifer with passively-dispersed cones, whereas *Fagus* is deciduous

130 with animal-dispersed seeds.

131 ***Occurrence data***

132 We used FIA data to describe the current distribution of *Fagus* and *Tsuga*. In every 2428 ha of

133 land in the United States classified "forested", there is one permanent FIA plot, each containing

134 four 7.2 m fixed-radius subplots (Woudenberg *et al.*, 2010). In each subplot, all trees >12.7 cm

135 diameter at breast height have been measured periodically since the 1940s; consistent nationwide

136 annual inventories were initiated in 2001. We used data from the most recent full plot inventory

137 (2003 – 2008) to calibrate our models.

138   Historic distributions of *Fagus* and *Tsuga* were derived from fossil pollen data in the

139 Neotoma Paleoecology Database (<www.neotomadb.org>). Paleoclimate data (described below)

140 were available at 1 kiloannum before present (kaBP) intervals from 0–21 kaBP, so we focused on

141 millennial historic distributions of *Fagus* and *Tsuga*. Given the variation in temporal scale and

142 spatial resolution across study sites and uncertainties associated with radiocarbon aging of pollen

143 from sediment cores (Blauw *et al.*, 2007), we compiled pollen datasets in which *Fagus* and

144 *Tsuga* were counted as present at a site if their pollen percentages reached threshold levels at any

145 time within 500 years centered on each historic millennium (Appendix S1). We chose a 500 year

146 window because cross-validation analyses of biostratigraphic ages from recently revised age

147 models for all pollen sites suggested that 500 years is a conservative estimate of temporal

148 uncertainty for sites in the Neotoma database (Blois *et al.*, 2011). To determine the sensitivity of

149 historic tree distributions to the pollen percentage thresholds used to define a genera's presence

150 or absence at a site, we specified low and high thresholds for each genus (Pearman *et al.*, 2008a):

151 0.5% or 1% for *Fagus* and 1% or 2% for *Tsuga* (Davis, 1981).

152 ***Extent and resolution***

153 The extent of the study area was the portion of eastern North America with the highest density of

154 pollen data (Fig. 1). This region contained 75,251 FIA sites and up to 379 Neotoma locations,

155 depending on time period considered. Paciorek & McLachlan (2009) found that spatial patterns

156 relating current and past climates to abundances of pollen and trees were unreliable at resolutions

157 below ~50 km, so the climatic predictors for our model (see below) were downscaled to a

158   resolution of 0.5-degrees (~50-80 km depending on latitude). We upscaled the current tree

159   occurrence data for each grid cell in the climate spatial data layers, keeping track of the number

160   of FIA sites per 0.5-degree cell to be used as weights in the models (Appendix S2). Following

161   this aggregation there were a total of 1,419 FIA observations with presence/absence ratios for

162   *Fagus* and *Tsuga* of 706/713 and 380/1,039, respectively. The number of aggregated pollen

163   observations varied for each 1 kaBP time period (Fig. 2). Although both paleoclimatic and pollen

164   data extended back 21 kaBP, the total sample size and the number of pollen grains of each genus

165   declined rapidly beyond 8 kaBP (Fig. 2). Thus, our hindcast projections extend only from 1 to 8

166   kaBP, which allowed us to validate the models using a minimum of 200 grid cells containing

167   observations, and at least 50 of which contain presences for each genus.

168   *Climate data*

169   Modern climate data came from the observed dataset of the Climate Research Unit (CRU),

170   University of East Anglia (Brohan *et al.*, 2006). Paleoclimate data for this study came from a

171   recent transient simulation of the CCSM3 global circulation model (GCM) (Liu *et al.*, 2009).

172   The standard change-factor approach was employed to statistically downscale and reduce bias in

173   the climate data (Wilby *et al.*, 2004). For each climate variable at each millennial interval, the

174   difference between modeled paleoclimate and modeled modern climate was calculated and then

175   resampled to a 0.5×0.5-degrees grid to match the resolution of the CRU observed climate

176   dataset (Mitchell & Jones, 2005).

177   Decadal averages of seasonal variables were the highest temporal resolution data

178   available from the archived CCSM3 simulations. To get a 'snapshot' of climatic conditions at

179   each millennial time point, decadal averages of seasonal climate variables from the CRU or

180   CCSM3 simulations were calculated for the first 100 years of each millennium (e.g., 8.0 to 7.9

181   kaBP). Because summaries of modern observed climate are available at centennial scales, these

182   same centennial summaries of paleoclimate were derived to aid comparisons between paleo and

183   modern SDMs. Bioclimatic variables that captured precipitation and temperature averages and

184   seasonalities were used because response surface analyses for *Fagus* and *Tsuga* have shown that

185   climatic annual averages, annual ranges, and seasonality were important factors controlling the

186   Holocene migrations of these genera (Bartlein *et al.*, 1986). Specifically, we calculated six

187   bioclimatic variables (Hijmans *et al.*, 2005): annual mean temperature (BIO1), mean diurnal

188   range (BIO2), temperature seasonality (BIO4), temperature annual range (BIO7), annual

189   precipitation (BIO12), and precipitation seasonality (BIO15).

190   Two of the six calculated bioclimatic variables, temperature seasonality and temperature

191   annual range, had within-time correlations with the other bioclimatic variables ≥0.7, so they were

192   not included as explanatory variables in the models that included environmental predictors (see

193   Appendix S3). The correlations between mean diurnal range and annual precipitation varied

194   between modern and historic times (see Appendix S3), and such changing correlation structures

195   between times could be problematic when projecting models beyond the present (Elith *et al.*,

196   2010). To determine if sufficient variance in the current distribution was explained by the two

197   remaining variables with stable correlation structures over time (i.e.*,* annual mean temperature

198   and precipitation seasonality), we compared a model with annual mean temperature,

199   precipitation seasonality, mean diurnal range, and annual precipitation with another that included

200   only annual mean temperature and precipitation seasonality.

201   *Model calibration*

202   We used Bayesian generalized linear models (GLMs) to model genera occurrence. While

203   approaches such as neural networks and genetic algorithms have been used for SDMs and

204 although model projections can be sensitive to the type of statistical model employed (Elith *et*

205 *al.*, 2010), classical approaches do not provide the statistical inferences we sought. Even though

206 GLMs describe a central tendency and not a limiting effect (e.g., of temperature or precipitation

207 extremes), Bayesian spatial GLMs provide exact inference for the random model parameters,

208 including spatial random effects, by estimating entire posterior distributions at both observed and

209 unobserved geographic locations (Gelfand *et al.*, 2006). Because our goal was to compare

210 consistently SDMs with three different specifications (i.e.*,* spatially--varying intercept only

211 (SVI), climate only, and spatially-varying intercept plus climate), we adopted a Bayesian

212 approach in fitting all of the models. Model structure is detailed in Appendix S2; model code is

213 provided in Appendix S4.

214 Including the SVI has a potential for overfitting as it allows variable intercepts for every

215 location and thus a very flexible spatial fit to the FIA data. As a null model, we also fit a

216 multilevel B-Spline to the FIA data (Lee *et al.*, 1997) using the 'MBA' package of 'R' statistical

217 software to determine whether our hindcasting test for the inclusion of a SVI in the Bayesian

218 models was sufficient. As an exploratory analysis into the strength of the residual spatial

219 dependence in the FIA data, we calculated Moran's I from the residuals of the non-spatial GLMs.

220 This latter analysis was conducted using the Spatial Analyst Tool in ArcMap10 (ESRI, 2011).

221 ***Model fit to calibration data***

222 We fit the Bayesian models to 90% of the FIA data ($N = 1,277$) and randomly selected a

223 10% holdout dataset ($N = 142$) to assess predictive performance. We also used DIC to rank the

224 Bayesian models fit to the calibration data (Spiegelhalter *et al.*, 2002). DIC is the sum of the

225 Bayesian deviance (a measure of model fit) and the effective number of parameters (a penalty for

226 model complexity). Lower DIC values indicate better model fit. Models are compared using

227 ΔDIC:

$$\Delta DIC_i = DIC_i - \min(DIC), \tag{3}$$

229 where min(DIC) is the DIC value for the model with the best fit (i.e., lowest DIC value). In

230 general, $\Delta DIC < 2$ indicates weak evidence; $5 < \Delta DIC < 10$ indicates strong evidence, and ΔDIC

231 >10 indicates very strong evidence that one model is preferred over another (Spiegelhalter *et al.*,

232 2002).

233 *FIA hold-out dataset and pollen validations*

234 When projecting the spatial models back in time for the pollen validation, the random effects

235 serve to draw the projected distributions for each genus back toward that of the observed

236 distribution used for model calibration (i.e., the FIA data) in the new time period (Appendix S2).

237 To compare the performance of the models in predicting current and projecting past distributions,

238 three measures were calculated using the 'ROCR' package of 'R' statistical software: the Area

239 Under the Curve (AUC) of a Receiver Operating Curve (ROC), false negative rates (FNR), and

240 false positive rates (FPR). The calculation of FNRs and FPRs requires converting the continuous

241 outputs to a binary form using a threshold, in this case the value that maximizes the sum of

242 sensitivity and specificity (Liu *et al.*, 2005; Lobo *et al.*, 2008).

243   Differences in AUC, FNR, and FPR between models, genera, pollen percentage

244 thresholds, time, and the model × genus interaction were tested with three GLMs. To normalize

245 residuals and reduce heteroskedasticity, AUC, FNR, and FPR were all arcsin transformed.

246 Model, genera, pollen percentage threshold, and the model × genus interaction entered the GLM

247 as fixed factors, and time entered as a covariate. The model × genus interaction was of particular

248 interest as it tested whether or not different models performed better or worse in hindcasting the

249   presence-absence of the two genera. The data were analyzed with separate GLMs for AUC, FNR,

250   and FPR to facilitate the interpretation of Tukey's Honest Significant Differences post-hoc

251   comparisons at the expense of increasing Type II error rates. Bonferroni corrections of the *P*-

252   values from the tests did not alter the significance of any of the effects.

253   **Results**

254   *Parameter estimates and model fit to calibration data*

255   In non-spatial models with two climatic variables (i.e., annual mean temperature and

256   precipitation seasonality) or four climatic variables (i.e., annual mean temperature, mean diurnal

257   range, annual precipitation, and precipitation seasonality), all climatic variables were significant

258   predictors of presence/absence: none of the 95% credible intervals of the parameter estimates

259   included zero (Tables 1, 2). In contrast, in the spatial models some of the climatic explanatory

260   variables were not significant predictors of presence/absence (e.g., annual mean temperature in

261   the *Tsuga* models with two climatic variables and mean diurnal range in the *Fagus* model with

262   four climatic variables; Tables 1 & 2). Changes in the magnitude and sign of parameter estimates

263   between non-spatial and spatial models suggested that non-spatial models violated the

264   assumption of independent identically distributed residuals. The residuals of the non-spatial

265   models for both *Fagus* and *Tsuga* also exhibited significant positive spatial autocorrelation

266   (Moran's I $= 0.604$, $P < 1 \times 10^{-7}$ for *Fagus*; Moran's I $= 0.761$, $P < 1 \times 10^{-7}$ for *Tsuga*),

267   supporting the conclusion that non-spatial models were inappropriate for these data.

268   For *Fagus*, the SVI plus climate model with annual mean temperature and precipitation

269   seasonality had the lowest DIC value and ΔDIC > 10 relative to all other *Fagus* models (Table 3,

270   Fig. 3). In contrast, for *Tsuga,* the SVI model with no bioclimatic predictors had the lowest DIC

271   value and ΔDIC > 10 relative to all other *Tsuga* models (Table 3, Fig. 4).

272   The non-spatial SDMs for both *Fagus* and *Tsuga* that included only annual mean

273   temperature and precipitation seasonality had ΔDIC values >10 relative to the non-spatial models

274   that included annual mean temperature, precipitation seasonality, mean diurnal range, and annual

275   precipitation (Table 3). Given that the correlative relationship between mean diurnal range and

276   annual precipitation was unstable between modern and historic times (see Appendix S3) and that

277   the inclusion of them did not provide a large decrease in the ΔDIC, these two climatic variables

278   were excluded from the models used for prediction that were validated with the 10% holdout FIA

279   dataset and fossil pollen record.

280   *FIA hold-out dataset and pollen validations*

281   For the contemporary 10% hold-out FIA dataset for both genera, the non-spatial model

282   performed worse than the SVI, SVI plus climate, or multilevel B-Spline models in terms of

283   AUC, FNR, and FPR (Table 4; Appendix S5). However, the same was not true when models

284   were hindcasted. Based on AUC, there were significant main effects of model type (non-spatial,

285   SVI, SVI plus climate, FIA B-Spline; $F_{3,118} = 32.4$, $P = 2.4 \times 10^{-15}$), and a significant genus ×

286   model interaction ($F_{3,118} = 13.8$, $P = 8.8 \times 10^{-8}$) (Table 4, Appendix S5) on model performance.

287   For the *Fagus* hindcasts, on average the non-spatial model had higher AUC values than the

288   spatial models (i.e., SVI and SVI plus climate) and FIA multilevel B-spline models, but the

289   opposite was true for *Tsuga*. The FNRs in the hindcasting validation varied by model ($F_{3,118} =$

290   8.1, $P = 6.2 \times 10^{-5}$). The FIA data multilevel B-spline model had the highest FNR and post-hoc

291   comparisons showed that there were no significant differences between the non-spatial and

292   spatial models in FNRs (Table 4, Appendix S5). Similar to the FNRs, the FPRs also varied by

293   model ($F_{3,118} = 9.0$, $P = 1.95 \times 10^{-5}$) (Table 4, Appendix S5). The FIA data multilevel B-spline

294   and the non-spatial models had higher FPRs than the spatial models. There were no significant

295   genus $\times$ model interactions for FNRs ($F_{3,118} = 2.3$, $P = 0.08$) and FPRs ($F_{3,118} = 1.7$, $P = 0.18$).

296   Overall for the three measures, model performance worsened as models were projected further

297   back in time (AUC: $F_{1,118} = 118$, $P = 2.0 \times 10^{-6}$; FNR: $F_{1,118} = 98.7$, $P = 2.0 \times 10^{-16}$; FPR: $F_{1,\ 118} =$

298   109, $P = 2.0 \times 10^{-16}$). Also, model performance was better (i.e., higher AUC and lower FNR and

299   FPR) for *Tsuga* than for *Fagus* (AUC: $F_{1,118} = 10.0$, $P = 0.002$; FNR: $F_{1,118} = 65.5$, $P = 5.8 \times 10^{-}$

300   ^{13}; FPR: $F_{1,118} = 88$, $P = 6.3 \times 10^{-16}$) and for the low pollen percentage thresholds than for the

301   high pollen percentage thresholds (AUC: $F_{1,118} = 14.0$, $P = 2.8 \times 10^{-4}$; FNR: $F_{1,118} = 15.3$, $P = 1.5$

302     $\times 10^{-4}$; FPR: $F_{1,118} = 24.9$, 2.13×10^{-16}). For all three test metrics (i.e., AUC, FNR, FPR), the

303   multilevel B-spline fit to the FIA data, which we used as a 'perfectly fit' model to assess whether

304   or not the spatial models were overfit to the calibration data, performed the worst. This assured

305   us that the pollen validation test was stringent enough.

306   **Discussion**

307   A key question regarding the application of SDMs to predicting the response of species to

308   climate change is whether the failure to include ecological and evolutionary processes (e.g.,

309   dispersal, biotic interactions, readjustment lags) will prove to be problematic (reviewed by

310   Pearson & Dawson, 2003). Depending on the species and its life history, ecological and

311   evolutionary processes may (or may not) lead to its inability to track changes in climate. While

312   there is evidence that vagile organisms (e.g., butterflies) can track rapid climate change (Warren

313   *et al.*, 2001), sessile organisms (e.g., trees) may not readily disperse to newly suitable habitat

314   resulting in limited niche space filling (Svenning & Skov, 2004; Meier *et al.*, 2012). Species

315   undergoing climate driven range expansions coupled with enemy release are hypothesized to be

316   more capable of realizing their potential niche (Hellman, *et al.*, 2012), whereas species limited

317   by a particular resource (e.g., host availability) can be constrained to the spatial distribution of

318   the resource (Merrill *et al.*, 2007). There is evidence that shorter-lived taxa (e.g., insects and

319   herbaceous plants; Woodward, 1990; Thomas *et al.*, 2001) can evolve in response to rapid

320   climate change, but longer-lived taxa that cannot evolve as quickly may experience readjustment

321   lags (Pearson & Dawson, 2003).

322   For those taxa whose distributions do not shift over time as a result of ecological and

323   evolutionary processes, the inclusion of spatial random effects in SDMs could improve

324   projections by providing a more conservative prediction of distributional shifts, especially when

325   climatic variables do not explain much variability in their observed distributions. Alternatively,

326   when climatic variables explain most of the variability in a taxon's observed distribution and the

327   taxon is capable of tracking climate, then accounting for spatial autocorrelation in SDMs won't

328   provide better projections. In other words, the spatial random effects keep the projected

329   distribution similar to the data used for model calibration, unless the covariates (e.g., climatic

330   variables) suggest otherwise. Further, if the climate variables do not explain much of the

331   variability in the observed distribution and the genera's distribution shifts far from the observed

332   distribution over time, then none of the models defined here will perform well. The predictive

333   abilities of non-spatial and spatial SDMs have rarely been compared with temporally varying

334   validation datasets to test these assertions (Gelfand *et al.*, 2006).

335   In this study we tested the predictive abilities of non-spatial and spatial SDMs across

336   eight millennia using data from the pollen record (Appendix S1). We found that spatial SDMs

337   had better fits to the calibration data, higher predictive accuracy for a modern hold-out validation

338   dataset, and greater variance in their outputs than non-spatial SDMs (see also Gelfand *et al.*,

339   2006; Bahn & McGill, 2007). For *Fagus*, the SVI plus climate model provided a better fit to the

340   calibration data than the SVI model, but the opposite was true for *Tsuga*. Also for the two

341   climatic variable models, for *Fagus* there was no change in the sign of the climatic regression

342   coefficients between the non-spatial and spatial models (Table 1), but with *Tsuga* there was a

343   sign change in the regression coefficient for annual mean temperature between the non-spatial

344   and SVI plus climate models (Table 2). This result suggests that for *Tsuga* the spatial random

345   effect could be accounting for dependence in the model's residuals across space as several other

346   studies have found that parameter estimates are affected by spatial autocorrelation (Dormann,

347   2007; Kühn, 2007; Bini *et al.*, 2009; Hodges & Reich, 2010).

348   In the hindcasting analyses, the SVI and SVI plus climate models performed similarly.

349   This suggests that the climatic variables do not contribute much to explaining the variability of

350   occurrence relative to that explained by the spatial random effects. AUC values based on fossil

351   pollen indicated that the non-spatial model performed better for *Fagus* than either of the two

352   spatial models, but the opposite was true for *Tsuga*. However, FNR values did not differ among

353   the models for either genus, and FPR values were greater for non-spatial models for both genera.

354   We have more confidence in FNR and FPR values than in AUC values because the latter

355   describe portions of the ROC curve that are rarely encountered and weights omission and

356   commission errors equally (Lobo *et al.*, 2008). With the pollen record, equal weighting of

357   omission and commission errors may not be ideal; we have much more confidence in the

358   presence of pollen grains than in their absence (Blauww *et al.*, 2007; Blois *et al.* 2011) and false

359   negatives in the pollen record are more problematic than false positives. The lack of differences

360   in false negative rates between models shows that the non-spatial and spatial models have similar

361   FNRs.

362   Although we have less confidence in actual absences in the pollen data, the FPRs are

363   interesting when considering the ecological and evolutionary processes leading to conserved

364   spatial structure in the distributions of species. The greater FPRs of non-spatial models for both

365   genera suggest that spatial effects may account for smaller-scale climatic spatial structure that is

366   not otherwise estimated in large-scale or averaged temperature and precipitation values (Gelfand

367   *et al.*, 2006; Hawkins *et al.*, 2007). Evidence from the fossil pollen and paleoclimate records

368   suggests that climatic shifts can result in abrupt ecological changes in vegetation that are driven

369   by internal dynamics, such as site-specific environmental characteristics (*e.g.,* soil moisture) or

370   biotic interactions (*e.g.,* competition) that create geographically localized variation in vegetation

371   composition (Williams *et al.*, 2011). Taxon-specific responses to climate forcing also could

372   explain why the SVI model had the lowest DIC for *Tsuga* and why the two spatial models

373   performed better in regards to both AUC and FPR for *Tsuga*, but not for *Fagus.* Approximately

374   5.5 kaBP *Tsuga* experienced a range contraction known as the "hemlock decline" potentially due

375   to an abrupt change in climate, a phytophagous insect infestation, or both (Bhiry & Filion, 1996;

376   Foster *et al.*, 2006). If the hemlock decline was due to an abrupt change in climate, then localized

377   ecological changes could have resulted in stronger spatial structure in its distribution. However,

378   decoupling changes in distributions due to climate and spatial structure due to biotic interactions

379   or site-specific abiotic characteristics is difficult because observed spatial structure is (or was)

380   inherently linked to abrupt climate change.

381   Alternatively, the spatial random effects may have captured a missing covariate, such as

382   an ecological process that generates spatial structure (Clayton *et al.*, 1993; Paciorek, 2010). Such

383   processes could include dispersal, competitive interactions, land-use history, or underlying

384   features of the terrain. For example, if dispersal limitation prevents distributional shifts, then we

385   might expect that spatial SDMs would perform better for dispersal-limited taxa (e.g., *Tsuga*) that

386   cannot track changes in climate, but not necessarily for taxa with effective dispersal vectors (e.g.,

387   *Fagus*) that can gain dominance by migrating faster to climatically favorable sites (Pearman *et*

388   *al.*, 2008b). These taxon-specific differences in dispersal mode and degree of dominance could

389   explain why *Tsuga* seemed to be less responsive to climate over the past 8 millennia than *Fagus*

390   as evidenced by the better performance over time of the two spatial models in regards to both

391   AUC and FPR for *Tsuga*, but not for *Fagus*. Simulation experiments for European trees with

392   spatially explicit process models accounting for changing macroclimate, competition, and habitat

393   connectivity showed that some of the spatial autocorrelation between two time periods may be

394   due to very slow migration rates resulting in severe time lags that are not accounted for in non-

395   dynamic and non-spatial SDMs (Meier *et al.*, 2012). Also, Dobrowski *et al.* (2011) found that

396   non-spatial SDMs fit to widespread plants with more effective dispersal mechanisms had higher

397   predictive accuracy over 75 years of climate change in California than non-spatial SDMs fit to

398   dispersal-limited plants.

399   Given the results of this study, should researchers include spatial random effects in

400   SDMs? We found that for two long-lived eastern North American trees, spatial models provided

401   better fits to calibration data and lower FPRs, but not necessarily improvements in AUC or the

402   FNR. The better fits of the spatial SDMs may have resulted from the richness of the FIA data

403   used to calibrate the models. Large samples of evenly-dispersed data likely will capture any

404   spatial structure; consequently a spatial SDM should fit well. However, when sample sizes are

405   small, there is less of a chance that the spatial structure will be represented adequately.

406   Ultimately, whether to include spatial random effects in SDMs will depend on the taxon being

407   modeled, the cost of false positives, and the quality of the data.

408

409   **Acknowledgments**

418   **Supplementary material**

419   **Appendix S1** Presence-absence plots of historic pollen distributions.

420   **Appendix S2** Detailed description of the models.

421   **Appendix S3** Plots of between- and within-time correlations of paleoclimate data.

422   **Appendix S4** Code for analyses programmed in R and C++.

423   **Appendix S5** Results of the *Fagus* and *Tsuga* low pollen threshold analysis.

424   **Biosketch**

425   Sydne Record is a post-doctoral research fellow at Harvard University – Harvard Forest with

426   broad research interests in validating ecological models and Bayesian statistics.

427   Author contributions: M.C.F., S.R., and A.M.E. conceived the ideas for the study. S.R., M.C.F.,

428   and A.O.F. conducted statistical analyses. S.R. led the writing of the manuscript with critical

429   comments from all co-authors. S.V. provided the downscaled climate data.

430   **Literature Cited**

431   Bahn, V. & McGill, B.J. (2007) Can niche-based distribution models outperform spatial

432   interpolation? *Global Ecology and Biogeography*, **16**, 733-742.

433   Bartlein, P.J., Prentice, I.C. & Webb, T. (1986) Climatic response surfaces from pollen data for

434   some eastern North American taxa. *Journal of Biogeography*, **13**, 35-57.

435   Bhiry, N. & Filion, L. (1996) Mid-Holocene decline in eastern North America linked with

436   phytophagous insect activity. *Quaternary Research*, **45**, 312-320.

437   Bini, L.M., Diniz-Filho, J.A.F., Rangel, T.F.L.V.B., *et al.* (2009) Coefficient shifts in

438   geographical ecology: an empirical evaluation of spatial and non-spatial regression.

439   *Ecography*, **32**, 193-204.

440   Blauw, M., Christen, J.A., Mauqouy, D., van der Plicht, J. & Bennett, K.D. (2007) Testing the

441   timing of radiocarbon-dated events between proxy archives. *Holocene*, **17**, 283-288.

442   Blois, J.L., Williams, J.W., Grimm, E.C., Jackson, S.T. & Graham, R.W. (2011) A

443   methodological framework for assessing and reducing temporal uncertainty in

444   paleovegetation mapping from late-Quaternary pollen records. *Quaternary Science*

445   *Reviews,* **30**, 1926-1939.

446   Brohan, P., Kennedy, J.J., Harris, I., Tett, S.F.B. & Jones, P.D. (2006) Uncertainty estimates in

447   regional and global observed temperature changes: a new dataset from 1850. *Journal of*

448   *Geophysical Research*, **111**, D12106.

449   Clayton, J.S., Carlin, B.P. & Montomoli, C. (1993) Spatial correlation in ecological analysis.

450   *Journal of Epidemiology*, **22**, 1193-1201.

451   Davis, M.B. (1981) Quaternary history and the stability of forest communities. *Forest Succession*

452   (ed. by D.C. West, H.H. Shugart and D.B. Botkin), pp. 132-153. Springer, New York.

453   Diez, J.M. & Pulliam, H.R. (2007) Hierarchical analysis of species distributions and abundance

454   across environmental gradients. *Ecology*, **88**, 3144-3152.

455   Dobrowski, S.Z., Thorne, J.H., Greenberg, J.A., Safford, H.D., Mynsberge, A.R., Crimmins,

456   S.M. & Swanson, A.K. (2011) Modeling plant distributions over 75 years of measured

457   climate change in California, USA: relating temporal transferability to species traits.

458   *Ecological Monographs*, **81**, 241-257.

459   Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species

460   distribution data. *Global Ecology and Biogeography*, **16**, 129-138.

461   Elith, J., Kearney, M. & Phillips, S. (2010) The art of modeling range-shifting species. *Methods

462   in Ecology and Evolution*, **1**, 330-342.

463   Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of

464   prediction error: why does the introduced range of the fire ant over-predict its native

465   range? *Global Ecology and Biogeography,* **16**, 24-33.

466   Foster, D.R., Oswald, W.W., Faison, E.K., Doughty, E.D. & Hansen B.C.S. (2006) A climatic

467   driver for abrupt mid-Holocene vegetation dynamics and the hemlock decline in New

468   England. *Ecology*, **87**, 2959-2966.

469   Franklin, J. & J.A. Miller. (2009) *Mapping species distributions: spatial inference and

470   prediction.* Cambridge University Press, Cambridge.

471   Gelfand, A.E., Latimer, A., Wu, S. & Silander, J.A. (2006) Building statistical models to analyze

472   species distributions. *Hierarchical Modelling for the Environmental Sciences: Statistical

473   Methods and Applications.* (ed. by J.S. Clark and A.E. Gelfand), pp. 77-07. Oxford

474   University Press, Oxford.

475   Guisan, A. & W. Thuiller. (2005) Predicting species distribution: offering more than simple

476   habitat models. *Ecology Letters*, **8**, 993-1009.

477   Hawkins, B.A., Diniz-Filho, J.A.F., Bini, L.M., De Marco, P. & Blackburn, T.M. (2007) Red

478   herrings revisited: spatial autocorrelation and parameter estimation in geographical

479   ecology. *Ecography*, **30**, 375-384.

480   Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W. & Sykes, M.T. (2006)

481   Methods and uncertainties in bioclimatic envelope modeling under climate change.

482   *Progress in Physical Geography*, **30**, 751-777.

483   Hellman, J.J., Prior, K.M. & Pelini, S.L. (2012) The influence of species interactions on

484   geographic range change under climate change. *Annals of the New York Academy of*

485   *Sciences*, **1249,** 18-28.

486   Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution

487   interpolated climate surfaces for global land areas. *Journal of Climatology*, **25**, 1965-

488   1978.

489   Hodges, J.S. & Reich, B.J. (2010) Adding spatially-correlated errors can mess up the fixed effect

490   you love. *American Statistical Association*, **64**, 325-334.

491   Kühn, I. (2007) Effects of incorporating spatial autocorrelation into the analysis of species

492   distribution data. *Global Ecology and Biogeography*, **16**, 129-138.

493   Lee, S., Wolberg, G. & Shin, S.Y. (1997) Scattered data interpolation with multilevel B-splines.

494   *IEEE Transactions on Visualization and Computer Graphics*, **3**, 229-244.

495   Legendre, P. (1993) Spatial autocorrelation – trouble or new paradigm. *Ecology*, **74**, 1659-1673.

496   Lichstein, J.W., Simons, T.R., Shriner, S.A. & Franzreb, K.E. (2002) Spatial autocorrelation and

497   autoregressive models in ecology. *Ecological Monographs*, **72**, 445-463.

498   Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in

499   the prediction of species distributions. *Ecography*, **28**, 385-393.

500   Liu, Z., Otto-Bleisner, B.L., He, F., *et al.* (2009) Transient simulation of last deglaciation with a

501   new mechanism for Bolling-Allerod warming. *Science*, **325**, 310-314.

502   Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the

503   performance of predictive distribution models. *Global Ecology and Biogeography*, **17**,

504   145-151.

505   Meier, E.S., Lischke, H., Schmatz, D.R. & Zimmerman, N.E. (2012) Climate, competition and

506   connectivity affect future migration and ranges of European trees. *Global Ecology and*

507   *Biogeography*, **21**, 164-178.

508   Merrill, R.M., Gutierrez, D., Lewis, O.T., Gutierrez, J., Diez, S.B. & Wilson, R.J. (2007)

509   Combined effects of climate and biotic interactions on the elevational range of a

510   phytophagous insect. *Journal of Animal Ecology*, **77**, 145-155.

511   Mitchell, T.D. & Jones, P.D. (2005) An improved method of constructing a database of monthly

512   climate observations and associated high-resolution grids. *International Journal of*

513   *Climatology*, **25**, 305-314.

514   Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niches. *Global*

515   *Ecology and Biogeography*, **18**, 521-531.

516   Paciorek, C.J. (2010) The importance of scale for spatial-confounding bias and precision of

517   spatial regression estimators. *Statistical Science*, **25**, 107-125.

518   Paciorek, C.J. & McLachlan, J.S. (2009) Mapping ancient forests: Bayesian inference for spatio-

519   temporal trends in forest composition using the fossil pollen proxy record. *Journal of the*

520   *American Statistical Society*, **104**, 608-622.

521   Pearman, P.B., Randin, C.F., Broennimann, *et al.* (2008a) Prediction of plant species

522   distributions across six millennia. *Ecology Letters*, **11**, 357-369.

523   Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008b) Niche dynamics in space

524   and time. *Trends in Ecology and Evolution*, **23**, 149-158.

525   Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the

526   distribution of species: Are bioclimate envelope models useful? *Global Ecology and*

527   *Biogeography*, **12**, 361-371.

528   Segurado, P., Araújo, M.B. & Kunin, W.E. (2006) Consequences of spatial autocorrelation for

529   niche-based models. *Journal of Applied Ecology*, **43**, 433-444.

530   Spiegelhalter, D.J., Best, N., Carlin, B.P. & van der Linde, A. (2002) Bayesian measures of

531   model complexity and fit (with discussion). *Journal of the Royal Statistical Society Series*

532   *B*, **64**, 583-639.

533   Svenning, J.-C. & Skov, F. (2004) Limited filling of the potential range in European tree species.

534   *Ecology Letters*, **7**, 565-573.

535   Thomas, C.D., Bodsworth, E.J., Wilson, R.J., Simmons, A.D., Davis, Z.G., Musche, M. &

536   Conradt, L. (2001) Ecological and evolutionary processes at expanding range margins.

537   *Nature*, **411**, 577-581.

538   Veloz, S.D., Williams, J.W., Blois, J.L., He, F., Otto-Bliesner, B. & Liu, Z. (2012) No-analog

539   climates and shifting realized niches during the late Quaternary: implications for 21st-

540   century predictions by species distribution models. *Global Change Biology*, **18**, 1698-

541   1713.

542   Warren, M.S., Hill, J.K., Asher, T.J., Fox, R., Huntley, B., Roy, D.B., Telfer, M.G., Jeffcoate, S.,

543   Harding, P., Jeffcoate, G., Willis, S.G., Greatorex-Davies, J.N., Moss, D. & Thomas, C.D.

544   (2001) Rapid responses of British butterflies to opposing forces of climate and habitat

545   change. *Nature*, **414**, 65-69.

546   Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P. & Mearns, L.O. (2004) Guidelines

547   for use of climate scenarios developed from statistical downscaling methods. IPCC Task

548   Group on data and scenario support for Impact and Climate Analysis (TGICA), pp. 1-27.

549   Williams, J.W., Blois, J.L. & Shuman, B.N. (2011) Extrinsic and intrinsic forcing of abrupt

550   ecological change: case studies from the late Quaternary. *Journal of Ecology*, **99**, 664-

551   677.

552   Wilson, T.L., Odei, J.B., Hooten, M.B. & Edwards, T.C. (2010) Hierarchical spatial models for

553   predicting pygmy rabbit distribution and relative abundance. *Journal of Applied Ecology*,

554   **47**, 401-409.

555   Woodward, F.I. (1990) The impact of low temperatures in controlling the geographical

556   distribution of plants. *Philosophical Transactions of the Royal Society of London B,* **326**,

557   585-593.

558   Woudenberg, S.W., Conkling, B.L., O'Connell, B.M., LaPoint, E.B., Turner, J.A. & Waddell,

559   K.L. (2010) The Forest Inventory and Analysis Database: Database description and user's

560   manual version 4.0 for Phase 2. Gen. Tech. Rep. RMRS-GTR- 245. Fort Collins, CO:

561   U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 336

562   p.

563 **Table 1**. Parameter credible intervals (2.5%, 50.0%, and 97.5% percentiles) for the *Fagus*

564 spatially-varying intercept (SVI), non-spatial (NS2 and NS4) and SVI plus climate (SVI2 and

565 SVI4) models. The numbers two and four in the acronyms for the non-spatial and SVI plus

566 climate models indicate the number of bioclimatic explanatory variables included in the models.

567 The two climatic variables models included annual mean temperature (BIO1) and precipitation

568 seasonality (BIO15). The four climatic variables models included annual mean temperature

569 (BIO1), mean diurnal range (BIO2), annual precipitation (BIO12), and precipitation seasonality

570 (BIO15). For models with spatial random effects, the spatial random effect variance and spatial

571 decay parameter are denoted σ^2 and φ, respectively.

	Model	β Parameter	2.5%	50.0%	97.5%
572					
573					
574	SVI	Intercept	-7.23	-5.28	-2.72
575	SVI	σ^2	8.11	12.90	20.24
576	SVI	φ	$1.09 \ 10^{-6}$	$1.62 \ 10^{-6}$	$2.63 \ 10^{-6}$
577	NS2	Intercept	-3.06	-3.01	-2.96
578	NS2	BIO1	-0.48	-0.46	-0.43
579	NS2	BIO15	-1.83	-1.78	-1.72
580	NS4	Intercept	-3.11	-3.06	-3.01
581	NS4	BIO1	-0.62	-0.58	-0.54
582	NS4	BIO2	0.33	0.37	0.40
583	NS4	BIO12	-0.20	-0.14	-0.09
584	NS4	BIO15	-2.03	-1.96	-1.90
585	SVI2	Intercept	-7.49	-5.77	-4.41

586	SVI2	BIO1	-1.57	-1.25	-0.89
587	SVI2	BIO15	-0.97	-0.47	-0.08
588	SVI2	σ^2	6.35	10.32	17.25
589	SVI 2	φ	$1.15 \ 10^{-6}$	$1.90 \ 10^{-6}$	$3.20 \ 10^{-6}$
590	SVI 4	Intercept	-8.27	-5.47	-3.13
591	SVI 4	BIO1	-1.37	-0.83	-0.25
592	SVI 4	BIO2	-0.16	-0.03	-0.11
593	SVI 4	BIO12	-0.15	-0.47	0.80
594	SVI 4	BIO15	-0.89	-0.36	-0.12
595	SVI 4	σ^2	5.53	10.50	17.78
596	SVI 4	φ	$1.14 \ 10^{-6}$	$1.91 \ 10^{-6}$	$3.69 \ 10^{-6}$

597

598 **Table 2**. Parameter credible intervals (2.5%, 50%, and 97.5% percentiles) for the *Tsuga* spatially-

599 varying intercept (SVI), non-spatial (NS2 and NS4) and SVI plus climate (SVI 2 and SVI 4)

600 models. The numbers two and four in the acronyms for the non-spatial and SVI plus climate

601 models indicate the number of bioclimatic explanatory variables included in the models. The two

602 climatic variables models included annual mean temperature (BIO1) and precipitation

603 seasonality (BIO15). The four climatic variables models included annual mean temperature

604 (BIO1), mean diurnal range (BIO2), annual precipitation (BIO12), and precipitation seasonality

605 (BIO15). For models with spatial random effects, the spatial random effect variance and spatial

606 decay parameter are denoted σ^2 and φ, respectively.

	Model	β Parameter	2.5%	50%	97.5%
607	Model	β Parameter	2.5%	50%	97.5%
609	SVI	Intercept	-9.10	-7.68	-4.15
610	SVI	σ^2	12.6	22.3	36.4
611	SVI	φ	1.09 10^{-6}	2.23 10^{-6}	2.74 10^{-6}
612	NS2	Intercept	-3.50	-3.45	-3.40
613	NS2	BIO1	-1.14	-1.11	-1.07
614	NS2	BIO15	-1.20	-1.16	-1.12
615	NS4	Intercept	-3.55	-3.50	-3.45
616	NS4	BIO1	-1.34	-1.30	-1.25
617	NS4	BIO2	0.31	0.35	0.40
618	NS4	BIO12	0.07	0.14	0.21
619	NS4	BIO15	-1.25	-1.21	-1.12
620	SVI2	Intercept	-10.18	-8.38	-3.45

621	SVI 2	BIO1	0.07	0.48	0.89
622	SVI 2	BIO15	-1.09	-0.55	-0.05
623	SVI 2	σ^2	10.86	18.57	32.11
624	SVI 2	φ	$1.09\ 10^{-6}$	$1.68\ 10^{-6}$	$2.96\ 10^{-6}$
625	SVI 4	Intercept	-8.28	-5.73	-4.00
626	SVI 4	BIO1	-1.28	-0.85	-0.26
627	SVI 4	BIO2	-0.16	-0.03	0.11
628	SVI 4	BIO12	-0.15	0.47	0.80
629	SVI 4	BIO15	-0.81	-0.36	0.12
630	SVI 4	σ^2	5.94	10.58	17.86
631	SVI 4	φ	$1.14\ 10^{-6}$	$1.89\ 10^{-6}$	$3.43\ 10^{-6}$

632

633 **Table 3.** Fits of the spatially-varying intercept (SVI), non-spatial, and SVI plus climate SDMs to

634 the modern Forest Inventory and Analysis (FIA) occurrence data for *Fagus* and *Tsuga*.

635 Bioclimatic variables included in the models with climatic predictors were: annual mean

636 temperature (BIO1), mean diurnal range (BIO2), annual precipitation (BIO12), and precipitation

637 seasonality (BIO15). Model fit was evaluated with the Deviance Information Criterion (DIC),

638 which is the sum of P_D (the effective number of parameters) and the posterior mean of the

639 deviance. To facilitate model comparison, ΔDIC was also calculated, where the model with the

640 lowest DIC has a value of zero and all other models are compared to it.

	Model	Bioclimatic variable	Genus	P_D	DIC	ΔDIC
641						
642	SVI	None	*Fagus*	247	35893	81
643	Non-spatial	1, 15	*Fagus*	3	41497	5685
644	Non-spatial	1, 2, 12, 15	*Fagus*	5	41125	5313
645	SVI-climate	1, 15	*Fagus*	248	35812	0
646	SVI-climate	1, 2, 12, 15	*Fagus*	251	35826	14
647	SVI-climate	None	*Tsuga*	170	23685	0
648	Non-spatial	1, 15	*Tsuga*	3	30025	6340
649	Non-spatial	1, 2, 12, 15	*Tsuga*	5	29715	6030
650	SVI-climate	1, 15	*Tsuga*	164	23708	23
651	SVI-climate	1, 2, 12, 15	*Tsuga*	160	23727	42

Table 4. Model performance as measured by Area Under the Receiver Operating Curve (AUC), false negative rates (FNR), and false positive rates (FPR) for the non-spatial model, spatially-varying intercept (SVI) model, SVI plus climate, and multilevel B-spline fit to modern *Fagus* and *Tsuga* occurrence data from the Forest Inventory and Analysis (FIA) data. Predictions of the models for modern time were validated with a 10% hold-out dataset from the FIA data. Hindcasts were validated with data from the fossil pollen record provided by the Neotoma database using the "high" pollen thresholds for both genera. The numbers behind the AUC, FNR, and FPR values in parentheses for the Bayesian models represent the standard error calculated from 1000 random draws from the post burn-in MCMC iterations. For the FIA multilevel B-spline approximation there is no standard error as there were no MCMC iterations to draw from.

Genus	Performance Measure	Time (kaBP)	Non-spatial	SVI	SVI-climate	FIA
Fagus	AUC	0	0.87 (4×10^{-4})	0.91 (0.01)	0.92 (0.01)	0.91
		1	0.89 (5×10^{-4})	0.87 (0.02)	0.87 (0.02)	0.86
		2	0.90 (4×10^{-4})	0.88 (0.02)	0.88 (0.02)	0.86
		3	0.89 (6×10^{-4})	0.88 (0.01)	0.88 (0.02)	0.86
		4	0.88 (6×10^{-4})	0.87 (0.02)	0.87 (0.02)	0.84
		5	0.85 (1×10^{-3})	0.85 (0.02)	0.84 (0.02)	0.83
		6	0.84 (2×10^{-3})	0.84 (0.02)	0.83 (0.03)	0.83
		7	0.81 (1×10^{-3})	0.80 (0.02)	0.80 (0.03)	0.78
		8	0.73 (2×10^{-3})	0.76 (0.01)	0.74 (0.02)	0.71
Fagus	FNR	0	0.22 (0.01)	0.14 (0.04)	0.14 (0.03)	0.11
		1	0.20 (0.01)	0.23 (0.04)	0.22 (0.09)	0.26

	2	0.19 (0.02)	0.21 (0.05)	0.20 (0.09)	0.24
	3	0.19 (0.01)	0.19 (0.04)	0.21 (0.09)	0.23
	4	0.22 (0.01)	0.20 (0.04)	0.22 (0.09)	0.23
	5	0.28 (0.02)	0.24 (0.04)	0.25 (0.10)	0.26
	6	0.26 (0.01)	0.25 (0.05)	0.27 (0.10)	0.24
	7	0.30 (0.01)	0.31 (0.05)	0.32 (0.10)	0.31
	8	0.34 (0.03)	0.33 (0.04)	0.35 (0.07)	0.38
Fagus FPR	0	0.23 (0.01)	0.14 (0.02)	0.14 (0.02)	0.12
	1	0.21 (0.02)	0.23 (0.04)	0.22 (0.05)	0.23
	2	0.20 (0.02)	0.21 (0.04)	0.20 (0.06)	0.22
	3	0.20 (0.01)	0.19 (0.03)	0.21 (0.06)	0.22
	4	0.24 (0.02)	0.19 (0.04)	0.22 (0.07)	0.25
	5	0.28 (0.03)	0.24 (0.03)	0.25 (0.07)	0.26
	6	0.27 (0.02)	0.25 (0.04)	0.26 (0.07)	0.28
	7	0.26 (0.02)	0.31 (0.04)	0.30 (0.07)	0.29
	8	0.35 (0.01)	0.33 (0.04)	0.34 (0.07)	0.40
Tsuga AUC	0	$0.85\ (3\times10^{-3})$	0.95 (0.02)	$0.95\ (8\times10^{-3})$	0.97
	1	$0.85\ (3\times10^{-3})$	0.91 (0.01)	0.91 (0.02)	0.82
	2	$0.86\ (4\times10^{-4})$	0.89 (0.01)	0.89 (0.02)	0.81
	3	$0.87\ (4\times10^{-4})$	0.88 (0.01)	0.87 (0.02)	0.80
	4	$0.83\ (3\times10^{-3})$	0.86 (0.02)	0.85 (0.02)	0.80
	5	$0.84\ (3\times10^{-3})$	0.90 (0.02)	0.89 (0.02)	0.84
	6	$0.86\ (2\times10^{-3})$	0.91 (0.02)	0.90 (0.02)	0.80

	7	0.85 (5×10^{-3})	0.88 (0.02)	0.87 (0.02)	0.80
	8	0.76 (5×10^{-3})	0.89 (0.02)	0.88 (0.02)	0.79
Tsuga FNR	0	0.20 (0.03)	0.11 (0.04)	0.11 (0.04)	0.05
	1	0.16 (0.07)	0.16 (0.04)	0.18 (0.03)	0.20
	2	0.19 (0.02)	0.18 (0.04)	0.20 (0.03)	0.21
	3	0.19 (0.03)	0.18 (0.04)	0.20 (0.04)	0.20
	4	0.21 (3×10^{-3})	0.20 (0.05)	0.21 (0.04)	0.21
	5	0.25 (0.02)	0.17 (0.04)	0.18 (0.04)	0.20
	6	0.20 (0.02)	0.16 (0.04)	0.18 (0.03)	0.20
	7	0.25 (0.01)	0.18 (0.04)	0.19 (0.03)	0.24
	8	0.30 (0.01)	0.19 (0.05)	0.19 (0.04)	0.33
Tsuga FPR	0	0.22 (0.01)	0.11 (0.03)	0.11 (0.03)	0.09
	1	0.19 (0.03)	0.16 (0.04)	0.17 (0.04)	0.23
	2	0.16 (0.01)	0.18 (0.04)	0.19 (0.03)	0.20
	3	0.19 (1×10^{-3})	0.18 (0.04)	0.19 (0.03)	0.22
	4	0.23 (0.01)	0.20 (0.03)	0.21 (0.03)	0.26
	5	0.24 (0.02)	0.17 (0.04)	0.18 (0.04)	0.23
	6	0.19 (0.01)	0.16 (0.05)	0.17 (0.04)	0.18
	7	0.23 (0.01)	0.18 (0.04)	0.18 (0.03)	0.22
	8	0.32 (0.01)	0.20 (0.04)	0.21 (0.03)	0.31

Figure 1. Map of the study extent in the eastern United States showing Forest Inventory and Analysis (FIA) plots (hollow circles) and Neotoma pollen sites (solid triangles) snapped to a resolution of 0.5-degrees (Alber's Equal Area Conic projection).

Figure 2. Numbers of sites with presences (black fill) or absences (white fill) of *Fagus* (a and c) and *Tsuga* (b and d) based on the low and high pollen thresholds from present to 21 kiloannums before present (kaBP) based on fossil pollen data from the Neotoma database. Data extending beyond 8 kaBP were not used in the analyses due to the low number of presences of *Fagus* and *Tsuga* beyond that time.

Figure 3. Maps of a) a surface approximation of the probability of occurrence of *Fagus* generated by a multilevel B-spline fit to the raw FIA data and the predicted probability of presence of the b) non-spatial, c) spatially-varying intercept, and d) spatially-varying intercept plus climate SDMs to modern *Fagus* FIA data (Alber's Equal Area Conic Projection). The surface approximation in a) was calculated with the MBA package in R.
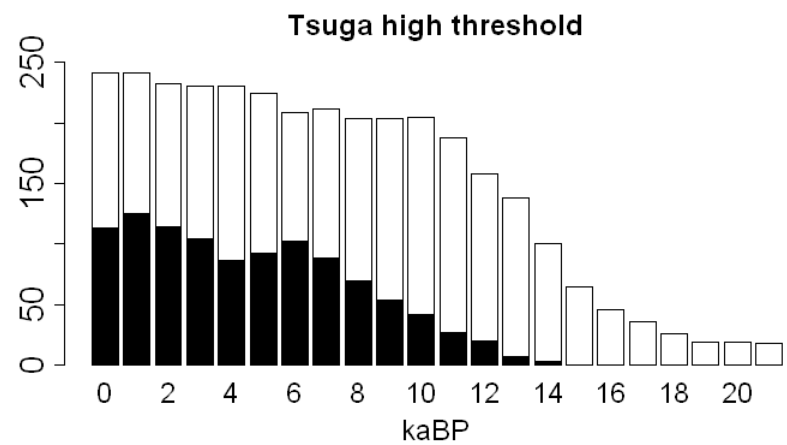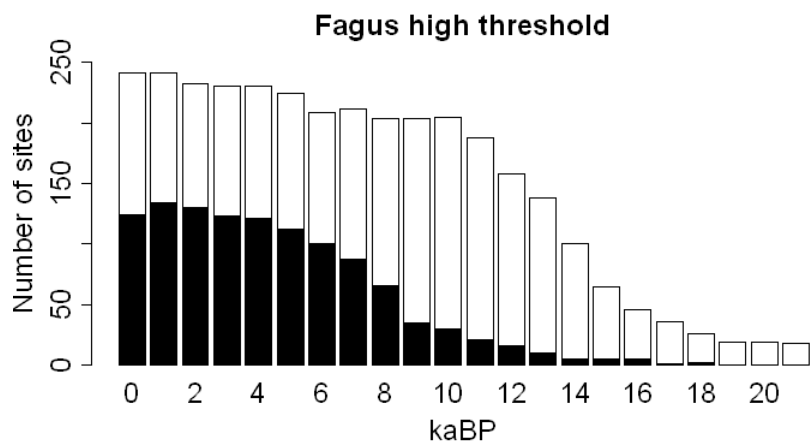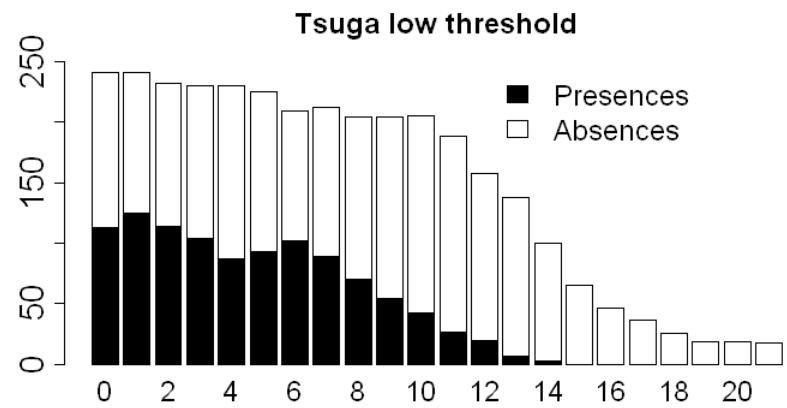
Figure 4. Maps of a) a surface approximation of the probability of occurrence of *Tsuga* generated by a multilevel B-spline fit to the raw FIA data and the predicted probability of presence of the b) non-spatial, c) spatially-varying intercept, and d) spatially-varying intercept plus climate SDMs to modern *Tsuga* FIA data (Alber's Equal Area Conic Projection). The surface approximation in a) was calculated with the MBA package in R.

Study extent
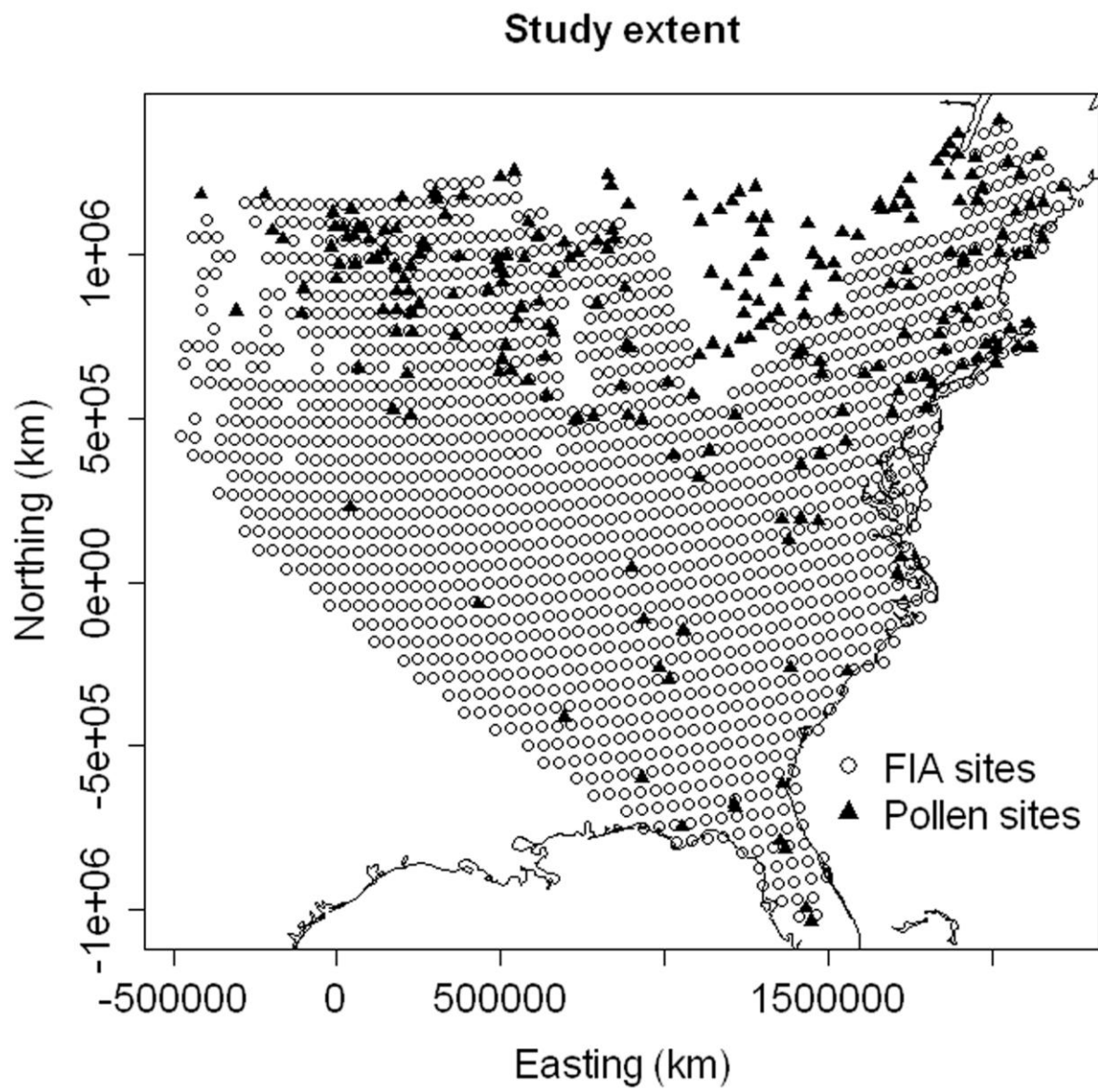
Figure 1

Figure 2

Figure 3

Figure 4