

Article

Comparing the Performance of CMCC-BioClimInd and WorldClim Datasets in Predicting Global Invasive Plant Distributions

Feixue Zhang ^{1,2}, Chunjing Wang ², Chunhui Zhang ^{1,*} and Jizhong Wan ^{1,*} ¹ State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 810016, China² College of Agriculture and Animal Husbandry, Qinghai University, Xining 810016, China

* Correspondence: zhangcqh@163.com (C.Z.); wan1276@163.com or wan1276@qhu.edu.cn (J.W.)

Simple Summary: The species distribution model has been widely used to predict the distribution of invasive plant species based on bioclimatic variables. However, the specific selection of bioclimate variables may affect the performance of the species distribution model. Here, we tested a new bioclimate variable dataset (i.e., CMCC BioClimInd) and used it in the species distribution model. We evaluated the predictive performance and explanatory power of WorldClim and CMCC-BioClimInd using AUC and omission rate, and also used the ODMAP protocol to record CMCC-BioClimInd to ensure reproducibility. The results indicate that CMCC BioClimInd can effectively simulate the distribution of invasive plant species. Based on the contribution rate of CMCC-BioClimInd to the distribution of invasive plant species, we inferred that the modified simplified continentality index and modified Kira warmth index from CMCC-BioClimInd had a strong explanatory power. Under the 35 bioclimatic variables of CMCC-BioClimInd, alien invasive plant species are mainly distributed in equatorial, tropical and subtropical regions. We tested a new bioclimate variable dataset to simulate the distribution of invasive plant species worldwide. Our research provides a new perspective for risk assessment and management of global invasive plant species.



Citation: Zhang, F.; Wang, C.; Zhang, C.; Wan, J. Comparing the Performance of CMCC-BioClimInd and WorldClim Datasets in Predicting Global Invasive Plant Distributions. *Biology* **2023**, *12*, 652. <https://doi.org/10.3390/biology12050652>

Academic Editors: Jianfan Sun and Xiao Guo

Received: 15 March 2023

Revised: 19 April 2023

Accepted: 23 April 2023

Published: 26 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Species distribution modeling (SDM) has been widely used to predict the distribution of invasive plant species based on bioclimatic variables. However, the specific selection of these variables may affect the performance of SDM. This investigation elucidates a new bioclimate variable dataset (i.e., CMCC-BioClimInd) for its use in SDM. The predictive performance of SDM that includes WorldClim and CMCC-BioClimInd was evaluated by AUC and omission rate and the explanatory power of both datasets was assessed by the jackknife method. Furthermore, the ODMAP protocol was used to record CMCC-BioClimInd to ensure reproducibility. The results indicated that CMCC-BioClimInd effectively simulates invasive plant species' distribution. Based on the contribution rate of CMCC-BioClimInd to the distribution of invasive plant species, it was inferred that the modified and simplified continentality and Kira warmth index from CMCC-BioClimInd had a strong explanatory power. Under the 35 bioclimatic variables of CMCC-BioClimInd, alien invasive plant species are mainly distributed in equatorial, tropical, and subtropical regions. We tested a new bioclimate variable dataset to simulate the distribution of invasive plant species worldwide. This method has great potential to improve the efficiency of species distribution modeling, thereby providing a new perspective for risk assessment and management of global invasive plant species.

Keywords: AUC; CMCC-BioClimInd; global scale; invasive plants; species distribution modelling; Maxent; WorldClim

1. Introduction

Invasive plant species (IPS) are a global problem affecting agriculture, forestry, fisheries, human health, and natural ecosystems [1–3]. Climate change affects the identified IPS niche, thereby affecting their regional and global distributions [4,5]. The global average

surface temperature has increased by about 0.6% in the past century [6,7]. The average climate model mainly represents the temperature and precipitation cycles within a year (month to season) and directly affects biological species' distribution, abundance, and interaction [6,8]. Furthermore, climatic change has damaged the ecological balance [9,10]. Climate change affects invasive plants' growth, development, physiological characteristics, and distribution pattern [1,5,6,11]. The growth and resistance of some species at the edge of the climate zone are weakened, and they will be more vulnerable to damage from invasive organisms [1,5,11]. Additionally, IPS often hinder local species from acquiring adequate nutrients, water, and light [12,13]. They also alter the invaded soils' chemistry, hydrology, and water-holding capacity, thereby altering local plant community dynamics [14,15]. Therefore, reliable intrusion prediction methods are urgently required to effectively monitor and formulate reasonable management policies to reduce IPS spread, expansion, and risk [16–21]. Climate data sets have long been the basis for describing the native distribution of species [22] to solve problems based on protection [22–25]. Climate variables are also used to simulate invasive alien species' potential distribution to better understand their evolution [21,26,27] and to assist biological control projects [19,28]. The global conformal data set is crucial to simulate the potential distribution of species in different continents [29].

Recently, multiple studies have outlined and compared species distribution modeling (SDM) methods [7,30]. SDM essentially predicts and describes the regional conditions suitable for species survival under climate change [31,32]. Environmental data especially climate data used for modeling, are among the least studied sources for SDM uncertainty [30]. With the increased availability of climate information, several global bioclimatic index datasets are being used as references by researchers [33,34]. WorldClim [29,35,36] is the most outstanding global climate dataset, often used in other applications, such as for ecological hydrology [37], comparisons of climatology [37,38], and the assessment of carbon stocks and their dynamics [39]. The WorldClim dataset includes 19 bioclimatic variables [40] calculated from monthly average temperature and total precipitation, including data from the global historical climate network dataset, which nearly meet distribution modeling requirements based on regression for most species [29].

The WorldClim dataset is valuable because of its refined resolution, including four spatial resolutions ranging from 30 s (~1 km²) to 10 min (~340 km²) [29,41]. However, it lacks the spatial details usually needed by models to evaluate impact, thereby fundamentally limiting its prediction for the essential spatial heterogeneity in steep mountain terrain [29,41]. In terms of spatial heterogeneity, a medium-resolution grid cell still spans the climate environment with an elevation difference of hundreds of meters [29,40]. Furthermore, it was also suggested that the 19 indicators of bioclimate variables were insufficient to simulate the distribution of plants and other organisms because plant growth requires more than precipitation and temperature [42]. For example, although there is a lot of precipitation in some areas, there is relatively increased evaporation; therefore, the water is transferred away from plants in different ways [43].

This paper describes the development and testing of a global dataset that can expand the availability of spatial information to regions [44–46] by testing a new indicator called the CMCC-BioClimInd [46], comprising 35 variables with a spatial resolution of 0.5°, including the historical period (1960–1999) and two future time ranges (2040–2079 and 2060–2099). The CMCC-BioClimInd data set is a collection of 11 CMIP5 climate simulations obtained by improving climate reanalysis. Research shows that the impact of climate change on the protection and management of wildlife, plants, and natural resources can be accurately estimated in different spatial ranges and research fields [46]. The CMCC-BioClimInd can more easily and quickly infer the relationship between the study's subject and climate variables [47] and prevent the variability caused by the future trajectory and the physical properties of different indicators [48].

Noce et al. (2020) analyzed correlations between climate variables in two models (i.e., CMCC-BioClimInd and WorldClim) and the cause of their differences. They proved

that CMCC-BioClimInd is accurate, comprehensive, and effective for predicting invasive species distributions and simulating climate change [46]. To our knowledge, no one has used CMCC-BioClimInd to predict invasive plant distributions nor applied it in practice. WorldClim is the most used dataset to predict species and potential invasive plant species distributions, which is greatly innovative (e.g., [2,15,32,49]).

Although species distribution models are widely used, the reproducibility of SDM methods is often limited due to the lack of reporting standards and the uncertainty of their predictions [7,30,31]. Therefore, here the ODMAP scheme was used to enhance the rationality and repeatability of this research [50,51]. The ODMAP (overview, data, model, evaluation, and prediction) reporting protocol provides a standardized way to communicate SDM results and outputs by describing objectives, model assumptions, scaling issues, data sources, model workflows, model predictions, and uncertainties [50,51]. The ODMAP protocol has two main purposes. First, it provides a checklist detailing the key steps of model construction and analysis to authors. Second, it introduces a standard documentation method to ensure transparency and repeatability [50,51]. Here, we tested the CMCC-BioClimInd dataset, described its basic elements, and detailed metadata based on the ODMAP (provided in Table S1).

We screened 11 most representative species of the 100 most dangerous alien invasive species in the world. We introduced a new global dataset of bioclimatic indicators to predict the distribution of invasive species and compared it with WorldClim (most used) to verify the prediction quality of the CMCC-BioClimInd data set and its results. This study specifically aimed to (a) compare the prediction performance of invasive species distributions by comparing WorldClim and CMCC-BioClimInd AUC values; (b) identify the most effective CMCC-BioClimInd variables affecting the distribution of invasive species; (c) determine the potential IPS distributions based on CMCC-BioClimInd; and (d) evaluate bioclimatic variables causing differences in IPS distributions for the same first 19 variables of WorldClim and CMCC-BioClimInd.

2. Materials and Methods

2.1. Occurrence Data

According to the expert group on invasive species, the world's most invasive non-native species were compiled, among which the most important 11 IPS [52] were *Ligustrum robustum*, *Cinchona pubescens*, *Morella faya*, *Miconia calvescens*, *Cecropia peltate*, *Spathodea campanulata*, *Melaleuca quinquenervia*, *Schinus terebinthifolia*, *Acacia mearnsii*, *Leucaena leucocephala*, and *Pinus pinaster*. The occurrence records of these species were downloaded from the Global Biodiversity Information Facility (GBIF; <https://www.gbif.org>; <https://doi.org/10.15468/dl.uxpqxy>, <https://doi.org/10.15468/dl.hrhg96>, <https://doi.org/10.15468/dl.9up5qj>, <https://doi.org/10.15468/dl.c2r3na>, <https://doi.org/10.15468/dl.qpgg5m>, <https://doi.org/10.15468/dl.fm3kec>, <https://doi.org/10.15468/dl.h77k3r>, <https://doi.org/10.15468/dl.8amej>, <https://doi.org/10.15468/dl.ftac38>, <https://doi.org/10.15468/dl.z84z4x> and <https://doi.org/10.15468/dl.h6h4fe> (accessed on 2 April 2022); Figure 1). We downloaded species distribution data from 1970 to 1999 on GBIF, because only between 1970 and 1999 did the time periods of the two datasets coincide. The data we downloaded were processed as follows: (1) Carefully checked and screened the inaccurate or heteronymous species; (2) Deleted records with the same longitude and latitude; (3) Deleted duplicate records in a specific spatial resolution area [2,51]. Finally, a total of 390,000 geographical coordinate points of these 11 species were included in our analysis [53,54].

2.2. Data of Bioclimatic Variables

The origin of CMCC-BioClimInd is based on the daily time series of temperature and precipitation available in the weather data set of Water and European Medium-Range Weather Forecast Centre (ECMWF) reanalysis (ERA-40) [55], which is described in detail for the historical period (http://www.eu-watch.org/data_availability, accessed on

4 April 2022). The CMCC-BioClimInd data set is from (<https://doi.org/10.1594/PANGAEA.904278>, accessed on 5 April 2022; [46]). We obtained a set of 35 bioclimatic variables with a spatial resolution of $0.5^\circ \times 0.5^\circ$ (1960–1999), covering the entire world (excluding Antarctica) [46]. WorldClim downloaded from (<https://www.worldclim.org>, accessed on 6 April 2022). A set of 5 arc minutes ($10 \times 10 \text{ km}^2$) spatial resolution for 19 bioclimatic variables (1970–2000) [29,35,36] required the authors to resample the WorldClim climate variable to 0.5° resolution in order to be consistent with the resolution of CMCC-BioClimInd climate variable. The details of bioclimatic variables were shown in Table S2.

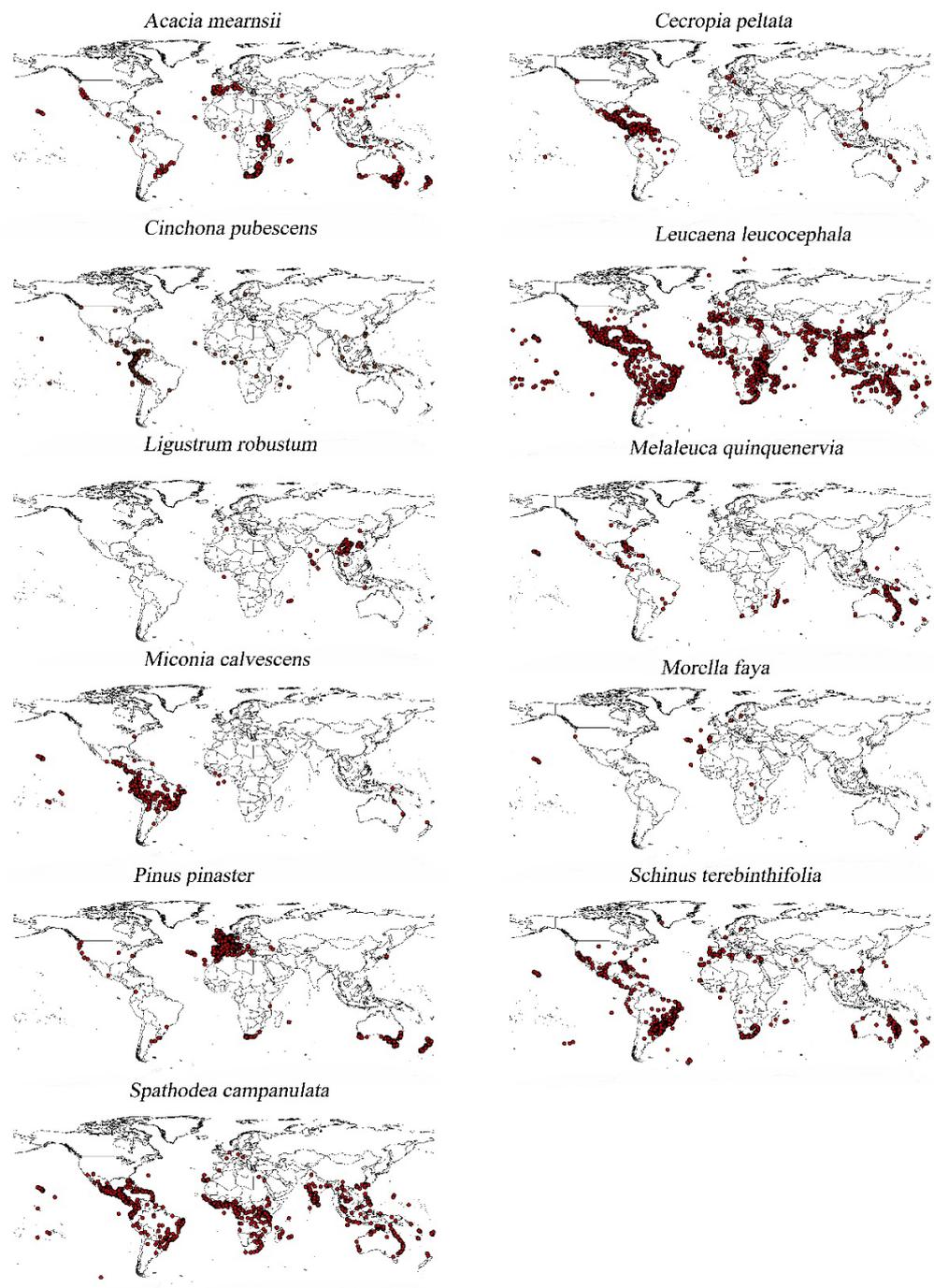


Figure 1. Geographical coordinates of the distribution of 11 IPS.

We ran the Maxent model four times, namely, CMCC-BioClimInd, CMCC-BioClimInd (bio1–bio19), CMCC-BioClimInd (bio20–bio35), and WorldClim. By running it four times,

we clearly compared CMCC-BioClimInd and WorldClim performance and verified the invasive species distribution prediction accuracy. Running CMCC-BioClimInd (bio1–bio19) compared the climate variable difference on invasive species distribution to WorldClim. Running CMCC-BioClimInd (bio20–bio35) revealed the impact of the new variables on invasive species distribution in addition to the first 19 variables, and more fully revealed the effectiveness of the new variables.

2.3. Modelling Approach and Evaluation

Based on species occurrence data and relevant environmental variables, the Maxent model is used to model species distribution under climate change [56,57]. Here, we established a logistic regression model with data from the 11 IPS distributions as response variables, and by running the Maxent model four times using the climate variables in the four climate data sets, namely, WorldClim and CMCC-BioClimInd (bio1–bio35), CMCC-BioClimInd (bio1–bio19), and CMCC-BioClimInd (bio20–bio35). The IPS distributional data were divided into a random training test set (auctest, 75%) and a test model set (auctrain, 25%). The regularization multiplier was set to two and the number of replicates to four [56,58,59].

We use the area under the curve (AUC) of the receiver's operating characteristics to evaluate the prediction accuracy of the species distribution model. The AUC takes each value of the prediction result as a possible threshold, and then obtains the corresponding sensitivity and specificity values to calculate the curve [58]. The greater the AUC value, the greater the deviation between species distribution and random distribution (i.e., $AUC = 0.5$; [57,59]). The greater the correlation between variables and models, the higher the accuracy of the models. An $AUC > 0.7$ indicates that the model is effective [57]. We have added the omission rate test metric. The omission rate refers to the proportion of evaluation areas that are not within the scope of the model once converted to binary prediction [60,61]. The omission rate provides information about discrimination and overfitting evaluated under specific thresholds. Generally speaking, the lower the omission rate, the higher the performance [60,61]. We evaluate the performance of the model through the AUC and omission, which has a certain degree of scientific accuracy.

2.4. Effects of Bioclimatic Variables on Global Invasive Plant Species Distributions

Firstly, we used a jackknife method to assess bioclimatic variable contribution to the species' distribution probability. The jackknife method output format showed the bioclimatic variables of each data set to the distribution probability, with values ranging from 0 (representing the smallest contribution) to 100% (representing the largest contribution) [61,62]. Secondly, we used an independent sample *t*-test [62] to compare the contribution rates of 19 bioclimatic variables in WorldClim (bio1–bio19) and CMCC-BioClimInd (bio1–bio19). We evaluated the difference of the average contribution rate of the first 19 bioclimatic variables to the distributions of IPS between the two models. Finally, after running Maxent, we generated ASCII files for both models (CCMC-BioClimInd and WorldClim). In GIS, we used mathematical analysis to subtract the WorldClim invasive plant distribution probability map from the CCMC-BioClimInd map [46]. We then obtained the distribution difference map for the 11 invasive plants. Positive values indicated that the predicted CMCC-BioClimInd distribution probability in a specific area was higher than WorldClim [46]. The opposite was true for negative values [46].

3. Results

3.1. Importance of Bioclimatic Variables for the Distribution of Invasive Species

Four different data sets were used, and the Maxent model was run four times. Their average AUC values were: CMCC-BioClimInd (bio20–bio35) > CMCC-BioClimInd (bio1–bio35) > CMCC-BioClimInd (bio1–bio19) > WorldClim (bio1–bio19) (Table 1). These four models performed well, with an average AUC above 0.940. Comparison of their average omission rates: CMCC-BioClimInd (bio20–bio35) < CMCC-BioClimInd (bio1–

bio35) < CMCC-BioClimInd (bio1–bio19) < WorldClim (bio1–bio19). The omission rates of all four datasets were less than 0.15, indicating that the CMCC-BioClimInd dataset effectively simulates invasive species distribution (Table 1).

Table 1. AUC values and omission rate of four climate data sets for modelling 11 invasive plant species distributions based on MaxEnt SDM. OR: omission rate; CMCC: CMCC-BioClimInd.

Species	WorldClim	OR	CMCC (bio1–bio19)	OR	CMCC (bio1–bio35)	OR	CMCC (bio20–bio35)	OR
<i>Acacia mearnsii</i>	0.953	0.102	0.963	0.069	0.985	0.052	0.984	0.051
<i>Cecropia peltata</i>	0.954	0.090	0.961	0.059	0.979	0.042	0.980	0.043
<i>Cinchona pubescens</i>	0.969	0.081	0.977	0.054	0.966	0.055	0.972	0.056
<i>Leucaena leucocephala</i>	0.861	0.213	0.892	0.164	0.958	0.079	0.956	0.081
<i>Ligustrum robustum</i>	0.980	0.052	0.974	0.054	0.980	0.050	0.977	0.049
<i>Melaleuca quinquenervia</i>	0.967	0.082	0.967	0.059	0.972	0.056	0.975	0.054
<i>Miconia calvescens</i>	0.953	0.099	0.963	0.063	0.978	0.041	0.975	0.044
<i>Morella faya</i>	0.977	0.075	0.927	0.124	0.863	0.162	0.927	0.124
<i>Pinus pinaster</i>	0.962	0.073	0.955	0.068	0.976	0.046	0.976	0.046
<i>Schinus terebinthifolia</i>	0.941	0.116	0.936	0.100	0.963	0.076	0.962	0.074
<i>Spathodea campanulata</i>	0.913	0.160	0.944	0.095	0.975	0.055	0.974	0.057
Average of all species	0.948	0.104	0.951	0.083	0.963	0.065	0.969	0.062

Among the 11 invasive plants, the average contribution of different climate variables, or different climate data sets, to the probability of invasive plants differed. First, we compared the average climate variable contribution rates for invasive plants between WorldClim (bio1–bio19) and CMCC-BioClimInd (bio1–bio19). In WorldClim, temperature seasonality (bio4), mean temperature of the coldest quarter (bio11), and isothermality (bio3) contributed more than 10% to the invasive species distribution (Table 2). In CMCC-BioClimInd (bio1–bio19), bio4 and annual mean temperature (bio1) had the greatest impact on invasive species distribution, with values of 23.9% and 14.131%, respectively (Table 2). Regardless of the model selection, bio4 was the most critical factor affecting invasive species (Table 2). Additionally, the average contribution rates of bio1, mean temperature of warmest quarter (bio10), and precipitation of the driest quarter (bio17) in predicting invasive species distribution probabilities in the WorldClim (bio1–bio19) and CMCC-BioClimInd (bio1–bio19) models differed significantly ($p < 0.05$; Table S3).

We used the complete CMCC-BioClimInd (bio1–bio35) to analyze the average contribution to the 11 invasive species. Surprisingly, the average contribution rate of 16 new variables (bio20–bio35) for invasive species distribution reached 56.732%, while the original 19 variables only reached 43.268% (Table 2). Here, the modified Kira warmth index (bio26), simplified continentality index (bio27), and bio4 all contributed more than 10% to invasive species distribution, including bio26 (13.883%), bio27 (12.322%), and bio4 (10.774%) (Table 2). To ensure the accuracy of these 16 variables for predicting the invasive species distribution and to make the results more intuitive, we ran CMCC-BioClimInd (bio20–bio35) separately. The results were consistent with the bioclimatic variable contribution rate of the complete CMCC-BioClimInd set. Moreover, bio26 and bio27 contributed markedly more to invasive species distribution (Table 2).

3.2. Distribution Probability of Invasive Species

The predicted species' distribution ranges were roughly similar as assessed by running the distribution probability map of the Maxent model four times with different climate variables. They were also concentrated in the same region. However, they were not completely consistent (Figure 2). For these 11 species, the CMCC-BioClimInd (bio1–bio19) distribution ranges were significantly larger than WorldClim (bio1–bio19) (Figure 2). Moreover, the predicted climate variable distribution ranges of WorldClim were larger than those of CMCC-BioClimInd; however, the differences were small (Figures 1 and 3). However,

CMCC-BioClimInd and CMCC-BioClimInd (bio20–bio35) similarly predicted the invasive plant distributions and distribution probabilities (Figure 2).

Table 2. Average contribution of climatic variables from four climate datasets to the invasive plant species distributions across 11 species. Bold values mean a large contribution to the invasive plant species distributions.

Climate Variables	WorldClim (%)	CMCC-BioclimInd bio1–bio19 (%)	CMCC-BioclimInd bio1–bio35 (%)	CMCC-BioclimInd bio20–bio35 (%)
bio1	5.412	14.131	2.758	
bio2	2.477	1.819	3.113	
bio3	12.991	7.391	2.776	
bio4	20.472	23.900	10.774	
bio5	1.406	1.223	0.094	
bio6	7.112	4.134	0.130	
bio7	2.877	3.721	1.736	
bio8	1.021	0.553	0.023	
bio9	0.225	0.147	0.001	
bio10	0.556	4.836	1.878	
bio11	14.020	3.116	1.189	
bio12	5.255	3.440	0.657	
bio13	2.541	1.652	0.484	
bio14	3.721	5.371	2.726	
bio15	1.944	4.471	2.554	
bio16	4.865	8.741	1.140	
bio17	0.401	3.025	2.390	
bio18	5.544	2.950	3.433	
bio19	7.161	5.381	5.411	
bio20			1.004	1.590
bio21			1.293	5.250
bio22			0.041	2.386
bio23			3.910	7.617
bio24			4.795	8.041
bio25			9.401	10.085
bio26			13.883	16.000
bio27			12.322	28.726
bio28			0.500	2.042
bio29			0.279	1.939
bio30			0.129	0.757
bio31			0.198	0.692
bio32			0.054	0.176
bio33			5.633	6.748
bio34			0.793	4.807
bio35			2.497	3.143

The complete CMCC-BioClimInd set was considered the criterion since it was more comprehensive and accurate for invasive species prediction. *Acacia mearnsii* is distributed mainly in western and eastern South America, southern Australia, eastern Africa, and the western Mediterranean (Figure 2). *Cecropia peltata* is primarily located near the equator, especially in northern South America, with a high distribution probability (Figure 2). *Cinchona pubescens* is distributed mainly in western South America and central Africa, and the IPS main distribution range is also near the tropics (Figure 2). *Leucaena leucocephala* is widely distributed in the southern hemisphere, specifically in central Africa, Maldives, northern and southern South America, southern Asia, northeastern and western Australia (Figure 2). *Ligustrum robustum* is mainly distributed in southeast China, Southeast Asia, and eastern India. *Melaleuca quinquenervia* is distributed mainly near the Tropic of Capricorn and Cancer, with a small distribution range primarily concentric in the Maldives and Brazil (Figure 2). *Miconia calvescens* is distributed mainly between the equator and the Tropic

of Cancer in Brazil, Peru, and central Africa (Figure 2). *Morella faya* showed a relatively scattered small distribution range (Figure 2) and only a few countries in the world are affected by it, being primarily distributed in southwest Spain, the Azores, Madeira, and the Canary Islands (Figure 2). *Pinus pinaster* is mainly distributed in the Mediterranean basin, southern Australia, and northeastern New Zealand, with a small distribution range (Figure 3). *Schinus terebinthifolia* is distributed near the Tropic of Cancer and the Tropic of Capricorn, mainly in eastern Brazil and eastern Africa (Figure 2). *Spathodea campanulata* is widely distributed and concentrated between the Tropic of Cancer and Capricorn, mainly in central Africa, north-central South America, southern Asia, and southern North America (Figure 2).

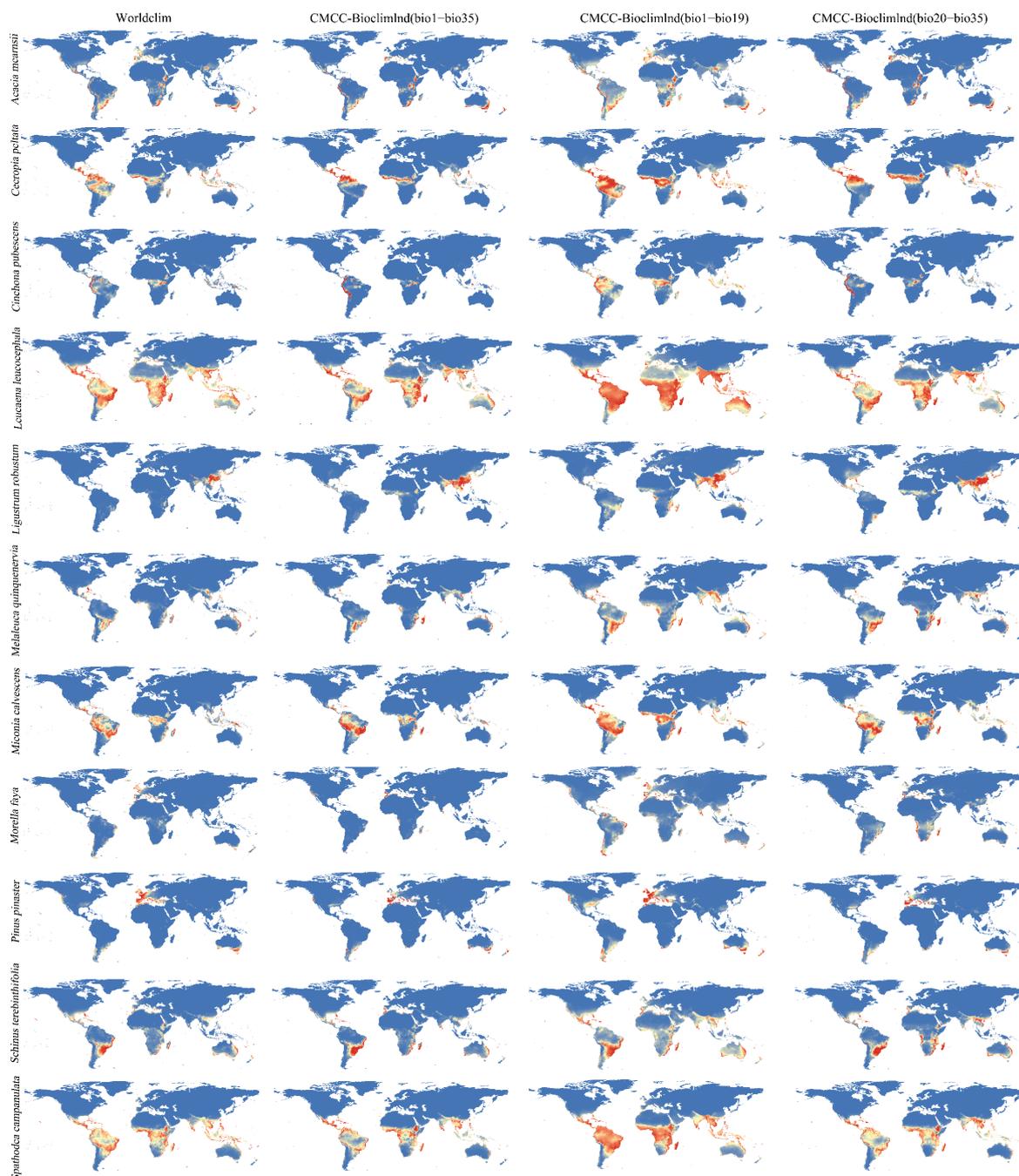


Figure 2. The impact of four model datasets (WorldClim, CMCC-BioclimInd bio1–bio35, CMCC-BioclimInd bio1–bio19, CMCC-BioclimInd bio20–bio35) on the distribution of 11 IPS. The distribution probability increases gradually from blue to red.

3.3. Differences in the Distribution Probability of Invasive Plant Species Predicted by the WorldClim and CMCC-BioClimInd Datasets

The differences in invasive plant distributions predicted by the two models for the 11 species were concentrated in the main distribution sites. Globally (except Antarctica), the invasive distribution probability maps predicted by the two models were roughly similar, and the area of difference was relatively small. The differences were concentrated in the Himalayas, Malaysia, and the Mediterranean (Figure 3). The species distribution probability predicted by CMCC-BioClimInd was greater than WorldClim in the Himalayas for *Leucaena leucocephala*, *Ligustrum robustum*, *Melaleuca quinquenervia*, and *Spathodea campanulata* (Figure 3). The distribution probability in Malaysia predicted by CMCC-BioClimInd was also higher than that of the WorldClim model for *Acacia mearnsii*, *Leucaena leucocephala*, *Melaleuca quinquenervia*, *Miconia calvenscens*, *Morella faya*, *Schinus terebinthifolia*, and *Spathodea campanulata* (Figure 3). However, for *Acacia mearnsii*, *Morella faya*, and *Pinus pinaster*, WorldClim was higher than CMCC-BioClimInd when predicting the potential Mediterranean basin distribution (Figure 3).

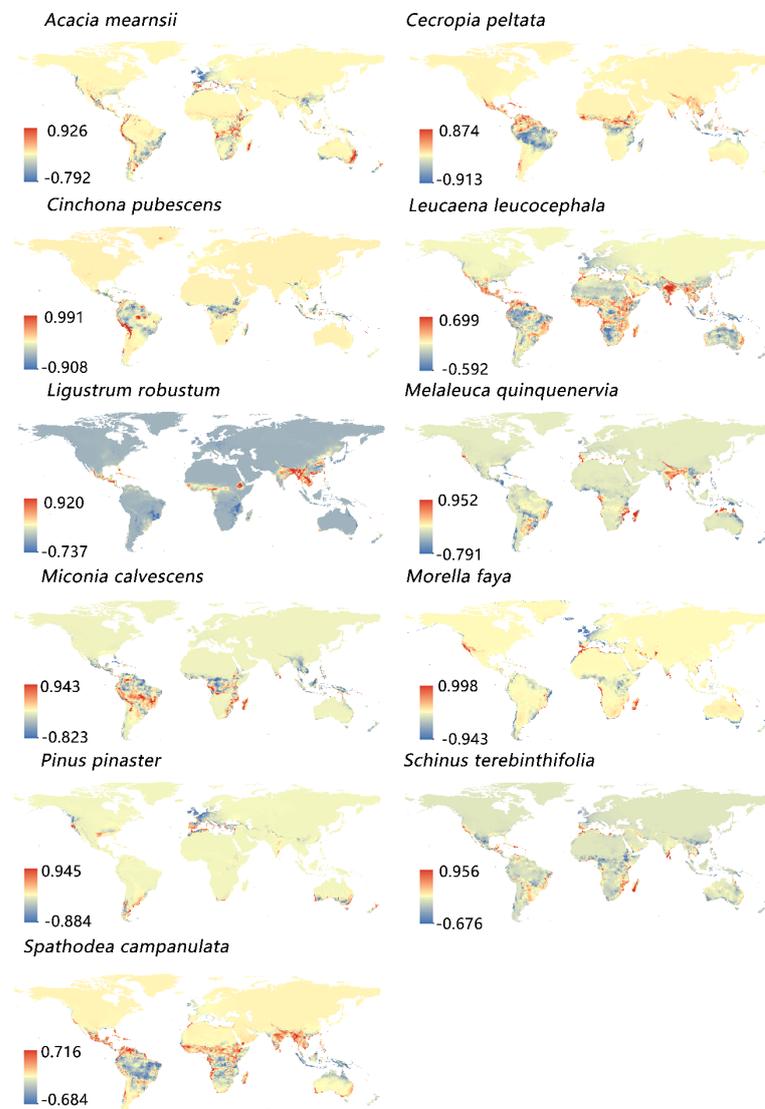


Figure 3. Maps of distribution probability differences calculated as CMCC-BioClimInd minus WorldClim. Reddish tones indicate higher values based on CMCC-BioClimInd dataset in relation to the WorldClim dataset; bluish tones indicate the opposite.

4. Discussion

This investigation introduced a novel global bioclimatic index dataset [46] to predict the distribution of 11 invasive species. The 35 climate variables of CMCC-BioClimInd were divided into three datasets, and WorldClim was added to run the model. By comparing the AUC and omission rates of the four models, it revealed that the average AUC of all four models was higher than 0.94%. The omission rate in CMCC-BioClimInd was less than 0.2, while the omission rate in WorldClim was less than 0.25. Therefore, CMCC-BioClimInd was very effective and accurate in predicting the distribution probability for IPS (Table 1).

WorldClim's climate dataset was derived by interpolating station data, while CMCC-BioClimInd's was derived from climate reanalysis and 11 CMIP5 climate simulations [45,63]. The comparison revealed that the following variables contributed greatly to the invasive species distributions: bio4, bio11, and bio3 from WorldClim, and bio4 and bio1 from CMCC-BioClimInd (bio1–bio19) (Table 2). The literature indicated an increased correlation between bio3 and bio11, and the contribution rates of these two variables to invasive species distribution in WorldClim were very high [50,64] and inappropriate. Among the climate variables of WorldClim and CMCC-BioClimInd (bio1–bio19), the contribution rates of bio1, bio10, and bio17 to invasive species distribution were quite different (Table S3). Notably, the differences mentioned above were mainly in areas where the variable estimates were less accurate due to the paucity of ground observations [65], and some artifacts may have arisen from the interpolation function used to create the spatial gridded dataset [46]. The possible explanation for the observed differences between WorldClim and CMCC-BioClimInd, might be the different weights given to observations concerning climate modeling data when creating the datasets [46].

CMCC-BioClimInd has 35 environmental variables [46,66]; it was noted that the first 19 CMCC-BioClimInd climate variables were similar to WorldClim, and had 16 additional climate variables compared to WorldClim [46]. Our results showed that among the 35 environmental variables, bio26, bio27, and bio4 had a higher contribution to invasive species distribution (Table 2). Furthermore, when the CMCC-BioClimInd (bio20–bio35) variables were compared with CMCC-BioClimInd (bio1–bio19) variables, it was found that the latter 16 variables dominated the invasive species distribution probabilities, with a 57% contribution rate, while the first 19 variables only contributed about 43% (Table 2). In CMCC-BioClimInd, bio27 and bio26 contributed significantly to invasive species distributions (Table 2), inferring the importance of the last 16 variables to invasive species distributions.

Although temperature and precipitation had different effects on invasive species distributions, inconsistent with previous studies [2,3], it was found that bio27 and bio26 had the most obvious effect on the distribution of invasive species (Table 2). Apart from how solar radiation varies with latitude, research has revealed that continentality is the most important factor controlling locality variation in Earth's climate and affecting plant growth [67]. Generally, various Earth's surface factors influence radiation fluxes, heat, and moisture at the air–land and air–water interfaces; these affect weather aspects such as temperature, precipitation, and cloudiness [67–69]. Additionally, previous studies have proved that the plant growth of the forest community is affected by the thermal climate to some extent; for example, there is a tendency for increased aboveground plant height, plant biomass, and the degree of forest canopy multi-layering toward warmer regions [70]. On the other hand, the diversity of component plant flora is extremely sensitive to changes in the thermal climate [70–72]. Plant growth requires more than just precipitation and temperature [73]. For example, studies have shown that although some areas receive sufficient precipitation and light, many factors, such as a high evaporation rate, serious water loss, and radiation, may result in poor species growth [43,49].

We consider that CMCC-BioClimInd can predict invasive plant distributions more comprehensively and accurately [46,73]. Other studies revealed that environmental variables were inconsistent for predicting species distributions [74,75]. The contribution rates of various species were different, associated with species habitat and to a certain extent, the mutual restriction between organisms [73]. For example, Dingle et al. (2000) found that

annual rainfall and soil moisture explained 90% and 62% of migratory butterfly species richness in the Australian drought center, respectively [76]. However, these two factors were not significant for butterflies in the rainy areas of eastern Australia [76]. In eastern Australia, temperature seasonality has become the best single climate predictor of butterfly species richness [76].

It was also observed that the 11 IPS mainly distributed in tropical rainforests and grasslands, subtropical evergreen broad-leaved forests, and the Mediterranean region (Figure 2). These climate regions have abundant species resources, sufficient rain, heat, and forest resources [77]. Biodiversity is also richer in places with abundant plant resources [77]. For example, the invasive *Morella faya* is highly scattered throughout a narrow distribution area [78]. The introduction of fruit-eating birds promoted the spread of this fast-growing plant, and these plants quickly formed dense stands, endangering local plant growth [78].

Temperature and precipitation are essential for IPS biology [2–5]. On a large spatial scale, tolerance of invasive plants is usually linked with climate and the main habitat [6,75]. With climate change, invasive species from adjacent areas may cross national borders and become new biota elements [2,3]. Invasive species threaten plant growth in local habitats by competing with local vegetation, replacing grassland communities, reducing local biodiversity, and increasing water loss in riparian zones [4,13]. Therefore, CMCC-BioClimInd can predict potential alterations for invasive species with respect to future climate according to the 35 climate variables and help to implement timely preventive measures [46]. Otherwise, invasive plants' economic loss and negative impacts on food security, biodiversity, and ecosystem services may soon sharply increase [1,77].

We found a significant difference in the response of IPS distribution probabilities between CMCC-BioClimInd and WorldClim near the Himalayas. According to the literature, in the Himalayas, the temperature and precipitation in CMCC-BioClimInd were higher than in WorldClim [46]. Therefore, the invasive plant distribution probabilities based on CMCC-BioClimInd were also higher in the Himalayas. Notably, a large distribution probability difference exists near the equator, Mediterranean, Malaysia, and western and eastern Australia (Figure 3). The effect of climate data's quality may also affect species distribution modelling [49]. To predict a reasonable future distribution, the indispensable bioclimatic variables used in a species distribution model must be reliable [49]. We have the following conjectures about the differences caused by the two data sets on the distribution of invasive species: (1) It may be because the CMCC-BioClimInd data set has 16 additional climate variables with greater impact compared to WorldClim for the distribution of invasive species [35,46]; (2) For bioclimatic variables, the sources of these variables are different [35,36,46]. Coarse-scale bioclimatic information may be insufficient or inconsistent for species distribution models derived from finer-scale species occurrence data [49]; (3) This may be because the distribution area for the main invasive species was in the region where the variable estimates were less accurate due to the paucity of ground observations. Furthermore, the observed differences between WorldClim and CMCC-BioClimInd might be because of the different observation's weights used with respect to climate modeling data when creating the datasets [35,46]. To limit IPS damage to global biodiversity, safety, and the economy, effective measures must be taken to prevent their further expansion [2,4,5]. Our research also provides a new global IPS risk assessment and management perspective.

Our research also has some limitations. (1) To some extent, the multiple linearities between variables in the CMCC-BioClimInd and WorldClim dataset were not addressed. Research has shown that correlation analysis refers to comparing and analyzing two or more related variables to measure the degree of correlation between variables [79]. There are various data indicators, and most overlap with each other, resulting in large redundancy in the data. Using a wide range of similar data cannot yield comprehensive information, resulting in a phenomenon where the amount of data and information is not proportional and potentially serious fitting [80,81]. In the selection and determination of MaxEnt environmental variables, the number of environmental variables can be changed

due to their different abilities to determine species distributions [82]. The number of environmental variables largely affects MaxEnt's ability to simulate the distribution of invasive plants by altering the model's complexity [82]. Therefore, in future research, the multicollinearity problem between variables should be addressed and environmental variables that are the most important for studying species distribution should be selected. (2) In this study, the number of replicates set in the species distribution model was too low. For example, recent research on species distribution modelling has shown that MaxEnt often produces better classification results when users choose the optimal parameters [83]. Research results indicate that model parameterization significantly impacts the prediction accuracy of MaxEnt; therefore, appropriate parameterization is highly correlated with good classification results [84]. Thus, in future research, the number of iterations for the Maxent model will be set to 10–100. (3) Using AUC and commission rates to determine MaxEnt's modeling performance does not provide the best model predictions. With imbalanced datasets, AUC may be misleading as the number of positive and negative samples is uneven. In this case, AUC may have overestimated the classifier's performance. In addition, the score of AUC ignores the actual probability value, making it insensitive to changes in the predicted probability of maintaining its ranking, and the testing performance of ROC in spatial regions is rarely successful [85,86]. Additional performance evaluation indicators (e.g., TSS, Kappa, the null model for significance testing, Boyce index, Still block cross-validation) should be included in future studies [87–90]. (4) Although MaxEnt is widely used to simulate plant invasion [91], research has shown that the number of species recorded, the number of environmental variables, and the spatial scale all affect the performance of the MaxEnt distribution model, indicating that these three inputs can lead to uncertainty in the invasive plant MaxEnt [80]. In future research, we can use models such as Maxlike and general linear models to evaluate invasive plants.

5. Conclusions

The CMCC-BioClimInd datasets improve existing global bioclimatic datasets used for SDM. This is a pragmatic compromise that addresses some of the limitations of the currently available products and is accurate for predicting invasive species distributions. In the rapidly changing global environment, bioclimatic species modelling has become an important tool for answering many conservation biology and invasion ecology questions. The CMCC-BioClimInd dataset can provide a wide range of core functions for these models. After combining 35 CMCC-BioClimInd climates, bio27 and bio26 were found to greatly impact invasive species distributions. Furthermore, it was revealed that the invasive species were mainly distributed in areas with sufficient rain and heat, such as tropical rainforests and grasslands, and subtropical evergreen broad-leaved forests. Therefore, policymakers must reinforce the management of areas vulnerable to the six kinds of invasion and formulate effective strategies to prevent invasive plant expansion.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biology12050652/s1>, Table S1: Description of the CMCC-BioClimInd dataset according to the ODMAP protocol; Table S2: Climate variables in CMCC-BioClimInd dataset including codes, full names and unIPS. bio1~bio19 were the same as WorldClim dataset; Table S3: The significance of differences in the average contribution rate (%) of bioclimatic variables to species distribution probability between WorldClim and CMCC-BioClimInd based on independent-sample *t* test.

Author Contributions: J.W. and C.Z. conceived the ideas and designed the study; F.Z. and J.W. managed the data; C.W. and J.W. conducted the fieldwork; C.W., F.Z. and J.W. collaborated with the statistical analysis and interpretation of data; F.Z. wrote the first version of the manuscript with substantial contribution from C.W., C.Z. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Project of Qinghai Science and Technology Department (grant number 2020-ZJ-733).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: We acknowledge De-Juan Xie for the lab support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ward, D.F. Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biol. Invasions* **2007**, *9*, 723–735. [[CrossRef](#)]
2. Beck, K.G.; Zimmerman, K.; Schardt, J.D.; Stone, J.; Lukens, R.R.; Reichard, S.; Randall, J.; Cangelosi, A.A.; Cooper, D.; Thompson, J.P. Invasive Species Defined in a Policy Context: Recommendations from the Federal Invasive Species Advisory Committee. *Invasive Plant Sci. Manag.* **2008**, *1*, 414–421. [[CrossRef](#)]
3. Bradley, B.A.; Wilcove, D.S.; Oppenheimer, M. Climate change increases risk of plant invasion in the Eastern United States. *Biol. Invasions* **2010**, *12*, 1855–1872. [[CrossRef](#)]
4. Early, R.; Bradley, B.A.; Dukes, J.S.; Lawler, J.J.; Olden, J.D.; Blumenthal, D.M.; Gonzalez, P.; Grosholz, E.D.; Ibanez, I.; Miller, L.P.; et al. Global threats from invasive alien species in the twenty-first century and national response capacities. *Nat. Commun.* **2016**, *7*, 12485. [[CrossRef](#)] [[PubMed](#)]
5. Hong, S.H.; Lee, Y.H.; Lee, G.; Lee, D.; Adhikari, P. Predicting Impacts of Climate Change on Northward Range Expansion of Invasive Weeds in South Korea. *Plants* **2021**, *10*, 1604. [[CrossRef](#)] [[PubMed](#)]
6. Walther, G.R.; Post, E.; Convey, P. Ecological responses to recent climate change. *Nature* **2002**, *416*, 389–395. [[CrossRef](#)]
7. Beaumont, L.J.; Hughes, L.; Poulsen, M. Predicting species distributions: Use of climatic parameters in BIOCLIM and IPS impact on predictions of species' current and future distributions. *Ecol. Modell.* **2005**, *186*, 250–269. [[CrossRef](#)]
8. Ehrlén, J.; Morris, W.F. Predicting changes in the distribution and abundance of species under environmental change. *Ecol. Lett.* **2015**, *18*, 3030–3314. [[CrossRef](#)]
9. Wagner, V.; Vecera, M.; Jimenez-Alfaro, B. Alien plant invasion hotspots and invasion debt in European woodlands. *J. Veg. Sci.* **2021**, *32*, 13014. [[CrossRef](#)]
10. Ciapała, B.; Jurasz, J.; Janowski, M.; Kepińska, B. Climate factors influencing effective use of geothermal resources in SE Poland: The Lublin. *Geotherm. Energy* **2021**, *9*, 3. [[CrossRef](#)]
11. Bacon, S.J.; Bacher, S.; Aebi, A. Gaps in Border Controls Are Related to Quarantine Alien Insect Invasions in Europe. *PLoS ONE* **2012**, *7*, e47689. [[CrossRef](#)] [[PubMed](#)]
12. Brundu, G.; Pauchard, A.; Pysek, P.; Pergl, J.; Bindewald, A.M.; Brunori, A.; Canavan, S.; Campagnaro, T.; Celesti-Grapow, L.; de Sá Dechoum, M.; et al. Global guidelines for the sustainable use of non-native trees to prevent tree invasions and mitigate their negative impacts. *Neobiota* **2020**, *61*, 65–116. [[CrossRef](#)]
13. Franzese, J.; Urrutia, J.; Garcia, R.A.; Taylor, K.; Pauchard, A. Pine invasion impacts on plant diversity in Patagonia: Invader size and invaded habitat matter. *Biol. Invasions* **2017**, *19*, 1015–1027. [[CrossRef](#)]
14. Milani, T.; Jobbagy, E.G.; Nunez, M.A.; Ferrero, M.E.; Baldi, G.; Teste, F.P. Stealth invasions on the rise: Rapid long-distance establishment of exotic pines in mountain grasslands of Argentina. *Biol. Invasions* **2020**, *22*, 2989–3001. [[CrossRef](#)]
15. Aranda, M.J.; Tognetti, P.M.; Mazia, N. Are field crops refuge for woody invaders? Rainfall, crop type and management shaped tree invasion in croplands. *Agric. Ecosyst. Environ.* **2021**, *319*, 107564. [[CrossRef](#)]
16. Pachauri, R.K.; Allen, M.R.; Barros, V.R.; Broome, J.; Cramer, W.; Christ, R.; Church, J.A.; Clarke, L.; Dahe, Q.; Dasgupta, P.; et al. *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*; Tech. Rep.; IPCC: Geneva, Switzerland, 2014.
17. Newbold, T. Future effects of climate and land–use change on terrestrial vertebrate community diversity under different scenarios. *Proc. R. Soc. B* **2018**, *285*, 20180792. [[CrossRef](#)]
18. Thuiller, W.; Araujo, M.B.; Lavorel, S. Do we need land–cover data to model species distributions in Europe? *J. Biogeogr.* **2004**, *31*, 353–361. [[CrossRef](#)]
19. Cerasoli, F.; D'Alessandro, P.; Biondi, M. Worldclim 2.1 versus Worldclim 1.4: Climatic niche and grid resolution affect between–version mismatches in Habitat Suitability Models predictions across Europe. *Ecol. Evol.* **2022**, *12*, 8430. [[CrossRef](#)]
20. Woodward, F.I.; Long, S.P. Temperature and the distribution of plant species. *Plants Temp. Symp. Soc. Exp. Biol.* **1988**, *42*, 59–75.
21. Kriticos, D.J.; Webber, B.L.; Leriche, A.; Ota, N.; Macadam, I.; Bathols, J.; Scott, J.K. CliMond: Global high–resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods Ecol. Evol.* **2012**, *3*, 53–64. [[CrossRef](#)]
22. Fitzpatrick, E.A.; Nix, H.A. *The Climatic Factor in Australian Grass–Lands*; Moore, R.M., Ed.; Australian National University Press: Canberra, Australia, 1970; pp. 3–26.
23. Graham, C.H.; Ferrier, S.; Huettman, F.; Moritz, C.P.; Peterson, A.T. New developments in museum–based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* **2004**, *19*, 497–503. [[CrossRef](#)] [[PubMed](#)]

24. Chefaoui, R.M.; Lobo, J.M. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecol. Modell.* **2008**, *210*, 478–486. [[CrossRef](#)]
25. Guisan, A.; Thuiller, W. Predicting species distribution: Offering more than simple habitat models. *Ecol. Lett.* **2005**, *8*, 993–1009. [[CrossRef](#)] [[PubMed](#)]
26. Baker, R.H.A.; Sansford, C.E.; Jarvis, C.H.; Cannon, R.J.C.; MacLeod, A.; Walters, K.F.A. The role of climatic mapping in predicting the potential geographical distribution of non-indigenous pests under current and future climates. *Agric. Ecosyst. Environ.* **2000**, *82*, 57–71. [[CrossRef](#)]
27. Kriticos, D.J.; Randall, R.P. *A Comparison of Systems to Analyse Potential Weed Distributions*; Groves, R.H., Panetta, F.D., Virtue, J.G., Eds.; CSIRO Publishing: Melbourne, Australia, 2001; pp. 61–79.
28. Julien, M.H.; Skarratt, B.; Maywald, G.F. Potential geographical distribution of alligator weed and its biological control by *Agasicles hygrophila*. *J. Aquat. Plant. Manag.* **1995**, *33*, 55–60.
29. Poggio, L.; Simonetti, E.; Gimona, A. Enhancing the WorldClim data set for national and regional applications. *Sci. Total Environ.* **2017**, *625*, 1628–1643. [[CrossRef](#)]
30. Soria-Auza, R.W.; Kessler, M.; Bach, K.; Barajas-Barbosa, P.M.; Lehnert, M.; Herzog, S.K.; Böhner, J. Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: A case study from Bolivia. *Ecol. Modell.* **2010**, *221*, 1221–1229. [[CrossRef](#)]
31. Kim, D. Mapping Species Distributions: Spatial Inference and Prediction. *Ann. Am. Assoc.* **2011**, *101*, 1384–1386. [[CrossRef](#)]
32. Booth, T.H.; Nix, H.A.; Busby, J.R.; Hutchinson, M.F. BIOCLIM: The first species distribution modelling package, its early applications and relevance to most current MAXENT. *Divers. Distrib.* **2014**, *20*, 1–9. [[CrossRef](#)]
33. Baker, D.J.; Hartley, A.J.; Butchart, S.H.; Willis, S.G. Choice of baseline climate data impacts projected species' responses to climate change. *Glob. Change Biol.* **2016**, *22*, 2392–2404. [[CrossRef](#)]
34. Morales-Barbero, J.; Vega-Álvarez, J. Input matters matter: Bioclimatic consistency to map more reliable species distribution models. *Methods Ecol. Evol.* **2019**, *10*, 212–224. [[CrossRef](#)]
35. 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](#)]
36. Fick, S.E.; Hijmans, R.J. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J.* **2017**, *37*, 4302–4315. [[CrossRef](#)]
37. Imbach, P.; Molina, L.; Locatelli, B.; Roupsard, O.; Ciais, P.; Corrales, L.; Mahe, G. Climatology-based regional modelling of potential vegetation and average annual long-term runoff for Mesoamerica. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1801–1817. [[CrossRef](#)]
38. Moreno, A.; Hasenauer, H. Spatial downscaling of european climate data. *Int. J. Climatol.* **2016**, *36*, 1444–1458. [[CrossRef](#)]
39. Yigini, Y.; Panagos, P. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. *Sci. Total Environ.* **2016**, *557*, 838–850. [[CrossRef](#)]
40. Busby, J.R. BIOCLIM—A bioclimatic analysis and prediction tool. *Plant Prot. Q.* **1991**, *6*, 8–9.
41. Daly, C.; Neilson, R.; Phillips, D. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteorol.* **1994**, *33*, 140–158. [[CrossRef](#)]
42. Killeen, T.J.; Douglas, M.; Consiglio, T.; Jorgensen, P.M.; Mejia, J. Dry spots and wet spots in the Andean hotspot. *J. Biogeogr.* **2007**, *34*, 1357–1373. [[CrossRef](#)]
43. Weedon, G.P.; Gomes, S.; Viterbo, P. Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. *J. Hydrometeorol.* **2011**, *12*, 823–848. [[CrossRef](#)]
44. 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. Change* **2011**, *109*, 5–31. [[CrossRef](#)]
45. Taylor, K.E.; Stouffer, R.J.; Meehl, G.A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 485–498. [[CrossRef](#)]
46. Noce, S.; Caporaso, L.; Santini, M. A new global dataset of bioclimatic indicators. *Sci. Data* **2020**, *7*, 398. [[CrossRef](#)] [[PubMed](#)]
47. Kenney, M.A.; Janetos, A.C.; Gerst, M.D. A framework for national climate indicators. *Clim. Change* **2018**, *163*, 1705–1718. [[CrossRef](#)]
48. Collins, M. Ensembles and probabilities: A new era in the prediction of climate change. *Phil. Trans. R. Soc. A* **2007**, *365*, 1957–1970. [[CrossRef](#)] [[PubMed](#)]
49. Liu, S.Y.; Liang, X.Z.; Gao, W.; Stohlgren, T.J. Regional climate model downscaling may improve the prediction of alien plant species distributions. *Front. Earth Sci.* **2014**, *8*, 457–471. [[CrossRef](#)]
50. Zurell, D.; Franklin, J.; König, C.; Bouchet, P.J.; Serra-Diaz, J.M.; Dormann, C.F.; Elith, J.; Fandos Guzman, G.; Feng, X.; Guillera-Arroita, G.; et al. A standard protocol for describing species distribution models. *Ecography* **2020**, *43*, 1261–1277. [[CrossRef](#)]
51. Chakraborty, D.; Moricz, N.; Rasztovits, E.; Dobor, L.; Schueler, S. Provisioning forest and conservation science with high-resolution maps of potential distribution of major European tree species under climate change. *Ann. For. Sci.* **2021**, *78*, 26. [[CrossRef](#)]
52. Cuthbert, R.N.; Diagne, C.; Haubrock, P.J.; Turbelin, A.J.; Courchamp, F. Are the “100 of the world's worst” invasive species also the costliest? *Biol. Invasion* **2021**, *24*, 1895–1904. [[CrossRef](#)]

53. Wisz, M.S.; Hijmans, R.J.; Li, J.; Peterson, A.T.; Graham, C.H.; Guisan, A. Effects of sample size on the performance of species distribution models. *Divers. Distrib.* **2008**, *14*, 763–773. [[CrossRef](#)]
54. Zhang, H.; Zhang, M. Insight into distribution patterns and conservation planning in relation to woody species diversity in Xinjiang, arid northwestern China. *Biol. Conserv.* **2014**, *177*, 165–173. [[CrossRef](#)]
55. Uppala, S.M.; Kallberg, P.W.; Simmons, A.J. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 2961–3012. [[CrossRef](#)]
56. Merow, C.; Smith, M.J.; Silander, J.A. A practical guide to Maxent for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **2013**, *36*, 1058–1069. [[CrossRef](#)]
57. 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](#)]
58. Elith, J.; Phillips, S.J.; Hastie, T.; Dudik, M.; Chee, Y.E.; Yates, C.J. A statistical explanation of Maxent for ecologists. *Divers. Distrib.* **2011**, *17*, 43–57. [[CrossRef](#)]
59. Saupe, E.E.; Hendricks, J.R.; Townsend Peterson, A.; Lieberman, B.S. Climate change and marine molluscs of the western North Atlantic: Future prospects and perils. *J. Biogeogr.* **2014**, *41*, 1352–1366. [[CrossRef](#)]
60. Boria, R.A.; Olson, L.E.; Goodman, S.M.; Anderson, R.P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol. Model.* **2013**, *275*, 73–77. [[CrossRef](#)]
61. Boria, R.A.; Blois, J.L. The effect of large sample sizes on ecological niche models: Analysis using a North American rodent, *Peromyscus maniculatus*. *Ecol. Model.* **2018**, *386*, 83–88. [[CrossRef](#)]
62. Wan, J.Z.; Zhang, Z.X.; Wang, C.J. Effects of ecoregional vulnerability on habitat suitability of invasive alien plants: An assessment using 13 species on a global scale. *Environ. Earth Sci.* **2019**, *78*, 180. [[CrossRef](#)]
63. Hempel, S.; Frieler, K.; Warszawski, L.; Schewe, J.; Piontek, F. A trend-preserving bias correction—The ISIMIP approach. *Earth Syst. Dyn.* **2013**, *4*, 219–236. [[CrossRef](#)]
64. Wittmann, M.E.; Barnes, M.A.; Jerde, C.L.; Jones, L.A.; Lodge, D.M. Confronting species distribution model predictions with species functional traits. *Ecol. Evol.* **2016**, *6*, 873–880. [[CrossRef](#)] [[PubMed](#)]
65. Harris, I.; Osborn, T.J.; Jones, P.; Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* **2020**, *7*, 109. [[CrossRef](#)] [[PubMed](#)]
66. O'Donnell, M.A.; Ignizio, D.A. *Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States*; US Geological Survey Data Series 691; U.S. Department of the Interior: Washington, DC, USA, 2012.
67. Driscoll, D.M.; Fong, J.M.Y. Continentality: A basic climatic parameter re-examined. *Int. J. Climatol.* **1992**, *12*, 185–192. [[CrossRef](#)]
68. Fobes, C. Continentality in New England. *Bull. Am. Meteorol. Soc.* **1954**, *35*, 197.
69. Kopec, R. Continentality around the Great Lakes. *Bull. Am. Meteorol. Soc.* **1965**, *46*, 54–57. [[CrossRef](#)]
70. Kira, T. Forest ecosystems of east and southeast—Asia in a global perspective. *Ecol. Res.* **1991**, *6*, 185–200. [[CrossRef](#)]
71. Yim, Y.J.; Kira, T. Distribution of forest vegetation and climate in the Korean Peninsula, I, Distribution of some indices of thermal climate. *Jpn. J. Ecol.* **1975**, *25*, 77–88.
72. Yim, Y.J. Distribution of forest vegetation and climate in the Korean Peninsula, IV, Zonal distribution of forest vegetation in relation to thermal climate. *Jpn. J. Ecol.* **1977**, *27*, 269–278.
73. Kriticos, D.J.; Jarošik, V.; Ota, N. Extending the suite of bioclim variables: A proposed registry system and case study using principal components analysis. *Methods. Ecol. Evol.* **2014**, *5*, 956–960. [[CrossRef](#)]
74. Davis, A.J.; Jenkinson, L.S.; Lawton, J.H.; Shorrocks, B.; Wood, S. Making mistakes when predicting shifts in species range in response to global warming. *Nature* **1998**, *391*, 783–786. [[CrossRef](#)]
75. Araújo, M.B.; Luoto, M. The importance of biotic interactions for modelling species distributions under climate change. *Glob. Ecol. Biogeogr.* **2007**, *16*, 743–753. [[CrossRef](#)]
76. Dingle, H.; Rochester, W.A.; Zalucki, M.P. Relationships among climate, latitude and migration: Australian butterflies are not temperate-zone birds. *Oecologia* **2000**, *124*, 196–207. [[CrossRef](#)] [[PubMed](#)]
77. Zhang, F.X.; Wang, C.J.; Wan, J.Z. Using Consensus Land Cover Data to Model Global Invasive Tree Species Distributions. *Plants* **2022**, *11*, 981. [[CrossRef](#)] [[PubMed](#)]
78. Loh, R.K.; Daehler, C.C. Influence of woody invader control methods and seed availability on native and invasive species establishment in a Hawaiian forest. *Biol. Invasions* **2008**, *10*, 805–819. [[CrossRef](#)]
79. Gallagher, R.V.; Hughes, L.; Leishman, M.R. Species loss and gain in communities under future climate change: Consequences for functional diversity. *Ecography* **2013**, *36*, 531–540. [[CrossRef](#)]
80. Wan, J.Z.; Wang, C.J.; Yu, F.H. Effects of occurrence record number, environmental variable number, and spatial scales on MaxEnt distribution modelling for invasive plants. *Biologia* **2019**, *74*, 757–766. [[CrossRef](#)]
81. Fernandez, I.C.; Morales, N.S. One-class land-cover classification using MaxEnt: The effect of modelling parameterization on classification accuracy. *PeerJ* **2019**, *7*, 7016. [[CrossRef](#)]
82. Suárez-Mota, M.E.; Ortiz, E.; Villaseñor, J.L.; Espinosa-García, F.J. Ecological niche modeling of invasive plant species according to invasion status and management needs: The case of *Chromolaena odorata* (Asteraceae) in South Africa. *Pol. J. Ecol.* **2016**, *64*, 369–383. [[CrossRef](#)]
83. Skowronek, S.; Asner, G.P.; Feilhauer, H. Performance of one-class classifiers for invasive species mapping using airborne imaging spectroscopy. *Ecol. Inform.* **2017**, *37*, 66–76. [[CrossRef](#)]

84. Mack, B.; Waske, B. In-depth comparisons of MaxEnt, biased SVM and one-class SVM for one-class classification of remote sensing data. *Remote Sens. Lett.* **2017**, *8*, 290–299. [[CrossRef](#)]
85. Austin, M.P. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecol. Model.* **2007**, *200*, 1–19. [[CrossRef](#)]
86. 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](#)]
87. Mouton, A.M.; De Baets, B.; Goethals, P.L.M. Ecological relevance of performance criteria for species distribution models. *Ecol. Model.* **2010**, *221*, 1995–2002. [[CrossRef](#)]
88. 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](#)]
89. Raes, N.; ter Steege, H. A null-model for significance testing of presence-only species distribution models. *Ecography* **2007**, *30*, 727–736. [[CrossRef](#)]
90. Valavi, R.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods Ecol. Evol.* **2019**, *10*, 225–232. [[CrossRef](#)]
91. Mainali, K.P.; Warren, D.L.; Dhileepan, K.; McConnachie, A.; Strathie, L.; Hassan, G.; Karki, D.; Shrestha, B.B.; Parmesan, C. Projecting future expansion of invasive species: Comparing and improving methodologies for species distribution modeling. *Glob. Change Biol.* **2015**, *21*, 4464–4480. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.