Bayesian model averaging of climate-dependent forest models using Expectation–Maximization

Nicolas Picard, Nikola Besic, Modeste Meliho, Julien Sainte-Marie, Frédéric Mortier, Myriam Legay

To cite this version:

Nicolas Picard, Nikola Besic, Modeste Meliho, Julien Sainte-Marie, Frédéric Mortier, et al.. Bayesian model averaging of climate-dependent forest models using Expectation–Maximization. Ecological Modelling, 2025, 510, pp.111355. ⟨10.1016/j.ecolmodel.2025.111355⟩. ⟨hal-05307661⟩

HAL Id: hal-05307661

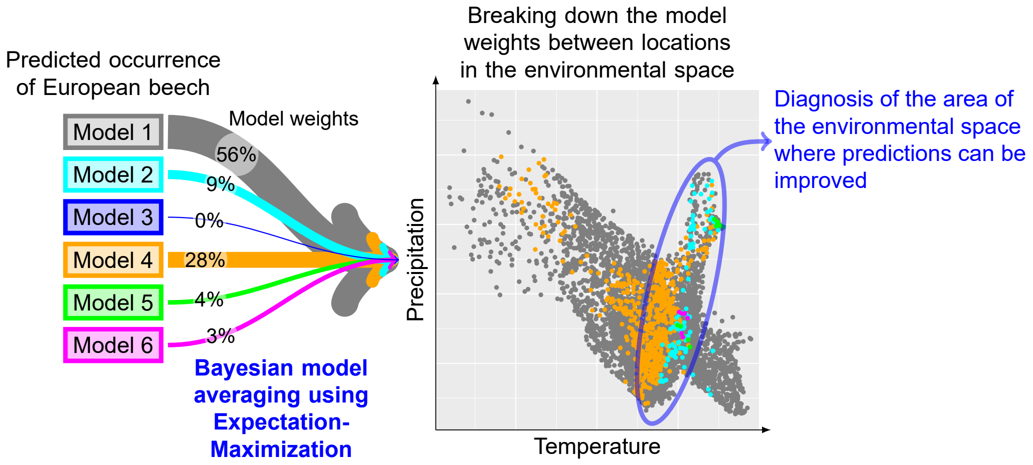
https://hal.inrae.fr/hal-05307661v1

Submitted on 22 Apr 2026

Graphical Abstract

Bayesian model averaging of climate-dependent forest models using Expectation-Maximization

Nicolas Picard, Nikola Besic, Modeste Meliho, Julien Sainte-Marie, Frédéric Mortier, Myriam Legay



Highlights

Bayesian model averaging of climate-dependent forest models using Expectation-Maximization

Nicolas Picard, Nikola Besic, Modeste Meliho, Julien Sainte-Marie, Frédéric Mortier, Myriam Legay

- Bayesian model averaging can be used to average species distribution models
- In contrast to simple and weighted model averaging, local model weights are derived
- Model weights can be mapped in the environmental and geographical spaces
- Differences in the predicted occurrence of beech in France were better understood

Bayesian model averaging of climate-dependent forest models using Expectation-Maximization

Nicolas Picard[a,*], Nikola Besic[b], Modeste Meliho[c,1], Julien Sainte-Marie[d], Frédéric Mortier[e], Myriam Legay[d]

[a] *GIP Ecofor, 42 rue Scheffer, Paris 75116, Île-de-France, France*
[b] *Laboratoire d'Inventaire Forestier, IGN, ENSG, 14 rue Girardet, Nancy 54000, Meurthe-et-Moselle, France*
[c] *Université du Québec à Rimouski, 300 allée des Ursulines, Rimouski, QC G5L 3A1, Québec, Canada*
[d] *UMR Silva, Université de Lorraine, AgroParisTech, INRAE, 14 rue Girardet, Nancy 54000, Meurthe-et-Moselle, France*
[e] *AMAP, Univ. Montpellier, IRD, CNRS, CIRAD, INRAE, UMR AMAP TA A-51/PS2, Boulevard de la Lironde, Montpellier 34398, Hérault, France*

Abstract

In the context of rapid climate change, climate-dependent models are essential for assessing species vulnerability. However, variation in model structure and divergence in their predictions introduce substantial uncertainty. Rather than selecting a single "best" model, a more robust strategy is to integrate predictions across models. Bayesian model averaging with Expectation–Maximization (BEM) provides an alternative to simple model averaging (SMA) and weighted model averaging (WMA) for combining ensemble predictions. To date, BEM has been rarely applied to tree species distribution models. We developed a BEM framework for models predicting either species occurrence or proxy variables linked to occurrence. The approach was applied to European beech (*Fagus sylvatica*) in France, using an ensemble of six models: four species distribution models, one model predicting the probability of hydraulic failure, and one model predicting juvenile produc-

---

[*] Corresponding author

Email addresses: `nicolas.picard@gip-ecofor.org` (Nicolas Picard), `nikola.besic@ign.fr` (Nikola Besic), `modestemeliho@yahoo.fr` (Modeste Meliho), `julien.sainte-marie@agroparistech.fr` (Julien Sainte-Marie), `frederic.mortier@cirad.fr` (Frédéric Mortier), `myriam.legay@agroparistech.fr` (Myriam Legay)

[1] Present address: CERFO, 2440 Chemin de Sainte-Foy, Québec, QC G1V 1T2, Canada

tivity. In contrast to SMA and WMA, which assigned similar weights across models, BEM concentrated 85% of the weight on two models. Furthermore, BEM enabled spatially explicit decomposition of model weights, allowing us to identify regions where predictions diverged most strongly. The resulting probability maps revealed a specific zone in environmental space where model agreement on beech occurrence was particularly limited. Focusing on this zone may help refine projections and shed light on the ecological mechanisms that enable local persistence.

Keywords: consensus prediction, ensemble modelling, European beech, finite mixture modelling, forest modelling, species distribution model, weighted model averaging

---

1. Introduction

The accelerating pace of climate change poses major challenges for forest management and conservation. Because trees are long-lived organisms, forest management necessarily operates over long temporal horizons. Consequently, evaluating the vulnerability of current tree species to future climatic conditions is essential, as is assessing the suitability of potential new species. Climate-dependent models provide relevant tools for addressing these questions. Over the past two decades, species distribution models have been widely developed to quantify relationships between species distributions and climate (Guisan et al., 2017). In parallel, climate-dependent forest dynamics models have been established, integrating the climate sensitivity of processes such as growth, mortality, recruitment, and water stress (Fontes et al., 2010). Beyond ecological dynamics, climate-dependent economic models of forestry have also been developed to evaluate the impacts of climate change on the forest sector and associated value chains, or conversely, to design new silvicultural strategies aligned with climate objectives (Aaheim et al., 2011).

Each of these models sheds light on a particular aspect of forest adaptation. A model focused on the hydraulics of tree (e.g. Cochard et al., 2021), for instance, may successfully predict drought-induced mortality but overlook other important processes such as the climate-dependent allocation of carbon between tree compartments (Guillemot et al., 2014), the climate-dependent shift in species composition in mixed stands (Morin et al., 2018), or climate-dependent regeneration dynamics (König et al., 2022). These processes are better captured by other models. Faced with this diversity of climate-dependent models, one option for building decision-making tools can be to select the "best" model. Here, "best" would refer to the predictive performances of models, assuming that (*i*) there is a hierarchy of processes to predict tree species adaptation, and (*ii*) this hierarchy is conservative along climate change. Yet, the uncertainty due to the model choice is often considerable and may even result in inconsistent outcomes (Pearson et al., 2006; Lindner et al., 2014; Cheaib et al., 2012; Bircher et al., 2015). This issue is not specific to forest models but has been noticed in many different contexts, e.g. for crop models (Tao et al., 2018), for regional climate models (Vidale et al., 2003), or for allometric models to predict tree biomass (Picard et al., 2015).

Given that many processes can be critical to predict the adaptation of tree species to climate change, rather than selecting a single "best" model, a more robust strategy is to integrte predictions across models. Building 'consensus' (*sensu* Araújo and New, 2007) from different models is part of ensemble modelling. Ensemble modelling consists in exploring the prediction range and building consensus from multiple model predictions. The multiple predictions reflect different sources of uncertainty. They may result from (*i*) different model choices, (*ii*) different values for the model coefficients, (*iii*) different values for the input variables of the model (as would result from measurement errors or from different climate models for input climate variables), or (*iv*) different initial conditions or different sites for model calibration. In this study, we focus on the first source of uncertainty, *viz.* the model choice.

How consensus is built from an ensemble of model predictions depends on the availability of validation data (Araújo and New, 2007). When no validation data is available and the prediction is numerical, a simple consensus is given by the average or median of predictions (see Martre et al., 2015, for an example). When the prediction is categorical (such as the presence or absence of a species), the consensus can be built by considering the distribution of predictions across categories (e.g. Mellert et al., 2015). The consensus then determines the probability of observing each category. Another approach is to use a multivariate analysis to summarize the different predictions according to consensus axes (Thuiller, 2004). When validation data is available and the prediction is numerical, a weighted mean of the predictions can be used as a consensus. The weights typically reflect how well each model predicts the observations in the validation dataset (Marmion et al., 2009). When the prediction is categorical, the frequencies of categories in the validation dataset conditional on each combination of predicted categories can be used as a consensus (e.g. Gritti et al., 2013).

Bayesian model averaging using Expectation-Maximization (BEM) is another method that can be used to build a consensus from an ensemble of model predictions (Hoeting et al., 1999; Raftery et al., 1997; Dormann et al., 2018). BEM is close to a frequentist approach and differs from the truly Bayesian approach proposed by Corani and Mignatti (2015) for presence-absence data. Like other consensus methods, BEM can be seen as weighted mean of predictions with weights reflecting the predictive performance of models. However, unlike other methods, BEM can also be seen as a finite mixture model (Raftery et al., 1997). It estimates as a latent variable the probability that each observation of the validation dataset is generated by one of the competing model in the ensemble of predictions. This latter interpretation of BEM is particularly fruitful to explore the range of predictions. BEM thus provides a way to diagnose differences in prediction between models that other consensus methods do not provide. Although BEM has been used in various contexts such as climate forecast (Smith et al., 2009), hydrology (Zhang et al., 2009) or forest biomass estimation (Li et al., 2008; Picard et al., 2012), it has been less used than other averaging methods in ecology (Dormann et al., 2018). We are not aware of its use with climate-dependent forest models.

The objective of this study was to assess how BEM can be used to understand differences in the predictions of climate-dependent forest models, so as to diagnose which model provided relevant information for the adaptation of tree species to climate change depending on the context. The different models were assessed on the basis of their capacity to predict the occurrence of a given tree species (European beech in our case). BEM was compared to simple model averaging (SMA) and weighted model averaging (WMA). SMA and WMA are two commonly used methods to average the predictions of species distribution models (Thuiller et al., 2009). We assumed that there was no single "best" model and that the performance of each model was context-dependent.

2. Materials and methods

2.1. Ensemble of models

We focused on the occurrence of European beech (*Fagus sylvatica* L.) in France. The autecology of this tree species has been extensively studied (San-Miguel-Ayanz et al., 2016, p.94-95). Observations of the occurrence of beech in France were obtained at a 8 km × 8 km resolution from the data of the French National Forest Inventory (Robert et al., 2010). This dataset of beech occurrence was the one used to train BEM and get the weights in WMA. Six climate-dependent models were included in the BEM: DIGI-SDM, EU-Trees4F, MOD-BJ, NBM, Sureau-Ecos and Sustree. The first four models were correlative species-distribution models, whereas Sureau-Ecos was a process-based model and Sustree was an empirical model.

DIGI-SDM predicted the occurrence of beech from six environmental variables: mean annual temperature, soil water deficit, soil pH, C/N ratio of the soil A horizon, temporary waterlogging, and permanent waterlogging (Piedallu et al., 2016). DIGI-SDM was fitted to the data of the inventory plots of the French National Forest Inventory (NFI) using a generalized additive model with smooth splines.

EU-Trees4F predicted the occurrence of beech from thirteen environmental variables: winter and summer temperature and precipitation, precipitation seasonality, mean annual temperature, mean temperature of the coldest month, total annual precipitation, continentality, a humidity index, growing degree days above 5 °C, soil pH and organic carbon content (Mauri et al., 2022). EU-Trees4F was fitted at pan-European scale using the EU-Forest dataset (Mauri et al., 2017) complemented with data from the monitoring plots of the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests). The predicted occurrence was obtained using the BIOMOD2 platform (Thuiller et al., 2009) as the average of six fitted models: generalized linear model, generalized additive model, generalized boosting model, multiple adaptive regression splines, maximum entropy and Random Forest.

MOD-BJ predicted the occurrence of beech from five environmental variables: mean annual temperature, water balance (i.e. rainfall minus Turc reference evapotranspiration) for July, average temperature in January, soil pH, and the C/N ratio of the soil A horizon (Javaux, 2011). MOD-BJ was fitted to the data of the EcoPlant phyto-ecological database (Gégout et al., 2005) using logistic regression with a second-degree polynomial response to each environmental variable.

NBM predicted the occurrence of beech from a wide set of environmental variables: minimum, maximum and mean annual temperature; annual rainfall; wind speed; relative humidity and vapour pressure deficit; atmospheric pressure; global solar irradiance; potential evapotranspiration (based on Turc's formula, Penman's formula and Penman-Monteith's formula corrected by the CO_2 concentrations); growing degree days ($> 0°C$, $> 5°C$, $> 10°C$); number of days of frost ($< 0°C$, $< -5°C$, $< -10°C$); water stress intensity; the number of days of water stress; and the start date of water stress (Badeau et al., 2010; Cheaib et al., 2012). NBM was fitted to the data of the inventory plots of the French NFI using logistic regression.

Sureau-Ecos predicted the probability of xylem hydraulic failure under drought (Cochard et al., 2021; Ruffault et al., 2022). Sureau-Ecos was based on a mechanistic modelling of the hydraulic (water flows) and hydric (water pools) functioning of the tree. Sureau-Ecos relied on four types of parameters: species-specific parameters (leaf, stomatal and hydraulic traits) that were derived from direct ecophysiological measurements or traits databases; plant (or stand) morphological parameters that determined the overall leaf area index and the plant internal water stores, which were derived from forest inventory and species-specific allometries; soil parameters that were derived from soil databases; and parameters linked to hydraulic conductance.

Sustree predicted the productivity of juvenile beech stands from climate, soil characteristics and seed provenance (Chakraborty et al., 2019). The productivity was expressed as the carbon sink per hectare. Soil characteristics included 39 soil physical and chemical properties (pH, total content of carbon and nitrogen, cation exchange capacity, soil water holding capacity, etc.). Climate variables were annual rainfall, rainfall minimum and maximum, temperature mean, minimum and maximum, length of vegetation period, and growing degree days. Sustree was fitted to provenance trials data using structural equation modelling.

Model predictions and the results of model averaging could be shown in the geographic space and in the environmental space. The former may be more meaningful for those familiar with the geography of the study area, but the later was easier to interpret from an autecological standpoint. A geographic location was mapped to a single location in the environmental space but, conversely, a given location in the environmental space could correspond to several geographic locations. Therefore, we performed the model averaging in the geographic space, then projected the results in the environmental space. Results in the geographic space were still given in Appendix S2 and S3. Following San-Miguel-Ayanz et al. (2016), the environmental space was summarized by three planes that consisted of the following pairs of environmental variables: annual precipitation *versus* annual average temperature; average temperature of the coldest month *versus* potential spring-summer solar irradiation; sum of precipitation of the driest month *versus* seasonal variation of monthly precipitation.

2.2. Homogenizing the responses of models

Model averaging in its standard formulation requires that all models in the ensemble have the same response variable (Hoeting et al., 1999). However, climate-dependent forest models may have different outputs, as is the case for the six models studied here. Following Cheaib et al. (2012), we homogenized the responses of models by relating their output variables to the presence/absence of the subject species. If the model already predicted the probability of presence of the species as is the case with species distribution models (here, DIGI-SDM, EU-Trees4F and MOD-BJ), no further step was needed. If the model did not predict the probability of presence (Sureau-Ecos and Sustree), logistic regression was used to predict it from the model output variables $\mathbf{x}_k$. As for the NBM model, we observed a sharp variation of its predicted probability of presence and a clear improvement of the relationship with the observed presence when using a logistic transformation. Therefore, although NBM predicted the probability of presence of beech, we also used the logistic transformation of its predictions in the model averaging.

The logistic regression used to homogenize the responses of models had a different fit for BEM on the one hand, and for SMA and WMA on the other hand. For SMA and WMA, the logistic regression was fitted prior to model averaging. All model predictions had the same weight in the regression. For BEM, the fit of the logistic regression was iterative and simultaneous with model weighting, as explained in the next section. The different predictions of a model then had different weights in the regression.

We considered that the output variables $\mathbf{x}_k$ were proxies for the species presence for all climate conditions. In other words, we assumed that the relationship between $\mathbf{x}_k$ and species presence was more straightforward than the relationship between species presence and climate. Hence, implicitly, we assumed that the complex response of the tree species to climate was well captured by model k through its output variables $\mathbf{x}_k$.

2.3. Bayesian model averaging using Expectation-Maximization

Let K be the number of models in the ensemble, including $Q \leq K$ models that directly predict the probability of presence of the species (here, $K = 5$ and $Q = 2$). Without loss of generality, these latter models were numbered from 1 to Q, while $Q+1$ to K were the numbers of the models that predicted the probability of presence from proxy variables. Let $\mathbf{P} = \{P_i : i = 1, \ldots, n\}$ be a dataset of n observations indicating if the species is present $(P_i = 1)$ or absent ($P_i = 0$) at n locations. For models 1 to Q, let q_{ik} be the probability of presence predicted by model k at location i. For models $Q+1$ to K, this probability was computed from the set $\mathbf{x}_{ik}$ of the output variables of model k at location i. Quantities q_{ik} for $1 \leq k \leq Q$ and $\mathbf{x}_{ik}$ for $Q+1 \leq k \leq K$ were the input to BEM.

Given observations $\mathbf{P}$, BEM modeled the Bernoulli distribution of the species occurrence O_i at location i (with $O_i = 1$ if the species was present and 0 if it was absent) as a weighted mean of the models' predictions:

$$\Pr(O_i|\mathbf{P}) = \sum_{k=1}^{K} \Pr(O_i|M = k, \mathbf{P})\ \Pr(M = k|\mathbf{P}) \tag{1}$$

where M was the categorical random variable indicating which model in the ensemble of K models generated event O_i. The probability $w_k \equiv \Pr(M = k|\mathbf{P})$ was the posterior probability of model k being the one generating observations. It reflected how well model k fited observations. The posterior probabilities added up to one: $\sum_{k=1}^{K} w_k = 1$. Finally,

$$\Pr(O_i|M = k, \mathbf{P}) = \begin{cases} q_{ik} & \text{for } k \leq Q \\ \text{logis}(\mathbf{x}_{ik}\boldsymbol{\beta}_k) & \text{for } k > Q \end{cases} \tag{2}$$

where $\boldsymbol{\beta}_k$ was the unknown vector of coefficients and logis the logistic function: $\text{logis}(x) = 1/[1 + \exp(-x)]$. In the present case, each model had a single output x_{ik} so that $\mathbf{x}_{ik}\boldsymbol{\beta}_k$ simplified to $\beta_{0k} + x_{ik}\beta_{1k}$, where β_{0k} was the intercept and β_{1k} the rate. Hence, BEM could be interpreted as a finite mixture of Bernoulli distributions and logistic regressions. When the parameters of the Bernoulli distributions are to estimated, the finite mixture of Bernoulli distributions is known to be an unidentifiable problem (cf. §20.1.5 in Shalizi, 2014). However, in our case, the parameters were either known (for $1 \leq k \leq Q$) or resulted from fitted logistic regressions (for $k > Q$). The parameters to be estimated were the logistic coefficients $\boldsymbol{\beta}_{Q+1}$, …, $\boldsymbol{\beta}_K$ and the weights w_1, …, w_K, with likelihood:

$$\mathcal{L}(w_1, \ldots, w_K, \boldsymbol{\beta}_{Q+1}, \ldots, \boldsymbol{\beta}_K) = \prod_{i=1}^{n} \left\{ \sum_{k=1}^{Q} w_k q_{ik}^{P_i} (1 - q_{ik})^{1-P_i} \right. \\ \left. + \sum_{k=Q+1}^{K} w_k [\text{logis}(\mathbf{x}_{ik}\boldsymbol{\beta}_k)]^{P_i} [1 - \text{logis}(\mathbf{x}_{ik}\boldsymbol{\beta}_k)]^{1-P_i} \right\}$$

The direct maximization of the likelihood of the mixture was difficult, but the Expectation-Maximization (EM) algorithm offered an easy alternative to fit BEM (McLachlan and Krishnan, 2008).

The EM algorithm introduced “missing data” z_{ik}, where z_{ik} was the posterior probability that model k generated observation i. The EM algorithm was iterative and alternated between the E and M steps. Let the subscript (j) indicate for any parameter its estimate at the jth iteration of the EM algorithm. All models were *a priori* equiprobable, so that w_k and z_{ik} were initialized as: $w_{k(0)} = z_{ik(0)} = 1/K$, while the coefficients $\boldsymbol{\beta}_k$ were initialized by fitting the logistic regressions to all observations equally. In the E step, the z_{ik} were estimated given the current estimates of the parameters as:

$$z_{ik(j)} = \frac{w_{k(j-1)}\, q_{ik(j-1)}^{P_i} (1 - q_{ik(j-1)})^{1-P_i}}{\sum_{l=1}^{K} w_{l(j-1)}\, q_{il(j-1)}^{P_i} (1 - q_{il(j-1)})^{1-P_i}}$$

for $j \geq 1$. The numerator of $z_{ik(j)}$ corresponded to the weighted likelihood of observation P_i according to a Bernoulli distribution with parameter $q_{ik(j-1)}$, so the ratio in $z_{ik(j)}$ corresponded to a ratio of the likelihood for model k over the likelihood for all models. In the M step, the model weights w_k and the model coefficients $\boldsymbol{\beta}_k$ were estimated using the current estimates of z_{ik}

as observation weights:

$$w_{k(j)} = \frac{1}{n}\sum_{i=1}^{n} z_{ik(j)},$$

$\boldsymbol{\beta}_{k(j)}$ was estimated by fitting a logistic regression to weighted observations using $z_{ik(j)}$ as the weight for observation i, and:

$$q_{ik(j)} = \begin{cases} q_{ik} & (1 \leq k \leq Q) \\ \text{logis}(\mathbf{x}_{ik}\boldsymbol{\beta}_{k(j)}) & (Q < k \leq K) \end{cases}.$$

The E and M steps were iterated until the absolute change in w_k and z_{ik} for all i and k in one iteration did not exceed 10^{-6}.

The BEM prediction of species occurrence was the conditional expectation:

$$\text{E}(O_i|\mathbf{P}) = \text{E}[\text{E}(O_i|M, \mathbf{P})] = \sum_{k=1}^{K} w_k q_{ik} \equiv \bar{q}_i$$

where q_{ik} for $k > Q$ equaled $\text{logis}(\mathbf{x}_{ik}\boldsymbol{\beta}_k)$, and the BEM predictive variance was:

$$\begin{aligned}\text{Var}(O_i|\mathbf{P}) &= \text{E}[\text{Var}(O_i|M, \mathbf{P})] + \text{Var}[\text{E}(O_i|M, \mathbf{P})] \\ &= \sum_{k=1}^{K} w_k q_{ik}(1 - q_{ik}) + \sum_{k=1}^{K} w_k (q_{ik} - \bar{q}_i)^2 \qquad (3)\end{aligned}$$

The first sum in the right-hand side of the predictive variance was the within-model variance while the second sum was the between-model variance. This between-model variance corresponded to the ensemble spread. If one model outcompetes the others (i.e. $w_1 \simeq 1$ and $w_k \simeq 0$ for $k > 1$), the between-model variance will be close to zero and the within-model variance will be close to the variance of the Bernoulli distribution for that model. In contrast, the between-model variance will be maximum when the models are equiprobable and their predicted probabilities of presence contrasted.

All computations were made using the R software. The R code for BEM is available in Appendix S4.

2.4. Simple and weighted model averaging

In contrast to BEM where all models had to be jointly considered to determine the posterior probabilities, the model weights in SMA and WMA were prior weights that could be determined for each model independently of the others. SMA was obtained by fitting a logistic regression to training data for models $Q + 1$ to K, then averaging the probabilities of presence predicted by each model using the same weight $1/K$ for all models. Because the iterative EM algorithm was initiated with equiprobable models ($w_{k(0)} = 1/K$ for all k), SMA could be computed using the same algorithm as BEM by stopping it before the first EM iteration.

WMA used the same predicted probabilities of presence as SMA, but weighted these predictions differently for each model. Following Thuiller et al. (2009), the weights of the models were proportional to their predictive performance of species occurrence. The predictive performance was assessed as the area under the receiver operating characteristic curve (AUC ROC) (Allouche et al., 2006).

The SMA and WMA predictive variances could be calculated using the same formula (3) as for BEM, by replacing the posterior probability w_k of model k by its weight. Thus, within-model variance and between-model variance were also defined for SMA and WMA. However, the posterior probability z_{ik} that model k generated observation i was not defined in SMA and WMA. Therefore, unlike BEM, SMA and WMA did not indicate how the model weights broke down between locations.

3. Results

3.1. Prediction of the occurrence of beech

BEM, SMA and WMA had almost the same overall predictive performance (Table 1). Any of these three model averages outperformed each of its component models. However, this similarity in predictive ability arose from contrasting model fits. The logistic regression used to homogenize the responses of models had a different fit for BEM on the one hand, and for SMA and WMA on the other hand (Table 2). In the latter case, the logistic regression showed a smooth transition between beech absence and beech presence (Appendix S1). In contrast, in the former case, the logistic regression showed a sharp transition between absence and presence for two of the models (Sureau-Ecos and Sustree). With the exception of the Sureau-Ecos model, all coefficients were significantly different from zero. For Sureau-Ecos, the sharp transition occurred at a probability of hydraulic failure of 0.015; for the Sustree model, it occurred at a carbon sink of 99.0 ha^{-1}.

The BEM prediction of the occurrence of beech in France was similar to that of the SMA and WMA. It was also similar to the prediction of the DIGI-SDM model (Fig. 1a and Appendix S1). However, the patterns of between-model and within-model predictive variances differed among BEM, SMA and WMA (Fig. 1b, c). The plot of between-model variance showed the areas in the environmental space where the disagreement between the six models was greatest. In particular, for BEM, there was strong agreement between the six models when the annual average temperature was high ($\gtrsim$ 12.5°C, left column in Fig. 1). This agreement was on a very low probability of occurrence of beech. In most of the environmental space, the within-model

Table 1: Predictive performance of six models that predicted the occurrence of beech in France, either directly (DIGI-SDM, EU-Trees4F, MOD-BJ) or indirectly (NBM, Sureau-Ecos, Sustree), and predictive performance of their average. The average of the six models was made in three different ways: Bayesian model averaging using Expectation-Maximization (BEM), simple model averaging (SMA), and weighted model averaging (WMA, where the weights are given by the predictive performance of each model). The predictive performance was given by the area under the receiver operating characteristic curve (AUC ROC): the closer this score to 1, the better the prediction of beech occurrence.

Averaging	Model						Model
method	DIGI-SDM	MOD-BJ	NBM	Sureau-Ecos	Sustree	EU-Trees4F	average
BEM	0.867	0.840	0.821	0.669	0.654	0.819	0.879
SMA	0.867	0.840	0.827	0.691	0.691	0.819	0.882
WMA	0.867	0.840	0.827	0.691	0.691	0.819	0.882

Table 2: Logistic regressions that predicted the probability of presence of beech in France from the output variable of a model. In simple and weighted model averaging (SMA, WMA), the values of these output variables had the same weight in the regression. In Bayesian model averaging using Expectation-Maximization (BEM), they had different weights.

	Averaging	Intercept			Rate parameter		
Model	method	Estimate	Std. Err.	p-value	Estimate	Std. Err.	p-value
NBM	BEM	-5.308	0.135	<0.001	9.723	0.210	<0.001
Sureau-Ecos	BEM	$4.122 \cdot 10^3$	$4.62 \cdot 10^4$	0.93	$-2.800 \cdot 10^5$	$3.14 \cdot 10^6$	0.93
Sustree	BEM	$5.379 \cdot 10^5$	$1.54 \cdot 10^5$	<0.001	$-5.431 \cdot 10^3$	$1.55 \cdot 10^3$	<0.001
NBM	SMA, WMA	-1.907	0.045	<0.001	3.032	0.062	<0.001
Sureau-Ecos	SMA, WMA	-0.0194	0.050	0.70	-13.046	1.623	<0.001
Sustree	SMA, WMA	5.927	0.226	<0.001	-0.0637	0.002	<0.001

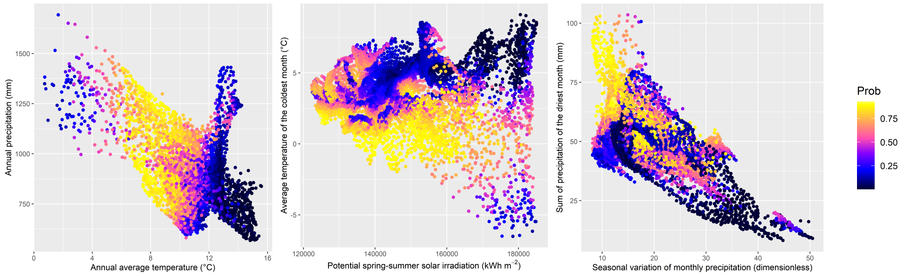
variance of BEM was greater than its between-model variance (Fig. 1d). In the environmental subspace defined by the annual average temperature and the annual precipitation, the main area where the within-model variance of BEM was less than its between-model variance corresponded to that zone with an annual average temperature close to 12.5–13°C and a large gradient in annual precipitation (from 700 to 1450 mm; left panel of Fig. 1d). Another area corresponded to the smallest annual average temperature ($\lesssim 2.5$°C).

SMA and WMA had very similar within-model and between-model variances (Appendix S1). The SMA/WMA pattern of within-model variance in the environmental space was similar to that of BEM, but shifted to higher variance values. Also, the SMA/WMA pattern of between-model variance in the environmental space was similar to that of BEM, but shifted to smaller variance values. Unlike BEM, the within-model variance of SMA/WMA was almost always greater than its between-model variance.

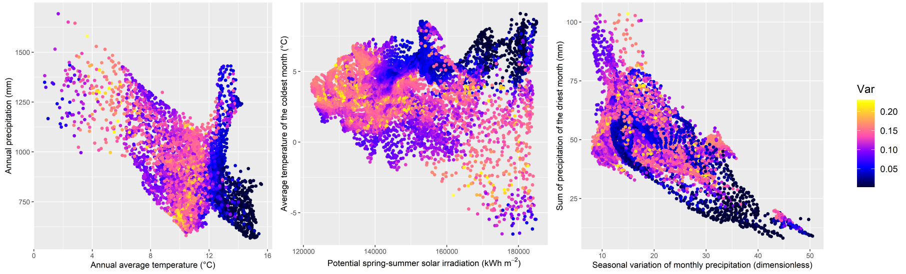
3.2. Posterior probabilities of models

Model weights in WMA were close to the constant weights of $1/K$ (here with $K = 6$ models) used in SMA (Table 3). In contrast, the posterior probabilities of the models in BEM were much more contrasted among models. Two models alone, DIGI-SDM and NBM, accumulated 85% of the posterior probabilities. The remaining 15% were shared between three other models, EU-Trees4F, Sustree and Sureau-Ecos. The last model, MOD-BJ, had an almost null contribution to the posterior probabilities. This very low posterior probability of MOD-BJ did not mean that this model had a poor predictive performance of the occurrence of beech, but rather that the information it provided was completely redundant with that provided by the DIGI-SDM

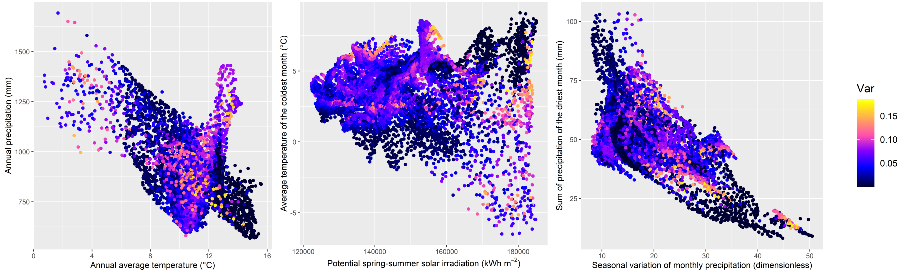
a. Prediction



b. Within-model variance



c. Between-model variance



d. Within − between variance

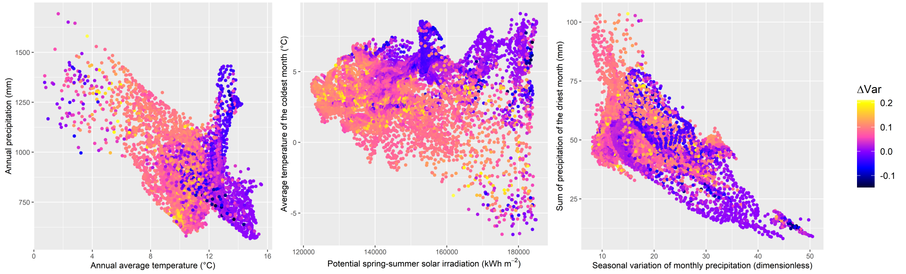


Figure 1: Prediction (a) and predictive variance (b-c) of the occurrence of beech using the Bayesian model averaging using Expectation-Maximization of six models: DIGI-SDM, EU-Trees4F, MOD-BJ, NBM, Sureau-Ecos or Sustree. The predictive variance was split into two components, the within-model (b) and the between-model variance (c). The difference between these two components is also shown (d). Predictions and variances are shown in three planes of the environmental space, defined by the following pairs of environmental variables: annual precipitation *versus* annual average temperature (left column), average temperature of the coldest month *versus* potential spring-summer solar irradiation (middle column), sum of precipitation of the driest month *versus* seasonal variation of monthly precipitation (right column).

model. The predictions by DIGI-SDM and MOD-BJ were indeed very similar, so that the former completely shadowed the latter.

Unlike SMA and WMA, BEM further allowed us to decompose model weights by location, making it possible to explore spatial variation in model dominance. When examining the model with the highest posterior probability at each location, model DIGI-SDM was the most likely model at most locations in the environmental space (Fig. 2). Model MOD-BJ was nowhere the most likely model. The locations where EU-Trees4F, Sustree or Sureau-Ecos were the most likely concentrated in that zone with an annual average temperature close to 12.5–13°C and a 700–14500 mm gradient in annual precipitation (left panel of Fig. 2). That zone was precisely the one where the between-model variance of BEM overcame its within-model variance.

Building on this spatial perspective, the maps of posterior probabilities provided further insight into how individual models contributed "corrections" to the DIGI-SDM predictions (Fig. 3). The NBM model introduced corrections broadly across the environmental space, whereas the EU-Trees4F, Sustree and Sureau-Ecos models brought additional "corrections" again in that zone of the environmental space with an annual average temperature close to 12.5–13°C and a 700–1450 mm gradient in annual precipitation (left column of Fig. 3). The Sureau-Ecos model predicted a high mortality rate when the annual average temperate was high and a very low mortality rate elsewhere (Appendix S1). Its capacity to predict the occurrence of beech in specific places where other models failed to do so was actually related to its poor discrimination across these large areas of the environmental space.

Table 3: Weights used to average the predictions of six models that predicted the occurrence of beech in France, when using Bayesian model averaging using Expectation-Maximization (BEM), simple model averaging (SMA), and weighted model averaging (WMA). For BEM, the weights were the posterior model probabilities. For WMA, the weights were given by the predictive performance of the models, assessed as the area under the receiver operating characteristic curve (AUC ROC).

Model	BEM	SMA	WMA
DIGI-SDM	0.563	0.166	0.183
EU-Trees4F	0.086	0.166	0.173
MOD-BJ	<0.001	0.166	0.177
NBM	0.282	0.166	0.175
Sureau-Ecos	0.037	0.166	0.146
Sustree	0.032	0.166	0.146

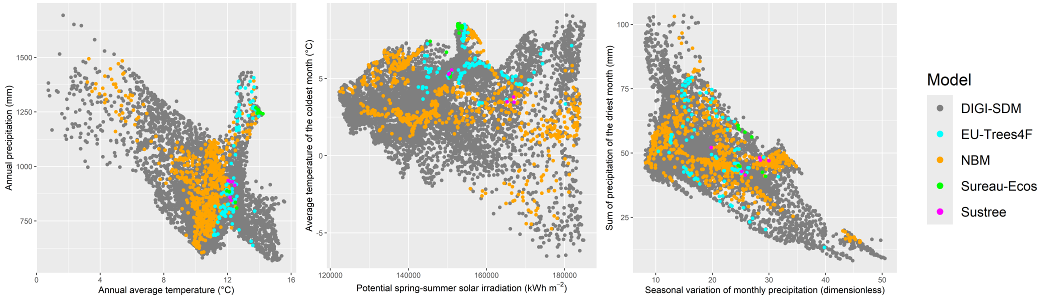


Figure 2: Indicator variable showing which model among DIGI-SDM, EU-Trees4F, MOD-BJ, NBM, Sureau-Ecos and Sustree had the highest posterior probability to be the one generating the observed occurrence of beech. The indicator is shown in three planes of the environmental space, defined by the following pairs of environmental variables: annual precipitation *versus* annual average temperature (left panel), average temperature of the coldest month *versus* potential spring-summer solar irradiation (middle panel), sum of precipitation of the driest month *versus* seasonal variation of monthly precipitation (right panel).

4. Discussion

4.1. Averaged prediction and model weights

BEM produced a map of the averaged prediction of the probability of presence of a subject tree species among different models. This averaged prediction was an alternative to model selection and made the best of the predictions of the different models (Araújo and New, 2007; Fragoso et al., 2018). For beech in France, the average improved the prediction of every model taken separately, even the one with the “best” predictive performance. The capacity of ensemble modelling to improve the prediction of single models had already been highlighted (Mahnken et al., 2022). BEM has been used to better represent modelling uncertainties in forest growth (van Oijen et al., 2013; Van Oijen, 2017), tree mortality (Lu et al., 2019), tree diameter

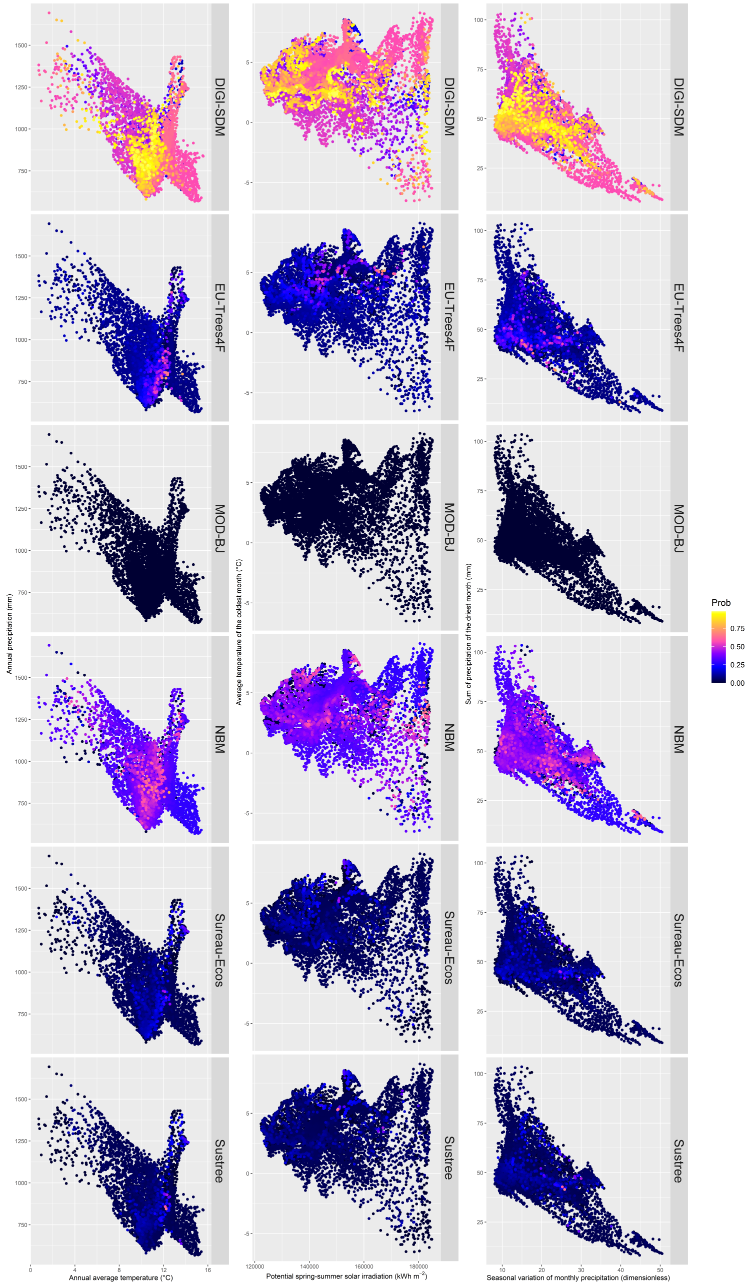


Figure 3: Posterior probability that the occurrence of beech was generated by one of the six models: DIGI-SDM, EU-Trees4F, MOD-BJ, NBM, Sureau-Ecos or Sustree. The probabilities are shown in three planes of the environmental space, defined by the following pairs of environmental variables: annual precipitation *versus* annual average temperature (left column), average temperature of the coldest month *versus* potential spring-summer solar irradiation (middle column), sum of precipitation of the driest month *versus* seasonal variation of monthly precipitation (right column). At each location in each of these environmental planes, the sum of probabilities over the six models was equal to one.

distributions (Bullock and Boone, 2007), or conservation planning (Wintle et al., 2003). However, as far as we know, this is the first time it was used to predict the probability of presence of a tree species, using species distribution models or models whose output was taken as a proxy of the probability of presence

The BEM prediction of the probability of presence of beech in France gave contrasted weights to the different models. A low weight did not necessarily mean that the model makes poor predictions. It could also result from the redundancy of the model with another, as was the case with the MOD-BJ model that was shadowed by DIGI-SDM. In a different context, think of the share of variance explained by collinear predictors in multiple regression, which is an identifiability problem (Brun et al., 2001). Due to this lack of identifiability, small changes in the regression can lead to a shift in the predictor explaining the largest share of variance. Dormann et al. (2018) pointed out the drawback of high covariance between model predictions in BEM. However, they were focusing on the error of model-averaged predictions, which is of secondary interest in our case. The identifiability problem is more evident if we consider the BEM as a finite mixture of distributions. If two models make similar predictions, they bring similar distributions in the mixture, and so they become unidentifiable in the mixture. Yet, the EM algorithm can still bring interesting results in this case (Carreira-Perpinán and Renals, 2000).

The predictive performance of an individual model was not only reflected by its weight in the averaged prediction but also by the way its predictions were transformed by the BEM. For the NBM model, the coefficients of the logistic regression were of the same order whether the logistic regression was directly fitted to the training data set of observed beech presence, or whether they were iteratively estimated through the BEM (Table 2). In contrast, for the Sureau-Ecos and Sustree models, the BEM and the SMA brought very different logistic coefficients (Table 2). The direct fit of the logistic regression led to a moderate transformation of the predictions of these models, while the BEM led to a sharp thresholding of their predictions. Using thresholds to convert model predictions into predicted presence-absence of the species was the approach followed by Cheaib et al. (2012). The proposed method can thus also be seen as an alternative to the binary classifier used by these authors.

4.2. Classification of map pixels and variance maps

Regarding the uncertainty related to the model choice among climate-dependent forest models, we believe that understanding differences among model predictions is the key issue to address (Besic et al., 2024). In this respect, the averaged prediction remains a secondary result of BEM. This is also the reason why we have preferred the EM approach to Bayesian model averaging proposed by Raftery et al. (1997) to truly Bayesian approaches such as those proposed by Corani and Mignatti (2015). The EM approach to BEM had the advantage to introduce the latent variable z_{ik} that could be interpreted as the probability that model k generated observation i. This latent variable enabled us to break down the overall model weight between locations. Only BEM offered this feature; neither SMA nor WMA provided it. When observations were the presence-absence of a species on a raster map, each pixel of the map could thus be assigned to the model with the highest

z_{ik} (Fig. 2). In other words, the BEM can also be seen as a classification of pixels into models (McNicholas, 2016).

The maps of the latent variables z_{ik} are useful tools to decipher the pros and cons of each model to support decision on the adaptation of tree species to climate change. In particular, we can investigate the relationships between these latent variables z_{ik} and any other variable that can be mapped. In the case of beech in France, z_{ik} indicated a discrepancy between models in that zone of the environmental space with an annual average temperature close to 12.5–13°C and a 700–1450 mm gradient in annual precipitation (left panels in Fig. 1d and 2). If we return to the geographic space, this area corresponded the west facade of France (specifically the Landes in the south-west and Périgord), the Mediterranean Arc margin, the southwestern Alps, and the Rhône corridor (Appendix S2). Strikingly, these geographic areas were precisely those where the presence of beech has been described as paradoxical in light of the species' autecology (Dugelay, 1958; Comps, 1972).

Such areas of model disagreement are particularly informative for anticipating the future distribution of beech under climate change. They highlight atypical environmental conditions under which the species is still able to persist. While European beech is generally expected to decline in its historical range across Europe (Klesse et al., 2024), focusing on these contested zones may help refine projections and shed light on the ecological mechanisms that enable local persistence.

Climatic gradients further clarify the source of these discrepancies. Across France, annual mean temperature and precipitation follow two main gradients: a broad gradient where the two variables are negatively correlated, and a secondary gradient where they are positively correlated. This second gradient was embodied in that zone of the environmental space where discrepancies between model predictions were observed (i.e., an annual average temperature close to 12.5–13°C and annual precipitation between 700–1450 mm). Coincidentally, this second gradient crossed the temperate threshold that separated the beech's presence from its absence (Appendix S1). We can therefore diagnose that the species distribution models (mainly DIGI-SDM and NBM) succeeded in capturing the variation in beech occurrence along the largest gradient of temperature and precipitation, but had more difficulty doing so along the other climatic gradient. This diagnosis will be helpful to improve the predictive performance of the models in the future.

Another interest of BEM was the production of maps of within- and between-model variance (Fig. 1b, c). With the growing number of models available to predict the same quantities, one interest of ensemble modelling is to assess the uncertainty resulting from different modelling choices (Zurell et al., 2023). The between-model variance is often interpreted as the uncertainty that results from the modelling choices (e.g., Mo et al., 2023). As in an analysis of variance, the smaller the between-model variance is compared to the within-model variance, the more confident one can be that the predictions are not related to some specific modelling choices. For beech in France, BEM indicated that there were areas where the choice of the model indeed mattered (Fig. 1d). These areas corresponded again to that climatic gradient with an annual average temperature close to 12.5–13°C and annual precipitation between 700–1450 mm. In contrast, the between-model variance of SMA or WMA never overcame their within-model variance (Appendix S1), suggesting that the former was underestimated. Thus, in contrast to BEM, SMA and WMA did not help much to understand the discrepancies in predictive performance among models.

4.3. Homogenizing the responses of models

To apply BEM, we basically need, for a given tree species, (1) a map of the observed species occurrence (to be used as a validation dataset to estimate the predictive performance of models) and (2) the maps of the output variables of each model. Such output variable can be species presence (as is the case with species distribution models). However, there can be other types of output variables (such as growth, standing volume or biomass, survival, recruitment rate, fitness, etc.), as long as such variables can be interpreted as proxies for species presence in a changing climate. Basing the BEM on the species occurrence favours those models whose output is already the species occurrence. In our case, the species distribution models DIGI-SDM and NBM, whose predictive performance was based on beech occurence, logically got the greatest weights in the BEM. Nevertheless, the logistic transformation of the output of the other models made in the BEM was consistent with the way these outputs inform the occurrence of the species. Thus, a high probability of xylem hydraulic failure indicated that the tree species could not be present, but a low probability of hydraulic failure told little about the presence of the species. Accordingly, the predictions of the Sureau-Ecos model for beech were transformed by a sharp logistic relationship but with a high uncertainty on the threshold value of the probability of hydraulic failure (Table 2; Appendix S1 and S2).

Another possible limitation of our method to homogenize model responses is related to the use of the logistic regression. Logistic regression assumes independent observations, whereas species occurrence is spatially structured. Moreover, when observations are discretized by dividing a geographic area into pixels, the coefficients of the logistic regression do depend on pixel size (Baddeley et al., 2010). This limitation can be solved by fitting an inhomogeneous Poisson process on species presence rather than a logistic regression on species occurrence (Warton and Shepherd, 2010; Baddeley et al., 2015, §9.10). In practice, little difference between the two methods is found when the pixel size is small enough. In future work, we shall propose the alternative of fitting a Poisson process within BEM.

Turning back to the more general question of tree species adaptation to climate change, what matters is to predict the site suitability for a tree species depending on climate. Species presence is one dimension of suitability. Growth or productivity is another one. It may be difficult to reconcile species occurrence and growth in a single assessment of suitability (Dolos et al., 2015). Märkel and Dolos (2017) proposed an approach where different scores of suitability, each for a different model output (presence, growth, etc.), were normalized, then averaged. Similarly, to restore symmetry between the different dimensions of suitability, one could make a separate BEM for each model output (species occurrence, species growth, etc.), then combine the different BEMs. Alternately, a multivariate BEM could also be investigated to jointly address the different model outputs in a common BEM framework. Multivariate BEM has already been used for weather or hydrologic forecast (Möller et al., 2013; Madadgar and Moradkhani, 2014). The current study can be seen as a first step towards such a multivariate approach.

CRediT authorship contribution statement

Nicolas Picard: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Nikola Besic:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Modeste Meliho:** Conceptualization, Methodology, Writing – review & editing. **Julien Sainte-Marie:** Methodology, Writing – review & editing. **Frédéric Mortier:** Methodology, Writing – review & editing. **Myriam Legay:** Funding acquisition, Writing – review & editing.

Conflict of interest statement

The authors have no conflicts of interests to declare.

Acknowledgments

This work was supported by a grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (ANR-11-LABX-0002-01, Lab of Excellence ARBRE), as well as by grants overseen by the RMT Aforce and the French region Grand Est. The authors would like to thank Christian Piedallu (INRAE, UMR Silva), Vincent Badeau (INRAE, UMR Silva), Alexandre Piboule (ONF), Céline Perrier (CNPF) and Hedi Kebli (CNPF) for constructive discussions on the subject addressed in this work. The authors are grateful to Silvio Schüler and Chakraborty Debojyoti (Austrian Research Center for Forests – BFW) for providing the data of the Sustree model.

Data availability

Script files to replicate the analysis are available in Appendix S4. Data on the occurrence of European beech in France is publicly available from the DataIFN portal at https://inventaire-forestier.ign.fr/dataifn/ (accessed 8 April 2025). Data on the Sustree model is available on the Sustree web site at https://programme2014-20.interreg-central.eu/Content.Node/SUSTREE.html (accessed 8 April 2025). The other models have been published. The repositories where their codes can be found are indicated in the publications relating to these models.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.xxxx.xxxxxx.

References

Aaheim, A., Chaturvedi, R.K., Sagadevan, A.A., 2011. Integrated modelling approaches to analysis of climate change impacts on forests and forest management. Mitig. Adapt. Strat. Glob. Change 16, 247–266. doi:10.1007/s11027-010-9254-x.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. doi:10.1111/j.1365-2664.2006.01214.x.

Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22, 42–47. doi:10.1016/j.tree.2006.09.010.

Baddeley, A., Berman, M., Fisher, N.I., Hardegen, A., Milne, R.K., Schuhmacher, D., Shah, R., Turner, R., 2010. Spatial logistic regression and change-of-support in Poisson point processes. Electron. J. Stat. 4, 1151–1201. doi:10.1214/10-EJS581.

Baddeley, A., Rubak, E., Turner, R., 2015. Spatial point patterns: methodology and applications with R. Interdisciplinary Statistics Series, Chapman & Hall/CRC press, Boca Raton.

Badeau, V., Dupouey, J.L., Cluzeau, C., Drapier, J., Le Bas, C., 2010. Climate change and the biogeography of french tree species: first results and perspectives, in: Loustau, D. (Ed.), Forests, carbon cycle and climate change, Éditions Quae, Versailles, France. pp. 231–252. URL: https://library.oapen.org/bitstream/handle/20.500.12657/24749/9782759203857.pdf.

Besic, N., Picard, N., Sainte-Marie, J., Meliho, M., Piedallu, C., Legay, M., 2024. A novel framework and a new score for the comparative analysis of forest models accounting for the impact of climate change. J. Agric. Biol. Environ. Stat. 29, 73–91. doi:10.1007/s13253-023-00557-y.

Bircher, N., Cailleret, M., Bugmann, H., 2015. The agony of choice: different empirical mortality models lead to sharply different future forest dynamics. Ecol. Appl. 25, 1303–1318. doi:10.1890/14-1462.1.

Brun, R., Reichert, P., Künsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. Water Resour. Res. 37, 1015–1030. doi:10.1029/2000WR900350.

Bullock, B.P., Boone, E.L., 2007. Deriving tree diameter distributions using Bayesian model averaging. For. Ecol. Manag. 242, 127–132. doi:10.1016/j.foreco.2007.01.024.

Carreira-Perpinán, M.A., Renals, S., 2000. Practical identifiability of finite mixtures of multivariate Bernoulli distributions. Neural Comput. 12, 141–152. doi:10.1162/089976600300015925.

Chakraborty, D., Jandl, R., Kapeller, S., Schueler, S., 2019. Disentangling the role of climate and soil on tree growth and its interaction with seed origin. Sci. Total Environ. 654, 393–401. doi:10.1016/j.scitotenv.2018.11.093.

Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S., Legay, M., Pagé, C., Thuiller, W., Viovy, N., Leadley, P., 2012. Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. Ecol. Lett. 15, 533–544. doi:10.1111/j.1461-0248.2012.01764.x.

Cochard, H., Pimont, F., Ruffault, J., Martin-StPaul, N., 2021. SurEau: a mechanistic model of plant water relations under extreme drought. Ann. For. Sci. 78, 55. doi:10.1007/s13595-021-01067-y.

Comps, B., 1972. Essai sur le déterminisme écologique du hêtre (*Fagus sylvatica* L.) en Aquitaine. Ph.D. thesis. Université de Bordeaux.

Corani, G., Mignatti, A., 2015. Robust Bayesian model averaging for the analysis of presence–absence data. Environ. Ecol. Stat. 22, 513–534. doi:10.1007/s10651-014-0308-1.

Dolos, K., Bauer, A., Albrecht, S., 2015. Site suitability for tree species: Is there a positive relation between a tree species' occurrence and its growth? Eur. J. For. Res. 134, 609–621. doi:10.1007/s10342-015-0876-0.

Dormann, C.F., Calabrese, J.M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C.M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J.J., Pollock, L.J., Reineking, B., Roberts, D.R., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Wood, S.N., Wüest, R.O., Hartig, F., 2018. Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. Ecol. Monogr. 88, 485–504. doi:10.1002/ecm.1309.

Dugelay, A., 1958. La hêtraie de la Sainte-Baume. Rev. For. Fr. 10, 1–26. doi:10.4267/2042/27370.

Fontes, L., Bontemps, J.D., Bugmann, H., Van Oijen, M., Gracia, C., Kramer, K., Lindner, M., Rötzer, T., Skovsgaard, J.P., 2010. Models for supporting forest management in a changing environment. For. Syst. 19, 8–29. doi:10.5424/fs/201019S-9315.

Fragoso, T.M., Bertoli, W., Louzada, F., 2018. Bayesian model averaging: A systematic review and conceptual classification. Int. Stat. Rev. 86, 1–28. doi:10.1111/insr.12243.

Gégout, J.C., Coudun, C., Bailly, G., Jabiol, B., 2005. EcoPlant: a forest site database linking floristic data with soil and climate variables. J. Veg. Sci. 16, 257–260. doi:10.1111/j.1654-1103.2005.tb02363.x.

Gritti, E.S., Duputie, A., Massol, F., Chuine, I., 2013. Estimating consensus and associated uncertainty between inherently different species distribution models. Method. Ecol. Evol. 4, 442–452. doi:10.1111/2041-210X.12032.

Guillemot, J., Delpierre, N., Vallet, P., François, C., Martin-StPaul, N.K., Soudani, K., Nicolas, M., Badeau, V., Dufrêne, E., 2014. Assessing the effects of management on forest growth across France: insights from a new functional–structural model. Ann. Bot. 114, 779–793. doi:10.1093/aob/mcu059.

Guisan, A., Thuiller, W., Zimmermann, N.E., 2017. Habitat suitability and distribution models: with applications in R. Ecology, Biodiversity and Conservation, Cambridge University Press, Cambridge. doi:10.1017/9781139028271.

Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T., 1999. Bayesian model averaging: a tutorial. Stat. Sci. 14, 382–417. doi:10.1214/ss/1009212519.

Javaux, B., 2011. Vulnérabilité des écosystèmes terrestres à l'azote atmosphérique : étude de faisabilité pour le développement d'un modèle de charges critiques intégrant la protection de la biodiversité végétale en France métropolitaine. Mémoire de fin d'études. AgroParisTech. Nancy.

Klesse, S., Peters, R.L., Alfaro-Sánchez, R., Badeau, V., Baittinger, C., Battipaglia, G., Bert, D., Biondi, F., Bosela, M., Budeanu, M., Cada, V., Camarero, J., Cavin, L., Claessens, H., Cretan, A.M., Cufar, K., de Luis, M., Dorado-Liñán, I., Dulamsuren, C., Espelta, J., Garamszegi, B., Grabner, M., Gricar, J., Hacket-Pain, A., Hansen, J., Hartl, C., Hevia, A., Hobi, M., Janda, P., Jump, A., Kašpar, J., Kazimirovic, M., Keren, S., Kreyling, J., Land, A., Latte, N., Lebourgeois, F., Leuschner, C., Lévesque, M., Longares, L., del Castillo, E., Menzel, A., Merela, M., Mikoláš, M., Motta, R., Muffler, L., Neycken, A., Nola, P., Panayotov, M., Petritan, A., Petritan, I., Popa, I., Prislan, P., Levanic, T., Roibu, C.C., Rubio-Cuadrado, Á., Sánchez-Salguero, R., Šamonil, P., Stajic, B., Svoboda, M., Tognetti, R., Toromani, E., Trotsiuk, V., van der Maaten, E., van der Maaten-Theunissen, M., Vannoppen, A., Vašícková, I., von Arx, G., Wilmking, M., Weigel, R., Zlatanov, T., Zang, C., Buras, A., 2024. No future growth enhancement expected at the northern edge for European beech due to continued water limitation. Global Change Biol. 30, e17546. doi:10.1111/gcb.17546.

König, L.A., Mohren, F., Schelhaas, M.J., Bugmann, H., Nabuurs, G.J., 2022. Tree regeneration in models of forest dynamics–Suitability to assess climate change impacts on European forests. For. Ecol. Manag. 520, 120390. doi:10.1016/j.foreco.2022.120390.

Li, Y., Andersen, H.E., McGaughey, R., 2008. A comparison of statistical methods for estimating forest biomass from light detection and ranging data. West. J. Appl. For. 23, 223–231. doi:10.1093/wjaf/23.4.223.

Lindner, M., Fitzgerald, J.B., Zimmermann, N.E., Reyer, C., Delzon, S., van Der Maaten, E., Schelhaas, M.J., Lasch, P., Eggers, J., van Der Maaten-Theunissen, M., Suckow, F., Psomas, A., Poulter, B., Hanewinkel, M., 2014. Climate change and European forests: what do we know, what are the uncertainties, and what are the implications for forest management? J. Environ. Manag. 146, 69–83. doi:10.1016/j.jenvman.2014.07.030.

Lu, L., Wang, H., Chhin, S., Duan, A., Zhang, J., Zhang, X., 2019. A Bayesian Model Averaging approach for modelling tree mortality in relation to site, competition and climatic factors for Chinese fir plantations. For. Ecol. Manag. 440, 169–177. doi:10.1016/j.foreco.2019.03.003.

Madadgar, S., Moradkhani, H., 2014. Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging. Water Resour. Res. 50, 9586–9603. doi:10.1002/2014WR015965.

Mahnken, M., Cailleret, M., Collalti, A., Trotta, C., Biondo, C., D'Andrea, E., Dalmonech, D., Marano, G., Mäkelä, A., Minunno, F., Peltoniemi, M., Trotsiuk, V., Nadal-Sala, D., Sabaté, S., Vallet, P., Aussenac, R., Cameron, D.R., Bohn, F.J., Grote, R., Augustynczik, A.L.D., Yousefpour, R., Huber, N., Bugmann, H., Merganicová, K., Merganic, J., Valent, P., Lasch-Born, P., Hartig, F., del Valle, I.D.V., Volkholz, J., Gutsch, M., Matteucci, G., Krejza, J., Ibrom, A., Meesenburg, H., Rötzer, T., van der Maaten-Theunissen, M., van der Maaten, E., Reyer, C.P.O., 2022. Accuracy, realism and general applicability of European forest models. Global Change Biol. 28, 6921–6943. doi:10.1111/gcb.16384.

Märkel, U., Dolos, K., 2017. Tree species site suitability as a combination of occurrence probability and growth and derivation of priority regions for climate change adaptation. Forests 8, 181. doi:10.3390/f8060181.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modelling. Divers. Distrib. 15, 59–69. doi:10.1111/j.1472-4642.2008.00491.x.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D., Hatfield, J.L., Rosenzweig, C., Aggarwal, P.K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R.F., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Kumar, S.N., Nendel, C., O'leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., White, J.W., Wolf, J., 2015. Multimodel ensembles of wheat growth: many models are better than one. Global Change Biol. 21, 911–925. doi:10.1111/gcb.12768.

Mauri, A., Girardello, M., Strona, G., Beck, P.S.A., Forzieri, G., Caudullo, G., Manca, F., Cescatti, A., 2022. EU-Trees4F, a dataset on the future distribution of European tree species. Sci. Data 9, 1–12. doi:10.1038/s41597-022-01128-5.

Mauri, A., Strona, G., San-Miguel-Ayanz, J., 2017. EU-Forest, a high-resolution tree occurrence dataset for Europe. Sci. Data 4, 1–8. doi:10.1038/sdata.2016.123.

McLachlan, G.J., Krishnan, T., 2008. The EM Algorithm and Extensions. Wiley Series in Probability and Statistics. 2 ed., John Wiley & Sons, Hoboken, NJ. doi:10.1002/9780470191613.

McNicholas, P.D., 2016. Mixture model-based classification. 1 ed., Chapman and Hall/CRC, New York. doi:10.1201/9781315373577.

Mellert, K.H., Deffner, V., Küchenhoff, H., Kölling, C., 2015. Modeling sensitivity to climate change and estimating the uncertainty of its impact: a probabilistic concept for risk assessment in forestry. Ecol. Model. 316, 211–216. doi:10.1016/j.ecolmodel.2015.08.014.

Mo, L., Zohner, C.M., Reich, P.B., Liang, J., de Miguel, S., Nabuurs, G.J., Renner, S.S., van den Hoogen, J., Araza, A., Herold, M., Mirzagholi, L., Ma, H., Averill, C., Phillips, O.L., Gamarra, J.G.P., Hordijk, I., Routh, D., Abegg, M., Adou Yao, Y.C., Alberti, G., Almeyda Zambrano, A.M., Alvarado, B.V., Alvarez-Dávila, E., Alvarez-Loayza, P., Alves, L.F., Amaral, I., Ammer, C., Antón-Fernández, C., Araujo-Murakami, A., Arroyo, L., Avitabile, V., Aymard, G.A., Baker, T.R., Balazy, R., Banki, O., Barroso, J.G., Bastian, M.L., Bastin, J.F., Birigazzi, L., Birnbaum, P., Bitariho, R., Boeckx, P., Bongers, F., Bouriaud, O., Brancalion, P.H.S., Brandl, S., Brearley, F.Q., Brienen, R., Broadbent, E.N., Bruelheide, H., Bussotti, F., Cazzolla Gatti, R., César, R.G., Cesljar, G., Chazdon, R.L., Chen, H.Y.H., Chisholm, C., Cho, H., Cienciala, E., Clark, C., Clark, D., Colletta, G.D., Coomes, D.A., Cornejo Valverde, F., Corral-Rivas, J.J., Crim, P.M., Cumming, J.R., Dayanandan, S., de Gasper, A.L., Decuyper, M., Derroire, G., DeVries, B., Djordjevic, I., Dolezal, J., Dourdain, A., Engone Obiang, N.L., Enquist, B.J., Eyre, T.J., Fandohan, A.B., Fayle, T.M., Feldpausch, T.R., Ferreira, L.V., Finér, L., Fischer, M., Fletcher, C., Frizzera, L., Gianelle, D., Glick, H.B., Harris, D.J., Hector, A., Hemp, A., Hengeveld, G., Hérault, B., Herbohn, J.L., Hillers, A., Honorio Coronado, E.N., Hui, C., Ibanez, T., Imai, N., Jagodzinski, A.M., Jaroszewicz, B., Johannsen, V.K., Joly, C.A., Jucker, T., Jung, I., Karminov, V., Kartawinata, K., Kearsley, E., Kenfack, D., Kennard, D.K., Kepfer-Rojas, S., Keppel, G., Khan, M.L., Killeen, T.J., Kim, H.S., Kitayama, K., Köhl, M., Korjus, H., Kraxner, F., Kucher, D., Laarmann, D., Lang, M., Lu, H., Lukina, N.V., Maitner, B.S., Malhi, Y., Marcon, E., Marimon, B.S., Marimon-Junior, B.H., Marshall, A.R., Martin, E.H., Meave, J.A., Melo-Cruz, O., Mendoza, C., Mendoza-Polo, I., Miscicki, S., Merow, C., Monteagudo Mendoza, A., Moreno, V.S., Mukul, S.A., Mundhenk, P., Nava-Miranda, M.G., Neill, D., Neldner, V.J., Nevenic, R.V., Ngugi, M.R., Niklaus, P.A., Oleksyn, J., Ontikov, P., Ortiz-Malavasi, E., Pan, Y., Paquette, A., Parada-Gutierrez, A., Parfenova, E.I., Park, M., Parren, M., Parthasarathy, N., Peri, P.L., Pfautsch, S., Picard, N., Piedade, M.T.F., Piotto, D., Pitman, N.C.A., Poulsen, A.D., Poulsen, J.R., Pretzsch, H., Ramirez Arevalo, F., Restrepo-Correa, Z., Rodeghiero, M., Rolim, S.G., Roopsind, A., Rovero, F., Rutishauser, E., Saikia, P., Salas-Eljatib, C., Saner, P., Schall, P., Schelhaas, M.J., Schepaschenko, D., Scherer-Lorenzen, M., Schmid, B., Schöngart, J., Searle, E.B., Seben, V., Serra-Diaz, J.M., Sheil, D., Shvidenko, A.Z., Silva-Espejo, J.E., Silveira, M., Singh, J., Sist, P., Slik, F., Sonké, B., Souza, A.F., Sterenczak, K.J., Svenning, J.C., Svoboda, M., Swanepoel, B., Targhetta, N., Tchebakova, N., ter Steege, H., Thomas, R., Tikhonova, E., Umunay, P.M., Usoltsev, V.A., Valencia, R., Valladares, F., van der Plas, F., Van Do, T., van Nuland, M.E., Vasquez, R.M., Verbeeck, H., Viana, H., Vibrans, A.C., Vieira, S., von Gadow, K., Wang, H.F., Watson, J.V., Werner, G.D.A., Wiser, S.K., Wittmann, F., Woell, H., Wortel, V., Zagt, R., Zawila-Niedzwiecki, T., Zhang, C., Zhao, X., Zhou, M., Zhu, Z.X., Zo-Bi, I.C., Gann, G.D., Crowther, T.W., 2023. Integrated global assessment of the natural forest carbon potential. Nature 624, 92–101. doi:10.1038/s41586-023-06723-z.

Möller, A., Lenkoski, A., Thorarinsdottir, T.L., 2013. Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas. Quart. J. Royal Meteorol. Soc. 139, 982–991. doi:10.1002/qj.2009.

Morin, X., Fahse, L., Jactel, H., Scherer-Lorenzen, M., García-Valdés, R., Bugmann, H., 2018. Long-term response of forest productivity to climate change is mostly driven by change in tree species composition. Sci. Rep. 8, 1–12. doi:10.1038/s41598-018-23763-y.

van Oijen, M., Reyer, C., Bohn, F., Cameron, D., Deckmyn, G., Flechsig, M., Härkönen, S., Hartig, F., Huth, A., Kiviste, A., Lasch, P., Mäkelä, A., Mette, T., Minunno, F., Rammer, W., 2013. Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. For. Ecol. Manag. 289, 255–268. doi:10.1016/j.foreco.2012.09.043.

Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P., Lees, D.C., 2006. Model-based uncertainty in species range prediction. J. Biogeogr. 33, 1704–1711. doi:10.1111/j.1365-2699.2006.01460.x.

Picard, N., Boyemba Bosela, F., Rossi, V., 2015. Reducing the error in biomass estimates strongly depends on model selection. Ann. For. Sci. 72, 811–823. doi:10.1007/s13595-014-0434-9.

Picard, N., Henry, M., Mortier, F., Trotta, C., Saint-André, L., 2012. Using Bayesian model averaging to predict tree aboveground biomass. For. Sci. 58, 15–23. doi:10.5849/forsci.10-083.

Piedallu, C., Gégout, J.C., Lebourgeois, F., Seynave, I., 2016. Soil aeration, water deficit, nitrogen availability, acidity and temperature all contribute to shaping tree species distribution in temperate forests. J. Veg. Sci. 27, 387–399. doi:10.1111/jvs.12370.

Raftery, A.E., Madigan, D., Hoeting, J.A., 1997. Bayesian model averaging for linear regression models. J. Am. Statist. Assoc. 92, 179–191. doi:10.1080/01621459.1997.10473615.

Robert, N., Vidal, C., Colin, A., Hervé, J.C., Hamza, N., Cluzeau, C., 2010. French national forest inventory, in: Tomppo, E., Gschwantner, T., Lawrence, M., McRoberts, R.E. (Eds.), National forest inventories: pathways for common reporting, Springer, Heidelberg, Germany. pp. 207–221. doi:10.1007/978-90-481-3233-1.

Ruffault, J., Pimont, F., Cochard, H., Dupuy, J.L., Martin-StPaul, N., 2022. SurEau-Ecos v2.0: a trait-based plant hydraulics model for simulations of plant water status and drought-induced mortality at the ecosystem level. Geosci. Model Dev. 15, 5593–5626. doi:10.5194/gmd-15-5593-2022.

San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), 2016. European Atlas of Forest Tree Species. Publication Office of the European Union, Luxembourg. doi:10.2788/038466.

Shalizi, C.R., 2014. Advanced data analysis from an elementary point of view. Cambridge University Press. URL: https://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/ADAfaEPoV.pdf. draft version.

Smith, R.L., Tebaldi, C., Nychka, D., Mearns, L.O., 2009. Bayesian modeling of uncertainty in ensembles of climate models. J. Am. Statist. Assoc. 104, 97–116. doi:10.1198/jasa.2009.0007.

Tao, F., Rötter, R.P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M.I., Semenov, M.A., Kersebaum, K.C., Nendel, C., Specka, X., Hoffmann, H., Ewert, F., Dambreville, A., Martre, P., Rodríguez, L., Ruiz-Ramos, M., Gaiser, T., Höhn, J.G., Salo, T., Ferrise, R., Bindi, M., Cammarano, D., Schulman, A.H., 2018. Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. Global Change Biol. 24, 1291–1307. doi:10.1111/gcb.14019.

Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. Global Change Biol. 10, 2020–2027. doi:10.1111/j.1365-2486.2004.00859.x.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD–a platform for ensemble forecasting of species distributions. Ecography 32, 369–373. doi:10.1111/j.1600-0587.2008.05742.x.

Van Oijen, M., 2017. Bayesian methods for quantifying and reducing uncertainty and error in forest models. Curr. For. Rep. 3, 269–280. doi:10.1007/s40725-017-0069-9.

Vidale, P.L., Lüthi, D., Frei, C., Seneviratne, S.I., Schär, C., 2003. Predictability and uncertainty in a regional climate model. J. Geophys. Res. Atmos. 108. doi:10.1029/2002JD002810.

Warton, D.I., Shepherd, L.C., 2010. Poisson point process models solve the "pseudo-absence problem" for presence-only data in ecology. Ann. Appl. Stat. 4, 1383–1402. doi:10.1214/10-AOAS331.

Wintle, B.A., McCarthy, M.A., Volinsky, C.T., Kavanagh, R.P., 2003. The use of Bayesian model averaging to better represent uncertainty in ecological models. Conserv. Biol. 17, 1579–1590. doi:10.1111/j.1523-1739.2003.00614.x.

Zhang, X., Srinivasan, R., Bosch, D., 2009. Calibration and uncertainty analysis of the SWAT model using genetic algorithms and bayesian model averaging. J. Hydrol. 374, 307–317. doi:10.1016/j.jhydrol.2009.06.023.

Zurell, D., Fritz, S.A., Rönnfeldt, A., Steinbauer, M.J., 2023. Predicting extinctions with species distribution models. Camb. Prisms Extinct. 1, e8. doi:10.1017/ext.2023.5.