



**HAL**  
open science

# A Novel Framework and a New Score for the Comparative Analysis of Forest Models Accounting for the Impact of Climate Change

Nikola Besic, Nicolas Picard, Julien Sainte-Marie, Modeste Meliho, Christian Piedallu, Myriam Legay

► **To cite this version:**

Nikola Besic, Nicolas Picard, Julien Sainte-Marie, Modeste Meliho, Christian Piedallu, et al.. A Novel Framework and a New Score for the Comparative Analysis of Forest Models Accounting for the Impact of Climate Change. *Journal of Agricultural, Biological, and Environmental Statistics*, 2023, 29 (1), pp.73-91. 10.1007/s13253-023-00557-y . hal-04518775

**HAL Id: hal-04518775**

**<https://hal.science/hal-04518775>**

Submitted on 24 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# A novel framework and a new score for the comparative analysis of forest models accounting for the impact of climate change

NIKOLA BESIC<sup>\*1,2</sup>, NICOLAS PICARD<sup>3</sup>, JULIEN SAINTE-MARIE<sup>1</sup>, MODESTE MELIHO<sup>1</sup>,  
CHRISTIAN PIEDALLU<sup>1</sup>, and MYRIAM LEGAY<sup>1</sup>

<sup>1</sup>Université de Lorraine, AgroParisTech, INRAE, UMR SILVA, Nancy 54000, France

<sup>2</sup>IGN, ENSG, Laboratoire d'Inventaire Forestier, Nancy 54000, France

<sup>3</sup>Groupement d'Intérêt Public (GIP) Ecofor, Paris 75116, France

\*Corresponding author: [nikola.besic@ign.fr](mailto:nikola.besic@ign.fr)

## Abstract

A broad consensus has been reached on the need to adapt the management of our forests to the context of the rapidly changing climate, which resulted in the development of numerous models capable of simulating the impact of the climate change on the forest. The primary goal of this specific endeavor is to propose a novel framework of comparative analysis which could lead to the unique and universal description and mapping of these models. This framework is based on the reduction of the model output to the relatively simplistic information about the presence of the tree species suitable for the forest management i.e. - a binary classifier, making it comparable with the largely available tree presence observations. The framework we propose comes along with a new score, based on the joint use of the Principal Component Analysis and the Co-inertia Analysis, which evaluates the model vis-à-vis the corresponding observations with the focus on its phase space dynamics i.e. its dependence on external environmental variables, rather than its spatial precision. The pertinence of the proposed multi-scale approach, suitable for the multi-scale analysis, is demonstrated by conjointly using prototype binary classifiers, designed for this purpose, and two different examples of binary classifiers used in the forest management - climate-dependent tree species distribution models. This work has the ambition to serve as the basis for a potential combination of different models at different spatial scales in order to improve the decision making process in the forest management.

**Keywords:** comparative analysis, score, adaptation, forest, climate change.

## 1 Introduction

Climate change has a major impact on the functioning of forest ecosystems, whether directly through the physiological response of trees to changing climatic parameters, or indirectly through the impact of this change on the biotic and abiotic components

interacting with the trees (Parmesan et al., 2022). This impact is manifested by droughts and heat episodes having consequences on the growth, mortality and recruitment rates of tree species, as well as on their genetic structure; or by the increased risk of forest fires, storm damage, insect pests and diseases, which finally results in niche distribution shifts (Dale et al., 2001; Bréda et al., 2006; Taccoen et al., 2019; Frank et al., 2015).

The climatic projections as well as the ongoing manifestations of the changing climate we are witnessing (IPCC, 2021) remind us that aside the urgency of mitigating further changes, there is also the absolute and imminent necessity for the well-timed adaptation to the consequences, which are to a certain extent inevitable due to the inertia of the system. Adapting forest to the changing climate is of major importance for the preservation of its multiple ecosystem services (Keenan et al., 2015). Furthermore, due to the crucial role of forests as both carbon sinks and sources of renewable materials and energy, the adaptation efforts in the domain of forest management are of particular importance even for the climate change mitigation (Ciais et al., 2005; Nabuurs et al., 2022).

In order to adapt forest to climate change, we ought to be able to simulate the response of forests to changing climate. There are different kinds of climate sensitive models for forest and vegetation, depending on the modeling approach, the type of process modeled, the data used, the response variable predicted, and the level of description of the forest (Porté and Bartelink, 2002). A common categorization distinguishes between statistical models (which look for correlations between forest characteristics and covariates), process-based models (which focus on the processes behind the functioning of the forest ecosystem), and mixed models (which combine the two). The first category includes species distribution models (Guisan and Thuiller, 2005, e.g. model BIOMOD, Thuiller et al., 2009) and growth models (e.g. model STASH, Sykes et al., 1996). The second category includes ecophysiological models (e.g. model CASTANEA, Dufrêne et al., 2005) and dynamic global vegetation models (e.g. model Orchidee, Krinner et al., 2005).

Given the variety of climate-dependent forest models currently available, a pending question is the selection of the models to use for decision-making. A straightforward approach would consist in comparing different models based on their prediction performance and to declare a “winner model” which should be used for decision-making. Another approach would consist in combining the different models for decision-making based on consensus prediction from the ensemble of available models (Bhat et al., 2011). In line with the latter approach, this study contributes to making the best of all available forest models in order to support relevant decision-making in forest management for the adaptation of forests to climate change (Yousefpour et al., 2017). Our goal is to provide tools to assess and open the way to the potential interpretation of differences between model predictions at different spatial scales, in order to help ensuring that the use of any model is well suited to the particular circumstances and issues addressed (Segurado and Araújo, 2004).

Our model assessment relies on a new comparative analysis framework that consists in back-projecting both model predictions and independent observations into the space of the covariates of the model. Indeed, the different models that predict the climate dependence of tree species operate in different spaces, i.e. have different “output” variables and covariates (including climatic variables). This makes their direct inter-comparison difficult or even impossible. We therefore envisaged an approach where the output of every model is reduced to the binary variable reflecting the potential presence of the species – a variable nearly equivalent to the information that can be derived from a national forest inventory or from a floristic survey. In other words, in order to be considered

in the proposed framework a model has to be transformed into a binary classifier. For each model, predicted species presence is then back-projected into the space of model covariates. Ditto, observed species presence originating from a national forest inventory is projected in the same space of the model covariates. The two clouds of presence points are finally compared using a similarity score. We established this score using principal component analysis and co-inertia analysis at different spatial scales. This score quantifies the mismatch between the cloud of points originating from model predictions and that originating from observations and, more importantly, could give us a key to understanding this mismatch in terms of model covariates, notably the climatic ones. The evaluation of the presented idea aims to demonstrate the advantages of such an approach having a focus on the phase space dynamics with respect to the scores concentrated on the spatial precision, the latter being commonly employed in the field of forest modeling and not being able to properly separate the ecological potential and the impact of the human management.

The article is organized as follows: in Section 2 we explain the challenging context of the comparative analysis between forest models accounting for the climatic impact, while simultaneously defining the basis of the mathematical structure of the article. In Section 3 we propose a novel framework that is the first part of our response to the challenge raised in Section 2. In Section 4 we elaborate the concept of a new score. Section 5 illustrates the proposed methods with real and synthesized data, and also applies them to a real case consisting of two climate-dependent species distribution models. The results are confronted to those obtained with conventional model comparison tools. Section 6 concludes the article and briefly evokes the perspectives of the work.

## 2 A common set of observations to compare different models

Let  $\mathbf{y}_{st'} = g(\mathbf{y}_{.t}, \mathbf{y}_{st}^e, \theta)$  be a dynamical system that predicts the state  $\mathbf{y}$  of a system at location  $s$  and time  $t'$  from its state at time  $t < t'$  and external variables  $\mathbf{y}^e$  whose dynamics are not determined by the system in question. The dot in notation  $\mathbf{y}_{.t}$  represents the set of all locations:  $\mathbf{y}_{.t} = (\mathbf{y}_{st})_{s \in \mathcal{S}}$  where  $\mathcal{S}$  is the study region. The function  $g$  that defines the dynamics also depends on parameters  $\theta$ . We hereafter suppose that parameters  $\theta$  are known, i.e. we do not consider here the issue of model calibration that consists in estimating  $\theta$  from data. In the forestry context of interest here, state variables  $\mathbf{y}$  characterize a forest stand and consist of variables such as tree density, basal area, canopy height, volume, biomass, species diversity, species abundances, species frequencies, genetic structure, economic value. External variables  $\mathbf{y}^e$  characterize the environment that influences the forest stand and consist of variables such as climatic variables, soil variables, topographical variables.

We are specifically interested here in assessing the response of a forest stand to climate change, i.e. assessing the change in  $\mathbf{y}$  when the climatic variables included in  $\mathbf{y}^e$  vary. The dynamical system defined by  $g$  may not only include the dependence of the forest dynamics on its environment (niche dependence) but also internal processes such as competition, dispersal of seeds, etc. (through the dependence of  $\mathbf{y}_{st'}$  on  $\mathbf{y}_{.t}$ ). The dynamical system may thus include spatial effects due to dispersal limitations and possibly differentiate the fundamental niche of a species from its realized niche, depending on the past history of the species.

Let  $\mathbf{z}_{st'} = h(\mathbf{z}_{.t}, \mathbf{z}_{st}^e, \theta^*)$  be a second dynamical system that predicts the state of a forest stand. We do not make any assumption about the phase spaces to which  $\mathbf{y}$  and  $\mathbf{z}$  belong, nor about the spaces to which  $\mathbf{y}^e$  and  $\mathbf{z}^e$  belong: these spaces may coincide, or share a common subspace, or be completely disjointed. For instance, when  $\mathbf{y}$  and  $\mathbf{z}$  give both the probability of presence of a subject tree species, the two phase spaces are the same. When  $\mathbf{y}$  gives the probability of presence of the subject species and  $\mathbf{z}$  its growth rate, the two phase spaces are disjointed. When  $\mathbf{y}$  gives the growth rate and genetic diversity of the subject species and  $\mathbf{z}$  gives its growth rate only, the two phase spaces share a common subspace (in this specific example, the phase space of  $\mathbf{z}$  is embedded in that of  $\mathbf{y}$ ).

The question addressed in this paper is to comparatively assess the performance of  $g$  and  $h$  to predict a subject tree species response to climate change. We assume that a dataset of observations at a given time  $t_1$  is available:  $\{(\mathbf{x}_{st_1}, \mathbf{x}_{st_1}^e) : s \in \mathcal{A}\}$ , where  $\mathcal{A}$  is the geographical region where observations have been made. We assume that the frequency of the subject species, denoted  $f$ , is among the observed characteristics of the forest stand, i.e.  $f$  is one element of vector  $\mathbf{x}$ . We also assume that all external variables  $\mathbf{y}^e$  and  $\mathbf{z}^e$  that are required to run models  $g$  and  $h$  are included in the observed variables  $\mathbf{x}^e$ , i.e.  $\mathbf{y}^e$  is a subvector of  $\mathbf{x}^e$ , and so is  $\mathbf{z}^e$ . Hence, starting from some initial conditions, we assume that models  $g$  and  $h$  allow us to predict  $\mathbf{y}_{st_1}$  and  $\mathbf{z}_{st_1}$  for all  $s \in \mathcal{A}$ . The main difficulty to overcome to compare models  $g$  and  $h$  is that the phase space  $\mathcal{Y}$  of  $\mathbf{y}$ , the phase space  $\mathcal{Z}$  of  $\mathbf{z}$  and the phase space  $\mathcal{X}$  of  $\mathbf{x}$ , may overlap or be completely disjointed. In other words, the variables that are predicted by any model may completely differ from those predicted by the other model or from observed variables. Therefore, model comparison is not simply a matter of comparing predicted variables.

If the phase spaces share a common subspace (*a fortiori* if they are embedded), i.e. if  $\mathcal{X} \cap \mathcal{Y} \cap \mathcal{Z} = \mathcal{F}$  with  $\mathcal{F} \neq \emptyset$ , the models may be compared on the basis on this common set of variables in  $\mathcal{F}$ . If  $\alpha$  denotes the projection operator that associates to any vector  $\mathbf{y} \in \mathcal{Y}$  its subvector  $\alpha(\mathbf{y}) \in \mathcal{F}$  and similarly  $\beta$  projects  $\mathcal{Z}$  onto  $\mathcal{F}$  and  $\gamma$  projects  $\mathcal{X}$  onto  $\mathcal{F}$ , models may be compared based on their capacity to predict observations  $\gamma(\mathbf{x})$ . For any model, comparing its predictions to observations corresponds to model validation. Let  $u : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}^+$  be a goodness-of-fit metric. The ‘‘best’’ model among  $g$  and  $h$  is the one with the greatest goodness-of-fit among  $u[\alpha(\mathbf{y}_{.t_1}), \gamma(\mathbf{x}_{.t_1})]$  and  $u[\beta(\mathbf{z}_{.t_1}), \gamma(\mathbf{x}_{.t_1})]$  (e.g. Zurell et al., 2016).

However, even in the case where phase spaces share a common subspace, assessing models based on this common subspace may not be the most relevant strategy. In the most general case where variables can be of any kind, the largest subset of variables that are in common to all models and observations may not be the most relevant one to assess the species response to climate change. Imagine for instance a species that has a very homogeneous genetic structure and whose dynamics hardly depend on its genetic diversity. Imagine that two models for this species are to be compared: a model  $g$  that predicts the species growth and genetic diversity, and a model  $h$  that predicts the species mortality and genetic diversity. Even if growth and mortality can be relevant variables to assess the species response to climate change, the variable in common to the two models is the genetic diversity that is an irrelevant variable in this theoretical case. Hence the projections of  $\mathcal{Z}$  and  $\mathcal{Y}$  onto  $\mathcal{F}$  may cause relevant information to be lost.

Moreover, even when the models have the same phase space, comparing them on the basis of a goodness-of-fit metric tells us nothing about the reasons why the predictions of the two models differ or not. To develop tools for decision making, beyond comparing

the prediction performance of the models, it is important to map the differences between their predictions in the space of external variables.

As explained in the subsequent sections, the issue of possibly disjointed phase spaces and the issue of mapping model predictions are both solved using the dataset of observations that is common to all models (Fig. 1). In the next sections (from Section 3 to 5.2), we focus on one of the models and the dataset of observations. In other words, we focus on the upper part of the diagram in Fig. 1 that comprises model  $g$  and observations, or equivalently on its lower part. We shall revert to the question of inter-model comparison in Section 5.3.

### 3 A new framework for model evaluation

In order to account for the most general case where the phase spaces are possibly disjointed, we here propose an approach that relies on the predicted species occurrences  $f_{t_1}^g$ , where occurrence is a binary variable indicating species presence ( $f_{t_1}^g = 1$ ) or absence ( $f_{t_1}^g = 0$ ). Unless species occurrence is already part of the model output variables (i.e.  $f_{t_1}^g \subset \mathbf{y}_{t_1}$ ), it is predicted from the output variables  $\mathbf{y}_{t_1}$  of the model and the observed species occurrence  $f_{t_1}^o \subset \mathbf{x}_{t_1}$ . Mapping differences between model predictions then consists in comparing the subspaces of the external variables that correspond to species presence (Fig. 1). Using the output variables of model  $g$  as predictors of the species occurrence allows us to compare models with disjointed phase spaces in a common space. The species occurrence plays a particular role that differentiates it from the other observed variables. It is assumed to be a relevant indicator of a species response to climate change.

It is also implicitly assumed that the complex response of the species to climate change is already captured by the model  $g$ , so that the relationship between the species occurrence and  $\mathbf{y}_{t_1}$  is more straightforward than the relationship between the species occurrence and climate variables. For instance, Cheaib et al. (2012) used an output variable of the model and a threshold value to define a binary classifier that predicts the species occurrence. The predictive performance of this classifier was assessed using the true skill statistic (TSS), and the estimate of the threshold value was the value that maximized the TSS. Hence, models were compared on the basis of their maximum TSS. An alternative method would consist in fitting a model that predicts the species occurrence, e.g. using logistic regression or a machine learning method. Models can then be directly compared on the basis of their goodness-of-fit to observed species occurrences.

Let  $p_{t_1} = \Pr(f_{t_1}^g = 1 | \mathbf{y}_{t_1})$  be the probability of presence of the subject species as predicted from the predictions of model  $g$ . To bring all models on a common ground, we move back to the subspace of external variables and compare two clouds of points: the one formed by the external variables where the species is predicted to be present ( $\mathbf{y}^e | p_{t_1} > u$ , where  $u$  is a given threshold value), and the one formed by the same external variables ( $\mathbf{y}^e$  is a subvector of  $\mathbf{x}^e$ ) where the species is observed to be present ( $\mathbf{y}^e | f_{t_1}^o = 1$ ). In Fig 1, we do not limit the comparison to the time  $t_1$  in order to make the framework eligible for the evaluation of external variables that have temporal rather than spatial influence (e.g. CO<sub>2</sub> concentration).

Spatial coordinates may be among the external variables  $\mathbf{y}^e$  that define the space where species presence is projected. If these external variables are exactly the spatial coordinates, then our framework boils down to comparing the geographical maps of predicted and observed species presence, which is a common approach to understanding

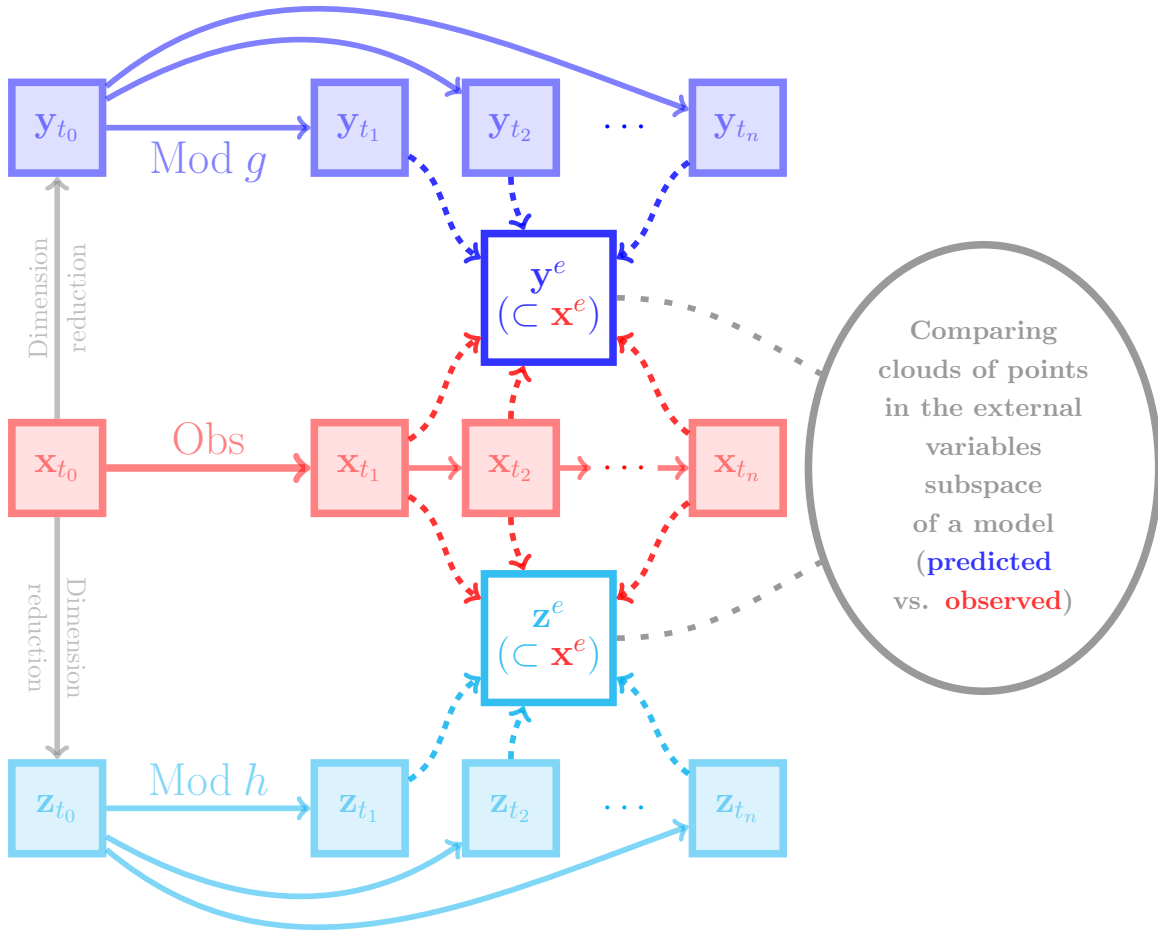


Figure 1: Schematic illustration of a new framework.

differences between model predictions. Our framework can thus be seen as an extension of the common approach to the mapping of species presence in a space that may also include environmental covariates.

We shall illustrate the framework with the global prototype land surface binary classifier model designed specifically for this purpose (Fig. 2). Its input are three Gaussian random fields simulated with Müller et al. (2022), having the same variance but different correlation lengths (Fig. 2a–c). These three variables form the vector  $\mathbf{y}_{t_1}^e$ . The output of the prototype model is the binary information, corresponding to the modeled species presence, obtained as the thresholded product of three random fields (Fig. 2d). It thus corresponds to a simple situation where species occurrence does not have to be predicted from the state variables  $\mathbf{y}_{t_1}$  of the model.

Projecting the predicted species presence in the three-dimensional space formed by the external variables  $\mathbf{y}^e$  gives the first cloud of points (Fig. 2e). The reference dataset, representing  $f_{t_1}^o \subset \mathbf{x}_{t_1}$  and corresponding to the observed species presence, is obtained by adding a Gaussian noise to the product of three stochastic random fields and thresholding it with a higher value than the prototype model (Fig. 2f). Because external variables are also mapped at this scale ( $\mathbf{y}^e \subset \mathbf{x}^e$ ), we can equally project the observed species presence

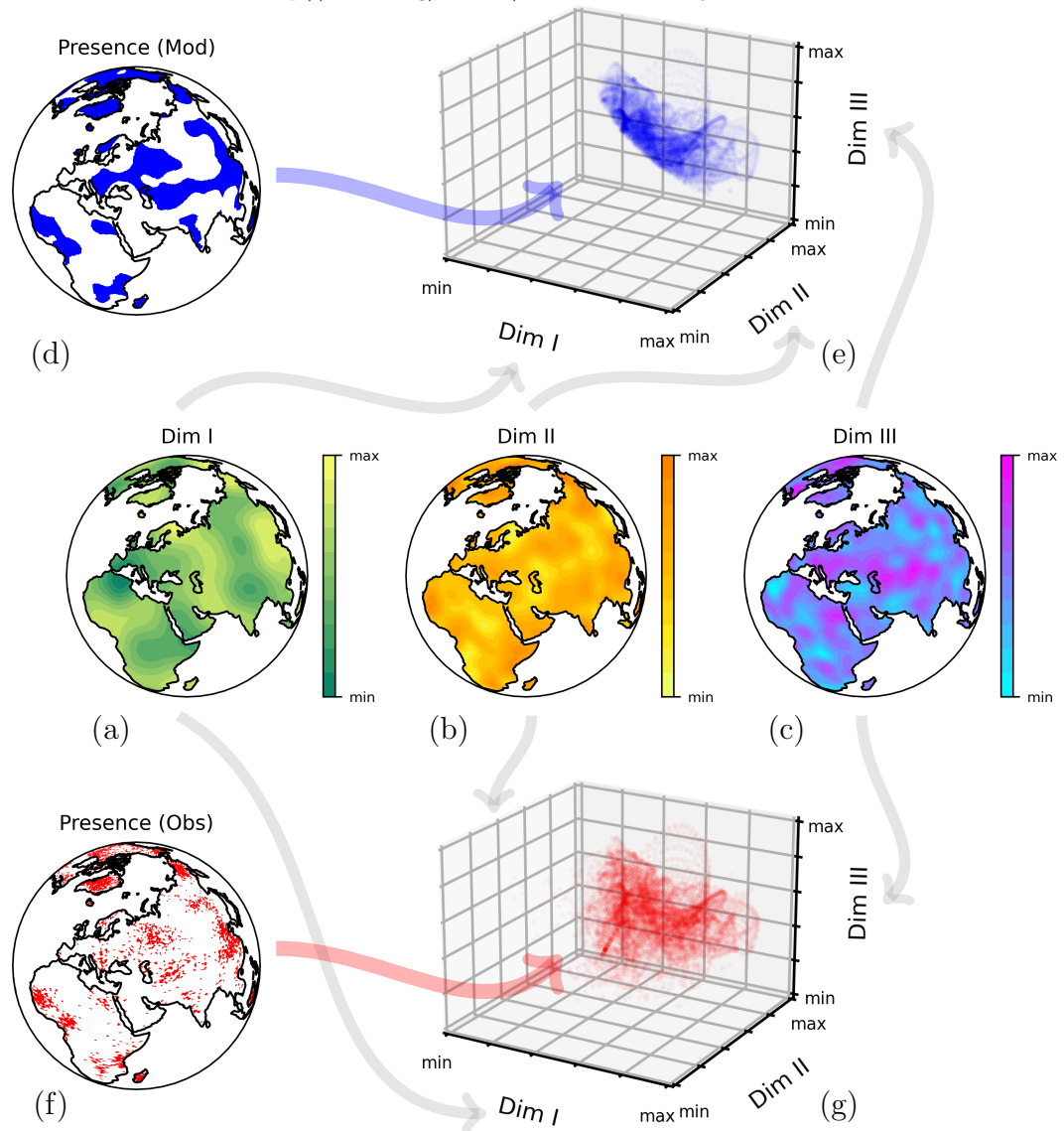


Figure 2: Illustration of the new framework using the prototype binary classifier model designed specifically for this purpose - part I: (a)–(c) external variables, (d) map of the predicted species presence, (e) cloud of points of the predicted species presence in the space formed by the prototype model external variables, (f) map of observed species presence, (g) cloud of points of the observed species presence in the space formed by the prototype model external variables.

in the space formed by the external variables, which gives us the second cloud of presence points (Fig. 2g). The model evaluation then boils down to estimating the similarity between these two clouds of points. This approach is more informative than a goodness-of-fit metric based on a  $2 \times 2$  contingency table with counts of true positives, true negatives, false positives and false negatives, which ignores the structure of predictions and observations in the geographical and environmental spaces.



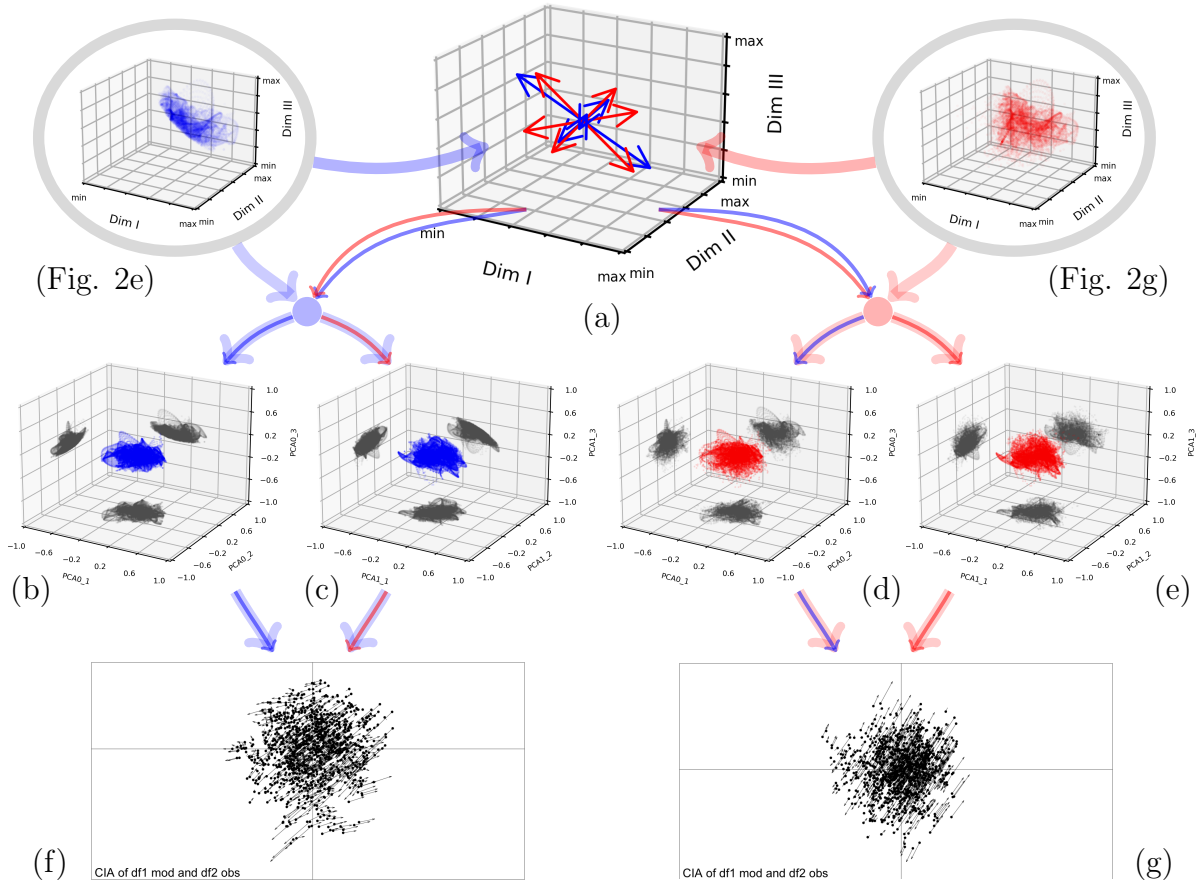


Figure 3: Illustration of the new framework - part II: (a) principal components of the clouds of points presented in Fig. 2e (model predictions, blue arrows) and Fig. 2g (observations, red arrows), (b) predicted presences projected in the space of the PCA of model predictions, (c) predicted presences projected in the space of the PCA of observations, (d) observed presences projected in the space of the PCA of model predictions, (e) observed presences projected in the space of the PCA of observations, (f) predicted presences projected in the space of the co-inertia between model predictions and observations, (g) observed presences projected in the space of the co-inertia between model predictions and observations.

## 4 A new score for point cloud similarity

The framework introduced in the previous section, which reduces the model evaluation to the comparison of clouds of points, must be completed by a similarity index between clouds of points, capable of summarising the different dimensions of the phase space into a one-dimensional metric. We here propose a new score which is to a large extent based on the combination of the principal component analysis (PCA) (Jolliffe, 2002) and coinertia analysis (CIA) (Dolédec and Chessel, 1994). When employing a term "score" we intend a quantity which informs on the similarity between the model and the observations.

PCA is a commonly used statistical technique depicting an orthogonal linear transformation which basically transforms the original space of the analyzed data into a new space constructed around the direction of maximal variance (Papoulis, 1991). The direction of maximal variance corresponds to the first eigenvector of the multivariate data covariance matrix ( $\mathbf{C}$ ), and all the subsequent axes are orthogonal and descending in

terms of explained variance. The principal axes of the predicted and observed clouds of points give us a first clue of the way external (in this case, climatic) variables drive model predictions on the one hand (Fig. 3a, blue arrows), and observations on the other hand (Fig. 3a, red arrows). Let  $D$  be the number of external variables of the model. Let  $N_g$  be the number of predicted presences according to model  $g$ , and let  $\mathbf{Y}^g$  be the matrix with  $N_g$  rows and  $D$  columns that contains the external variables  $\mathbf{y}^e | f_{t_1}^g = 1$  at all predicted species presences. Ditto, let  $N_o$  be the number of observed presences and  $\mathbf{Y}^o$  the  $N_o \times D$  matrix that contains the external variables  $\mathbf{y}^e | f_{t_1}^o = 1$  at all observed species presences. The principal axes  $\mathbf{Q}^g$  and  $\mathbf{Q}^o$  of the model predictions and observations are  $D \times D$  matrices that follow from:

$$\mathbf{C}^g \mathbf{Q}^g = \mathbf{Q}^g \mathbf{\Lambda}^g, \text{ where } \mathbf{C}^g = \frac{1}{N_g} (\mathbf{Y}^{g\top} \mathbf{Y}^g), \quad (1)$$

$$\mathbf{C}^o \mathbf{Q}^o = \mathbf{Q}^o \mathbf{\Lambda}^o, \text{ where } \mathbf{C}^o = \frac{1}{N_o} (\mathbf{Y}^{o\top} \mathbf{Y}^o), \quad (2)$$

with  $\mathbf{\Lambda}^g$  and  $\mathbf{\Lambda}^o$  being the diagonal matrices containing eigenvalues and designing the proportion of variance explained by the component obtained using the corresponding column of  $\mathbf{Q}$ , i.e. by the corresponding principal component. The number of principal components is equal to the number  $D$  of external variables, but to have a score whose formal conception is independent of this dimensionality which varies from model to model, we propose to keep, if possible, only the first three axes  $\mathbf{Q}_{[D \times 3]}$  - considered by the authors to be the minimal dimensionality in order for a model to be envisaged for applications this methodology is destined to (e.g. defining public policies). We specified if possible because if the substantial part of the variance is not explained by the first three axes, keeping more axes should be envisaged, even if that could mean that the score formal conception differs between different models involved in the comparison. The score we are looking for is related to the transformation that aligns the principal axes  $\mathbf{Q}_{[D \times 3]}^g$  and  $\mathbf{Q}_{[D \times 3]}^o$  (depending on the angle between them), which leads us to the co-inertia analysis.

While there are many statistical ways to measure the relation between two variables, there are fewer solutions available to measure the relation between two matrices like  $\mathbf{Q}_{[D \times 3]}^g$  and  $\mathbf{Q}_{[D \times 3]}^o$  (Indahl et al., 2018). Co-inertia analysis is a statistical multivariate method which, based on inertia as a measure of data variability, can be used to measure the concordance between two datasets (Dray et al., 2003a,b). Co-inertia analysis requires to have coupled set of points, i.e. the same points in two spaces, meaning that we have to project predicted presences  $\mathbf{Y}^g$  or observed presences  $\mathbf{Y}^o$  onto  $\mathbf{Q}_{[D \times 3]}^g$  or  $\mathbf{Q}_{[D \times 3]}^o$  bases:

$$\mathbf{Y}_{[N_g \times 3]}^{gg} = \mathbf{Y}_{[N_g \times D]}^g \mathbf{Q}_{[D \times 3]}^g, \quad (3)$$

$$\mathbf{Y}_{[N_g \times 3]}^{go} = \mathbf{Y}_{[N_g \times D]}^g \mathbf{Q}_{[D \times 3]}^o, \quad (4)$$

$$\mathbf{Y}_{[N_o \times 3]}^{og} = \mathbf{Y}_{[N_o \times D]}^o \mathbf{Q}_{[D \times 3]}^g, \quad (5)$$

$$\mathbf{Y}_{[N_o \times 3]}^{oo} = \mathbf{Y}_{[N_o \times D]}^o \mathbf{Q}_{[D \times 3]}^o, \quad (6)$$

with  $\mathbf{Y}^{gg}$ ,  $\mathbf{Y}^{go}$ ,  $\mathbf{Y}^{og}$  and  $\mathbf{Y}^{oo}$  being illustrated in Fig. 3b, 3c, 3d, and 3e respectively. Once we have the same sites in two hyperspaces (for example the predicted presences  $\mathbf{Y}^{gg}$  and  $\mathbf{Y}^{go}$ ), the co-inertia analysis maximizes the square covariance between the projection of  $\mathbf{Y}^{gg}$  on the vector maximizing its proper inertia and the projection  $\mathbf{Y}^{go}$  on its own vector corresponding to maximal inertia, meaning that the co-inertia between two hyperspaces is the sum of squares of the covariances between all the variable pairs (Fig. 3f). Our score

is finally defined as the normalized co-inertia or RV-coefficient, based either on predicted presences:

$$RV_g = \frac{\text{Co-inertia}(\mathbf{Y}^{gg}, \mathbf{Y}^{go})}{\sqrt{\text{Inertia}(\mathbf{Y}^{gg}) \text{Inertia}(\mathbf{Y}^{go})}}, \quad (7)$$

or on observed presences (Fig. 3g):

$$RV_o = \frac{\text{Co-inertia}(\mathbf{Y}^{og}, \mathbf{Y}^{oo})}{\sqrt{\text{Inertia}(\mathbf{Y}^{og}) \text{Inertia}(\mathbf{Y}^{oo})}}. \quad (8)$$

It is important to mention here that values  $N_g$  and  $N_o$  are reduced through the random sampling to the same value (sample size) before calculating  $RV_g$  or  $RV_o$  coefficients.

These two coefficients which are from now considered as the two facets of the new score are proportional to the illustrated distance between the points from two clouds (or one cloud in two spaces) we can observe in Fig. 3f and 3g, and represent finally the measure of similarity between the cloud of modeled presences and the cloud of observed presences.

It is significant to state that these two coefficients are suited to depict the similarity evoked above in case of a deterministic model (output does not change if the input does not change either), as it is the case with the vast majority of models in three target groups evoked in the introduction: statistical models, process-based models and mixed models. Considering stochastic models (output can change even if the input does not change) would require a certain adaptation of the presented framework-score couple.

## 5 Application to climate-dependent forest models

To assess the relevance of the proposed framework and score, we now apply them to two different examples of binary classifiers used in the forest management - climate dependent tree species distribution models, as well as to the prototype binary classifier which is specifically designed for this section. The results thus obtained are compared with those obtained with two commonly used goodness-of-fit scores, namely Cohen's kappa coefficient (Cohen, 1960) and the True Skill Statistic (TSS) (Alluue et al., 2006).

### 5.1 Multi-scale comparison of scores with IKS model

IKS is a tree species distribution model (RMT-Aforce, 2022), resembling quite a bit the prototype binary classifier designed for the illustration of the framework and the score (Sections 3 and 4). Namely, the IKS output variable is the species occurrence, which is predicted from three external variables only: annual water deficit ( $mm$ ), annual sum of degree-days ( $^{\circ}C$ ) and minimum annual temperature ( $^{\circ}C$ ), or more precisely the normal temperature of the coldest month. Given the difference in units, all the inputs are normalized to the range  $[0, 1]$  through the "MinMax" scaling -  $y_{scaled}^e = (y^e - y_{min}^e) / (y_{max}^e - y_{min}^e)$ .

It can be used to model the presence of 61 difference species - here we apply it to European beech (*Fagus sylvatica*). Moreover, the French national forest inventory provided us with observations on species occurrence which can be used as a reference dataset. The species distribution model IKS for European beech can be analyzed at the scale of France, but by compiling observations from the forest inventories of the different European countries (Mauri et al., 2017) or by splitting the observations of the French

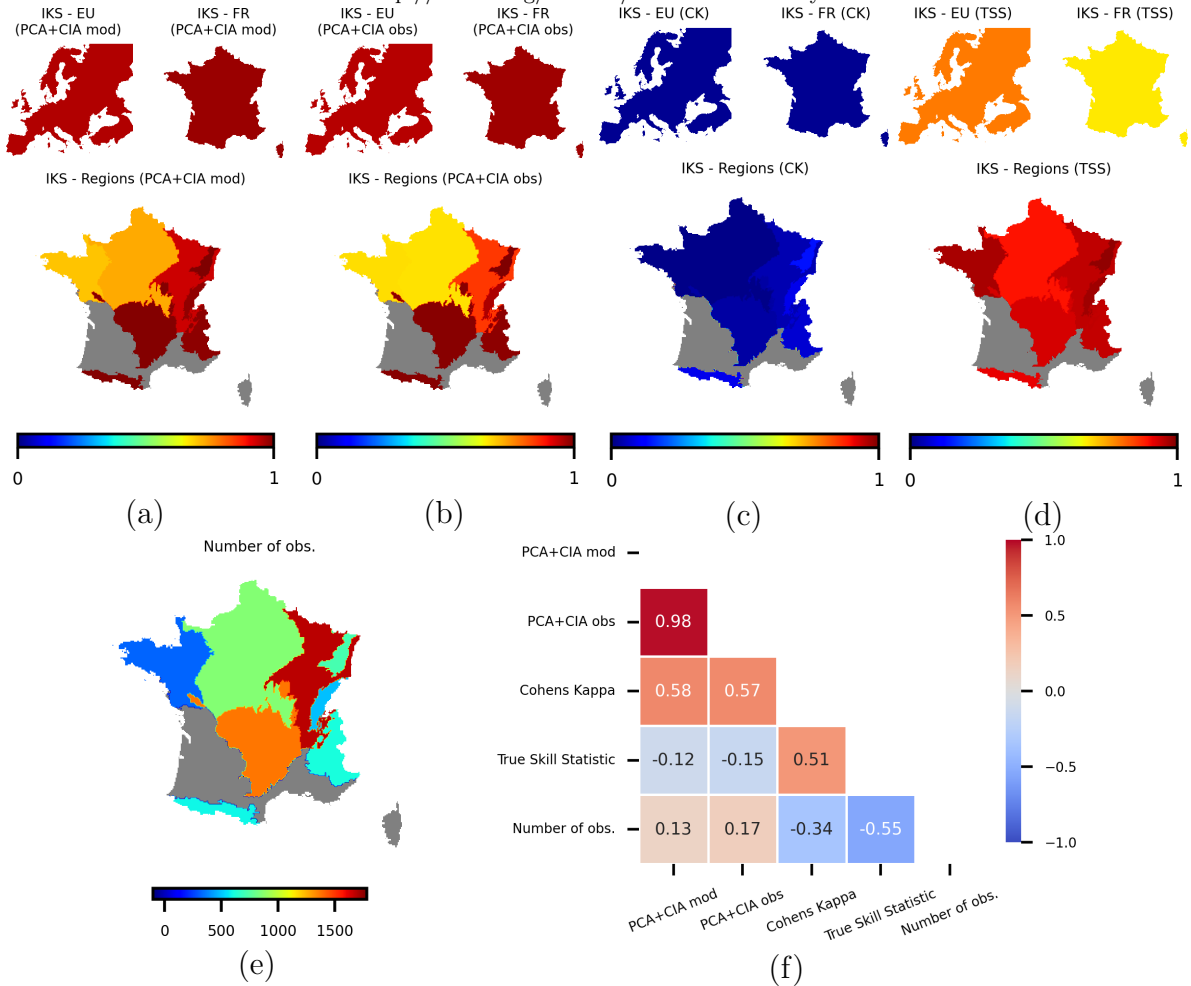


Figure 4: IKS model evaluation for European beech at three scales (Europe, France, French ecological regions): (a)  $RV_g$  coefficient of the proposed score, (b)  $RV_o$  coefficient of the proposed score, (c) Cohen's kappa coefficient (CK), (d) True Skill Statistic (TSS), (e) number of observations by region, (f) Pearson correlation matrix between the different scores at the regional scale and number of observations, where correlation is computed across different scales and regions.

forest inventory across ecological regions, it can also be analyzed at the continental (i.e. over the entire range of European beech) and regional levels.

The two coefficients  $RV_g$  and  $RV_o$  that compose our score have a high value at the continental scale (Fig. 4a and b). Because the IKS model was calibrated using the European dataset, this high score is expected and reflects the amplitude of the gradients of climatic variables and the corresponding amplitude of the beech response at the European scale. High values of the RV coefficients were also obtained at the scale of France due to the good environmental representativity of France for this tree species. However, when scaling down to the regional level, the  $RV_g$  and  $RV_o$  values vary with ecological regions, showing that the model calibrated at the continental scale may not explain properly the phase space dynamics locally. The RV values at the regional scale are not significantly correlated with the number observations by region (Fig. 4e and 4f).

Cohen's kappa coefficient and the True Skill Statistic (TSS) bring quite different results (Fig. 4c and d). Cohen's kappa coefficient is weak at all scales, including the European one despite the calibration of the model at this scale, indicating its low utility

within the proposed framework. TSS shows an interesting behavior which is somehow opposite to the one observed with the new score in terms of scale dependence. Namely, the statistic is rather good at the European scale, slightly worse at the French one, and very good at the regional scale, varying to a certain extent among regions. This can be partly explained by the dependence of the TSS on the ratio between presences and absences in the observations (prevalence), which can become important with the decrease in the number of observations (Somodi et al., 2017).

## 5.2 Scores comparison with a prototype binary classifier model

In order to demonstrate the difference between the proposed score and conventional ones in a controlled setting, we also relied on a prototype binary classifier model. This prototype is slightly different from the one conceived for the purpose of illustration of the approach, because in this controlled setting, the true probability of presence of the species is envisaged as a weighted product of the species ecological potential and of a human modifier. The ecological potential is defined as the product of three environmental variables that are conceptualized as three simple spatial gradients (horizontal, vertical and diagonal; Fig. 5a–c). The human modifier corresponds to an area where trees are planted by humans. It is equal to one in a square located at the maximum of the ecological potential and zero elsewhere. Observations of species occurrence are simulated according to a binomial law using the true probability of presence and four different weightings of the ecological potential and of the human modifier: in scenario I, the human modifier has a null weight and the true probability of presence of the species is entirely determined by its ecological potential (Fig. 5d); scenarios II to IV correspond to an increasing weight of the human modifier (Fig. 5e–g).

The model analyzed is similar to IKS in the sense that it is based on the three environmental variables only and provides the species presence as output. The model is supposed to exactly capture the ecological potential of the species but ignores the human modifier. In other words, scenarios I to IV correspond to an increasing divergence between the modeled probability of presence and the true one. If the three environmental variables are denoted as  $y_1^e$ ,  $y_2^e$ ,  $y_3^e$  (no need for scaling), the model prediction thus is:  $f^g = \mathbf{1}(y_1^e \times y_2^e \times y_3^e > u)$ , where  $\mathbf{1}(p)$  is the indicator function of proposition  $p$  ( $= 1$  if  $p$  is true and  $0$  otherwise) and  $u$  is a threshold. Three model realizations corresponding to three increasing value of  $u$  are considered (see Fig. 5h, j and l for the maps of their predictions).

Conventional scores like Cohen's kappa and TSS are found to be sensitive to the threshold value  $u$  (compare Fig. 5i, k and m for a given scenario I–IV). In contrast, the two RV coefficients of the proposed score are almost independent of the threshold value  $u$ . This result confirms that conventional scores are sensitive to the spatial distribution of the predicted presence (the red area in Fig. 5h, j, l), whereas the proposed ones, which focus on the correlation between environmental variables where the species is present, are less sensitive to the correspondence between the map of predicted presences and the map of observed ones. Area under the Curve (AUC) (Hanley and McNeil, 1982), the score we adjoin here to the analysis shows a similar behavior, which is neither very surprising considering that this score can only compare the non-thresholded probability of predicted presence, implying its independence to the threshold value. The AUC score was not employed in the previous chapter for precisely this reason, in view of the fact that the IKS model does not provide the probability of presence.

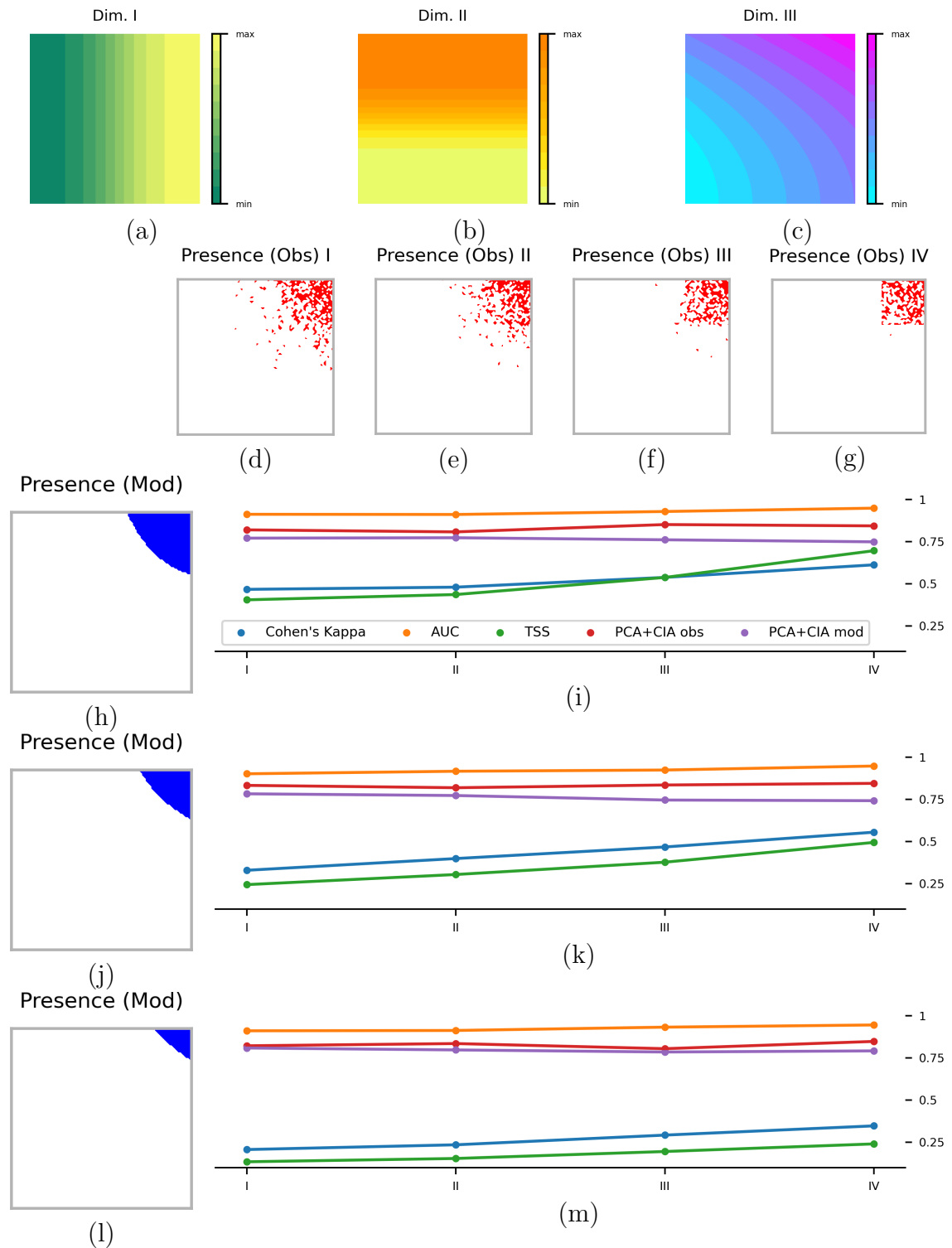


Figure 5: Evaluation of a prototype binary classifier model in a controlled setting: (a)-(c) simulated environmental variables, (d)-(g) observations randomly drawn from a true probability of presence that combines the environmental variables and a human modifier, with an increasing weight of the later from left to right (scenario I to IV), (h) optimistic model predictions of the species occurrence, (i) evaluation scores of the optimistic model depending on the scenario for the probability of presence, (j) neutral model predictions of the species occurrence, (k) evaluation scores of the neutral model, (l) pessimistic model predictions of the species occurrence, (m) evaluation scores of the pessimistic model.

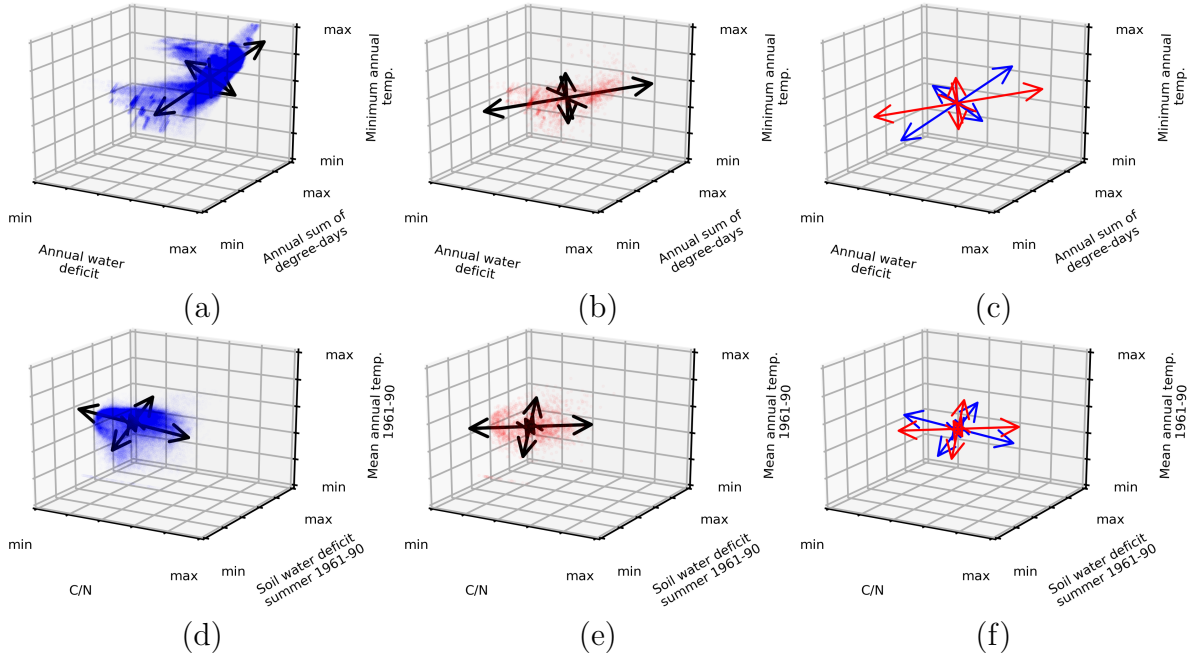


Figure 6: Inter-model comparison between IKS (panels a–c) and DIGI-SDM (panels d–f) for white fir (*Abies alba*) in France: (a) and (d) predicted species presences in the space of model external variables, (b) and (e) observed species presences in the space of model external variables, (c) principal components of the cloud of points shown in (a) (depicted here in blue) and (b) (depicted here in red), (f) principal components of the cloud of points shown in (d) (depicted here in blue) and (e) (depicted here in red).

The greatest potential advantage of the new score can be observed if we focus on the difference between scores as a function of the observation scenario. Cohen's kappa coefficient and TSS are also found to be sensitive to the observation scenario - whose different realizations translate different “human-induced” divergences between the modeled probability of presence and the true one (shown as lines in Fig. 5i, k and m). In contrast, the two RV coefficients of the proposed score are almost independent of the latter.

### 5.3 Inter-model comparison

We finally revert to our initial question of confronting the capability of two models to predict the impact of the climate on the forest, which motivated the development of the new framework for model evaluation and the new score for point cloud similarity. We illustrate the approach with two tree species distribution models applied to white fir (*Abies alba*) in France. One of these models is IKS. The other one is the DIGI-SDM model by Piedallu et al. (2016). Like IKS, DIGI-SDM depends exclusively on external variables, out of which only: mean annual temperature 1961-1990 ( $^{\circ}C$ ), summer soil water deficit 1961-1990 ( $mm$ ) and nitrogen supply ( $C/N$  - Carbon / Nitrogen ratio) appear to be pertinent in case of white fir. The input variables are normalized to the range  $[0, 1]$ . Its output is the probability of presence of the species that is then converted into an occurrence by comparing it to a threshold probability value (Liu et al., 2005).

Following the same steps as before, we first project the predicted presences for each model in their respective space of external variables (Fig. 6a and d). Secondly, we project the observed presences issuing from the data of the national forest inventory (Strona et al.,

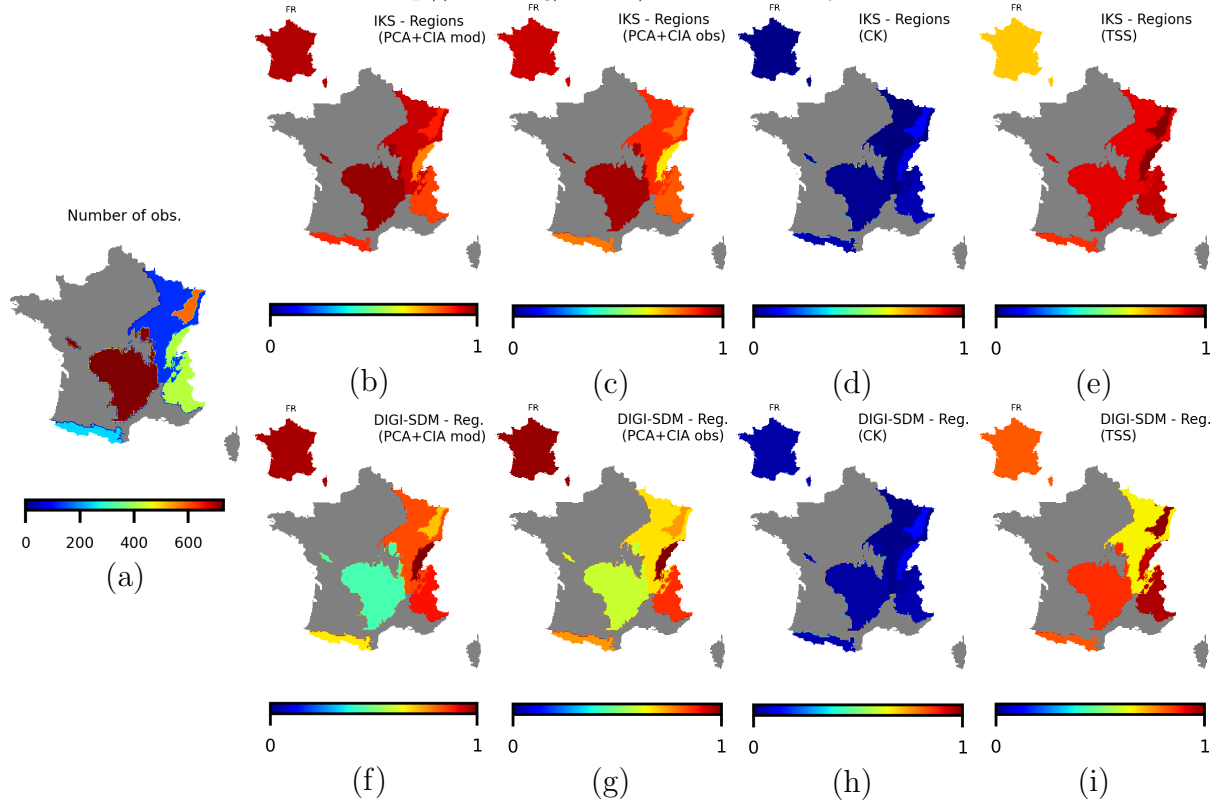


Figure 7: Inter-model comparison between IKS (panels b–e) and DIGI-SDM (panels f–i) for white fir (*Abies alba*) at two scales which are France and its ecological regions: (a) number of observations by ecological region, (b) and (f)  $RV_g$  score, (c) and (g)  $RV_o$  score, (d) and (h) Cohen's kappa coefficient, (e) and (i) True Skill Statistic.

2016) to these two different spaces (Fig. 6b and e), allowing us to have a first view of the concordance between the two pairs of clouds after applying the PCA (Fig. 6c and f).

However, the real added value of the proposed approach comes from the RV coefficients at the different geographical scales (Fig. 7). Unlike Cohen's kappa coefficient which appears to make almost no difference between the two models, despite the non-negligible difference in the external variables they rely on, the TSS and the two RV coefficients of the newly proposed score unveil some differences. As already noticed and commented (prevalence) in Section 5.1, the TSS score mostly increases as we scale down to the regional level.

As for the new score, despite a high value at the national scale (which is the scale at which DIGI-SDM was calibrated and IKS, as explained in section 5.1, maintains the properties of the calibration at the European level), values vary substantially between regions, and not in the same fashion for the two models.

For example, IKS shows significantly higher score for white fir in the region corresponding to the Massif Central (south-central France), suggesting that the IKS combination of environmental variables (annual: water deficit, sum of degree-days and minimum temperature) represents a good proxy for the complexity of natural dimensions influencing the white fir distribution in this region. Combination of environmental variables retained by DIGI-SDM (mean annual temperature 1961-1990, summer soil water deficit 1961-1990 and nitrogen supply) is on the other side not sufficient to replicate the natural tendency in this region, and at this scale. The opposite example is the region correspond-



ing to Jura Mountains (French-Swiss border) where environmental variables selected by DIGI-SDM appear to be almost a perfect proxy for the natural complexity impacting the distribution of the species in question.

This possibly indicates that the modeling approach and the choice of the external variables can be more or less relevant depending on the scale, and depending on the ecological regions too. Coming back to the bottom line motivation of our work, the adaptation decisions should not be based on one model only.

This section is also somehow mostly focusing on the advantages and features of the new score compared to conventional ones as Cohen's Kappa coefficient or TSS, commonly employed in the field of forest modeling. This is due to the conviction of the authors that by being based on the evaluation of the phase space behavior, the score we propose should be, unlike the evoked reference ones, less sensitive on the human decisional impact, i.e. on the human decision to plant/keep the species or not, the latter being an important challenge in the community of forest modelers and managers. It should therefore better reflect the capability of model to reproduce the essential natural tendency, i.e. the ecological potential.

## 6 Conclusion and perspectives

In this article we presented a new framework, containing a new score, for the comparative analysis of forest models integrating the impact of external climatic variables. This framework addresses the most general case of disjointed phase spaces between models, reducing the comparison to the evaluation of the predicted vs. observed presences in the space of external variables of each of the models. The evaluation itself is performed using a score that combines principal component and co-inertia analyses. This score provides a measure of the pertinence of the dynamics of a model as opposed to the one of observations projected to the space of variables of the former. The added value of the approach is illustrated using prototype binary classifier models designed for this purpose, but also using a multi-scale evaluation as well as a multi-scale comparison of two binary classifier models used in the forest management - tree species distribution models.

The most obvious perspective of this work would be the analytical exercise of comparing the multitude of forest models sensitive to climatic parameters, both statistical and process-based ones, and trying to interpret the differences assessed by the proposed approach. The most ambitious perspective would be to build upon the presented analytical approach a consensus method which would combine information coming from different models at multiple scales (Picard et al., 2012), assuring that forest adaptation measures are based on the best of our knowledge to predict natural processes. The presented framework and score could be also adapted and exploited by modelers during the model development process, e.g. to check the relevance of the variables considered, or to compare different sets of environmental variable (e.g. climate variables issuing from different climate models). We as well believe that the presented contribution has a certain universality and can be thus used in domains other than the forest modeling.

## Acknowledgment

The authors would like to thank to Vincent Badeau (INRAE, UMR Silva), Alexandre Piboule (ONF), as well as to Céline Perrier and Hedi Kebli (CNPF) for some very

constructive discussions around the problematic dealt with in this work.

The authors would equally like to thank to two anonymous reviewers, to the associate editor in charge of the manuscript, as well as to the editor-in-chief for their valuable contributions in improving the quality of the manuscript.

## Declarations

## Funding

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.

## Conflicts of interests/Competing interests

The authors declare that they have no conflict of interest.

## References

- O. Alluue, A. Tsoar, and R. Kadmon. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (tss). *Journal of Applied Ecology*, 43(6):1223–1232, 2006. doi: <https://doi.org/10.1111/j.1365-2664.2006.01214.x>.
- K. S. Bhat, M. Haran, A. Terando, and K. Keller. Climate projections using Bayesian model averaging and space–time dependence. *Journal of Agricultural, Biological, and Environmental Statistics*, 16(4):606–628, 2011. doi: [10.1007/s13253-011-0069-3](https://doi.org/10.1007/s13253-011-0069-3).
- N. Bréda, R. Huc, A. Granier, and E. Dreyer. Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. *Ann. For. Sci.*, 63(6):625–644, 2006. doi: [10.1051/forest:2006042](https://doi.org/10.1051/forest:2006042).
- A. Cheaib, V. Badeau, J. Boe, I. Chuine, C. Delire, E. Dufrêne, C. François, E. S. Gritti, M. Legay, C. Pagé, W. Thuiller, N. Viovy, and P. Leadley. Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. *Ecology Letters*, 15(6):533–544, 2012. doi: <https://doi.org/10.1111/j.1461-0248.2012.01764.x>.
- P. Ciais, M. Reichstein, N. Viovy, A. Granier, J. Ogee, V. Allard, M. Aubinet, N. Buchmann, C. Bernhofer, A. Carrara, F. Chevallier, N. De Noblet, A. D. Friend, P. Friedlingstein, T. Grünwald, B. Heinesch, P. Keronen, A. Knohl, G. Krinner, D. Loustau, G. Manca, G. Matteucci, F. Miglietta, J. M. Ourcival, D. Papale, K. Pilegaard, S. Rambal, G. Seufert, J. F. Soussana, M. J. Sanz, E. D. Schulze, T. Vesala, and R. Valentini. Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437(7058):529–533, Sep 2005. ISSN 1476-4687. doi: [10.1038/nature03972](https://doi.org/10.1038/nature03972). URL <https://doi.org/10.1038/nature03972>.
- J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960. doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104).

- V. H. Dale, L. A. Joyce, S. McNulty, R. P. Neilson, M. P. Ayres, M. D. Flannigan, P. J. Hanson, L. C. Irland, A. E. Lugo, C. J. Peterson, D. Simberloff, F. J. Swanson, B. J. Stocks, and B. M. Wotton. Climate Change and Forest Disturbances: Climate change can affect forests by altering the frequency, intensity, duration, and timing of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes, windstorms, ice storms, or landslides. *BioScience*, 51(9):723–734, 09 2001. doi: 10.1641/0006-3568(2001)051[0723:CCAFD]2.0.CO;2.
- S. Dolédec and D. Chessel. Co-inertia analysis: an alternative method for studying species–environment relationships. *Freshwater Biology*, 31(3):277–294, 1994. doi: <https://doi.org/10.1111/j.1365-2427.1994.tb01741.x>.
- S. Dray, D. Chessel, and J. Thioulouse. Procrustean co-inertia analysis for the linking of multivariate datasets. *Écoscience*, 10(1):110–119, 2003a. doi: 10.1080/11956860.2003.11682757.
- S. Dray, D. Chessel, and J. Thioulouse. Co-inertia analysis and the linking of ecological data tables. *Ecology*, 84(11):3078–3089, 2003b. doi: <https://doi.org/10.1890/03-0178>.
- E. Dufrêne, H. Davi, C. François, G. le Maire, V. Le Dantec, and A. Granier. Modelling carbon and water cycles in a beech forest: Part i: Model description and uncertainty analysis on modelled net. *Ecological Modelling*, 185(2):407–436, 2005. doi: <https://doi.org/10.1016/j.ecolmodel.2005.01.004>.
- D. Frank, M. Reichstein, M. Bahn, K. Thonicke, D. Frank, M. D. Mahecha, P. Smith, M. van der Velde, S. Vicca, F. Babst, C. Beer, N. Buchmann, J. G. Canadell, P. Ciais, W. Cramer, A. Ibrom, F. Miglietta, B. Poulter, A. Rammig, S. I. Seneviratne, A. Walz, M. Wattenbach, M. A. Zavala, and J. Zscheischler. Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts. *Global Change Biology*, 21(8):2861–2880, 2015. doi: <https://doi.org/10.1111/gcb.12916>.
- A. Guisan and W. Thuiller. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9):993–1009, 2005. doi: <https://doi.org/10.1111/j.1461-0248.2005.00792.x>.
- J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 143(1):29–36, 1982. doi: <https://doi.org/10.1148/radiology.143.1.7063747>.
- U. G. Indahl, T. Næs, and K. H. Liland. A similarity index for comparing coupled matrices. *Journal of Chemometrics*, 32(10):e3049, 2018. doi: 10.1002/cem.3049.
- IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, volume In Press. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021. doi: 10.1017/9781009157896.
- I. T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics, 2nd edition, 2002.

- R. J. Keenan, G. A. Reams, F. Achard, J. V. de Freitas, A. Grainger, and E. Lindquist. Dynamics of global forest area: Results from the fao global forest resources assessment 2015. *Forest Ecology and Management*, 352:9–20, 2015. doi: <https://doi.org/10.1016/j.foreco.2015.06.014>. Changes in Global Forest Resources from 1990 to 2015.
- G. Krinner, N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19(1), 2005. doi: <https://doi.org/10.1029/2003GB002199>.
- C. Liu, P. M. Berry, T. P. Dawson, and R. G. Pearson. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3):385–393, 2005. doi: <https://doi.org/10.1111/j.0906-7590.2005.03957.x>.
- A. Mauri, G. Strona, and J. San-Miguel-Ayanz. Eu-forest, a high-resolution tree occurrence dataset for europe. *Scientific Data*, 4(1):160123, 2017. doi: [10.1038/sdata.2016.123](https://doi.org/10.1038/sdata.2016.123).
- S. Müller, L. Schüler, A. Zech, and F. Heße. **GSTools v1.3**: a toolbox for geostatistical modelling in python. *Geoscientific Model Development*, 15(7):3161–3182, 2022. doi: [10.5194/gmd-15-3161-2022](https://doi.org/10.5194/gmd-15-3161-2022). URL <https://gmd.copernicus.org/articles/15/3161/2022/>.
- G.-J. Nabuurs, R. Mrabet, A. Abu Hatab, M. Bustamante, H. Clark, P. Havlík, J. House, C. Mbow, K. Ninan, A. Popp, S. Roe, B. Sohngen, and S. Towprayoon. *Agriculture, Forestry and Other Land Uses (AFOLU)*, book section 7. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2022. doi: [10.1017/9781009157926.009](https://doi.org/10.1017/9781009157926.009).
- A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, NY, USA, 3rd edition, 1991.
- C. Parmesan, M. Morecroft, Y. Trisurat, R. Adrian, G. Anshari, A. Arneth, Q. Gao, P. Gonzalez, R. Harris, J. Price, N. Stevens, and G. Talukdarr. *Terrestrial and Freshwater Ecosystems and their Services*, book section 2. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2022. In press.
- N. Picard, M. Henry, F. Mortier, C. Trotta, and L. Saint-André. Using Bayesian Model Averaging to Predict Tree Aboveground Biomass in Tropical Moist Forests. *Forest Science*, 58(1):15–23, 02 2012. doi: [10.5849/forsci.10-083](https://doi.org/10.5849/forsci.10-083).
- C. Piedallu, J.-C. Gégout, F. Lebourgeois, and I. Seynave. Soil aeration, water deficit, nitrogen availability, acidity and temperature all contribute to shaping tree species distribution in temperate forests. *Journal of Vegetation Science*, 27(2):387–399, 2016. doi: <https://doi.org/10.1111/jvs.12370>.
- A. Porté and H. Bartelink. Modelling mixed forest growth: a review of models for forest management. *Ecological Modelling*, 150(1):141–188, 2002. doi: [https://doi.org/10.1016/S0304-3800\(01\)00476-8](https://doi.org/10.1016/S0304-3800(01)00476-8).
- RMT-Aforce. Climesences. <https://climesences.fr/>, June 2022.

- P. Segurado and M. B. Araújo. An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31(10):1555–1568, 2004. doi: <https://doi.org/10.1111/j.1365-2699.2004.01076.x>.
- I. Somodi, N. Lepesi, and Z. Botta-Dukát. Prevalence dependence in model goodness measures with special emphasis on true skill statistics. *Ecology and evolution*, 7(3): 863–872, Jan 2017. doi: [10.1002/ece3.2654](https://doi.org/10.1002/ece3.2654).
- G. Strona, A. Mauri, and J. San-Miguel-Ayanz. A high-resolution pan-european tree occurrence dataset, Dec 2016.
- M. T. Sykes, I. C. Prentice, and W. Cramer. A bioclimatic model for the potential distributions of north european tree species under present and future climates. *Journal of Biogeography*, 23(2):203–233, 1996.
- A. Taccoen, C. Piedallu, I. Seynave, V. Perez, A. Gégout-Petit, L.-M. Nageleisen, J.-D. Bontemps, and J.-C. Gégout. Background mortality drivers of european tree species: climate change matters. *Proceedings of the Royal Society B: Biological Sciences*, 286(1900):20190386, 2019. doi: [10.1098/rspb.2019.0386](https://doi.org/10.1098/rspb.2019.0386).
- W. Thuiller, B. Lafourcade, R. Engler, and M. B. Araújo. Biomod – a platform for ensemble forecasting of species distributions. *Ecography*, 32(3):369–373, 2009. doi: <https://doi.org/10.1111/j.1600-0587.2008.05742.x>.
- R. Yousefpour, C. Temperli, J. B. Jacobsen, B. J. Thorsen, H. Meilby, M. J. Lexer, M. Lindner, H. Bugmann, J. G. Borges, J. H. N. Palma, D. Ray, N. E. Zimmermann, S. Delzon, A. Kremer, K. Kramer, C. P. O. Reyer, P. Lasch-Born, J. Garcia-Gonzalo, and M. Hanewinkel. A framework for modeling adaptive forest management and decision making under climate change. *Ecology and Society*, 22(4), 2017. doi: [10.5751/ES-09614-220440](https://doi.org/10.5751/ES-09614-220440). 40.
- D. Zurell, W. Thuiller, J. Pagel, J. S. Cabral, T. Münkemüller, D. Gravel, S. Dullinger, S. Normand, K. H. Schiffers, K. A. Moore, and N. E. Zimmermann. Benchmarking novel approaches for modelling species range dynamics. *Global Change Biology*, 22(8): 2651–2664, 2016. doi: <https://doi.org/10.1111/gcb.13251>.