ecologies

MDPI

Review

# Uncertainties in Plant Species Niche Modeling under Climate Change Scenarios

**Isabel Passos [1,2,*], Albano Figueiredo [1], Alice Maria Almeida [3] and Maria Margarida Ribeiro [2,3,4]**

1 CEGOT—Centre of Studies in Geography and Spatial Planning, Department of Geography and Tourism—University of Coimbra, St. Jerónimo's College, 3004-530 Coimbra, Portugal; geofiguc@gmail.com
2 CERNAS-IPCB—Research Centre for Natural Resources, Environment and Society, Polytechnic University, Polytechnic Institute of Castelo Branco, Quinta Sra. de Mércules, 6001-909 Castelo Branco, Portugal; mataide@ipcb.pt
3 IPCB-ESA—School of Agriculture, Polytechnic University, Polytechnic Institute of Castelo Branco, Quinta Sra. de Mércules, 6001-909 Castelo Branco, Portugal; alicemalmeida@ipcb.pt
4 CEF—Forest Research Centre, TERRA Associated Laboratory, School of Agriculture, University of Lisbon, Tapada da Ajuda, 1349-017 Lisboa, Portugal
* Correspondence: ipassos.uc@gmail.com

**Abstract:** Species distribution models (SDMs) have been used to forecast the impact of climate change on species' potential distribution, with results that might support decisions for conservation and biodiversity management. Despite their vulnerability to parameterization and data quality input, SDM use has been increasing in the last decades. In fact, inappropriate inputs and the lack of awareness about the effects of methodological decisions on results can lead to potential unreliability in results, a problem that might gain relevance when SDMs are used to predict climate change impacts on species-suitable areas. Aiming to assess how far such a topic is considered, an analysis of the calibration data and methodological decisions was conducted for recent publications (2018 to 2022) that include SDMs in this context, aiming to identify putative deviations from the consensual best practices. Results show that the parameters presented more consistently are the algorithm in use (MaxEnt was used in 98% of the studies), the accuracy measures, and the time windows. But many papers fail to specify other parameters, limiting the reproducibility of the studies. Some papers fail to provide information about calibration procedures, others consider only a fraction of the species' range, and others provide no justification for including specific variables in the model. These options can decrease reliability in predictions under future scenarios, since data provided to the model are inaccurate from the start or there is insufficient information for output discussion.

**Keywords:** climate change; plant species; range shift; species distribution models

check for updates

**Citation:** Passos, I.; Figueiredo, A.; Almeida, A.M.; Ribeiro, M.M. Uncertainties in Plant Species Niche Modeling under Climate Change Scenarios. *Ecologies* **2024**, *5*, 402–419. https://doi.org/10.3390/ecologies5030025

Academic Editor: José Ramón Arévalo Sierra

Received: 6 July 2024
Revised: 19 August 2024
Accepted: 24 August 2024
Published: 27 August 2024



## 1. Introduction

The need for spatially explicit results when assessing climate change impacts on species distribution promotes the search for a deep understanding about abiotic factors' influence on species distribution patterns, a task facilitated by the increasing availability of environmental and species occurrence data with high resolution, namely for climatic scenarios, and dedicated tools, such as species distribution modeling techniques based on a wide array of algorithms based on correlation [1,2].

Species distribution models (SDMs) are widely used to predict species ranges and environmental niches, and their use has been increasing over the last two decades [1]. Models of a correlative nature are more common, since they relate species occurrence data and environmental variables, generating maps predicting past, present, and future species distributions [2–4].

SDMs have been used for species conservation purposes and biodiversity management, like selecting locations for protected areas, habitat restoration actions, and/or species translocation, especially in the context of global climate change [5–12]. Under climate change scenarios, such an approach is used to assess possible impacts on biodiversity [12,13], aiming to assess potential changes in species-suitable areas, from expansion [14,15] to contraction [16–19] and sometimes even extinction [20,21].

The choices made during the modeling process can significantly affect model predictive performance, and predictive results may vary greatly due to those choices [1,22,23]; thus, models must be fitted for the purpose, and options should be carefully considered [12]. Possible inaccuracies or uncertainties can arise in different steps [24–27] and from different sources. This includes data sources involving occurrences and environmental data, including future climate change scenarios; spatial niche truncation, including geographical and ecological range fractions; the clamping effect, comprising new conditions outside the range used to calibrate models; parametrization in the modeling process, including variable selection and variable correlation; using only climatic variables; evaluation strategies; and limited model discussion. Several authors have looked into these issues, addressing different errors that can lead to inaccurate results [2,28–39].

Unreliable species occurrence data can produce models that underestimate suitable areas [34], affecting their quality [5]. Data for SDMs can be collected from various sources, such as museums or other natural history collections, as well as from bibliographies, field surveys, and databases. Data coming exclusively from museums or other natural history collections can be incomplete or biased concerning the actual range of the species, since they were probably collected in more accessible locations [24]. Otherwise, collecting data from systematic field surveys can lead to oversampling in some areas as compared to others [31]. Ideally, systematic surveys should be performed in the species total range area [5]. These surveys can be feasible for species with small range sizes but highly demanding for wide-range species [36,40]. Online platforms (e.g., GBIF) currently provide occurrence data, commonly used to estimate climate change impacts on species distributions. However, differences in funding between nations and data sharing lead to differences in contribution, creating spatial bias due to uneven sampling efforts [28,34]. Also, data collected by the general public may have several errors, such as misidentification and georeferencing errors or sampling bias across more accessible areas, near cities and roads [41], as well as data storage and mobilization issues [28,34].

Future climate scenarios are based on emissions and development scenarios, established by the Intergovernmental Panel on Climate Change (IPCC). The most recent were released in their Assessment Report (AR6) [42]. These scenarios are projected based on possible development scenarios, which consider different levels of greenhouse gas emissions, population growth, economic and technological development, and land use [43–47]. Although these scenarios are now robust projections and essential in climate change research and assessment [45], they are still scenarios prone to errors and uncertainties, as are the models based on them [29,48,49].

Study area limits are critical when modeling species suitability. When data for a fraction of the entire area of the species range is used, not all the abiotic conditions endured by the species may be considered, compromising the models' ability to capture the full range of suitable areas [2,50]. Leaving out marginal areas and marginal populations may also compromise results, since these populations may be adapted to more extreme conditions [51,52]. In these situations, called spatial niche truncation, only a subset of species' ecological niches is considered, which can lead to incorrect forecasts when projecting future suitability [53,54]. Species occurrence data should be as comprehensive as possible to improve SDM results and represent environments and geographical areas where the species can live and disperse [5]. In fact, in studies that assess climate change effects, it might be critical to consider areas beyond the species' present range, accounting for locations that may reflect potential future environmental conditions [55]. In fact, it is essential to include areas around current distribution, which might be suitable in the future [32,33,39,52,55].

Climate variables greatly influence plant species' spatial (and temporal) distribution. However, these are not the only variables that explain their distribution, especially when dealing with restricted ranges and high-resolution data. Other environmental and abiotic variables (e.g., soil, topography, fire frequency) are also important when modeling distributions and range shifts [35,38,56,57], and the rejection of non-climatic environmental variables must be based on dedicated selection methods. The inclusion of such variables might also support the identification of other restrictive factors, namely associated with land use, since areas with greater slopes present lower human pressure [38] and register a higher number of occurrences, or might act as limiting factor themselves, like soil conditions, as it is unlikely that the species will be able to establish itself on unsuitable soil conditions even under appropriate climatic conditions [56–59]. So, the exclusive use of climate data can erroneously estimate a species' range, often producing overpredictions [57]. However, not all available variables should be blindly included in the model, since these variables may be highly correlated [60,61], sharing high amounts of information [30]. In this case, variables with indirect effects (e.g., altitude) should be discarded, and correlated variables with direct influence (temperature or precipitation) [62,63] should remain, namely those with high ecological significance for the species under analysis, making it possible to (i) simplify the interpretation of the model [64]; (ii) avoid over-fitted results; and (iii) eliminate crossed effects on the response curves of each variable, as inaccuracies caused by interactions with other variables will remain when correlated variables are in use [2,65], making it difficult to disentangle the influence of each variable [60]. This might be a severe drawback when a model is fitted on data from one area or time and projected to another area or period with a different or unknown structure of collinearity, since collinearity between environmental variables is not constant in space and time [30]. It is impossible to eliminate collinearity, but it can be reduced [30]. There are several methods to quantify collinearity. One of the most effective is to select variables using a threshold under a specific value of correlation coefficients (e.g., $|r| < 0.7$) [30,60]. Ignoring environmental variables that are determinant to tackle the species' ecology can lead to unlikely predictions of species' responses to climate change [32]. Therefore, it is crucial to know the species' ecological preferences to select the most meaningful variables to include in the model and to use a model that is as reliable as possible [24,30,60,66,67].

There are many techniques and modeling algorithms available to perform SDMs; these belong to different categories of models, such as regression methods—generalized linear models (GLMs), generalized additive models (GAMs), and multivariate adaptive regression spline (MARS); classification methods—classification tree analysis (CTA) and Flexible Discriminant Analysis (FDA); machine learning algorithms—random forest (RF), Boosted Regression Tree (BRT), and Maximum Entropy (MaxEnt); [37,68,69], and others, including Support Vector Machine (SVM) [70]. No single model is superior in all situations [70,71], so the algorithm's choice depends on the data specificities and the study objective [72].

Evaluation strategies or performance metrics are important to assess the discriminatory capacity of a model or its ability to distinguish suitable from unsuitable conditions. There are several ways to assess model performance, such as sensitivity (the proportion of presences correctly predicted), specificity (the proportion of absences correctly identified), Cohen's Kappa Statistic (kappa), true skill statistic (TSS), percentage of correct classification rate (CCR), Area Under the ROC Curve (AUC), and error rate (ER) [22,73]. The most widely used evaluation metrics are AUC and TSS [2,74], but even the most widely used performance metrics have important limitations for ecological studies [74–76]. They are designed to reflect the trade-off between sensitivity and specificity and generally weigh sensitivity and specificity equally [77]. A single use of AUC can identify well-fitting and strongly predictive over-fitted models [48]. The AUC value depends on the size of the study area: if the area is large enough to comprehend habitats different from those occupied by the species, the AUC will be higher, even if the model is not that good, since more points with correct predictions of low suitability are considered [75,77]. The same occurs with the TSS, which tends to be correlated with the AUC. Also, TSS depends on species prevalence and may lead to misleading results [78].

These common and recurrent mistakes during SDM application have led to the publication of several works that intend to standardize SDM procedures, improving their quality and reproducibility [5,6,12,64,79].

In this context, the main objective of this study is to analyze the available recent literature [80] dedicated to assessing climate change effects on plant distribution, based on niche modeling, to determine

- Which are the most used data and methodologies, namely those related to model calibration;
- Which are the most common deviations from consensual best practices and what information is most omitted from methodological descriptions;
- How far the faults referred to above are identified and discussed;
- New recommendations to improve SDM results, making them clearer and more comprehensive.

The analysis considers the methodologies used in recent papers, from species occurrence data to abiotic variable data sources, and the implications for models' accuracy and the potential reproduction by pairs. Key aspects of the SDM elements were registered for each paper and assembled into a database, including (i) the source of species occurrence data; (ii) the geographical range analyzed; (iii) the type of occurrence data (presence only, pseudo-absence, and absence data); (iv) the abiotic variables; (v) the variables' selection methods; (vi) the used algorithm(s) for modeling; (vii) the model performance metrics; (viii) the use of an ensemble model; (ix) the climatic scenario(s); (x) the source of climatic models (databases and GMCs); and, ultimately (xi) the lack of details considering methodological decisions.

## 2. Materials and Methods

This study aims to identify if best practices are followed when assessing changes in plant species distribution under climate change scenarios based on niche modeling in recent papers. The article search was conducted in November 2022 in two databases (DBs): Web of Science (WOS) and Scopus. These DBs may have some overlapping results but are considered the two most comprehensive sources of bibliographic resources [81,82]. The following search equation was included, using Boolean search strategies: ("climate change" OR "global change") AND ("model*" OR "ecological niche model*" OR "species distribution model*" OR "habitat suitability model*" OR "range shift") AND ("R software" OR "maxent" OR "Biomod*" OR "GLM" OR "average model*" OR "ensemble*") AND ("flora" OR "plant*"). The search was carried out for the item "Topic" in the WOS Core Collection and the "Article title, Abstract, Keywords" topics in Scopus. Since modeling methodologies are constantly changing, a time limit was imposed on the research, considering only scientific papers published from 2018 to 2022. Only original articles were considered, and other document types, such as review articles, books, and book chapters, were removed. The search was based on PRISMA guidelines [82,83], and the flow chart (Figure 1) shows the different steps undertaken in the current study.

The screening of the articles was conducted until January of 2023. Duplicate records and articles in other languages besides English were initially removed. Unavailable documents were also excluded. The title and abstract of the remaining documents were thoroughly screened and evaluated for inclusion in the study (Figure 1) according to pre-established exclusion criteria: (a) not exclusively focused on terrestrial vascular plant species; (b) dedicated to agricultural species and their production, such as vines, rice, and corn; (c) focused on invasive flora; (d) considering aquatic environments or islands; (e) focused on the evaluation of modeling methods rather than assessing climate change effects on species distributions; and (f) lack of modeling for the future. Some studies, 3 in total, were excluded after a full analysis of the documents (Figure 1).

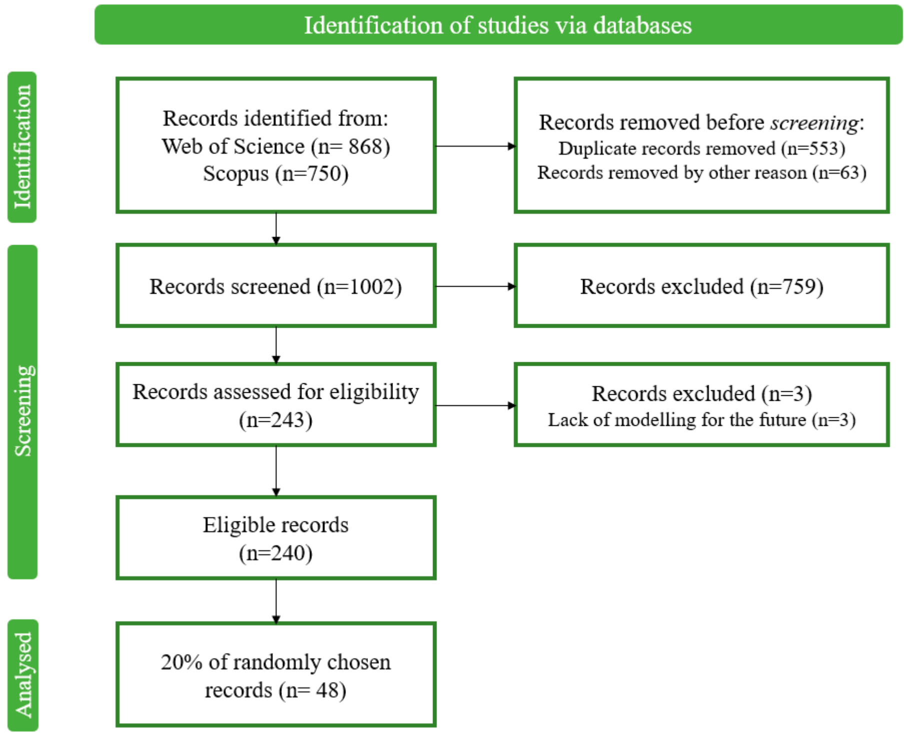


**Figure 1.** Flow diagram of the selection process, based on PRISMA [83,84].

The databases' screening found 240 documents complying with the selection criteria. According to Amobonye et al. [80], in general, no comprehensive instructional framework exists to guide scientists on how to analyze and synthesize the literature in terms of their niches in publishable review articles. Indeed, the amount of data to be analyzed is fairly massive, comprising hundreds of and tens of thousands of items. Therefore, pruning of the obtained information is needed for a comprehensible review. Thus, a representative and more manageable number of articles was randomly selected (20%) in the current study. Lastly, the 48 selected articles, after pruning the initial number of documents (240), were analyzed (Supplementary Materials, File S1).

The key aspects of the SDM elements were noted in each selected publication and assembled into a database that included (i) the source of species occurrence data; (ii) the geographical range analyzed; (iii) the type of occurrence data (presence only, pseudo-absence, and absence data); (iv) the abiotic variables; (v) the variables' selection; (vi) the used algorithm(s) for modeling; (vii) the model performance metrics; (viii) the use of an ensemble model; (ix) the climate scenario studied; (x) the source of climatic models (databases and GMCs); and, ultimately (xi) the lack of details for methodological decisions.

## 3. Results

### *3.1. Species Occurrence Data*

The studies use different data sources for the species occurrence data, and only in one case was this information missing. The use of one or two combined sources was the most common situation (35.4%), but the combination of three (18.8%) or four (8.3%) data sources was also identified. The most used data source was a field survey (62.5%), followed by the use of online databases, such as the Global Biodiversity Information Facility (GBIF) (58.3%) (Figure 2).

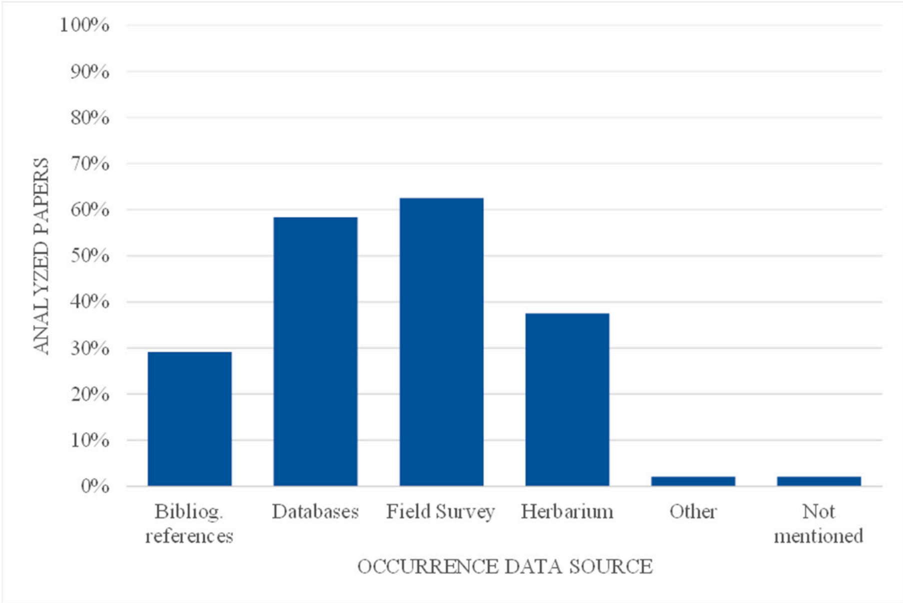


**Figure 2.** Percentage of analyzed papers considering occurrence data sources.

For the analyzed studies, 34% considered the total geographic range of the target species, while 66% considered only a fraction, usually delimited by political borders. In almost 44% of the articles, the species' natural geographical range was not presented, and it was unclear if the work considered the species' total range and all the conditions it could endure.

Presence-only data were used in 27.1% of the papers, while only 4.2% referred to absence data and 12.5% to pseudo-absence points. In the remaining 56.2% of the documents, the data type was not clarified.

## 3.2. Abiotic Variables

### 3.2.1. Climate Variables

Most of the papers (97.9%) mentioned the source of environmental variables used on models' calibration, and only in one case was the data source not identified. The WorldClim (WorldClim) [85,86] was the most commonly used database for climate data, namely bioclimatic variables, included in 83.3% of the documents. Usually, the bioclimatic variables were downloaded in two different spatial resolutions: 2.5 min (approximately 5 $km^2$) (28.3%) or 30 s (approximately 1 $km^2$) (69.6%). Some studies (14.6%) used other climatic variables sources, such as Climate Change, Agriculture and Food Security (CCAFS Data—CCAFS Climate (ccafs-climate.org); 8.3% of the cases), the Africlim [87] (4.2%), or the ClimateAP (ClimateAP_Map) (2.1%).

### 3.2.2. Other Environmental Variables

Up to half of the papers (56%) included other variables besides climatic variables. The altitude/elevation was the most common (74%), followed by the slope (63%) and aspect (59%) (Figure 3). In some cases, the authors used more specific variables, such as distance to different features, including rivers, dwellings [88], and fresh or salty water bodies [89].

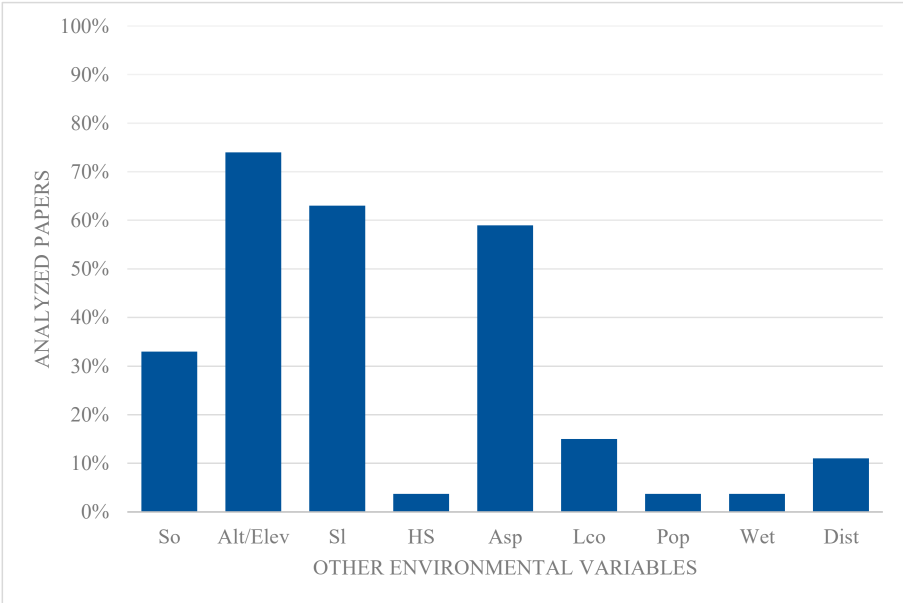


**Figure 3.** Percentage of papers using other environmental variables: soil (So), altitude/elevation (Alt/Elev), slope (Sl), hillshade (HS), aspect (Asp), land cover (Lco), human populations (Pop), distance to wetland (Wet), and distance to (. . .) (Dist).

#### 3.2.3. Variable Selection

Correlation analysis between abiotic variables was performed in 79% of the papers. In comparison, in the remaining papers (21%), no reference was made to the correlation between variables or the methodology used to perform correlation analysis and the variable selection.

In papers using correlation analysis, 54.2% opted for Pearson's correlation test, 10.5% for the Variance Inflation Factor (VIF), and 8.3% for the ArcGIS/ArcMap. In 6.3% of the cases, the method used for correlation analysis was omitted.

### *3.3. Modeling Algorithm*

In the analyzed papers, eleven different modeling algorithms were used. The MaxEnt was the most popular one, being used in 98% of the cases and being the sole algorithm in use 83% of the time. Besides MaxEnt, none of the other algorithms was used as the sole algorithm, with 16% of the papers using several different methods. In these cases, the researchers chose to use between four (2%) and ten (4%) algorithms.

### *3.4. Model Performance*

The model's predictive performance was evaluated using four different accuracy measures: the Area Under the Curve of the Receiver Operating Characteristic/(AUC of ROC), Akaike's Information Criterion (AICc), the True Skill Statistic (TSS) test, and Cohen's Kappa coefficient. The most used method was the AUC of ROC, mentioned in 93.8% of the articles, the TSS was used in 31.3%, the AICc in 10.4%, and Cohen's kappa coefficient in 4.2% (Figure 4).

All articles mentioned at least one method to measure model performance. In 62.5% of the cases, only one method was used; in 35.4% the authors opted to use two methods; and in only 2.1% of the cases, three different methods were used.

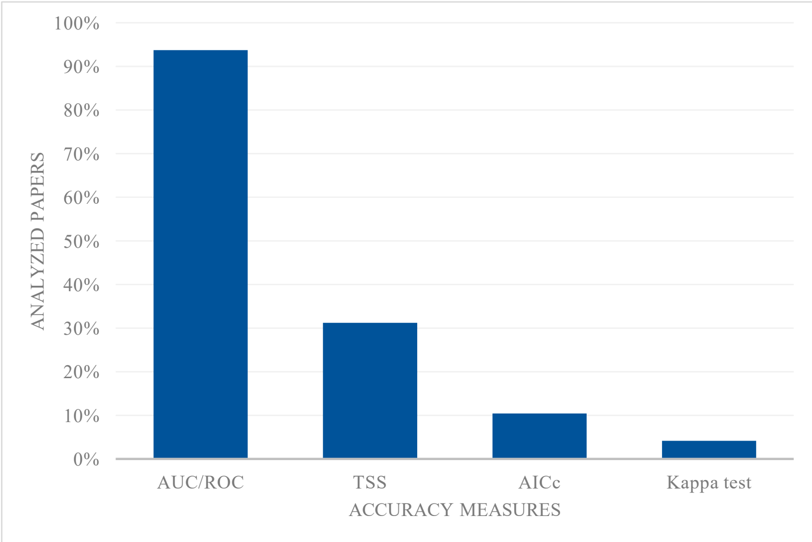


**Figure 4.** Accuracy measures used to assess model performance.

### *3.5. Ensemble Models*

In 42% of the papers, only one algorithm and a GCM were used; thus, no ensemble model was produced. In the remaining 58%, where more than one modeling algorithm and/or GCM was used, only 32% created an ensemble model (18.8% of total analyzed papers). The Biomod2 was the package used to perform the ensemble model in 67% of these cases (12.5% of total analyzed papers), with some works using a threshold to choose which models should be considered for the ensemble model.

### *3.6. Future Climate Projections*

#### Climate Scenarios

The analyzed papers mainly used two (59.6%) or four different climatic scenarios (19.1%). The use of one or three different scenarios was less common and occurred in 8.5% and 12.8% of the cases, respectively. The most popular scenarios were the Representative Concentration Pathways (RCPs): the RCP8.5 (70.8% of the documents), the RCP4.5 (56.3%), and RCP2.6 (45.8%) (Figure 5).

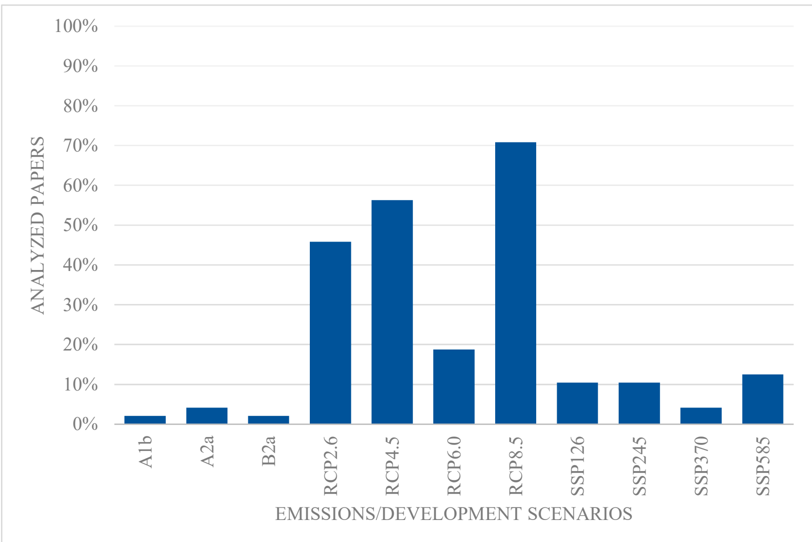


**Figure 5.** Emission/development scenarios used in analyzed papers: A1b, A2a, and B2a from CMIP3 [90]; RCP2.6, RCP4.5, RCP6.0, and RCP8.5 from CMIP5 [91]; SSP126, SSP245, SSP370, and SSP585 from CMIP6 [44].

Some future time windows are more popular among SDM papers, namely 2050 (average for 2041–2060) and 2070 (average for 2061–2080), which appear in 91.7% and 81.2% of the documents, respectively (Figure 6). Papers used from one to five different time intervals, with 75% using two different time intervals, 12.5% using only one, and 12.5% using 3 to 5.

Summing up, thirty-two Global Circulation Models (GCMs) were used in the analyzed documents, considering all versions of several models. In 6% of the cases, the used GCM was not described, and when the used GCM was mentioned, the number ranged from one to eight GCMs, with most papers (64%) using only one GCM to perform the model.

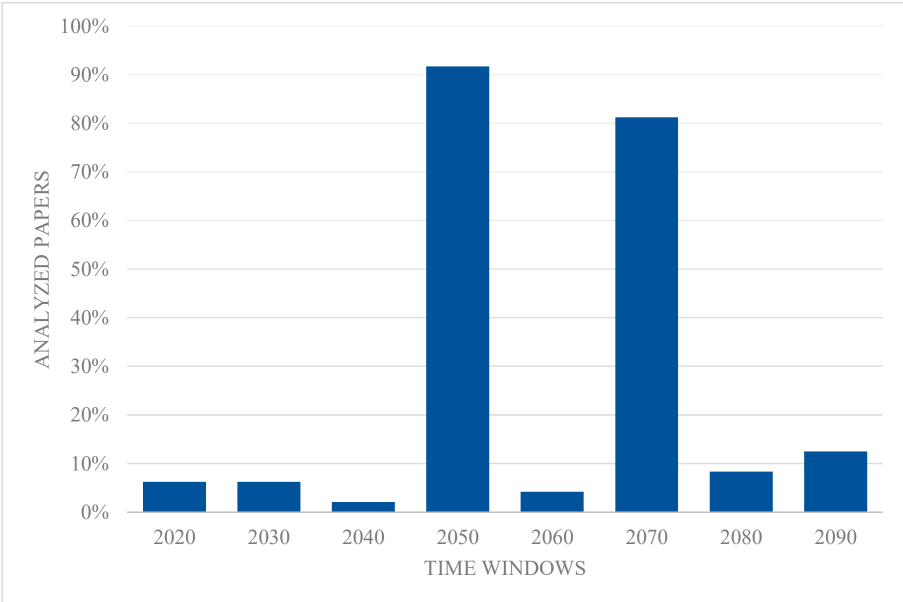


**Figure 6.** Time windows used in analyzed papers.

The CCSM4, developed by the National Science Foundation (NSF) and National Centre for Atmospheric Research (NCAR), and the HadGEM2-ES, developed by the UK Meteorological Office, was used in more than 29.8% and 25.5% of the papers, respectively (Table 1). The GCMs developed by the UK Meteorological Office seem to be the most popular (40.4%), followed by the National Science Foundation (NSF), the National Centre for Atmospheric Research (NCAR) (29.8%), and the Beijing Climate Centre Climate System Model (25.5%) (Table 1).

**Table 1.** Global Circulation Model (GCM) used in the analyzed studies and the independent Climate Research Centers (CRCs) that developed them. Largest percentages are in bold.

| Global Circulation Model (GCM) | Climate Research Centres (CRCs) | Country | Number of Documents by GCM, % | Number of Documents by CRC, % |
|---|---|---|---|---|
| ACCESS1-0 | Australian Community Climate and Earth System Simulator Coupled Model | Australia | 2.13 | 2.13 |
| AFRICLIM | York Institute for Tropical Ecosystems (KITE) and Kenya Meteorological Service | Kenya | 4.26 | 4.26 |
| BCC-CSM1.1 | Beijing Climate Centre Climate System Model | China | 12.77 | **25.54** |
| BCC-CSM2-MR | | | 12.77 | |
| CanESM5 | Canadian Earth System Model | Canada | 2.13 | 2.13 |
| CCAFS | CCAFS-Climate Statistically Downscaled Delta Method | Colombia | 6.38 | 6.38 |

**Table 1.** *Cont.*

| Global Circulation Model (GCM) | Climate Research Centres (CRCs) | Country | Number of Documents by GCM, % | Number of Documents by CRC, % |
|---|---|---|---|---|
| CCCMA | Canadian Centre for Climate Modelling and Analysis | Canada | 2.13 | 2.13 |
| CCSM4 | National Science Foundation (NSF) and National Centre for Atmospheric Research (NCAR) | United States | **29.79** | **31.92** |
| CCSM5 | | | 2.13 | |
| CGCM3.1-T63 | Canadian Centre for Climate Modelling and Analysis | Canada | 2.13 | 2.13 |
| CNRM-CM5–1 | CNRM (Centre National de Recherches Météorologiques—Groupe d'études de l'Atmosphère Météorologique) and Cerfacs (Centre Européen de Recherche et de Formation Avancée | France | 2.13 | 12.77 |
| CNRM-CM6–1 | | | 4.26 | |
| CNRM-ESM2–1 | | | 6.38 | |
| CSIRO | Commonwealth Scientific and Industrial Research Organisation | Australia | 2.13 | 6.39 |
| CSIRO-MK3.6 | | | 4.26 | |
| GFDL-CM3 | Geophysical Fluid Dynamics Laboratory (GFDL) | United States | 4.26 | 4.26 |
| GISS-E2-R | Goddard Institute for Space Studies (GISS—NASA) | United States | 2.13 | 2.13 |
| HadCM3 | UK Meteorological Office | United Kingdom | 2.13 | **40.43** |
| HadGEM2-AO | | | 6.38 | |
| HadGEM2-ES | | | **25.53** | |
| HadGEM-CC | | | 4.26 | |
| HadGEM-IS | | | 2.13 | |
| IPSL-CM5A-LR | Institut Pierre-Simon Laplace (IPSL) | France | 2.13 | 4.26 |
| IPSL-CM6A-LR | | | 2.13 | |
| MIROC5 | Center for Climate System Research (CCSR), National Institute for Environmental Studies (NIES) and Japan Agency for Marine-Earth Science and Technology | Japan | 6.38 | 14.9 |
| MIROC6 | | | 2.13 | |
| MIROC-ES2L | | | 4.26 | |
| MIROC-ESM | | | 2.13 | |
| MPI-ESM-LR | Max Planck Institute for Meteorology | Germany | 2.13 | 2.13 |
| MRI-CGCM3 | Meteorological Research Institute (MRI) | Japan | 8.51 | 12.77 |
| MRI-ESM2-0 | | | 4.26 | |
| NorESM1-M | Norwegian Earth System Model (NorESM) | Norway | 2.13 | 2.13 |

## 4. Discussion

The information about the methodology used in each work is not always clear and complete. Some parameters are described more consistently, e.g., the origin of the data. However, many articles fail to specify other parameters, such as the use or not of pseudo-absence points and ensemble modeling techniques, or even the GCM used. The same tendency, which limits the reproducibility of the studies, was noticed by other authors [1,6,71]. This is a problem that has been addressed in the recent literature by several authors, aiming to provide guidelines/checklists for future publications [2,5,6,12]. In addition to the gaps in the description of adopted methodologies, common and recurrent mistakes during SDM application have also been pointed out by recent studies [36,37]. These poor modeling practices can lead to inaccurate conclusions and poor planning of conservation actions [64]. The examined studies had many similarities concerning the different elements analyzed. Target species distribution areas could have been more clearly stated, either in whole or in part. Over a third of the papers used the total range of the species, while the rest only considered a fraction. This is an important point, since models that rely on partial distributions may not be able to capture the full range of abiotic conditions in which a species can survive [2], and marginal populations can have adaptations to more extreme situations [51]. It is also essential to include areas outside a species' current geographical range, to produce spatial predictive models (e.g., using buffer zones), so that the model can assess the suitability of these conditions in the future. Ignoring this can promote inaccurate projections for future climatic conditions [32], [33]. However, this seems common in ecological modeling [53–55]. For this reason, niche truncation and clamping can lead to incorrect predictions when projecting future climatic conditions, since future conditions may be unavailable in the calibration area but may be suitable for the species [39,53,55]. This can result in predicting false local extinctions or extirpations and, hence, inaccurate predictions of future species suitability, especially at range margins [50]. However, excluding areas under a climate that will no longer exist in the future, e.g., the northern range limit of a European species, may not be problematic, since those conditions will no longer be present [50]. The explanation of the species' range and the study area should be well specified, together with the reasons for those choices [25], which is not always the case. Nevertheless, only one work addressed superficial niche truncation.

Field surveys were the most popular data source, but performing models only with field data can lead to problems related to some areas being oversampled, especially in broad-range area species [31]. Although systematically designed surveys covering major species ranges are recommended [5], systematic surveys along all species range areas in occupied major environmental gradients can be source-demanding, expensive, and time-consuming [36,40]. On the other hand, opportunistic sampling (e.g., GBIF) can have other problems, such as the misidentification of species and spatial bias in records due to uneven efforts in sampling [28,34], but larger sample sizes for these types of data seem to compensate for this and outperform systematic sampling [92–94]. These biases and inaccuracies in distributional data can place heavy limitations on SDM studies and affect the quality of final results [5]. About two-thirds of the studies used more than one source for occurrence data gathering, from field work and large databases to locations mentioned in specific studies or herbaria. This can be a good strategy, since the more information that is given to the model, the better it will perform [95], and data from different sources might complement each other [91]. Also, when sample data are collected from broad geographical areas, including different environmental gradients, a higher possibility exists that environmental conditions limiting species distribution will be well sampled [24].

The climate variables were used in all studies, and the most common source was WorldClim, which was included in a large majority of the papers. The models were performed mainly at a 30 s spatial resolution (approximately ~1 km$^2$), the highest resolution used. Although, depending on the study goal or for small-range species, a finer special scale should be used [58]. The 30 s scale is often the finer available scale, which limits the possibility of performing finer-scale models. Larger-scale models may detect less variation in topography or soil conditions compared with finer-scale data, resulting in a lower ability for the models to discern topographic and soil variation within the landscape [58].

However, non-climactic factors might also influence plant species distribution [35,56,57]. About half of the analyzed studies used climatic variables only. Other environmental variables were not included in the model, which can overestimate habitat suitability for many plant species, both for present and future scenarios, since climate-based projections might integrate areas with unsuitable soil conditions [57]. Some of these studies highlight this fact, pointing to this issue as a limitation [96,97], and others use a lack of reliable data on a scale that would allow their inclusion in the model. Yet, including all climate and non-climate variables in the same models may not always be suitable [6], since these variables may be highly correlated [61], and their correlation can change through time [37], making future projections less reliable.

Indeed, variable selection is a crucial step in SDM, but one-fifth of the analyzed articles fail to mention variable selection or do not describe the method used. Some simply use all the variables to perform models, without considering possible correlations between them. However, most modeling algorithms are sensitive to high levels of correlation between variables. MaxEnt, the most used algorithm in the analyzed papers, seems to be capable of dealing with redundant variables and the independence between the degree of predictor collinearity and collinearity shift [60]. So, the strategy of removing highly correlated variables seems to have a small impact on MaxEnt model performance [60]. The articles that do not refer to variable selection mainly use MaxEnt. However, in those using other algorithms (BRT, RF, GLM, GAM, MARS and CTA) no justification is given for the absence of correlation analysis and variable selection. The variable selection based on correlation should be performed to simplify the interpretation of the model [64]. Additionally, the species' ecological preferences should be considered, to select the most meaningful variables to be included in the model [24,30,66,67,98].

Several methods are available to perform SDMs; no single one is superior in all situations [70,71], and they seem to have similar performance [96]. BRT, MaxEnt, and RF were reported to be the best-performing modeling algorithms, while parametric and semi-parametric regression models (like GLM and GAM) can be adequate choices when the number of occurrences is very low [70]. In accordance with other similar ones [71,99], MaxEnt was by far the most used algorithm in the screened studies, as previously said. However, the percentage of papers using this algorithm was larger in our review than in others [71,100], and it was the only used algorithm in most papers. MaxEnt is a machine learning method [99,101], and some of its features can contribute to its popularity compared to other algorithms: it is user-friendly, even for a beginner user; outputs are easy to access and read; it is very accessible, as it can be used in open-source software or free software R programming packages; there is no need to provide absence points; and it generates significant results with a small number of points and spatially biased presence points, delivering good outcomes [2,58,70,100–103]. Despite this, in climate change assessments and future projections, it seems advisable to use more than one algorithm to produce a final model, according to consensual best practices [5].

Most papers use more than one climate scenario and more than one time interval. The Shared Socioeconomic Pathways (SSPs) [42] are notably less used than RCPs [91], probably because they are more recent and were unavailable when some of these works were developed. On the other hand, the scenarios provided in 2007 [90] had shallow usage, which makes sense, as more robust scenarios were available when these papers were published. The RCP8.5 was the most used scenario, although it describes a situation with very high anthropogenic greenhouse gas emissions without additional efforts to constrain them [91]. RCP8.5's popularity is likely related to its role as a "worst case scenario", making it a benchmark for comparison with other scenarios, since no study has used only the RCP8.5 for projecting possible future species range shift. Papers using this scenario also used at least other intermediate scenarios. By using the RCP8.5 scenario, researchers can exemplify the most extreme possible impacts for informing climate change mitigation and adaptation policies. Most screened papers displayed two different future time intervals, and a preference existed for more distant temporal periods. This makes sense and might be helpful when the goal is to plan management actions, especially for long-living species. Adaptive and management strategies require a longer-term perspective, since areas managed nowadays must cope with the future climate conditions of at least several decades [104,105]. However, many species may not yet be able to be established in places that will only be suitable in a few decades [106–109]. Therefore, not-so-distant periods might also provide meaningful information about transition areas.

The verified studies used a wide range of GCMs, with a total of 32 considering all versions of the models, with most articles using only one GCM to perform the analysis. Since GCMs are projections and prone to uncertainties, using more than one GCM has been emphasized to reduce uncertainty when projecting species distribution in time [29,49]. Still, more than one-third of the papers used more than one GCM, ranging from two to eight. Some GCMs are more used than others. Those developed by the UK Meteorological Office are the most popular, followed by the National Science Foundation (NSF) and National

Centre for Atmospheric Research (NCAR) from the United States and the Beijing Climate Centre Climate System Model from the People's Republic of China.

When several algorithms or GCMs were used, ensemble models were often performed, in less than a fifth of the articles. Ensemble modeling is often considered to have better predictive results and to be more reliable than single models, and it is often used to reduce the degree of uncertainty in the model selection [1,70,71]. Still, performing an ensemble using models with good and bad predictive capacity may not result in a good final model [2]. Similar to other works [1] and likewise to other analyzed parameters, the methodology for performing the ensemble is sparsely described in the analyzed works. Only two-thirds of the articles performing ensembles clearly stated the use of Biomod2, and only one-third described the choice of the best models to include in the ensemble, using a threshold, based on AUC or TSS.

A large majority of the papers used ROC/AUC to measure model performance. Although over a third of the studies used more than one method, often in a complementary way, the use of AUC stands out. This holds true in another study [74] and is possibly related to the non-threshold dependency of the ROC/AUC, which is a metric provided by MaxEnt and is used in a wide range of applications related to producing predictions. Despite its wider use, the single use of AUC, or another single metric, can misidentify over-fitted models as well-fitting and strongly predictive [48]; therefore, models should be carefully evaluated by specialists, to check whether they make sense ecologically for the target species [110].

## 5. Conclusions

The current review identified 240 papers modeling plant species niches and possible future range shifts under climate change scenarios, 48 of which were randomly selected and analyzed. Despite published standards for the use of niche models, recent studies focused on climate change still exhibit uncertainty related to inconsistent methodological decisions. Although modeling strategies and data sources are pretty consistent, clear methodological decisions are sometimes missing, which hinders the reproducibility of SDM studies and increases uncertainty considering the discussion of results.

Species occurrence data mainly comprise the part of the species range and use more than just one source, with field surveys being the most popular choice. All papers used climate data, but other environmental variables were used in over half of the documents.

The choice of modeling algorithm was quite homogeneous; almost all documents use MaxEnt, which is often the only used algorithm. Using only one GCM was a popular choice, although it is best practice to use more than one; no clear preference was found for a particular GCM.

The parameter analysis indicates that several articles base their models on choices that may lead to inaccurate and possibly unreliable results. The definition of a study area that does not include the species' entire natural range, leaving out areas and environments in which the species can live, and not including areas having climatic conditions that might be more usual in the future were common, since over half of the studies only considered a part of the species range. Also, ignoring species' ecological preferences when choosing the variables to use in the model, both at the outset and after variable selection, is another error that appears to be common and which can lead to putative inaccuracies in the results.

Overall, there is a need to make the information clearer and more comprehensive in the SDM studies. In this paper, we emphasize that the information regarding the species being studied and the modeling process is often missing. Therefore, besides the best practices referred to in guideline papers previously cited, it is considered pertinent in future modeling studies to include and state the following information:

- Target species' natural range;
- The species' total range in the study area, including a buffer to ensure the inclusion of different environmental conditions;

- Comparison of the study area and the natural range of the species, as well as justification of the exclusion of certain areas from the model, if this is the case;
- Species' ecological preferences according to the bibliography, to support the selection of variables.

Whatever the author's options, there should be a greater criticism of the obtained results, identifying putative constraints that may influence final results and the points can be improved in future studies.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ecologies5030025/s1, File S1.

**Author Contributions:** I.P.: conceptualization, methodology, writing—original draft. A.F.: conceptualization, writing—review and editing, supervision. A.M.A.: writing—review and editing; M.M.R.; writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by national funds through the Portuguese Foundation for Science and Technology (FCT) under the grant UI/BD/152853/2022, projects CERNAS-IPCB (UIDB/00681/2020) and CEF (UIDB/00239/2020), and the Centre of Studies in Geography and Spatial Planning (CEGOT) (UIDB/GEO/04084/2020_UC).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data cannot yet be provided due to the restricted access of our repository but will be archived upon acceptance.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Hao, T.; Elith, J.; Guillera-Arroita, G.; Lahoz-Monfort, J.J. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Divers. Distrib.* **2019**, *25*, 839–852. [CrossRef]
2. Sillero, N.; Arenas-Castro, S.; Enriquez-Urzelai, U.; Vale, C.G.; Sousa-Guedes, D.; Martínez-Freiría, F.; Real, R.; Barbosa, A. Want to model aspecies niche? A step-by-step guideline on correlative ecological niche modelling. *Ecol. Modell.* **2021**, *456*, 109671. [CrossRef]
3. Elith, J.; Leathwick, J.R. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 677–697. [CrossRef]
4. Guillera-Arroita, G. Modelling of species distributions, range dynamics and communities under imperfect detection: Advances, challenges and opportunities. *Ecography* **2017**, *40*, 281–295. [CrossRef]
5. Araújo, M.B.; Anderson, R.P.; Barbosa, A.M.; Beale, C.M.; Dormann, C.F.; Early, R.; Garcia, R.A.; Guisan, A.; Maiorano, L.; Naimi, B.; et al. Standards for distribution models in biodiversity assessments. *Sci. Adv.* **2019**, *5*, eaat4858. [CrossRef]
6. Feng, X.; Park, D.S.; Walker, C.; Peterson, A.T.; Merow, C.; Papeş, M. A checklist for maximizing reproducibility of ecological niche models. *Nat. Ecol. Evol.* **2019**, *3*, 1382–1395. [CrossRef]
7. Gómez-Pineda, E.; Blanco-García, A.; Lindig-Cisneros, R.; O'Neill, G.A.; Lopez-Toledo, L.; Sáenz-Romero, C. Pinus pseudostrobus assisted migration trial with rain exclusion: Maintaining Monarch Butterfly Biosphere Reserve forest cover in an environment affected by climate change. *New For.* **2021**, *52*, 995–1010. [CrossRef]
8. Guo, C.; Lek, S.; Ye, S.; Li, W.; Liu, J.; Li, Z. Uncertainty in ensemble modelling of large-scale species distribution: Effects from species characteristics and model techniques. *Ecol. Modell.* **2015**, *306*, 67–75. [CrossRef]
9. Merow, C.; Smith, M.J.; Edwards, T.C., Jr.; Guisan, A.; McMahon. S.M.; Normand, S.; Wilfried Thuiller, W.; Wüest, R.O.; Zimmermann, N.E.; Elith, J. What do we gain from simplicity versus complexity in species distribution models? *Ecography* **2014**, *37*, 1267–1281. [CrossRef]
10. Shen, Y.; Tu, Z.; Zhang, Y.; Zhong, W.; Xia, H.; Hao, Z.; Zhang, C.; Li, H. Predicting the impact of climate change on the distribution of two relict Liriodendron species by coupling the MaxEnt model and actual physiological indicators in relation to stress tolerance. *J. Environ. Manag.* **2022**, *322*, 116024. [CrossRef]
11. Xiao, F.; She, Y.; She, J.; Zhang, J.; Zhang, X.; Luo, C. Assessing habitat suitability and selecting optimal habitats for relict tree Cathaya argyrophylla in Hunan, China: Integrating pollen size, environmental factors, and niche modeling for conservation. *Ecol. Indic.* **2022**, *145*, 109669. [CrossRef]
12. Zurell, D.; Franklin, J.; König, C.; Bouchet, P.J.; Dormann, C.F.; Elith, J.; Fandos, G.; Feng, X.; Guillera-Arroita, G.; Guisan, A.; et al. A standard protocol for reporting species distribution models. *Ecography* **2020**, *43*, 1261–1277. [CrossRef]

13. Pecchi, M.; Marchi, M.; Burton, V.; Giannetti, F.; Moriondo, M.; Bernetti, I.; Bindi, M.; Chirici, G. Species distribution modelling to support forest management. A literature review. *Ecol. Model.* **2019**, *411*, 108817. [CrossRef]
14. Bogawski, P.; Damen, T.; Nowak, M.M.; Pędziwiatr, K.; Wilkin, P.; Mwachala, G.; Pierzchalska, J.; Wiland-Szymańska, J. Current and future potential distributions of three *Dracaena* Vand. ex L. species under two contrasting climate change scenarios in Africa. *Ecol. Evol.* **2019**, *9*, 6833–6848. [CrossRef] [PubMed]
15. Jian, S.; Zhu, T.; Wang, J.; Yan, D. The Current and Future Potential Geographical Distribution and Evolution Process of Catalpa bungei in China. *Forests* **2022**, *13*, 96. [CrossRef]
16. Almeida, A.M.; Martins, M.J.; Campagnolo, M.L.; Fernandez, P.; Albuquerque, T.; Gerassis, S.; Gonçalves, J.C.; Ribeiro, M.M. Prediction scenarios of past, present, and future environmental suitability for the Mediterranean species *Arbutus unedo* L. *Sci. Rep.* **2022**, *12*, 84. [CrossRef]
17. Dimobe, K.; Ouédraogo, A.; Ouédraogo, K.; Goetze, D.; Stein, K.; Schmidt, M.; Nacoulma, B.M.I.; Gnoumou, A.; Traoré, L.; Porembski, S.; et al. Climate change reduces the distribution area of the shea tree (*Vitellaria paradoxa* C.F. Gaertn.) in Burkina Faso. *J. Arid. Environ.* **2020**, *181*, 104237. [CrossRef]
18. Farahat, E.A.; Refaat, A.M. Predicting the impacts of climate change on the distribution of *Moringa peregrina* (Forssk.) Fiori—A conservation approach. *J. Mt. Sci.* **2021**, *18*, 1235–1245. [CrossRef]
19. Kumar, D.; Rawat, S.; Joshi, R. Predicting the current and future suitable habitat distribution of the medicinal tree *Oroxylum indicum* (L.) Kurz in India. *J. Appl. Res. Med. Aromat. Plants* **2021**, *39*, 100309. [CrossRef]
20. Mendoza-Fernández, A.J.; Martínez-Hernández, F.; Salmerón-Sánchez, E.; Pérez-García, F.J.; Teruel, B.; Merlo, M.E.; Mota, J.F. The relict ecosystem of maytenus senegalensis subsp. europaea in an agricultural landscape: Past, present and future scenarios. *Land* **2021**, *10*, 1. [CrossRef]
21. Rana, S.K.; Rana, H.K.; Stöcklin, J.; Ranjitkar, S.; Sun, H.; Song, B. Global warming pushes the distribution range of the two alpine 'glasshouse' Rheum species north- and upwards in the Eastern Himalayas and the Hengduan Mountains. *Front. Plant Sci.* **2022**, *13*, 925296. [CrossRef]
22. Tessarolo, G.; Lobo, J.M.; Rangel, T.F.; Hortal, J. High uncertainty in the effects of data characteristics on the performance of species distribution models. *Ecol. Indic.* **2021**, *121*, 107147. [CrossRef]
23. Warren, D.L.; Dornburg, A.; Zapfe, K.; Iglesias, T.L. The effects of climate change on Australia's only endemic Pokémon: Measuring bias in species distribution models. *Methods Ecol. Evol.* **2021**, *12*, 985–995. [CrossRef]
24. Araújo, M.B.; Guisan, A. Five (or so) challenges for species distribution modelling. *J. Biogeogr.* **2006**, *33*, 1677–1688. [CrossRef]
25. Araújo, M.B.; Peterson, A.T. Uses and misuses of bioclimatic envelope models. modeling. *Ecology* **2012**, *93*, 1527–1539. [CrossRef]
26. Fernandez, M.; Hamilton, H.; Kueppers, L.M. Characterizing uncertainty in species distribution models derived from interpolated weather station data. *Ecosphere* **2013**, *4*, 1–17. [CrossRef]
27. Heikkinen, R.K.; Luoto, M.; Araújo, M.B.; Virkkala, R.; Thuiller, W.; Sykes, M.T. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Prog. Phys. Geogr.* **2006**, *30*, 751–777. [CrossRef]
28. Beck, J.; Böller, M.; Erhardt, A.; Schwanghart, W. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecol. Inform.* **2014**, *19*, 10–15. [CrossRef]
29. Casajus, N.; Périé, C.; Logan, T.; Lambert, M.C.; de Blois, S.; Berteaux, D. An objective approach to select climate scenarios when projecting species distribution under climate change. *PLoS ONE* **2016**, *11*, e0152495. [CrossRef]
30. Dormann, C.F.; Schymanski, S.J.; Cabral, J.; Chuine, I.; Graham, C.; Hartig, F.; Kearney, M.; Morin, X.; Römermann, C.; Schröder, B.; et al. Correlation and process in species distribution models: Bridging a dichotomy. *J. Biogeogr.* **2013**, *39*, 2119–2131. [CrossRef]
31. Fourcade, Y.; Engler, J.O.; Rödder, D.; Secondi, J. Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS ONE* **2014**, *9*, e97122. [CrossRef] [PubMed]
32. Guevara, L.; Gerstner, B.E.; Kass, J.M.; Anderson, R.P. Toward ecologically realistic predictions of species distributions: A cross-time example from tropical montane cloud forests. *Glob. Chang. Biol.* **2018**, *24*, 1511–1522. [CrossRef]
33. Guevara, L.; León-Paniagua, L. How to survive a glaciation: The challenge of estimating biologically realistic potential distributions under freezing conditions. *Ecography* **2019**, *42*, 1237–1245. [CrossRef]
34. Marshall, B.M.; Strine, C.T. Exploring snake occurrence records: Spatial biases and marginal gains from accessible social media. *PeerJ* **2019**, *2019*, e8059. [CrossRef]
35. Moonlight, P.W.; de Miranda, P.L.S.; Cardoso, D.; Dexter, K.G.; Oliveira-Filho, A.T.; Pennington, R.T.; Ramos, G.; Särkinen, T.E. The strengths and weaknesses of species distribution models in biome delimitation. *Glob. Ecol. Biogeogr.* **2020**, *29*, 1770–1784. [CrossRef]
36. Rocchini, D.; Garzon-Lopez, C.X.; Marcantonio, M.; Amici, V.; Bacaro, G.; Bastin, L.; Brummitt, N.; Chiarucci, A.; Foody, G.M.; Hauffe, H.C.; et al. Anticipating species distributions: Handling sampling effort bias under a Bayesian framework. *Sci. Total Environ.* **2017**, *584–585*, 282–290. [CrossRef]
37. Sillero, N.; Barbosa, A.M. Common mistakes in ecological niche models. *Int. J. Geogr. Inf. Sci.* **2020**, *35*, 213–226. [CrossRef]
38. Silva, L.D.; de Azevedo, E.B.; Reis, F.V.; Elias, R.B.; Silva, L. Limitations of species distribution models based on available climate change data: A case study in the azorean forest. *Forests* **2019**, *10*, 575. [CrossRef]
39. Stohlgren, T.J.; Jarnevich, C.S.; Esaias, W.E.; Morisette, J.T. Bounding Species Distribution Models. 2011. Available online: http://groups.google.com/group/maxent/ (accessed on 7 October 2019).

40. ElQadi, M.M.; Dorin, A.; Dyer, A.; Burd, M.; Bukovac, Z.; Shrestha, M. Mapping species distributions with social media geo-tagged images: Case studies of bees and flowering plants in Australia. *Ecol. Inform.* **2017**, *39*, 23–31. [CrossRef]
41. Anderson, R.P.; Araújo, M.; Guisan, A.; Martínez-Meyer, E.; Lobo, J.M.; Peterson, A.T.; Soberón, J. Report of the Task Group on GBIF Data Fitness for Use in Distribution Modelling. 2016. Available online: http://www.gbif.org (accessed on 6 December 2023).
42. IPCC. *Climate Change 2022—Impacts, Adaptation and Vulnerability*; Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2022. [CrossRef]
43. O'Neill, B.C.; Tebaldi, C.; van Vuuren, D.P.; Eyring, V.; Friedlingstein, P.; Hurtt, G.; Knutti, R.; Kriegler, E.; Lamarque, J.-F.; Lowe, J.; et al. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model. Dev.* **2016**, *9*, 3461–3482. [CrossRef]
44. Moss, R.H.; Edmonds, J.A.; Hibbard, K.A.; Manning, M.R.; Rose, S.K.; Van Vuuren, D.P.; Carter, T.R.; Emori, S.; Kainuma, M.; Kram, T. The next generation of scenarios for climate change research and assessment. *Nature* **2010**, *463*, 747–756. [CrossRef] [PubMed]
45. Riahi, K.; Van Vuuren, D.P.; Kriegler, E.; Edmonds, J.; O'Neill, B.C.; Fujimori, S.; Bauer, N.; Calvin, K.; Dellink, R.; Fricko, O.; et al. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob. Environ. Chang.* **2017**, *42*, 153–168. [CrossRef]
46. van Vuuren, D.P.; Edmonds, J.; Kainuma, M.; Riahi, K.; Thomson, A.; Hibbard, K.; Hurtt, G.C.; Kram, T.; Krey, V.; Lamarque, J.-F.; et al. The representative concentration pathways: An overview. *Clim. Chang.* **2011**, *109*, 5–31. [CrossRef]
47. O'Neill, B.C.; Carter, T.R.; Ebi, K.; Harrison, P.A.; Kemp-Benedict, E.; Kok, K.; Kriegler, E.; Preston, B.L.; Riahi, K.; Sillmann, J.; et al. Achievements and needs for the climate change scenario framework. *Nat. Clim. Chang.* **2020**, *10*, 1074–1084. [CrossRef]
48. Beale, C.M.; Lennon, J.J. Incorporating uncertainty in predictive species distribution modelling. *Philos. Trans. R. Soc. B Biol. Sci.* **2012**, *367*, 247–258. [CrossRef]
49. McSweeney, C.F.; Jones, R.G.; Lee, R.W.; Rowell, D.P. Selecting CMIP5 GCMs for downscaling over multiple regions. *Clim. Dyn.* **2015**, *44*, 3237–3260. [CrossRef]
50. Barbet-Massin, M.; Thuiller, W.; Jiguet, F. How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography* **2010**, *33*, 878–886. [CrossRef]
51. Lázaro-Nogal, A.; Matesanz, S.; Hallik, L.; Krasnova, A.; Traveset, A.; Valladares, F. Population differentiation in a Mediterranean relict shrub: The potential role of local adaptation for coping with climate change. *Oecologia* **2016**, *180*, 1075–1090. [CrossRef]
52. Titeux, N.; Maes, D.; Van Daele, T.; Onkelinx, T.; Heikkinen, R.K.; Romo, H.; García-Barros, E.; Munguira, M.L.; Thuiller, W.; van Swaay, C.A.M.; et al. The need for large-scale distribution data to estimate regional changes in species richness under future climate change. *Divers. Distrib.* **2017**, *23*, 1393–1407. [CrossRef]
53. Chevalier, M.; Broennimann, O.; Cornuault, J.; Guisan, A. Data integration methods to account for spatial niche truncation effects in regional projections of species distribution. *Ecol. Appl.* **2021**, *31*, e02427. [CrossRef]
54. Chevalier, M.; Zarzo-Arias, A.; Guélat, J.; Mateo, R.G.; Guisan, A. Accounting for niche truncation to improve spatial and temporal predictions of species distributions. *Front. Ecol. Evol.* **2022**, *10*, 944116. [CrossRef]
55. Scherrer, D.; Esperon-Rodriguez, M.; Beaumont, L.J.; Barradas, V.L.; Guisan, A. National assessments of species vulnerability to climate change strongly depend on selected data sources. *Divers. Distrib.* **2021**, *27*, 1367–1382. [CrossRef]
56. Figueiredo, F.O.G.; Zuquim, G.; Tuomisto, H.; Moulatlet, G.M.; Balslev, H.; Costa, F.R.C. Beyond climate control on species range: The importance of soil data to predict distribution of Amazonian plant species. *J. Biogeogr.* **2018**, *45*, 190–200. [CrossRef]
57. Zuquim, G.; Costa, F.R.C.; Tuomisto, H.; Moulatlet, G.M.; Figueiredo, F.O.G. The importance of soils in predicting the future of plant habitat suitability in a tropical forest. *Plant Soil* **2020**, *450*, 151–170. [CrossRef]
58. Rovzar, C.; Gillespie, T.W.; Kawelo, K. Landscape to site variations in species distribution models for endangered plants. *For. Ecol. Manag.* **2016**, *369*, 20–28. [CrossRef]
59. Velazco, S.J.E.; Galvão, F.; Villalobos, F.; De Marco, P. Using worldwide edaphic data to model plant species niches: An assessment at a continental extent. *PLoS ONE* **2017**, *12*, e0186025. [CrossRef]
60. Feng, X.; Park, D.S.; Liang, Y.; Pandey, R.; Papeş, M. Collinearity in ecological niche modeling: Confusions and challenges. *Ecol. Evol.* **2019**, *9*, 10365–10376. [CrossRef] [PubMed]
61. Sehler, R.; Li, J.; Reager, J.; Ye, H. Investigating Relationship Between Soil Moisture and Precipitation Globally Using Remote Sensing Observations. *J. Contemp. Water Res. Educ.* **2019**, *168*, 106–118. [CrossRef]
62. de Tomás Marín, S.; Rodríguez-Calcerrada, J.; Arenas-Castro, S.; Prieto, I.; González, G.; Gil, L.; de la Riva, E.G. Fagus sylvatica and Quercus pyrenaica: Two neighbors with few things in common. *For. Ecosyst.* **2023**, *10*, 100097. [CrossRef]
63. Vitasse, Y.; Bresson, C.C.; Kremer, A.; Michalet, R.; Delzon, S. Quantifying phenological plasticity to temperature in two temperate tree species. *Funct. Ecol.* **2010**, *24*, 1211–1218. [CrossRef]
64. García-Díaz, P.; Prowse, T.A.A.; Anderson, D.P.; Lurgi, M.; Binny, R.N.; Cassey, P. A concise guide to developing and using quantitative models in conservation management. *Conserv. Sci. Pract.* **2019**, *1*, e11. [CrossRef] [PubMed]
65. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]

66. Petitpierre, B.; Broennimann, O.; Kueffer, C.; Daehler, C.; Guisan, A. Selecting predictors to maximize the transferability of species distribution models: Lessons from cross-continental plant invasions. *Glob. Ecol. Biogeogr.* **2017**, *26*, 275–287. [CrossRef]
67. Tanner, E.P.; Papeş, M.; Elmore, R.D.; Fuhlendorf, S.D.; Davis, C.A. Incorporating abundance information and guiding variable selection for climate-based ensemble forecasting of species' distributional shifts. *PLoS ONE* **2017**, *12*, e0184316. [CrossRef]
68. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol. Evol.* **2012**, *3*, 327–338. [CrossRef]
69. Guerrina, M.; Conti, E.; Minuto, L.; Casazza, G. Knowing the past to forecast the future: A case study on a relictual, endemic species of the SW Alps, Berardia subacaulis. *Reg. Environ. Chang.* **2015**, *16*, 1035–1045. [CrossRef]
70. Valavi, R.; Guillera-Arroita, G.; Lahoz-Monfort, J.J.; Elith, J. Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecol. Monogr.* **2022**, *92*, e01486. [CrossRef]
71. Qazi, A.W.; Saqib, Z.; Zaman-ul-Haq, M. Trends in species distribution modelling in context of rare and endemic plants: A systematic review. *Ecol. Process.* **2022**, *11*, 40. [CrossRef]
72. Alvarado-Serrano, D.F.; Knowles, L.L. Ecological niche models in phylogeographic studies: Applications, advances and precautions. *Mol. Ecol. Resour.* **2014**, *14*, 233–248. [CrossRef]
73. Sor, R.; Park, Y.S.; Boets, P.; Goethals, P.L.M.; Lek, S. Effects of species prevalence on the performance of predictive models. *Ecol. Modell.* **2017**, *354*, 11–19. [CrossRef]
74. Lawson, C.R.; Hodgson, J.A.; Wilson, R.J.; Richards, S.A. Prevalence, thresholds and the performance of presence-absence models. *Methods Ecol. Evol.* **2014**, *5*, 54–64. [CrossRef]
75. Lobo, J.M.; Jiménez-valverde, A.; Real, R. AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **2008**, *17*, 145–151. [CrossRef]
76. Sofaer, H.R.; Hoeting, J.A.; Jarnevich, C.S. The area under the precision-recall curve as a performance metric for rare binary events. *Methods Ecol. Evol.* **2019**, *10*, 565–577. [CrossRef]
77. Jiménez-Valverde, A. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Glob. Ecol. Biogeogr.* **2012**, *21*, 498–507. [CrossRef]
78. Leroy, B.; Delsol, R.; Hugueny, B.; Meynard, C.N.; Barbet-Massin, M.; Bellard, C. Title: Without quality presence-absence data, discrimination metrics such as TSS can be misleading measures of model performance. *J. Biogeogr.* **2018**, *45*, 1994–2002. [CrossRef]
79. Rapacciuolo, G. Strengthening the contribution of macroecological models to conservation practice. *Glob. Ecol. Biogeogr.* **2019**, *28*, 54–60. [CrossRef]
80. Amobonye, A.; Lalung, J.; Mheta, G.; Pillai, S. *Writing a Scientific Review Article: Comprehensive Insights for Beginners*; Hindawi Limited: London, UK, 2024. [CrossRef]
81. Zhu, J.; Liu, W. A tale of two databases: The use of Web of Science and Scopus in academic papers. *Scientometrics* **2020**, *123*, 321–335. [CrossRef]
82. Pranckutė, R. Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications* **2021**, *9*, 12. [CrossRef]
83. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Med.* **2021**, *18*, e1003583. [CrossRef]
84. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *29*, 372. [CrossRef]
85. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]
86. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2005**, *25*, 1965–1978. [CrossRef]
87. Platts, P.J.; Omeny, P.A.; Marchant, R. AFRICLIM: High-resolution climate projections for ecological applications in Africa. *Afr. J. Ecol.* **2015**, *53*, 103–108. [CrossRef]
88. Wouyou, H.G.; Lokonon, B.E.; Idohou, R.; Zossou-Akete, A.G.; Assogbadjo, A.E.; Kakaï, R.G. Predicting the potential impacts of climate change on the endangered *Caesalpinia bonduc* (L.) Roxb in Benin (West Africa). *Heliyon* **2022**, *8*, e09022. [CrossRef] [PubMed]
89. Xiao, J.; Eziz, A.; Zhang, H.; Wang, Z.; Tang, Z.; Fang, J. Responses of four dominant dryland plant species to climate change in the Junggar Basin, northwest China. *Ecol. Evol.* **2019**, *9*, 13596–13607. [CrossRef] [PubMed]
90. IPCC. *Climate Change 2007: Synthesis Report*; Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change; Intergovernmental Panel on Climate Change: Geneva, Switzerland, 2007.
91. IPCC. *Climate Change 2014: Synthesis Report*; Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change; Intergovernmental Panel on Climate Change: Geneva, Switzerland, 2014.
92. Henckel, L.; Bradter, U.; Jönsson, M.; Isaac, N.J.B.; Snäll, T. Assessing the usefulness of citizen science data for habitat suitability modelling: Opportunistic reporting versus sampling based on a systematic protocol. *Divers. Distrib.* **2020**, *26*, 1276–1290. [CrossRef]

93. Mateo, R.G.; Gastón, A.; Aroca-Fernández, M.J.; Saura, S.; García-Viñas, J.I. Optimization of forest sampling strategies for woody plant species distribution modelling at the landscape scale. *For. Ecol. Manag.* **2018**, *410*, 104–113. [CrossRef]
94. Seoane, J.; Estrada, A.; Jones, M.M.; Ovaskainen, O. A case study on joint species distribution modelling with bird atlas data: Revealing limits to species' niches. *Ecol. Inform.* **2023**, *77*, 102202. [CrossRef]
95. Ochoa-Ochoa, L.M.; Flores-Villela, O.A.; Bezaury-Creel, J.E. Using one vs. many, sensitivity and uncertainty analyses of species distribution models with focus on conservation area networks. *Ecol. Modell.* **2016**, *320*, 372–382. [CrossRef]
96. Xu, W.; Zhu, S.; Yang, T.; Cheng, J.; Jin, J. Maximum Entropy Niche-Based Modeling for Predicting the Potential Suitable Habitats of a Traditional Medicinal Plant (Rheum nanum) in Asia under Climate Change Conditions. *Agriculture* **2022**, *12*, 610. [CrossRef]
97. Zhao, Q.; Zhang, Y.; Li, W.-N.; Hu, B.-W.; Zou, J.-B.; Wang, S.-Q.; Niu, J.-F.; Wang, Z.-Z. Predicting the potential distribution of perennial plant coptis chinensis franch. In china under multiple climate change scenarios. *Forests* **2021**, *12*, 1464. [CrossRef]
98. Bell, D.M.; Schlaepfer, D.R. On the dangers of model complexity without ecological justification in species distribution modeling. *Ecol. Modell.* **2016**, *330*, 50–59. [CrossRef]
99. Kopsco, H.L.; Smith, R.L.; Halsey, S.J. A Scoping Review of Species Distribution Modeling Methods for Tick Vectors. *Front. Ecol. Evol.* **2022**, *10*, 893016. [CrossRef]
100. Phillips, S.B.; Aneja, V.P.; Kang, D.; Arya, S.P. Modelling and analysis of the atmospheric nitrogen deposition in North Carolina. *Ecol. Model.* **2006**, *190*, 231–252. [CrossRef]
101. Phillips, S.J.; Anderson, R.P.; Dudík, M.; Schapire, R.E.; Blair, M.E. Opening the black box: An open-source release of Maxent. *Ecography* **2017**, *40*, 887–893. [CrossRef]
102. Elith, J.; Graham, C.P.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A.; et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151. [CrossRef]
103. Elith, J.; Kearney, M.; Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **2010**, *1*, 330–342. [CrossRef]
104. Mauri, A.; Girardello, M.; Forzieri, G.; Manca, F.; Beck, P.S.; Cescatti, A.; Strona, G. Assisted tree migration can reduce but not avert the decline of forest ecosystem services in Europe. *Glob. Environ. Chang.* **2023**, *80*, 102676. [CrossRef]
105. Sousa-Silva, R.; Verbist, B.; Lomba, Â.; Valent, P.; Suškevičs, M.; Picard, O.; Hoogstra-Klein, M.A.; Cosofret, V.-C.; Bouriaud, L.; Ponette, Q.; et al. Adapting forest management to climate change in Europe: Linking perceptions to adaptive responses. *For. Policy Econ.* **2018**, *90*, 22–30. [CrossRef]
106. Kumar, D.; Singh, M.; Sharma, S. Fate of important medicinal plants in the eastern himalaya in changing climate scenarios: A case of Panax pseudoginseng wall. *Appl. Ecol. Environ. Res.* **2019**, *17*, 13493–13511. [CrossRef]
107. Shrestha, B.; Tsiftsis, S.; Chapagain, D.J.; Khadka, C.; Bhattarai, P.; Shrestha, N.K.; Kolanowska, M.A.; Kindlmann, P. Suitability of habitats in nepal for dactylorhiza hatagirea now and under predicted future changes in climate. *Plants* **2021**, *10*, 467. [CrossRef] [PubMed]
108. Wan, J.N.; Mbari, N.J.; Wang, S.-W.; Liu, B.; Mwangi, B.N.; Rasoarahona, J.R.; Xin, H.-P.; Zhou, Y.-D.; Wang, Q.-F. Modeling impacts of climate change on the potential distribution of six endemic baobab species in Madagascar. *Plant Divers.* **2021**, *43*, 117–124. [CrossRef] [PubMed]
109. Yadav, S.; Bhattacharya, P.; Areendran, G.; Sahana, M.; Raj, K.; Sajjad, H. Predicting impact of climate change on geographical distribution of major NTFP species in the Central India Region. *Model. Earth Syst. Environ.* **2022**, *8*, 449–468. [CrossRef]
110. Domisch, S.; Kuemmerlen, M.; Jähnig, S.C.; Haase, P. Choice of study area and predictors affect habitat suitability projections, but not the performance of species distribution models of stream biota. *Ecol. Modell.* **2013**, *257*, 1–10. [CrossRef]