

RESEARCH

Open Access



Predicting the effect of climate change on the spatiotemporal distribution of two endangered plant species, *Silene leucophylla* Boiss. and *Silene schimperiana* Boiss., using machine learning, in Saint Catherine Protected Area, Egypt

Aliaa Muhammad Refaat^{1*} , Ashraf Mohamed Youssef¹, Hosny Abdel-Aziz Mosallam¹ and Haitham Farouk²

Abstract

Background Climate change significantly influences the geographical distribution of plant species worldwide, especially endemics. Endemic species are plants that live in limited distribution ranges of unique ecology and, thus, are the most vulnerable species to climate change. Therefore, understanding the impacts of climate change on the distribution of these species can assist in developing appropriate plans for their conservation. In this study, we aimed to apply various species distribution models (SDMs) to predict the current potential distributions of two endangered plant species, *Silene leucophylla* (*S. leucophylla*, endemic) and *Silene schimperiana* (*S. schimperiana*, near-endemic), in Saint Catherine protected area (St. Catherine PA), Egypt. Then, using the best-fit model to project their future distribution under the maximum climate emission scenario (Representative Concentration Pathway 8.5 (RCP8.5)). Six different SDMs were constructed using different geospatial raster imagery sets of environmental factors. For each model, five machine learning (ML) algorithms were used. The results of these ML algorithms were then ensembled by calculating the weighted average of their predictions.

Results Based on the analysis of digital geospatial imageries produced by the best-fitting model, the predicted suitable areas of *S. leucophylla* and *S. schimperiana* were 23.1 km² and 125 km², respectively. These sites are located mainly in the high-elevation middle northern part of the study area. Annual precipitation, mean temperature of the driest quarter, altitude, and precipitation seasonality were the essential predictors of the distributions of both species. Future predictions of both species indicated opposing results between the studied species. Predictions in the 2050 and 2070 future conditions revealed significant range contraction for the distribution of *S. leucophylla*. For *S. schimperiana*, a range shift is predicted, with both range contraction and range expansion of its current suitable habitats, for the same future projections. Unfortunately, in 2080 predictions, both species could be projected to a complete loss from the entire area.

Conclusion This study highlights the importance of including diverse types of environmental variables in SDMs to produce more accurate predictions, rather than relying only on one variable type. It also revealed the potential

*Correspondence:

Aliaa Muhammad Refaat
aliaamuhammad@sci.asu.edu.eg

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

negative impacts of future climate change on the distributions of two endangered plant species, *S. leucophylla* and *S. schimperiana*, inhabiting St. Catherine PA. Consequently, we urgently recommend the initiation of different plans and strategies seeking their conservation.

Keywords Environmental variables, SDMs, Ensemble models, Distribution change detection, Range contraction, Range expansion

1 Background

We live in the era of global climate change; there are no doubts. The recent world has, perhaps, encountered no greater threat of such severe consequences. We are now witnessing times of rapid and massive habitat depletion, biodiversity loss, and species extinction at an unprecedented rate. Climate change is a significant driver of shifting the geographical distribution of many species, questioning their ability to survive in these new habitats [1]. For example, mountain plants are predicted to migrate to higher altitudinal ranges in response to climate change. Therefore, mountain ecosystems are highly sensitive to global warming due to the reduction in available areas as elevation increases [2].

Many endemic plants depend on mountain habitats for their growth and survival [3]. Endemic plants are species that naturally grow in unique and geographically limited regions. Therefore, climate change-induced shifts in the natural habitats of these geographically limited species can prevent them from adapting to unusual climatic conditions and thus becoming endangered or even extinct [4–6]. Consequently, it is essential to forecast the distribution of their current and future environmentally suitable areas to develop appropriate conservation strategies for such threatened and rare species.

Species distribution models (SDMs) are machine learning (ML) techniques that can significantly accelerate the processing and analysis of massive repositories of raw data, enhancing the performance of various ecological analyses [7–11]. These models are commonly used for estimating the ecological niches of plant species and forecasting the potential impacts of climate change on their distribution [12]. Therefore, these models can assist experts in environment, conservation, and resource management in developing conservation plans, leading to sustainable uses of species for the next generations. SDMs correlate species occurrence points with environmental variables to estimate the possible spatial distribution of a species through space and time [13]. However, understanding the limitations of these models is crucial to minimize the potential for inaccurate outputs. Among others, there are two major limitations of SDMs, which are the choice of modeling method and the selection of environmental variables for incorporation into the models.

There are several SDM methods. In this study, we used the ensemble modeling technique, where several single modeling algorithms are used for projecting species distributions, and then a consensus projection is built [14]. The main aim of ensemble methods is to decrease the predictive uncertainty of single models by combining their predictions; thus, this approach would greatly improve the accuracy of these predictions [15, 16].

Regarding the environmental factors to include in the models, plants respond to various environmental factors, including climate, topography, edaphic, hydrology, biotic, etc. Therefore, the appropriate selection of the ecological characteristics that might influence the distribution pattern of the target species is fundamental in the modeling process [17]. However, variable selection has received limited attention, with some studies relying solely on climate-derived variables and overlooking other potential factors [18, 19]. While this could be true, specifically when predicting the impact of climate change on species distributions across large regions, using only climate predictors should be enough to assess the main changes. Nonetheless, it is reasonable to consider other factors [20–22].

In addition to climate, many studies have noted the importance of various variables, especially topographic, edaphic, hydrologic, and biotic variables, that interfere with each other in governing the distribution of plants in mountainous habitats [2, 23, 24]. Climate influences all aspects of plant growth and productivity by affecting energy and water availability [25–27]. On the other hand, topography plays a crucial role in shaping plant distributions in mountain habitats, both directly and indirectly. Differences in the topographical features of mountains significantly impact mountain habitat properties (e.g., microclimate, soil temperature, soil stability, precipitation runoff, solar radiation, etc.) that influence plant growth along elevational gradients at a regional scale. In addition, topography forms geographical barriers, influencing plant dispersal across landscapes [2, 28–30].

Moreover, edaphic factors are known to control plant species distributions by providing plants with available water and nutrients, offering them physical support for growth, and affecting plant physiological processes. On a regional scale, soil properties can exhibit considerable variations with great complexity. These variations arise

from diverse ecological processes contributing to soil formation, such as different parent materials, weathering processes, and topographic positions [30–33]. These variations in soil chemical and physical properties further influence plant distribution, particularly for species associated with specific soil types, rocks, and bedrock, such as plant endemics [32, 34]. In addition, water availability and hydrologic processes in a habitat are vital in shaping plant distribution [35]. Furthermore, biotic interactions, including competition, facilitation, and herbivory, play critical roles in determining species distributions and community composition [36, 37].

Consequently, plant growth is influenced by complex interactions among various variables; thus, no single factor can predict their distribution. Therefore, incorporating a specific predictor for SDMs depends on the purpose of the modeling and its biological relevance to the studied species [25]. Considering the various scales of ecological processes associated with different environmental factors, climatic factors are assumed to play a central role in shaping plant distributions at the global scale. Other variables, such as topographic and edaphic factors, may hold greater importance at regional scales, such as mountains [30, 32].

According to the fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC), four representative concentration pathway climate change scenarios (RCPs) were designed to project future greenhouse gas concentrations. These RCPs are RCP2.6, RCP4.5, RCP6.0, and RCP8.5, named after the possible radiative forcing values at the end of the century (2100) relative to preindustrial values [38]. While the former scenario is the low-emission (firm-mitigation) scenario, the latter is the high-emission scenario, and the others are the intermediate ones. These RCPs are commonly used to characterize and predict the possible effects of future climate change on plant species distributions. However, many authors argue that historical and anticipated future total CO₂ emissions show more agreement with RCP8.5 [39] than other scenarios. Specifically, many studies have pointed out that countries in the Middle East and North Africa (MENA) region are heading to the RCP8.5 limit [40–45].

The study area is part of Egypt, a country in the MENA region. Based on the IPCC's last report, the observed average and seasonal surface temperatures have increased at twice the global rate [1.09 °C] over most regions in North Africa because of climate change. Since the 1970s, temperatures have increased between 0.2 and 0.4 °C per decade, especially in the summer. In the future, warming is projected to be stronger in terms of maximum and minimum temperatures, the number of hot days and hot nights, the duration of warm spells, and the

frequency and intensity of heat waves [46, 47]. According to an analysis by the German Climate Service Center (GERICS), Egypt is expected to experience a change in annual mean temperature from 1.8 to 5.2 °C by the 2080s. Over the same period, the maximum temperatures are expected to increase by 2.1 to 5.7 °C, with minimum temperatures rising from 1.5 to 4.6 °C. By the middle of the century, temperatures are expected to increase between 2 and 3 °C and will continue to rise in Egypt until the end of the century. Additionally, the analysis by GERICS indicates that the observed reduction in precipitation is expected to continue by the end of the century, with projections indicating a trend of even longer dry spells [48, 49]. According to the Berkeley Earth organization, Egypt has reached +1.8 °C in 2020 and is heading to +3.4 °C in 2100 [50]. Therefore, in this study, we focus on RCP8.5 only for future predictions [38].

Saint Catherine Protected Area (St. Catherine PA) is one of the largest PAs in Egypt. It is located in the southern part of the Sinai Peninsula and has expanded over its entire mountainous massifs (Fig. 1). It lies between the two longitudinal lines 33° 55' to 34° 30' E and the two latitudinal lines 28° 30' to 28° 35' N and covers an area of approximately 4350 km² with an altitudinal range up to 2640 m above sea level (a.s.l.) [51]. St. Catherine is recognized by its wide range of variations in air temperature and precipitation. It is also known as the coldest region in Egypt and the only city with snowfalls. For the period 1991–2020, the average minimum temperature ranged from 7.1 °C in winter to 21.2 °C in summer, whereas the average maximum temperature ranged from 18.3 °C in winter to 34.1 °C in summer [52]. The annual precipitation can reach 105 mm. However, rainfall does not occur annually; usually 2 to 3 successive years without rainfall are familiar. Rain takes the form of sporadic flash floods or limited local showers; thus, high spatial seasonality in the receiving moisture is normal [3].

Mountains exhibit complex interactions between the regional climate and the geophysical factors that establish their peculiar microhabitats [2, 23]. Since the landforms and geologic structures of St. Catherine PA are highly diverse, along with its heterogeneous climate, Khedr (2021) recognized six microhabitat types. These include wadis, terraces, slopes, gorges, farshes (basins), and caves; each has distinctive environmental conditions and unique flora [53]. Therefore, the study area is considered one of the floristically richest hot spots in the Mediterranean Basin, with many medicinal, rare, and endemic plants.

St. Catherine PA is home to 14 endemic vascular plants (~30% of the endemic plants in Egypt). Among them are three *Silene* spp., which are *Silene leucophylla* (*S. leucophylla*), *Silene schimperiana* (*S. schimperiana*)

[near-endemic (endemic to Egypt and Yemen)], and *Silene oreosinaica* (*S. oreosinaica*). These species have medical importance besides their economic value as fodder for domestic animals. For example, *S. leucophylla* is used to treat leprosy, heal cuts and inflamed wounds, and remedy diarrhea; moreover, its roots have hepato-protective function [54]. *S. schimperiana* showed significant cytotoxicity against colon and liver carcinoma cells when n-hexane and ethyl-acetate extracts of its aerial parts were used [55]. Unfortunately, recent field surveys and published literature indicate a continuous decline in their population size. According to the International Union for Conservation of Nature (IUCN), both *S. leucophylla* and *S. schimperiana* are classified as endangered threatened species, whereas *S. oreosinaica* is critically endangered [56]. However, in our field visits, *S. schimperiana* had a wider distribution than *S. leucophylla*, which inhabited a narrower ecological niche, and *S. oreosinaica* was rare and was not spotted during our surveys. Therefore, in this study, we focused on the first two species.

Like other endemic plants in St. Catherine PA, both species thrive in areas with moist conditions, low temperatures, and high elevations. Therefore, changes in temperature and precipitation patterns could impact their habitat suitability. Additionally, both are found to grow in rocky soils and are restricted to specific microhabitats, which are slopes, terraces, and gorges [3, 53, 57]. *S. leucophylla* is a perennial tufted herb with the vegetative season in December to June, the flowering season in April–June, and the fruiting season in

May–July with seed dispersal in August–September. In contrast, *S. schimperiana* is a perennial tall herb that grows vegetatively in March–June, the flowering season in May–July, and the fruiting season in June–August, with seed dispersal occurring in September–November. Genetically, it is found that *S. leucophylla* populations exhibited lower genetic diversity than populations of *S. schimperiana* [57]. To our knowledge, there have been no studies regarding the physiological traits of these two species, specifically. However, a study on their relative, *Silene capitata* Kom., in Korea found that *Silene* species exhibit changes in phenology (flowering, fruiting, and leafing) and growth under a changing climate. Traits such as transpiration rate, stomatal conductance, photosynthetic rate, and water use efficiency may be affected [58].

Generally, endemic *Silene* spp. are highly vulnerable to environmental threats (e.g., climate change) and human impacts, such as over-collection and overgrazing. These threats significantly influence the diversity and distribution of such endangered endemics in St. Catherine PA, pushing them to the extinction threshold.

In this study, we aim to (1) predict the current distribution of both *Silene* spp. using six sets of environmental variables, (2) define which environmental variable set is the best and accurately captures the ecological niches of both species, and (3) use the best-fit models to predict the possible future distributions (three time periods: 2050, 2070, and 2080) under the maximum climate emission scenario (RCP8.5).

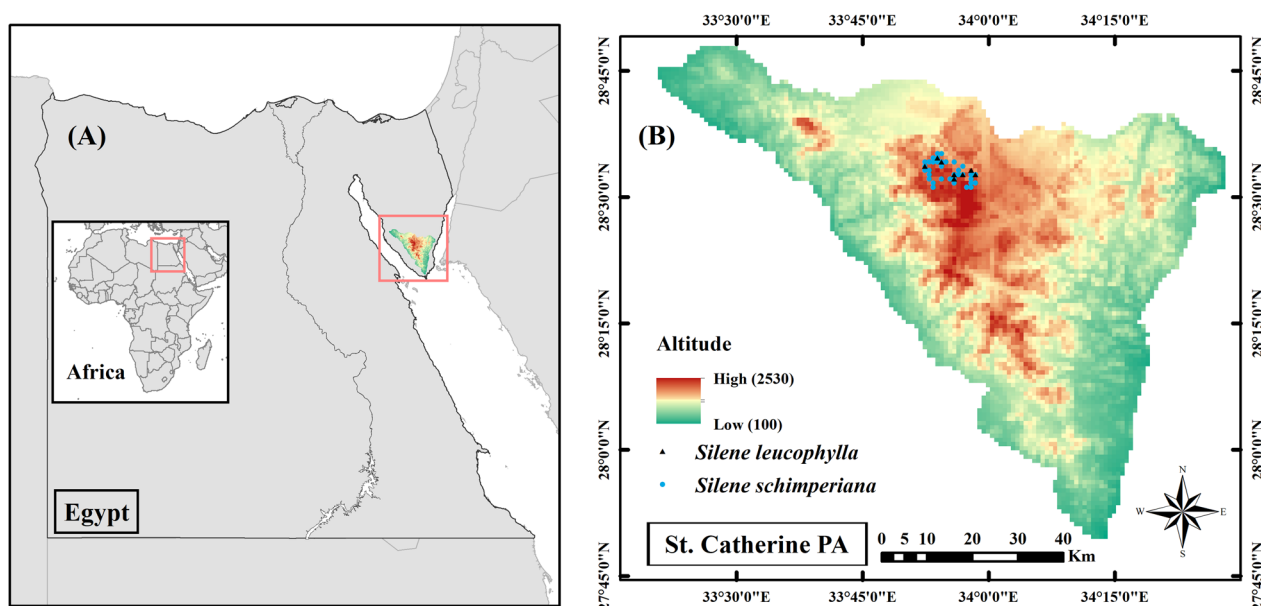


Fig. 1 **A** Location of the study area (Saint Catherine protected area) inside Egypt and Africa. **B** The elevation map (Altitude) of the study area, presenting the occurrence points of both *Silene leucophylla* and *Silene schimperiana*

2 Methods

2.1 Plant occurrences

Due to the lack of recorded lat./long. coordinates of focal species in published literature and national and international herbaria, the georeferenced location points of *S. leucophylla* and *S. schimperiana* were recorded during our field surveys in Spring 2020 and 2021, using a GPS (Garmin, model: *GPSmap*[®]76). In these surveys, we followed the main sites of their distribution as mentioned in the literature, with the help of St. Catherine PA's directors and the guidance of local Bedouins. Then, we filtered the points of each species, keeping one occurrence point only in each 1×1 km grid using ArcGIS ver. 10.8 [59], to avoid the spatial autocorrelation problem and prevent model overfitting. Consequently, *S. leucophylla* and *S. schimperiana* were represented by 8 and 35 occurrence points, respectively, to generate their models. The occurrence points of both species are shown together in Fig. 1.

2.2 Geospatial rasters of environmental parameters

In this study, we aimed to collect diverse geospatial environmental layers from trusted and reliable sources to develop the most fitting models. It is necessary to recognize that encompassing various types of environmental variables is essential for defining the fundamental niche of a species. Therefore, we derived 50 layers representing the current environmental conditions from online resources, classified under four categories: bioclimate, edaphic, hydrology, and topography (SD-Table 1).

2.2.1 Current environmental conditions

The bioclimatic variables were represented by 35 geospatial layers. We obtained the 19 WorldClim variables from the WorldClim (ver. 2.1) database [60] for the current period from 1970 to 2000 (<https://www.worldclim.org/>). The remaining 16 layers were downloaded from the ENVironmental Rasters for Ecological Modeling (ENVIREM—ver. 1.0) repository (<https://envirem.github.io/>) for the current period [61]. All layers were obtained with a spatial resolution of 30 arc-seconds (*ca.* 1 km²). These data are commonly used in modeling plant distributions since they are related to the ecology and physiology of plant species.

Eleven geospatial edaphic variables within the soil depth range of 0–30 cm were retrieved from the Soil-Grid (ver. 2.0) online dataset (<https://soilgrids.org/>), which is provided by the International Soil Reference and Information Centre (ISRIC)-World Soil Information [62]. These layers were obtained at 250 m spatial resolution and then resampled to the spatial resolution of 30 arc-seconds via ArcGIS (ver. 10.8).

We collected six geospatial topographical variables, as follows. The elevation layer (Altitude) was obtained from NASA's digital elevation model (DEM, <https://srtm.csi.cgiar.org/>) at a 90 m spatial resolution raster. The slope and aspect layers were derived from the elevation layer using ArcGIS (ver. 10.8). The landforms layer was obtained from the United States Geological Survey (USGS) website (<https://www.usgs.gov/>), which is available from the United States Department of the Interior. In addition, layers for the topographic wetness index (topoWet) and terrain roughness index (tri) were downloaded from the ENVIREM dataset [61].

Finally, three geospatial hydrology layers were derived from the elevation layer using ArcGIS (ver. 10.8). The elevation layer and all the layers derived from it were then resampled to reach a spatial resolution of 30 arc-seconds.

2.2.2 Future bioclimatic conditions

All future RCP8.5 geospatial bioclimatic variables were downloaded from the Community Climate System Model (CCSM4) [63], a global circulation model based on the Coupled Model Inter-Comparison Project Phase 5 (CMIP5), over three time periods: 2050 (average for 2041–2060), 2070 (average for 2061–2080), and 2080 (average for 2081–2100). We downloaded future geospatial bioclimatic variables for the same variable layers used in current conditions.

The 19 WorldClim variables for the 1st two future periods (2050 and 2070) were obtained from the WorldClim database, while those for the 2080 period were derived from the Climate Change, Agriculture, and Food Security (CCAFA) data portal (<http://www.ccafs-climate.org/>). We produced the ENVIREM variables for future conditions using the *envirem* package in R, following the procedures outlined by the ENVIREM developers [61]. The input datasets required for generating the ENVIREM layers are monthly rasters of minimum and maximum temperature, precipitation, and extraterrestrial solar radiation in the study area. Accordingly, we obtained the related temperature and precipitation layers from the CCSM4 model under the same emission scenario (RCP8.5) for the three future periods from the same resources. The solar radiation layers for the three future periods were generated using the *palinsol* package integrated into the *envirem* package in the R environment [64].

2.3 Species distribution modeling

In this section, we describe the steps involved in generating our ensemble SDMs. Figure 2 outlines the workflow of the developed SDMs.

Table 1 Environmental variables used to conduct different models in this study. Note, the M-All used all listed variables except the one with asterisk (*)

Type	Code/Unit	Parameter	Source
Bioclimate	Bio_2 [°C]	Mean Diurnal Range [Mean of monthly (Max temp-Min temp)]	WorldClim (ver. 2.1)
	Bio_3 [°C]	Isothermality (Bio_2/ Bio_7) (* 100)	
	Bio_4 (unitless)	Temperature Seasonality (SD *100)	
	Bio_7 [°C]	Temp Annual Range (Bio_5–Bio_6)	
	Bio_8 [°C]	Mean Temp. of Wettest Quarter	
	Bio_9 [°C]	Mean Temp. of Driest Quarter	
	Bio_12 [mm/year]	Annual Precipitation	
	Bio_14 [mm]	Precipitation of Driest Month	
	Bio_15 (unitless)	Precipitation Seasonality (Coefficient of Variation) = (Mean/SD*100)	
	Bio_17 [mm]	Precipitation of Driest Quarter	
	PET_Driest_Q [mm]	Mean monthly PET of driest quarter	
	PET_Wettest_Q [mm]	Mean monthly PET of wettest quarter	
	PET_Seasonality [mm]	Monthly variability in PET	
			ENVIREM
Edaphic	BD [kg/dm ³]	Bulk density of the fine earth fraction	SoilGrid (ver. 2.0)
	CEC [cmol(c)/kg]	Cation Exchange Capacity of the soil	
	CF [cm ³ /100cm ³ (vol%)]	Volumetric fraction of coarse fragments (> 2 mm)	
	Sand [g/100 g (%)]	Proportion of sand particles (> 0.05 mm) in soil	
	Silt [g/100 g (%)]	Proportion of silt particles (≥ 0.002 mm & ≤ 0.05 mm) in soil	
	Clay [g/100 g (%)]	Proportion of clay particles (< 0.002 mm) in soil	
	Soil_OC [g/kg]	Soil organic carbon content	
	OC_Density [kg/m ³]	Organic carbon density	
	OC_Stock [kg/m ²]	Organic carbon stocks	
	Soil_Nitrogen [g/kg]	Total nitrogen (N)	
Hydrology	Soil_pH	Soil pH	derived from Alt
	Hydr_F *	Hydrology filled	
	Hydr_FFD	Hydrology filled flow direction	
Topography	Hydr_FFA	Hydrology filled flow direction flow accumulation	derived from Alt
	Alt [m. asl]	Altitude	SRTM
	Aspect [degree]	Aspect	derived from Alt
	Slope [degree]	Slope	derived from Alt
	Landforms (unitless)	Land Surface Forms	USGS
	topoWet (unitless)	SAGA-GIS topographic wetness index	ENVIREM
	tri (unitless)	Terrain roughness index	ENVIREM

2.3.1 Testing for variables multicollinearity

We calculated the Pearson correlation coefficient (r) to test the multicollinearity between the variables using the SDMtoolbox (ver. 2.5) in ArcMap software [65]. After screening the results, we removed the variables with $(r) > |0.9|$. Correlation tests were performed for the variables under each category type separately, i.e., bioclimatic, edaphic, hydrologic, and topographic variables, and for all the variables together. The remaining variables were the variables that we incorporated into the models.

2.3.2 Current modeling

We developed six distinct models for each species. The first four models used variables under only one environmental category each, i.e., bioclimate, edaphic, hydrology, and topography. Hence, we named these models M-Bioclimate, M-Edaphic, M-Hydrology, and M-Topography. The fifth model, M-Predictors, utilized the variables with the highest contributions (> 5%) to the previously developed models as the environmental variables. The last model, the M-All model, incorporates the entire set of variables that are not highly

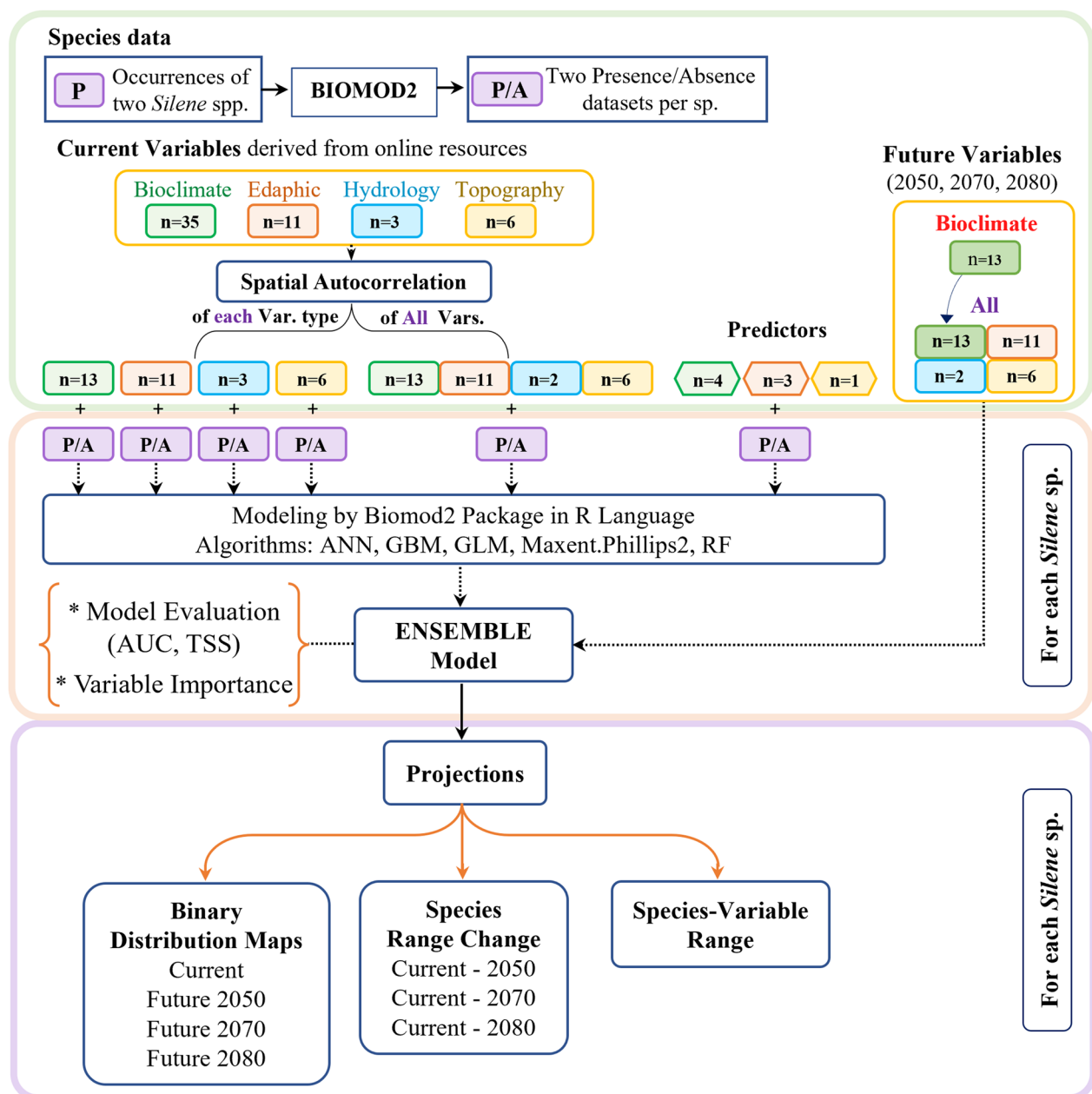


Fig. 2 Method flow chart of the study

correlated. We computed all the models using the *biomod2* Package [14, 66], implemented in the R environment [64].

As *biomod2* requires both the presence and absence data of the species to develop its models, we generated two presence/absence (P/A) datasets for each studied species [66]. Using the *Biomod_FormatingData()* function, we formatted the settings as follows: number of P/A replicates = 2, number of absences per dataset = 100, and absence simulation strategy = random.

2.3.2.1 Running machine learning (ML) algorithms (single models) In this context, we chose to perform five single modeling algorithms, which are Artificial Neural Network (ANN), Generalized Boosting Model (GBM), Generalized Linear Model (GLM), Maximum Entropy (Maxent), and Random Forest (RF). All these models were run with the default settings. For calibration, we employed two options: data splitting and model repetition. In the *Biomod_Modeling()* function, we applied both Data split = 80 and cross validation (CV) runs = 15.

2.3.2.2 Model validation The model's performance was assessed using the testing dataset to measure its precision in predicting known species distributions. We employed two accuracy meters for this evaluation: The Area Under the Curve (AUC) of the Receiver Operating Characteristic curve (ROC) [67] and the True Skill Statistics (TSS) [68]. The evaluation values were averaged across the CV runs. The AUC values range from 0 to 1, where values <0.7, 0.7–0.9, and >0.9 are considered poor, good and excellent performance, respectively [69, 70]. The TSS values range from −1 to +1, where values toward +1 indicate agreement between predictions and observations. TSS values <0.4 indicate poor performance, 0.4–0.8 indicate good performance, and >0.8 indicate excellent accuracy [71, 72].

2.3.3 Ensemble modeling and variable importance

In this step, we built the ensemble model by combining the predictions of the previously calibrated single models to decrease their predictive uncertainty [73]. *Biomod2* offers many methods for calculating ensemble models [66]. We chose to create them using the weighted average method of the prediction probabilities of all single models. Using the *Biomod_EnsembleModeling()* function, we constructed ensemble models with the following settings.

The previously developed models (5 models × 2 P/A datasets × 15 CV runs) were filtered using the evaluation metric quality threshold (TSS = 0.8). Subsequently, filtered models were used to create the ensemble. The evaluation of the accuracy and performance of the ensemble models was also done by calculating the same accuracy metrics: AUC and TSS. Another computed value for ensemble models is the variable importance (VI). It aims to evaluate the expected relationship between the response (the distribution) and the variables (the environment), i.e., assessing the variables' contributions to the model. Notably, in the Predictors model, we included all variables whose VI values exceeded 5% in the other five models as the environmental data.

2.3.4 Model projections

All previously developed models were applied to predict the species distribution across the entire area under current environmental conditions (projection over space). For future predictions, only the M-All and M-Bioclim models were projected into the future scenarios, spanning three time intervals: 2050, 2070, and 2080. In the M-All future projections, and despite its potential limitations to the models, we maintained uniformity by utilizing the same topography, edaphic, and hydrology variables employed in the current models because there are no projections of these variables in the future. Due to the importance of these variables in shaping species'

future distributions and to simplify the modeling process, we followed the methods of several authors who assumed that these variables would remain constant over time [2, 33, 41, 74–76]. Our aim was to assess the differences between the M-All (all variables) and M-Bioclim (climate only) models in projecting the future potential distributions of the studied species. In addition, we aim to explore how incorporating these non-climatic variables could influence the prediction of their distribution under future conditions.

2.3.5 Visualizing modeling output

The models of *biomod2* produce their predictions in the form of two types of maps: the suitability maps and binary maps, i.e., Presence/Absence (P/A) maps. *Biomod2* produces P/A maps by calculating the best threshold that maximizes the percentage of accurately predicted presence and absence for ROC curves of each species per model. Probability values above this threshold were considered present (1), and values below that threshold were considered absent (0). We presented only the binary maps produced by the ensemble models for the two species under current and future conditions.

2.3.6 Further analysis

2.3.6.1 Change detection in the species range The differences in the distribution area of a species between the current and future projections were calculated in *biomod2/R*. The *Biomod_RangeSize()* function in *biomod2* calculates the differences in suitable areas of a species pixel by pixel. This function produces two types of outputs. First, a classified map of four distribution change classes: gain, stable presences, stable absences, and loss. Second, a table in which *biomod2* calculates the areas of current range sizes (km² and % of the total St. Catherine PA), areas of loss and gain (km² and % of the predicted current range size), the percent of the total species range change (% gain—%loss), areas of future range size (no distribution) = [current range size (km²)—loss (km²)], and future range size (full distribution) = [(current range size (km²) + gain (km²)) − loss (km²)]. The areas of future range sizes were also calculated in both km² and % of the total St. Catherine PA.

2.3.6.2 Species-variable range We calculate the species-variable ranges of the essential variables that defined the preferred environmental conditions for each species to grow and distribute. From the binary maps, we extracted the values of every cell determined as a presence location (cell value = 1) of the species from all the environmental layers in the ArcMap platform. Then, using simple commands in Excel, we calculated the minimum (Min), mean ± standard deviation (SD), and maximum (Max) of

these points with respect to each environmental predictor in the present and future.

3 Results

3.1 The selected environmental variables

The application of Pearson's correlations to the 35 bioclimatic variables revealed the presence of 22 variables (11 pairs) with $(r) > |0.9|$. Removing these highly collinear variables keeps only 13 less correlated bioclimatic variables to use in the M-Bioclimatic models. However, the correlations between either the edaphic, hydrology, or topographic variables revealed that the variables of each set were not highly correlated. Therefore, the variables of each type were used in developing their corresponding models (i.e., the M-Edaphic, M-Hydrology, and M-Topographic models). Finally, the correlations between the complete variables set revealed a final set of 32 variables that were not highly correlated and were used to construct the M-All model (Table 1).

3.2 Model evaluation

Notably, the six ensemble models constructed for both *Silene* species consistently shown excellent levels of predictive performance in terms of both AUC (> 0.97) and TSS (> 0.91) scores (Table 2).

However, the results revealed variations in the evaluation values among models that used different algorithms/methods and input variables. Based on both evaluation metrics, the predictive ability of all the ensemble models outperformed that of the other single algorithms. Although most of our single models demonstrated excellent predictive performance, with values of AUC (> 0.9) and TSS (> 0.8), the accuracy scores of some models (mostly GLMs) indicated only good levels of accuracy (the evaluations of these single models are shown in SD-Table 2 & 3). Additionally, when comparing the predictive performances of the models that used different environmental variable sets, the M-All model outperforms the other models, followed by the M-Bioclimatic model, for both studied species. This result implies the positive effect of adding various predictors in enhancing

model performance. Regarding the accuracy metrics, the performance of the M-all model increased by 0.001 and 0.014 in terms of both the AUC and TSS values, respectively, for *S. schimperiana* and by 0.001 in terms of the TSS values for *S. leucophylla*.

3.3 Variable importance

Figure 3 illustrates the contributions of the variables to the different distribution models constructed for the two studied species.

In detail, for *S. leucophylla*, the M-All model indicated that annual precipitation (Bio_12—36.2%), mean temperature of the driest quarter (Bio_9—27.6%), elevation (Alt—7.8%) and precipitation seasonality (Bio_15—5%), contributed to 76.6% of the total. The cumulative contributions of the bioclimatic, soil, hydrology, and topography variables were 79.9%, 3.4%, 4.8%, and 11.9%, respectively. According to the M-Bioclimatic model, the strongest contributors to the *S. leucophylla* distribution were the mean monthly potential evapotranspiration (PET) of the driest quarter (PET_Driest_Q—45.2%), annual precipitation (Bio_12—21.5%), temperature seasonality (Bio_4—10.4%), and precipitation seasonality (Bio_15—9.8%).

On the other hand, for the *S. schimperiana* distribution, Bio_9 (48.3%) followed by Bio_15 (11.3%), Silt content (6.3%), and mean monthly potential evapotranspiration (PET) of the driest quarter (PET_Driest_Q—5.3%) exhibited the highest contributions to the M-All model, with a total contribution of 71.1%. The cumulative contributions of the different variable types were 86%, 8.9%, 1%, and 4.1% for the bioclimate, soil, hydrology, and topography, respectively. The M-Bioclimatic model indicated that PET_Driest_Q (38.8%), Bio_12 (18.3%), Bio_9 (13.5%), and Bio_15 (7.3%) were the most important predictors of *S. schimperiana* distribution.

3.4 Projections of species distribution

The predicted binary distribution maps of both *Silene* spp. under current and future conditions are shown in Figs. 4 and 5. The results highlight that the highly suitable

Table 2 AUC and TSS data evaluating the performance of the six models (ensemble) predicting the current distribution of the studied species

Species	Metric	Model					
		M-Bioclimatic	M-Edaphic	M-Hydrology	M-Topography	M-Predictors	M-All
(a) <i>S. leucophylla</i>	AUC	1.000	0.997	0.978	0.993	1.000	1.000
	TSS	0.999	0.995	0.920	0.973	0.995	1.000
(b) <i>S. schimperiana</i>	AUC	0.998	0.995	0.989	0.992	0.998	0.999
	TSS	0.970	0.960	0.915	0.929	0.965	0.984



Fig. 3 Percent of variables contributions to the different distribution models of the two studied species

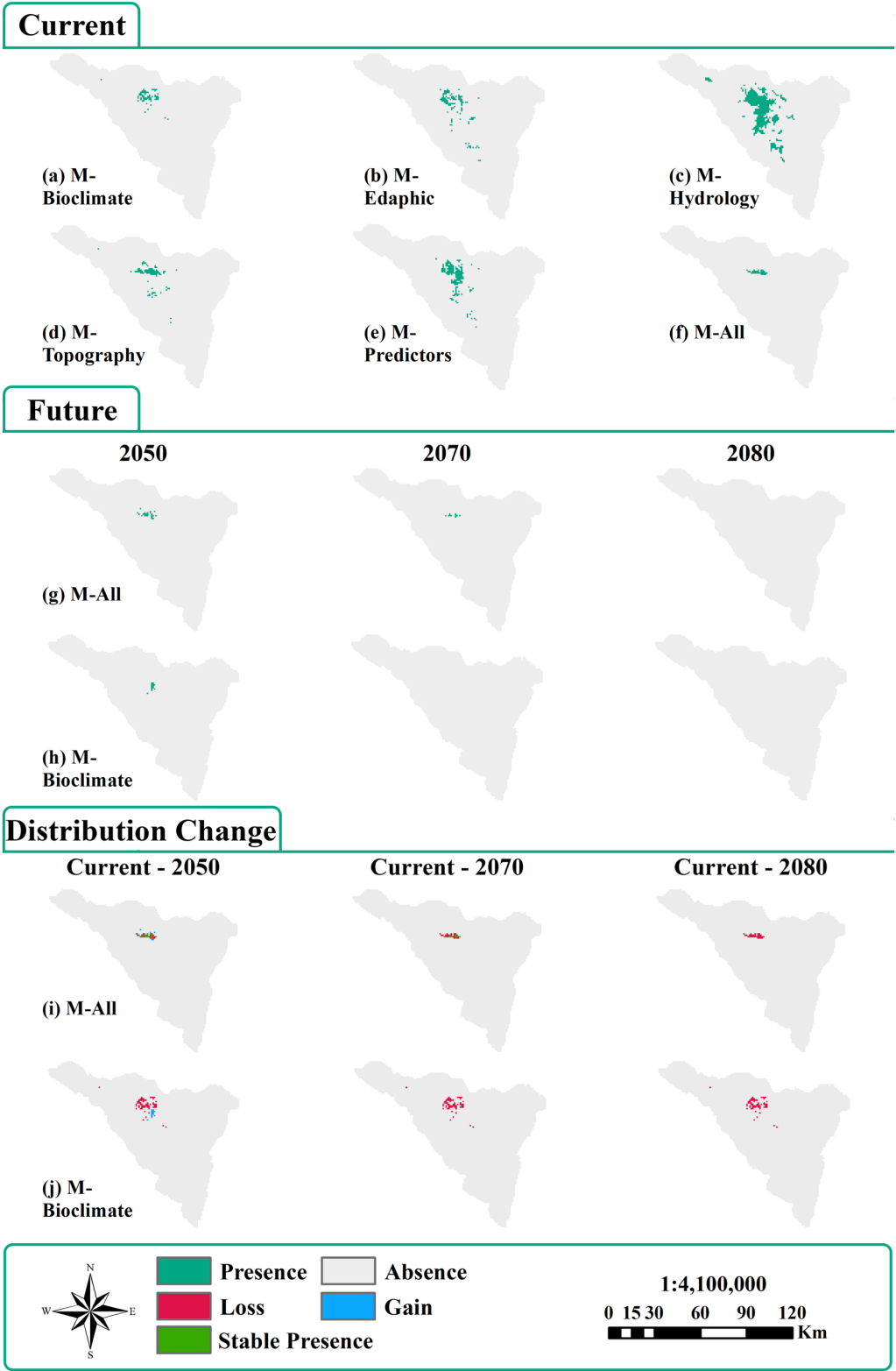


Fig. 4 Binary maps showing the predicted suitable areas of *S. leucophylla* in current (a–f) and future projections (g–h). Maps of changes in the suitable distribution areas of *S. leucophylla* between current and each of the projected future time periods per model (i–j) are also shown

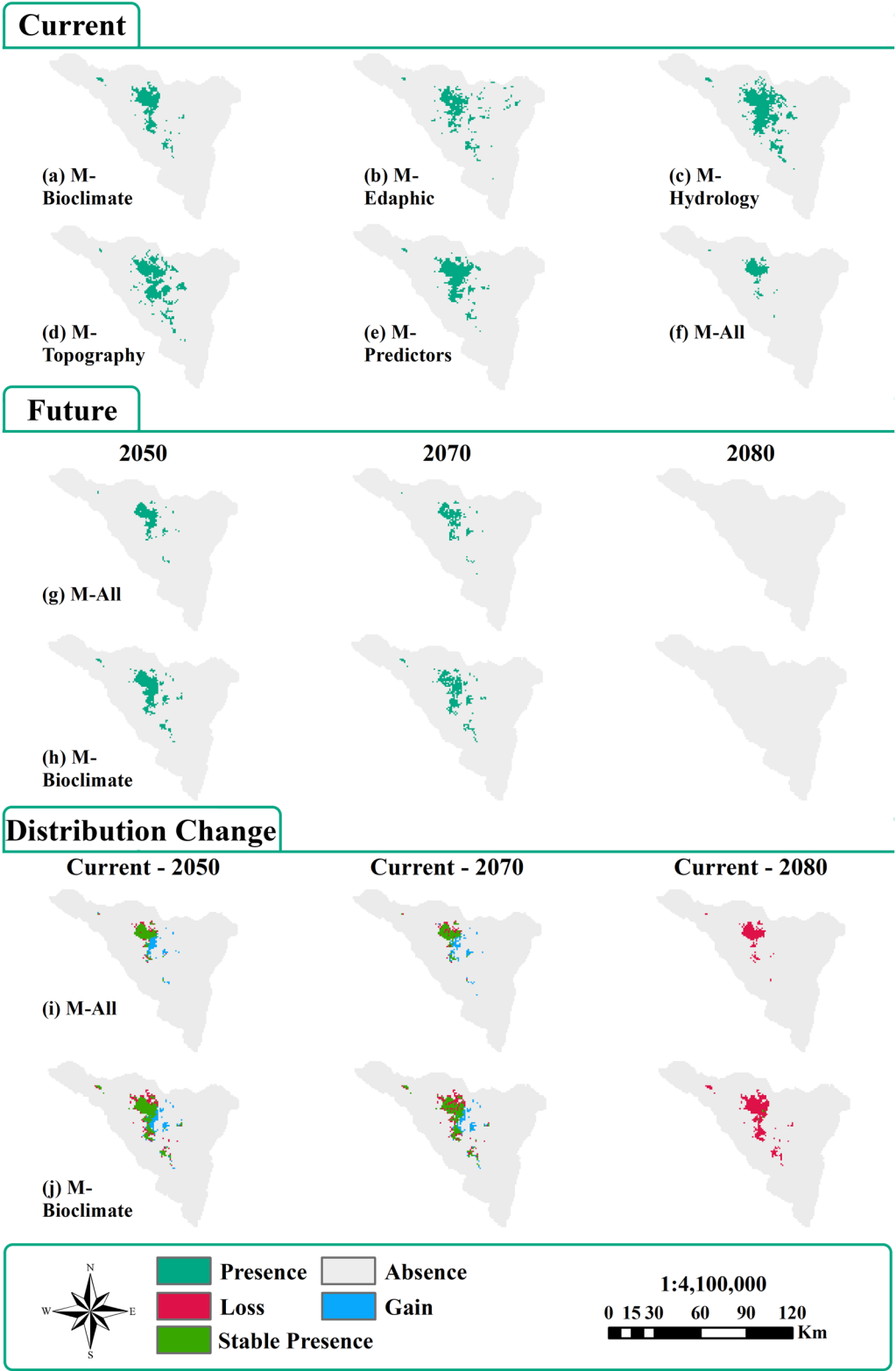


Fig. 5 Binary maps showing the predicted suitable areas of *S. schimperiana* in current (a–f) and future projections (g–h). Maps of changes in the suitable distribution areas of *S. schimperiana* between current and each of the projected future time periods per model (i–j) are also shown

Table 3 Areas of the current suitable range size [km² (%)] of both studied species predicted by the six constructed models

Species	Predicted current range size [km ² (%)]/Model					
	M-Bioclimate	M-Edaphic	M-Hydrology	M-Topography	M-Predictors	M-All
(a) <i>S. leucophylla</i>	46.9 (1.0%)	82.6 (1.7%)	325.8 (6.8%)	66.2 (1.4%)	123.5 (2.6%)	23.1 (0.5%)
(b) <i>S. schimperiana</i>	251.5 (5.3%)	297.6 (6.2%)	480.6 (10.1%)	381.6 (8.0%)	336.3 (7.0%)	125.0 (2.6%)

habitat for both species is concentrated in the high mountain areas in the northern middle sector of the study area. Additionally, the areas of current suitable habitats predicted by the six developed models are shown in Table 3. As shown, there are differences in the predicted suitable areas between different models, with an apparent decrease in the predicted areas in those using many explanatory variables compared with those using lower numbers of variables. Focusing on the M-All and M-Bioclimate models, there is a noticeable trend that the M-Bioclimate model's predicted areas were approximately twice the suitable areas predicted by the M-All model (Table 3).

3.4.1 Potential current and future geospatial distributions of *S. leucophylla*

Under current environmental conditions, the predicted distribution areas of *S. leucophylla* are 23.1 km² and 46.9 km² (0.5% and 1.0% of the St. Catherine PA), as projected by the M-All and M-Bioclimate models, respectively (Fig. 4). However, in the 2050 future projection, the results of both models indicated a significant reduction in the current suitable area for *S. leucophylla*, with losses of 11.2 km² (48.4%) and 46.9 km² (100%), and gains of only 5.2 km² (22.6%) and 9.7 km² (20.6%), respectively. Consequently, for *S. leucophylla*, the total size of suitable areas predicted by the M-All model in the mid-century would become 17.1 km² (0.4% of the St. Catherine PA) if the species successfully disperses, or 11.9 km² (0.25%) if dispersal did not occur (Table 4). On the other hand, the mid-century prediction of the M-Bioclimate model showed a possible complete loss of the species if dispersal fails and only 7.4 km² (0.2%) if dispersal could occur. Similarly, the M-All (2070) projection anticipated a continuous decrease in the current range size of *S. leucophylla* (15.6 km² (67.7%)), leaving only 7.4 km² (0.16% of the St. Catherine PA) as suitable areas for the species. Unfortunately, in the 2070 M-Bioclimate model's projection and 2080 projections of both models, the results predicted the same possible complete loss of the current suitable ranges of both species (Fig. 4 and Table 4).

3.4.2 Potential current and future geospatial distribution of *S. schimperiana*

The projected distribution areas of *S. schimperiana* under current conditions (Fig. 5) are 125 (2.6%) and 251.45

km² (5.3%), as predicted by the M-All and M-Bioclimate model, respectively. In both the 2050 and 2070 predictions, our models expected a notable expansion in the current suitable area of the species. For these future projections, the M-All model expected gains of 56.5 km² (45.2%) and 55.8 km² (44.6%), compared with losses of 26.8 km² (21.4%) and 32.0 km² (25.6%), respectively. The total range size of *S. schimperiana* will reach 154.7 km² and 148.8 km² if the species is fully dispersed or 98.2 km² and 93.0 km² if dispersal does not occur (Table 4). The predicted distribution in the 2070 forecast showed a slight decrease in the suitable areas compared with those in the 2050 projection. Similarly, in the 2050 and 2070 future projections, the M-Bioclimate model predicted both expansions and contractions of the current suitable range of *S. schimperiana*. *S. schimperiana* is subsequently predicted to be distributed over 250.7 km² and 148.8 km² if the dispersal of the species succeeds or 98.2 km² and 93.0 km² if dispersal fails. Unfortunately, in the 2080 forecasts, a substantial loss is predicted in the current suitable range of the species, with total species range changes of −100% and −98.5% for the M-All and M-Bioclimate models, respectively (Fig. 5 and Table 5).

3.5 Species-variable ranges

Our model predictions successfully delineated the preferred environmental conditions under the current conditions for both *S. leucophylla* and *S. schimperiana*. The variable ranges of the observed occurrences were included within the predicted suitable variable range, indicating alignment with the current suitable conditions for the species to grow and thrive. The observed and predicted suitable ranges of the most essential variables shaping the distributions of both species are shown in Table 5.

The M-All model revealed that the suitable elevation (Alt) of *S. leucophylla* ranged from 1643 to 2130 m a.s.l, with an average elevation of 2161 m a.s.l. under current environmental conditions. With time, the minimum suitable elevation increased to 1871 m a.s.l and 1958 m a.s.l in the 2050 and 2070 future conditions, respectively. It prefers to grow in soil with an average silt percentage between 26.4% and 31.7%, with north facing (N, NE, and NW) directions and steeper slopes. With respect to the bioclimatic variables, the optimum annual precipitation

Table 4 Changes in the predicted distribution areas between current and the three future projections of both *Silene* spp. Areas of current and future range sizes (km² and % of the total St. Catherine PA), and areas of loss, and gain (km² and % of the predicted current range size) and the percent of the total species range change (% gain–%loss) are shown

Species	Model	Current Range Size [km ² (%)]	Projection difference	Loss [km ² (%)]	Gain [km ² (%)]	Species Range change [%]	Future Range Size	
							No Dispersal [km ² (%)]	Full Dispersal [km ² (%)]
<i>S. leucophylla</i>	M-All	23.1 (0.5%)	Current—2050	11.2 (48.4%)	5.2 (22.6%)	–25.8	11.9 (0.25%)	17.1 (0.4%)
			Current—2070	15.6 (67.7%)	0.0	–67.7	7.4 (0.16%)	7.4 (0.16%)
			Current—2080	23.1 (100%)	0.0	–100	0.0	0.0
	M-Bioclimate	46.9 (1.0%)	Current—2050	46.9 (100%)	9.7 (20.6%)	–79.4	0.0	9.7 (0.2%)
			Current—2070	46.9 (100%)	0.0	–100	0.0	0.0
			Current—2080	46.9 (100%)	0.0	–100	0.0	0.0
	M-All	125.0 (2.6%)	Current—2050	26.8 (21.4%)	56.5 (45.2%)	23.8	98.2 (2.1%)	154.7 (3.2%)
			Current—2070	32.0 (25.6%)	55.8 (44.6%)	19.0	93.0 (1.9%)	148.8 (3.1%)
			Current—2080	125.0 (100%)	0.0	–100	0.0	0.0
<i>S. schimperiana</i>	M-Bioclimate	251.45 (5.3%)	Current—2050	69.9 (27.8%)	69.2 (27.5%)	–0.3	181.5 (3.8%)	250.7 (5.3%)
			Current—2070	90.0 (35.8%)	63.2 (25.1%)	–10.7	161.4 (3.4%)	224.7 (4.7%)
			Current—2080	247.7 (98.5%)	0.0	–98.5	3.7 (0.08%)	3.7 (0.1%)

(Bio_12) ranged from 62 to 105 mm at maximum, with a mean of 76.5 mm under the current conditions. In the 2050 and 2070 forecasts, the predicted average annual precipitation suitable for *S. leucophylla* increased to 87 mm and 96.6 mm, respectively. The suitable mean temperature of the driest quarter (Bio_9) for *S. leucophylla* ranged from 18 to 22.7 °C. According to the prediction, the species adapts to high precipitation seasonality (Bio_15) with averages of 66.5 at minimum and 79.5 at maximum. The suitable mean potential evapotranspiration values of the driest quarter (PET_Driest_Q) were 142.2 ± 1.8 , 152.6 ± 3.1 , and 158.1 ± 2.1 in current, 2050, and 2070, respectively. In comparison, the current and future predictions of the M-Bioclimate model were, to some extent, in alignment with those of the M-All with respect to the preferred bioclimatic variables. However, when we superimposed the layers of other variables on the current prediction, the M-Bioclimate model is found to predict wider ranges of suitability of topography and soil conditions, such as elevation, slope degrees, and silt content (Table 5).

For *S. schimperiana*, the model determined that, under current conditions, its optimal altitude range was between 1352 and 2441 m, with an average of 1956 m a.s.l. In contrast, for the 2050 and 2070 predictions, the suitable elevation increased to 1675 m a.s.l. at the lowest level and 2530 m a.s.l. at the highest level. The species is found to grow in all aspect directions and lower slopes. Considering the bioclimatic factors, *S. schimperiana* can adapt to a broad range of mean temperatures of the driest quarter (Bio_9), with suitable values ranging from 18 °C to 23.4 °C. The suitable average PET_Driest_Q is slightly greater than that of *S. leucophylla*, increasing from 144.6 ± 2.2 mm under the current conditions to 154.9 ± 2.8 mm and 160.8 ± 1.8 mm for the 2050 and 2070 forecasts, respectively. The models also revealed that *S. schimperiana* can grow under relatively low mean annual precipitation (Bio_12) under current and future conditions. Furthermore, the results showed that *S. schimperiana* prefers more silty soils than *S. leucophylla* does. A similar trend is found with the M-Bioclimate model in capturing the most suitable and

Table 5 (continued)

Variable	Obs./Pred	<i>S. leucophylla</i>						<i>S. schimperiana</i>					
		M-All			M-Bioclimate			M-All			M-Bioclimate		
		Min	Mean ± SD	Max	Min	Mean ± SD	Max	Min	Mean ± SD	Max	Min	Mean ± SD	Max
OC_ Stock [kg/m ²]	Obs	1.6	1.9 ± 0.2	2.1	1.6	1.9 ± 0.2	2.1	1.4	1.8 ± 0.2	2.1	1.4	1.8 ± 0.2	2.1
	Current	1.9	2 ± 0.1	2.3	1.6	1.9 ± 0.1	2.1	1.4	1.9 ± 0.2	2.2	1.3	1.8 ± 0.2	2.3
	2050	1.8	2 ± 0.1	2.2	–	–	–	1.4	1.9 ± 0.1	2.3	–	–	–
	2070	1.9	2.1 ± 0.1	2.2	–	–	–	1.6	1.9 ± 0.1	2.3	–	–	–
	2080	–	–	–	–	–	–	–	–	–	–	–	–

preferred bioclimatic conditions of *S. schimperiana* but in broader ranges of topo-edaphic conditions.

For more detailed suitable variable ranges for both species under different environmental conditions, see Table 5.

4 Discussion

The use of SDMs to forecast the potential distributions of plant species under climate change has become a familiar and valuable tool for ecological and biogeographical research. Ecologists and statisticians are increasingly interested in ML methods because of the abundance of available data to address the complex relationships between species and their environments [77–82]. However, it is necessary for researchers dealing with ML to understand these models in detail to reduce the potential for inaccurate results. This should be achieved by understanding aspects like model limitations, uncertainties, and the risk of potential model overfitting. Therefore, in this section, we address these aspects consecutively.

Many authors have highlighted that one of the main restrictions of SDMs is that model predictions can be extremely sensitive when various methodologies are employed [83, 84]. Because of the sensitivity of the prediction outcome to the method, a thorough comparison of the evaluations across different modeling methods is required to identify the best-fit one. In this study, we decided to build an ensemble modeling strategy, which several authors have suggested is a suitable approach that eliminates model uncertainties and limitations while outperforming other single models [15, 16, 85–87]. Moreover, many studies have shown that ensemble models are more accurate in predicting species habitat suitability [88–90]. Our findings were consistent, as all our ensemble models achieved remarkable performance with the highest prediction capacities based on AUC and TSS assessments.

Another key factor contributing to SDMs limitations and uncertainties is the selection of environmental

variables for incorporation into the models. Plants respond to the interaction of various factors that affect their survival and distribution. The utilization of different factors can lead to different SDMs predictions. Therefore, incorporating a specific predictor into SDMs depends on the purpose of the modeling and its biological relevance to the studied species [25]. Many studies have recommended including edaphic factors when modeling plant distributions [32, 41, 74, 91–96], whereas others have reported that using topography and hydrology is highly important [33, 97–99]. Therefore, it is recommended that SDMs be provided with relevant combinations of variables to decrease any uncertainties and to obtain better predictions.

In this study, the results were compatible, as the M-All model, which incorporated all the variable types (e.g., climate, topography, edaphic, and hydrology), outperformed the other models that used variables related to only one category. It showed excellent levels of predictive performance, with both AUC and TSS values ~ > 9.84. Although its predicted suitable areas were the smallest among the different models, it was the best model that captured the ecological niches of the species. This statistical tendency of projecting smaller ranges when using more predictors has been documented [74], as adding more environmental constraints can eliminate species' suitable habitats. Conversely, models relying only on edaphic, hydrologic, or topographic factors are insufficient to predict habitat suitability for both *Silene* spp. These models result in an unrealistic predictive map compared with the known defined habitats and the actual distribution range of both target species in the St. Catherine area.

On the other hand, the performance of the M-Bioclimate model was slightly lower than that of the M-All model in terms of both the AUC and TSS values. The M-Bioclimate model also succeeded in defining the suitable climatic niche of the studied species. However, it has failed to capture their ecological and geographical niches.

This may be attributed to the fact that models incorporating only climatic variables exclude the indirect effects of non-climatic variables in predicting species habitat suitability. In addition, the M-Bioclim model tends to predict wider ranges of species suitability (overprediction) than the M-All model does because it misses other important interactions. Consequently, including more relevant non-climatic variables can enhance model accuracy and fine-tune model predictions, yielding a more realistic distribution prediction that captures the significance of the complex interaction between climate and geophysical factors [2, 91, 100].

In addition, we believe that the full image of a species' ecological range is also defined by biotic interactions besides the abiotic factors. Biotic interactions, including species competition, herbivory, and facilitation, play critical roles in shaping species distributions [101–103]. However, very few studies have included biotic factors in their models. Therefore, we recommend integrating biotic interactions into SDMs, which can add more power in understanding and predicting species' habitat suitability. Finally, these findings highlight the additive benefits of including soil and topo-hydrological predictors when modeling species distributions. They also demonstrated the indirect effects of these variables, which may not be apparent in the variable contributions [33, 104].

Other issues to address when conducting SDMs are the spatial resolution of the environmental predictors used and the limited sample size of the species. In this study, we used environmental layers with a relatively coarse spatial resolution of approximately 1 km² (30 arc-seconds). We acknowledge that the relatively coarse spatial resolution could indeed limit the precision of predictions, particularly in regions with high environmental heterogeneity. However, due to the unavailability of consistent, high-quality data, particularly climate data, at a global scale, environmental variables of 1 km² have been extensively used in SDM studies. Another limitation of our study is the low sample size for *S. leucophylla*, which could introduce uncertainty into the model predictions for the species. Unfortunately, this low number of occurrences represents the actual current distribution of *S. leucophylla* which was revised by the protected area affairs office, expertise, and local Bedouins.

Additional limitation to SDMs is the potential for model overfitting. In ML, overfitting occurs when an algorithm fits too closely or even exactly to its training data, resulting in a model that cannot make accurate predictions from any other data and is unable to generalize well to new data. In this study, we took some regularizations to avoid this problem, such as reducing the spatial autocorrelation of occurrence data, which could only add noise to the models; selecting non-collinear sets of

environmental predictors; calibrating the models with data partitioning and CV runs; simplifying the modeling design; and using ensemble methods. These techniques are commonly used in SDM research [15, 21, 66, 89]. As a result, the best model achieved excellent performance in predicting test data and projecting to new conditions.

Our model results showed that, under current climatic conditions, the suitable habitat distribution of both *Silene* spp. is situated primarily within the high mountains of the central northern boundaries of St. Catherine PA. The predicted suitable habitats fit with the known documented distribution of both *Silene* species and other endemic plants in the study area, as observed through field surveys and reported in the literature [3, 5, 105–114].

S. leucophylla is known to occur at only eight sites, namely, Al-Jebel Al-Ahmar (Abu-Haman and Kahf Al-Ghoulah), Mount Mousa (Farsh Elias and Farsh Al-Saf-safah), and Wadi Jebal (Al-Tebah, Al-Ma'een, Shaq Saqr, and Teneiah). However, the M-All model projected four other suitable main sites for the species: Shaq-Mousa and Um-Sellah in Mount Catherine, and Al-Shaq and Abu Al-Yasaa' in Wadi Jebal. Only the Al-Shaq site was proven through our field validations. For M-Bioclim current prediction, the model was found to predict many suitable sites for *S. leucophylla* distribution, including several sites belonging to Mount Mousa, Wadi Jebal, Wadi Al-Arbeen, Wadi EL-Garagniah, and Shaq-Mousa. After field validation, we found only one occurrence point at two sites (Al-Shaq and Abu-Geefah). These new individuals were freshly inhabiting both sites.

For *S. schimperiana*, the known occurrences have been reported in large areas, such as Wadi Etlah, Wadi Talaa, Mount Catherine, Wadi Al-Arbeen, Shaq-Mousa, Al-Jebel Al-Ahmar, Mount Mousa, many sites inside Wadi Jebal (Abu-Geefah, Abu Al-Yasaa', Farsh Al-Rommanah, Al-Tebah, Al-Qalt Al-Azraq, Um-Sellah, and Abu-Twita), and others. However, the M-All and M-Bioclim models projected wider areas of potential distribution. After field validation, it was found at other sites, such as Shaq-Mousa, Um-Sellah of Mount Catherine, Sleetbat, Toboq, Wadi EL-Garagniah, and Al-Shaq. Unfortunately, we could not access the distal projected areas in the southern part of the study area.

Although we did not find the species present in some of the projected suitable areas, those areas were considered prominent for their distributions. In addition, those areas could be utilized as a conservation plan for both species through enforcement and reintroduction.

The results also revealed that the geographical distribution of *S. leucophylla* is narrow and highly fragmented, with a small population size. On the other hand, the geographical distribution of *S. schimperiana* was wider

but was also visibly fragmented. These distribution differences may be owing to the variations in the adaptive capabilities of the two species to different environmental conditions. A study conducted by Rabei et al. (2021) investigating the impacts of genetic diversity on the distribution of *Silene* species at St. Catherine PA revealed low genetic diversity among *S. leucophylla* populations, whereas *S. schimperiana* populations shown high genetic diversity. According to different phylogenetic tests, they found that *Silene* species inhabiting St. Catherine PA can be divided into three clades, where *S. leucophylla* and *S. schimperiana* are in two different clades. Also, according to the neighbor-joining evolutionary relationship, the branch length carrying *S. schimperiana* was distal to other branches. Accordingly, they suggested that *S. schimperiana* has high variation in the genetic constituents of its populations with recent evolutionary characteristics [57]. This indicated its high adaptive capacity to different environmental conditions and habitats, explaining its wide distribution compared with that of *S. leucophylla*. In addition, cluster analysis among *Silene* species based on their anatomical characteristics also placed the two species in separate clades [106].

Additionally, our models revealed that the predicted suitable habitats of both *Silene* species are in line with their known environmental requirements and suggested that their current distribution is controlled by elevation, precipitation, and dry season-related variables.

The suitable environmental conditions for *S. leucophylla* under current conditions are predicted to include elevations higher than 1600 m a.s.l., annual precipitation range of 62–105 mm, low mean temperature of the driest quarter of 19.7–22.4 °C, and optimum precipitation seasonality between ~65–69. The species was also found to dominate mountain slopes and was distributed in all aspect directions except South and Flat, with a preference for North-facing directions. These results are compatible with those of Omar & Elgamal (2021), who reported that *S. leucophylla* was primarily restricted to slopes of deep mountain cracks, gorges, and terraces (minor percentage) at elevations between 1800 and 2100 m a.s.l. and in aspects of the most frequent Northeast and North directions [3].

For *S. schimperiana*, the suitable current environment spans a wider altitudinal range between ~1350 m and 2440 m a.s.l., with a slightly warmer mean temperature of the driest quarter (19.7–23.4 °C) and a broader range of annual precipitation (40–105 mm). Thus, it can grow under conditions of greater precipitation seasonality (~61–71.5). These results are in accordance with those of Omar (2017), who reported that *S. schimperiana* naturally grows in four microhabitats, which are terraces, gorges, wadi beds, and slopes, at low temperatures and

rainfall zones and elevations higher than 1700 m a.s.l. [114].

Although both the M-All and M-Bioclimate models capture nearly the same climatic niche of both species, the M-Bioclimate model fails to detect their preferred topographic and edaphic conditions. For example, when we superimposed topo-edaphic layers on M-Bioclimate current predictions, wider ranges of suitable elevation, slope, and soil silt content were suggested. This is why it seems to predict larger suitability areas than the M-All model does. Overall, it seems likely that the effects of adding other non-climatic variables to the M-All model on plant distributions are conservative (not overestimates).

Among the main predictors that strongly influence our SDMs, annual precipitation, mean temperature of the driest quarter, altitude, and precipitation seasonality were the most important environmental variables affecting the *S. leucophylla* distribution, whereas the mean temperature of the driest quarter, and precipitation seasonality were the most influential on the *S. schimperiana* distribution. This finding is consistent with the fact that precipitation, particularly in mountain ecosystems, plays a crucial role in shaping the distribution of plant species and acts as a limiting factor for plant dispersal in arid regions [115, 116]. Although elevation did not gain much value in terms of variable contributions to the models, we also believe that it is an important factor that influences the predictions of SDMs, particularly in mountain ecosystems. The importance of elevation along with climatic factors has been confirmed for the spatial distribution of many plants in many studies [2, 117].

Model predictions for future conditions disclosed a massive gradual reduction in the suitable current geographic distribution of *S. leucophylla* under the 2050 and 2070 conditions. The only gain is predicted under the 2050 future climate by both the M-All and M-Bioclimate models. In contrast, the M-All future predictions of *S. schimperiana*, for both the 2050 and 2070 forecasts, expected a notable expansion in the species geographic distribution to new locations of higher elevations compared with smaller contractions in some locations of lower elevations, which would become less suitable with a changing climate. The future M-Bioclimate model projections showed a gradual contraction of suitable areas for *S. schimperiana* until the late century, along with little range expansion. These results could become true if *S. schimperiana* succeeds in fully dispersing to such new locations. If not, the species would be subjected to a slight decrease in its distribution area. Plants vary in their responses to future climate change, mainly due to their variations in physiological, genetic, or phenological characteristics [118]. In particular, plants with broad

ecological niches are expected to have a greater capacity to adapt to climate change than those with narrow ecological niches [119, 120]. As mentioned earlier, this may be because of the differences in genetic diversity between the two species, implying various adaptability of both species to different habitats and environmental conditions. In addition, ecological barriers, either natural (elevation and water availability) or human-induced, play important roles in restricting the ability of plants to distribute new habitats. These contrasting results are also found in the literature that models the effect of climate change on the distribution of mountain plants, with some studies predicting range expansions [121] and others predicting range contractions [122].

Unfortunately, in late century projections (2080), both models projected a complete disappearance of the suitable areas for the two species (except for *S. schimperiana*'s M-Bioclimatic prediction), suggesting full species extinction. Although this is a concerning result, these predictions are based on current data and assumptions and may not fully capture the complex dynamics of species responses to climate change. The actual response of a species to climate change could be influenced by other factors, such as adaptive evolution or phenotypic plasticity. In addition, potential mitigation efforts could influence the actual response of species in the future.

Finally, the IUCN Red List identified both *S. leucophylla* and *S. schimperiana* as endangered species due to many natural and anthropogenic threats they face in their geographically restricted habitats [107, 108]. In brief, the main natural threats are long-lasting drought and irregular rainfall during the year (precipitation seasonality), followed by natural topographic barriers that eliminate their dispersal to new suitable areas. The main anthropogenic threats include overgrazing, especially by domestic animals (e.g., camels and goats and feral donkeys), over-collection (for medicinal or other uses), and uncontrolled and continuous collection for scientific research. Consequently, the populations of both species are highly fragmented, and more efforts should be made to conserve these two important endangered endemic and near-endemic plants. This could be achieved by initiating programs of assisted migrations in the wild, including reinforcements and reintroductions. These activities should be accompanied by increasing public awareness and implementing policies to reduce human-related threats.

5 Conclusion

This study illustrated the importance of incorporating various environmental variables, including bioclimatic, edaphic, and topographic factors into SDMs to yield

more accurate predictions of plant species distributions. For both studied species, the M-All model was the best model that accurately predicted their current distributions and was consistent with their actual/observed distributions. Additionally, this study revealed that the studied species respond differently to future climate scenarios (2050 and 2070 predictions) owing to their different abilities to adapt to the new environmental conditions. Under both future climatic conditions, a significant range contraction was predicted for the distribution of *S. leucophylla* (a narrow ecological niche endemic), whereas both range expansion and contraction of the current suitable habitats were projected for *S. schimperiana* (a wider ecological niche near-endemic). Unfortunately, in 2080 predictions, both species might have gone extinct from the entire area. The study also disclosed that the populations of both species were highly fragmented due to many natural and anthropogenic threats. To mitigate the risk of extinction for *S. leucophylla* and *S. schimperiana* from their natural habitats, both ex-situ and in-situ conservation techniques are urgently needed.

Abbreviations

a.s.l.	Above sea level
ANN	Artificial Neural Network
ArcGIS	Aeronautical Reconnaissance Coverage Geographic Information System
AUC	Area Under the Curve
biomod2	BIOdiversity MODeling version 2
CCAFS	Climate Change, Agriculture, and Food Security
CCSM4	Community Climate System Model version 4
CMIP5	The Coupled Model Inter-Comparison Project Phase 5
CV	Cross Validation
DEM	Digital Elevation Model
ENVIREM	ENVironmental Rasters for Ecological Modeling
GBM	Generalized Boosting Model
GERICS	German Climate Service Center
GLM	Generalized Linear Model
GPS	Global Positioning System
IPCC	Intergovernmental Panel on Climate Change
ISRIC	International Soil Reference and Information Centre
IUCN	International Union for Conservation of Nature
Maxent	Maximum Entropy
MENA	The Middle East and North Africa region
ML	Machine Learning
NASA	National Aeronautics and Space Administration
Obs.	Observed
P/A	Presence/absence
PA(s)	Protected Area(s)
Pred.	Predicted
r	Pearson correlation coefficient
RCP(s)	Representative Concentration Pathway(s)
RF	Random Forest
ROC	The Receiver Operator Characteristic curve
SD	Standard Deviation
SDM(s)	Species Distribution Model(s)
spp.	Species (plural)
St. Catherine	Saint Catherine
TSS	True Skill Statistics
USGS	The United States Geological Survey
ver.	Version
VI	Variable Importance

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43088-024-00553-2>.

Additional file 1 (DOCX 54 kb)

Acknowledgements

The authors appreciate other colleagues and anonymous reviewers for their constructive suggestions.

Author contributions

AR designed the work, collected the data, conducted the models, analyzed, and interpreted the results, and wrote the manuscript. AY conceived the work, supervised the practical work, and revised the manuscript. HM conceived the work, followed-up and supervised the practical work. HF supervised the practical work and revised the manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All methods performed in this study complied with the relevant institutional, national, and international guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Botany, Faculty of Science, Ain Shams University, Cairo 11566, Egypt.

²Department of Computer Science, Faculty of Computers and Information, Suez University, Suez 43512, Egypt.

Received: 9 May 2024 Accepted: 13 September 2024

Published online: 30 September 2024

References

- Lawler JJ, Shafer SL, White D, Kareiva P, Maurer EP, Blaustein AR, Bartlein PJ (2009) Projected climate-induced faunal change in the Western Hemisphere. *Ecology* 90:588–597. <https://doi.org/10.1890/08-0823.1>
- Oke OA, Thompson KA (2015) Distribution models for mountain plant species: the value of elevation. *Ecol Modell* 301:72–77. <https://doi.org/10.1016/J.ECOLMODEL.2015.01.019>
- Omar K, Elgamel I (2021) Assess the extinction risk of mountain endemic plants in Egypt under the current climatic condition: a case study of endemic *Silene* species. *Eur J Biol Biotechnol* 2:34–47. <https://doi.org/10.24018/EJBIO.2021.2.5.261>
- Cuena-Lombraña A, Fois M, Fenu G, Cogoni D, Bacchetta G (2018) The impact of climatic variations on the reproductive success of *Gentiana lutea* L. in a Mediterranean mountain area. *Int J Biometeorol* 62:1283–1295. <https://doi.org/10.1007/S00484-018-1533-3>
- Abdelaal M, Fois M, Fenu G, Bacchetta G, Catherine S (2019) Using Max-Ent modeling to predict the potential distribution of the endemic plant *Rosa arabica* Crép. in Egypt. *Ecol Inform* 50:68–75. <https://doi.org/10.1016/j.ecoinf.2019.01.003>
- Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annu Rev Ecol Evol Syst* 37:637–669. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110100>
- Beery S, Cole E, Parker J, Perona P, Winner K (2021) Species distribution modeling for machine learning practitioners: a review. In: Proceedings of the 4th ACM SIGCAS conference on computing and sustainable societies. Association for Computing Machinery, Inc, pp 329–348
- Bonaccorso G (2018) Machine learning algorithms: popular algorithms for data science and machine learning. Packt Publishing Ltd
- Maresh B (2020) Machine learning algorithms—a review. *Int J Sci Res* 9:381–386
- Melo-Merino SM, Reyes-Bonilla H, Lira-Noriega A (2020) Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. *Ecol Modell* 415:108837. <https://doi.org/10.1016/J.ECOLMODEL.2019.108837>
- Ryo M (2024) Ecology with artificial intelligence and machine learning in Asia: a historical perspective and emerging trends. *Ecol Res* 39:5–14. <https://doi.org/10.1111/1440-1703.12425>
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Modell* 135:147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40:677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography* (Cop) 32:369–373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends Ecol Evol* 22:42–47. <https://doi.org/10.1016/j.tree.2006.09.010>
- Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. *Divers Distrib* 15:59–69. <https://doi.org/10.1111/j.1472-4642.2008.00491.x>
- Heikkinen RK, Luoto M, Araújo MB, Virkkala R, Thuiller W, Sykes MT (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. *Prog Phys Geogr* 30:751–777. <https://doi.org/10.1177/0309133306071957>
- Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. *J Biogeogr* 38:1–8. <https://doi.org/10.1111/J.1365-2699.2010.02416.X>
- Gábor L, Šimová P, Keil P, Zarzo-Arias A, Marsh CJ, Rocchini D, Malavasi M, Barták V, Moudrý V (2022) Habitats as predictors in species distribution models: Shall we use continuous or binary data? *Ecography* (Cop) 2022:e06022. <https://doi.org/10.1111/ECOG.06022>
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob Ecol Biogeogr* 12:361–371. <https://doi.org/10.1046/j.1466-822X.2003.00042.x>
- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *J Biogeogr* 33:1677–1688. <https://doi.org/10.1111/J.1365-2699.2006.01584.X>
- Gardner AS, Maclean IMD, Gaston KJ (2019) Climatic predictors of species distributions neglect biophysically meaningful variables. *Divers Distrib* 25:1318–1333. <https://doi.org/10.1111/DDI.12939>
- Randin CF, Engler R, Normand S, Zappa M, Zimmermann NE, Pearman PB, Vittoz P, Thuiller W, Guisan A (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Glob Chang Biol* 15:1557–1569. <https://doi.org/10.1111/J.1365-2486.2008.01766.X>
- Colwell RK, Lees DC (2000) The mid-domain effect: geometric constraints on the geography of species richness. *Trends Ecol Evol* 15:70–76. [https://doi.org/10.1016/S0169-5347\(99\)01767-X](https://doi.org/10.1016/S0169-5347(99)01767-X)
- Austin M (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecol Modell* 200:1–19. <https://doi.org/10.1016/J.ECOLMODEL.2006.07.005>
- Babst F, Bouriaud O, Poulter B, Trouet V, Girardin MP, Frank DC (2019) Twentieth century redistribution in climatic drivers of global tree growth. *Sci Adv* 5:eaat4313. https://doi.org/10.1126/SCIADV.AAT4313/SUPPL_FILE/AAT4313_SM.PDF

27. Hatfield JL, Prueger JH (2015) Temperature extremes: Effect on plant growth and development. *Weather Clim Extrem* 10:4–10. <https://doi.org/10.1016/J.WACE.2015.08.001>
28. Barry RG (2008) Mountain weather and climate. *Mt Weather Clim Third Ed*. <https://doi.org/10.1017/CBO9780511754753>
29. Cai S, Geng H, Pan B, Hong Y, Chen L (2020) Topographic controls on the annual runoff coefficient and implications for landscape evolution across semiarid Qilian Mountains, NE Tibetan Plateau. *J Mt Sci* 17:464–479. <https://doi.org/10.1007/S11629-019-5584-7/METRCS>
30. Huang E, Chen Y, Fang M, Zheng Y, Yu S (2021) Environmental drivers of plant distributions at global and regional scales. *Glob Ecol Biogeogr* 30:697–709. <https://doi.org/10.1111/GEGB.13251>
31. Austin MP (2002) Case studies of the use of environmental gradients in vegetation and fauna modelling: theory and practice in Australia and New Zealand. In: Scott J, Heglund P, Morrison M, Haufler J, Raphael M, Wall W, Samson F (eds) *Predicting species occurrences: Issues of accuracy and scale*. Island Press, London, pp 73–82
32. Velazco SJE, Galvão F, Villalobos F, De Marco P (2017) Using worldwide edaphic data to model plant species niches: an assessment at a continental extent. *PLoS ONE* 12:e0186025. <https://doi.org/10.1371/JOURNAL.PONE.0186025>
33. Zhong Y, Xue Z, Jiang M, Liu B, Wang G (2021) The application of species distribution modeling in wetland restoration: a case study in the Songnen Plain, Northeast China. *Ecol Indic* 121:107137. <https://doi.org/10.1016/J.ECOLIND.2020.107137>
34. Crawford RMM (2008) *Plants at the margin: Ecological limits and climate change*. Cambridge University Press, Cambridge
35. Perry LG, Jarnevich CS, Shafroth PB (2022) Models combining multiple scales of inference capture hydrologic and climatic drivers of riparian tree distributions. *Ecosphere* 13:e4305. <https://doi.org/10.1002/ECS2.4305>
36. Thompson K, Gilbert F (2015) Spatiotemporal variation in the endangered *Thymus decussatus* in a hyper-arid environment. *J Plant Ecol* 8:79–90. <https://doi.org/10.1093/JPE/RTU004>
37. Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dormann CF, Forchhammer MC, Grytnes JA, Guisan A, Heikkinen RK, Høye TT, Kühn I, Luoto M, Maiorano L, Nilsson MC, Normand S, Öckinger E, Schmidt NM, Termansen M, Timmermann A, Wardle DA, Aastrup P, Svenning JC (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol Rev* 88:15–30. <https://doi.org/10.1111/J.1469-185X.2012.00235.X>
38. IPCC (2014) *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*. IPCC, Geneva
39. Schwalm CR, Glendon S, Duffy PB (2020) RCP8.5 tracks cumulative CO₂ emissions. *Proc Natl Acad Sci U S A* 117:19656–19657. <https://doi.org/10.1073/PNAS.2007117117>
40. Hajat S, Proestos Y, Araya-Lopez JL, Economou T, Lelieveld J (2023) Current and future trends in heat-related mortality in the MENA region: a health impact assessment with bias-adjusted statistically downscaled CMIP6 (SSP-based) data and Bayesian inference. *Lancet Planet Heal* 7:e282–e290. [https://doi.org/10.1016/S2542-5196\(23\)00045-1](https://doi.org/10.1016/S2542-5196(23)00045-1)
41. Rose MB, Velazco SJE, Regan HM, Flint AL, Flint LE, Thorne JH, Franklin J (2024) Uncertainty in consensus predictions of plant species' vulnerability to climate change. *Divers Distrib* 30:e13898. <https://doi.org/10.1111/DDI.13898>
42. Zittis G, Hadjinicolaou P, Almazroui M, Bucchignani E, Driouech F, El Rhaz K, Kurnaz L, Nikulin G, Ntoumos A, Ozturk T, Proestos Y, Stenchikov G, Zaaboul R, Lelieveld J (2021) Business-as-usual will lead to super and ultra-extreme heatwaves in the Middle East and North Africa. *NPJ Clim Atmos Sci* 4:1–9. <https://doi.org/10.1038/s41612-021-00178-7>
43. Lelieveld J, Proestos Y, Hadjinicolaou P, Tanarhte M, Tyrllis E, Zittis G (2016) Strongly increasing heat extremes in the Middle East and North Africa (MENA) in the 21st century. *Clim Change* 137:245–260. <https://doi.org/10.1007/S10584-016-1665-6/FIGURES/5>
44. Francis D, Fonseca R (2024) Recent and projected changes in climate patterns in the Middle East and North Africa (MENA) region. *Sci Rep* 14:1–25. <https://doi.org/10.1038/s41598-024-60976-w>
45. Varela R, Rodríguez-Díaz L, DeCastro M (2020) Persistent heat waves projected for Middle East and North Africa by the end of the 21st century. *PLoS ONE* 15:e0242477. <https://doi.org/10.1371/JOURNAL.PONE.0242477>
46. IPCC (2023) *Climate change 2023: synthesis report. Contribution of working groups i, ii and iii to the sixth assessment report of the intergovernmental panel on climate change* (Core Writing Team, H Lee, J Romero (eds)). IPCC, Geneva
47. Trisos CH, Adekan IO, Totin E, Ayanlade A, Efitre J, Gameda A, Kalaba K, Lennard C, Masao C, Mgaya Y, Ngaruiya G, Olago D, Simpson NP, Zaki-deen S (2022) Africa. In: Pörtner H-O, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegria A, Craig M, Langsdorf S, Löschke S, Möller V, Okem A, Rama B (eds) *Climate change 2022: impacts, adaptation and vulnerability. Contribution of working group II to the 6th assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, pp 1285–1455
48. GERICS (2019) *Climate Fact Sheet – Egypt*. <https://www.climate-service-center.de/>
49. *Climate Risk Profile: Egypt* (2021) The World Bank Group. <https://www.worldbank.org/>
50. Berkeley Earth (2024) *Country-Level Warming Projections*. <https://berkeleyearth.org/>
51. Moustafa AA, Klopatek JM (1995) Vegetation and landforms of the Saint Catherine area, southern Sinai, Egypt. *J Arid Environ* 30:385–395. <https://doi.org/10.1006/JARE.1995.0033>
52. Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci Data* 7:1–18. <https://doi.org/10.1038/s41597-020-0453-3>
53. Khedr AHA (2021) Microhabitats supporting endemic plants in Sinai, Egypt. In: *Management and development of the endemic taxa in the Egyptian desert*. Springer Nature, pp 369–381
54. Mosallam HAM (2007) Assessment of target species in saint katherine protectorate, Sinai. *Egypt J Appl Sci Res* 3:456–469
55. El-Khalafy MM, Ahmed DAEA, Shaltout KH, Haroun SA, Al-Sodany YM (2023) Ethnobotanical importance of the endemic taxa in the Egyptian flora. *J Ecol Environ* 47:146–156. <https://doi.org/10.5141/JEE.23.044>
56. IUCN (2024) *The IUCN Red List of Threatened Species. Version 2024-1*. <https://www.iucnredlist.org>
57. Rabei SH, El Gamal IA, Nada RM (2021) Effect of genetic diversity on the distribution of endemic species of the genus *Silene* (Caryophyllaceae) in Saint Katherine Protectorate, Sinai, Egypt. *Plant Mol Biol Report* 39:364–375. <https://doi.org/10.1007/S1105-020-01253-8/FIGURES/7>
58. Park JH, Han YS, Lee EP, Lee SY, Jeong HM, You YH (2019) Studies on the changes in phenological, growth and physiological responses of *Silene capitata* Kom., an endangered plant in Korea, under climate change treatment. *J Ecol Environ* 43:1–9. <https://doi.org/10.1186/S41610-019-0128-1/FIGURES/5>
59. ESRI (2019) *ArcGIS Desktop: Release 10.8*. Environmental Systems Research Institute. Redlands, California
60. Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315. <https://doi.org/10.1002/JOC.5086>
61. Title PO, Bemmels JB (2018) ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* 41:291–307. <https://doi.org/10.1111/ECOG.02880>
62. Poggio L, De Sousa LM, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, Rossiter D (2021) SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *Soil* 7:217–240. <https://doi.org/10.5194/Soil-7-217-2021>
63. Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, Worley PH, Yang Z-L, Zhang M (2011) The community climate system model version 4. *J Clim* 24:4973–4991. <https://doi.org/10.1175/2011JCLI4083.1>
64. R Development Core Team (2017) *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria
65. Brown JL, Bennett JR, French CM (2017) SDMtoolbox 2.0: The next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ* 5:e4095. <https://doi.org/10.7717/peerj.4095>
66. Thuiller W, Georges D, Engler R, Breiner F (2016) 'biomod2': Ensemble platform for species distribution modeling.

67. Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49. <https://doi.org/10.1017/S0376892997000088>
68. Monserud RA, Leemans R (1992) Comparing global vegetation maps with the Kappa statistic. *Ecol Modell* 62:275–293. [https://doi.org/10.1016/0304-3800\(92\)90003-V](https://doi.org/10.1016/0304-3800(92)90003-V)
69. Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
70. Franklin J, Miller JA (2010) Mapping species distributions. Cambridge University Press, Cambridge
71. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
72. Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43:1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
73. Araújo MB, Whittaker RJ, Ladle RJ, Erhard M (2005) Reducing uncertainty in projections of extinction risk from climate change. *Glob Ecol Biogeogr* 14:529–538. <https://doi.org/10.1111/j.1466-822X.2005.00182.x>
74. Ni M, Vellend M (2024) Soil properties constrain predicted poleward migration of plants under climate change. *New Phytol* 241:131–141. <https://doi.org/10.1111/NPH.19164>
75. Rose MB, Velazco SJE, Regan HM, Franklin J (2023) Rarity, geography, and plant exposure to global change in the California Floristic Province. *Glob Ecol Biogeogr* 32:218–232. <https://doi.org/10.1111/GEB.13618>
76. Stralberg D, Wang X, Parisien MA, Robinne FON, Solymos P, Mahon CL, Nielsen SE, Bayne EM (2018) Wildfire-mediated vegetation change in boreal forests of Alberta. *Canada Ecosphere* 9:e02156. <https://doi.org/10.1002/ECS2.2156>
77. Cai L, Kreft H, Taylor A, Denelle P, Schrader J, Essl F, van Kleunen M, Pergl J, Pyšek P, Stein A, Winter M, Barcelona JF, Fuentes N, Inderjit KDN, Kartesz J, Kuprijanov A, Nishino M, Nickrent D, Nowak A, Patzelt A, Pelsler PB, Singh P, Wieringa J, Weigelt P (2023) Global models and predictions of plant diversity based on advanced machine learning techniques. *New Phytol* 237:1432–1445. <https://doi.org/10.1111/NPH.18533>
78. Mohammed M, Khan MB, Bashie EBM (2016) Machine learning: Algorithms and applications. CRC Press
79. Murphy KP (2023) Probabilistic machine learning: advanced topics. The MIT Press
80. Naeem S, Ali A, Anam S, Ahmed MM (2023) An unsupervised machine learning algorithms: comprehensive review. *Int J Comput Digit Syst* 13:911–921. <https://doi.org/10.12785/IJCDS/130172>
81. Randin CF, Ashcroft MB, Bolliger J, Cavender-Bares J, Coops NC, Dullinger S, Dirnböck T, Eckert S, Ellis E, Fernández N, Giuliani G, Guisan A, Jetz W, Joost S, Karger D, Lembrechts J, Lenoir J, Luoto M, Morin X, Price B, Rocchini D, Schaeppman M, Schmid B, Verburg P, Wilson A, Woodcock P, Yoccoz N, Payne D (2020) Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sens Environ* 239:111626. <https://doi.org/10.1016/j.RSE.2019.111626>
82. Ray S (2019) A quick review of machine learning algorithms. in: proceedings of the international conference on machine learning, big data, cloud and parallel computing: trends, perspectives and prospects, COMITCon 2019. Institute of Electrical and Electronics Engineers Inc., pp 35–39
83. Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods Ecol Evol* 1:330–342. <https://doi.org/10.1111/j.2041-210X.2010.00036.x>
84. Pearson RG, Thuiller W, Araújo MB, Martinez-Meyer E, Brotons L, McClean C, Miles L, Segurado P, Dawson TP, Lees DC (2006) Model-based uncertainty in species range prediction. *J Biogeogr* 33:1704–1711. <https://doi.org/10.1111/j.1365-2699.2006.01460.x>
85. Kaky E, Nolan V, Alatawi A, Gilbert F (2020) A comparison between Ensemble and MaxEnt species distribution modelling approaches for conservation: a case study with Egyptian medicinal plants. *Ecol Inform* 60:101150. <https://doi.org/10.1016/j.ECOINF.2020.101150>
86. Hao T, Elith J, Guillera-Aroita G, Lahoz-Monfort JJ (2019) A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Divers Distrib* 25:839–852. <https://doi.org/10.1111/DDI.12892>
87. Crimmins SM, Dobrowski SZ, Mynsberge AR (2013) Evaluating ensemble forecasts of plant species distributions under climate change. *Ecol Modell* 266:126–130. <https://doi.org/10.1016/J.ECOLMODEL.2013.07.006>
88. Hao T, Elith J, Lahoz-Monfort JJ, Guillera-Aroita G (2020) Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography* 43:549–558. <https://doi.org/10.1111/ecog.04890>
89. Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JMC, Townsend Peterson A, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
90. Valavi R, Guillera-Aroita G, Lahoz-Monfort JJ, Elith J (2022) Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecol Monogr* 92:e01486. <https://doi.org/10.1002/ECM.1486>
91. Abdelaal M, Fois M, Dakhil MA, Bacchetta G, El-Sherbeny GA (2020) Predicting the potential current and future distribution of the endangered endemic vascular plant *primula boveana* Decne ex. Duby in Egypt. *Plants* 9:957. <https://doi.org/10.3390/plants9080957>
92. Dubuis A, Giovanettina S, Pellissier L, Pottier J, Víttoz P, Guisan A (2013) Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *J Veg Sci* 24:593–606. <https://doi.org/10.1111/JVS.12002>
93. Diekmann M, Michaelis J, Pannek A (2015) Know your limits—the need for better data on species responses to soil variables. *Basic Appl Ecol* 16:563–572. <https://doi.org/10.1016/J.BAAE.2015.08.010>
94. Beauregard F, De Blois S (2014) Beyond a climate-centric view of plant distribution: edaphic variables add value to distribution models. *PLoS ONE* 9:e92642. <https://doi.org/10.1371/JOURNAL.PONE.0092642>
95. Bertrand R, Perez V, Gégout JC (2012) Disregarding the edaphic dimension in species distribution models leads to the omission of crucial spatial information under climate change: the case of *Quercus pubescens* in France. *Glob Chang Biol* 18:2648–2660. <https://doi.org/10.1111/J.1365-2486.2012.02679.X>
96. Buri A, Cianfrani C, Pinto-Figueroa E, Yashiro E, Spangenberg JE, Adatte T, Verrecchia E, Guisan A, Pradervand JN (2017) Soil factors improve predictions of plant species distribution in a mountain environment. *Prog Phys Geogr Earth Environ* 41:703–722. <https://doi.org/10.1177/0309133317738162>
97. Suggitt AJ, Wilson RJ, Isaac NJB, Beale CM, Auffret AG, August T, Bennie JJ, Crick HQP, Duffield S, Fox R, Hopkins JJ, Macgregor NA, Morecroft MD, Walker KJ, Maclean IMD (2018) Extinction risk from climate change is reduced by microclimatic buffering. *Nat Clim Chang* 8:713–717. <https://doi.org/10.1038/s41558-018-0231-9>
98. Liao C, Liu B, Xu Y, Li Y, Li H (2019) Effect of topography and protecting barriers on revegetation of sandy land. *Southern Tibetan Plateau Sci Rep* 9:6501. <https://doi.org/10.1038/s41598-019-43034-8>
99. Albano CM (2015) Identification of geophysically diverse locations that may facilitate species' persistence and adaptation to climate change in the southwestern United States. *Landsc Ecol* 30:1023–1037. <https://doi.org/10.1007/s10980-015-0167-7/TABLES/3>
100. Dobrowski SZ (2011) A climatic basis for microrefugia: the influence of terrain on climate. *Glob Chang Biol* 17:1022–1035. <https://doi.org/10.1111/J.1365-2486.2010.02263.X>
101. Marchioro CA, da Silva KF, dos Santos KL, Siminski A (2023) Biotic interactions under risk: climate change drives spatial mismatch between a critically endangered tree and its seed dispersers and predators. *Clim Change* 176:1–20. <https://doi.org/10.1007/s10584-023-03642-W/METRICS>
102. Von Holle B, Weber SE, Nickerson DM (2020) The influence of warming and biotic interactions on the potential for range expansion of native and nonnative species. *AoB Plants* 12:plaa040. <https://doi.org/10.1093/AOBPLA/PLAA040>
103. Blois JL, Zarnetske PL, Fitzpatrick MC, Finnegan S (2013) Climate change and the past, present, and future of biotic interactions. *Science* (80-) 341:499–504. <https://doi.org/10.1126/SCIENCE.1237184>
104. Niu Y, Zhou J, Yang S, Chu B, Ma S, Zhu H, Hua L (2019) The effects of topographical factors on the distribution of plant communities in a

- mountain meadow on the Tibetan Plateau as a foundation for target-oriented management. *Ecol Indic* 106:105532. <https://doi.org/10.1016/J.ECOLIND.2019.105532>
105. El-Khalafy MM, Ahmed DAEA, Shaltout KH, Al-Sodany YM, Haroun SA (2021) Re-assessment of the endemic taxa in the Egyptian Flora. *Afr J Ecol* 59:784–796. <https://doi.org/10.1111/AJE.12880>
 106. Rabei SH, Nada RM, El Gamal I (2020) Studies on five *Silene* L. Taxa in Saint Catherine Protectorate, South Sinai, Egypt. *Jordan J Biol Sci* 13:59–67
 107. Omar K, Elgamal I (2021) *Silene leucophylla*. IUCN Red List Threat Species 2021 eT184588883A184588988. <https://doi.org/10.2305/IUCN.UK.2021-2.RLTS.T184588883A184588988.en>
 108. Omar K, Elgamal I (2021) *Silene schimperiana*. IUCN Red List Threat Species 2021 eT184589181A184589187. <https://doi.org/10.2305/IUCN.UK.2021-2.RLTS.T184589181A184589187.en>
 109. Täckholm V (1974) Student's flora of Egypt, 2nd edn. Cairo University Press, Cairo
 110. Boulos L (2009) Flora of Egypt Checklist - Revised Annotated Edition. Al Hadara Publishing, Cairo, Egypt
 111. El-Ghani MMA, Fahmy AG (1994) Studies on the threatened woody perennial taxa in the flora of Egypt II. Extinct and endemic taxa Feddes Rept 105:243–250. <https://doi.org/10.1002/FEDR.19941050318>
 112. Moustafa ARA, Kamel M (1995) Ecological notes on the floristic composition and endemic species of Saint Catherine area, South Sinai. *Egypt Egypt J Bot* 35:179–200
 113. Ayyad MA, Fakhry AM, Moustafa ARA (2000) Plant biodiversity in the Saint Catherine area of the Sinai peninsula. *Egypt Biodivers Conserv* 9:265–281. <https://doi.org/10.1023/A:1008973906522>
 114. Omar K (2017) Community based conservation of threatened plants *Silene schimperiana* and *Polygala sinaica* in South Sinai, Egypt. www.Rufford.org
 115. Moustafa ARA, Zayed A (1996) Effect of environmental factors on the flora of alluvial fans in southern Sinai. *J Arid Environ* 32:431–443. <https://doi.org/10.1006/JARE.1996.0036>
 116. Zohary M (1973) Geobotanical foundations of the Middle East. Gustav Fischer Verlag Press, Stuttgart
 117. Kaky E, Gilbert F (2017) Predicting the distributions of Egypt's medicinal plants and their potential shifts under future climate change. *PLoS ONE* 12:e0187714. <https://doi.org/10.1371/JOURNAL.PONE.0187714>
 118. Zhao Q, Li R, Gao Y, Yao Q, Guo X, Wang W (2018) Modeling impacts of climate change on the geographic distribution of medicinal plant *Fritillaria cirrhosa* D. Don. *Plant Biosyst* 152:349–355. <https://doi.org/10.1080/11263504.2017.1289273>
 119. Abolmaali SMR, Tarkesh M, Bashari H (2018) MaxEnt modeling for predicting suitable habitats and identifying the effects of climate change on a threatened species, *Daphne mucronata*, in central Iran. *Ecol Inform* 43:116–123. <https://doi.org/10.1016/j.ecoinf.2017.10.002>
 120. Khanum R, Mumtaz AS, Kumar S (2013) Predicting impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. *Acta Oecologica* 49:23–31. <https://doi.org/10.1016/j.actao.2013.02.007>
 121. Gottfried M, Pauli H, Reiter K, Grabherr G (1999) A fine-scaled predictive model for changes in species distribution patterns of high mountain plants induced by climate warming. *Divers Distrib* 5:241–251. <https://doi.org/10.1046/J.1472-4642.1999.00058.X>
 122. Guisan A, Theurillat J-P (2000) Equilibrium modeling of alpine plant distribution: how far can we go? *Phytocoenologia* 30:353–384. <https://doi.org/10.1127/PHTO/30/2000/353>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.