INTL J TROP DRYLANDS
Volume 9, Number 2, December 2025 E-ISSN: 2775-6130
Pages: 99-110 DOI: 10.13057/tropdrylands/t090201

Modelling drought-tolerant *Sorghum bicolor* distributions in Eastern Indonesia using machine learning approaches

SUYUD WARNO UTOMO[1], ANDRIO A. WIBOWO[2,♥], FATMA LESTARI[2]
[1]School of Environmental Science, Universitas Indonesia. Jl. Salemba Raya Kampus UI Salemba, Kenari, Senen, Central Jakarta 10430, Jakarta, Indonesia
[2]Disaster Risk Reduction Center, Universitas Indonesia. Integrated Laboratory and Research Centre Building, 2nd Floor Kampus Baru UI, Depok 16424, West Java, Indonesia. ♥email: awbio2021@gmail.com

Manuscript received: 6 May 2025. Revision accepted: 17 July 2025.

Abstract. *Utomo SW, Wibowo AA, Lestari F. 2025. Modelling drought-tolerant* Sorghum bicolor *distributions in Eastern Indonesia using machine learning approaches. Intl J Trop Drylands 9: 99-110.* Tropical arid ecosystems require alternative species that are drought-tolerant. *Sorghum bicolor* (L.) Moench has been considered an alternative, drought-tolerant species. Despite its resistance to drought, information on the potential distribution of sorghum is very limited. This information is very important, especially in arid eastern Indonesia, where sorghum is nominated as an alternative species to sustain food security. This study aims to model the potential distribution of *S. bicolor* using machine learning (random forest), geoclimate (Bioclim), and statistical methods (GAM/GLM) on five arid islands, including Lombok, Sumba, Sumbawa, Flores, and the Timor Islands, Indonesia. Area Under Curve (AUC) was used to evaluate model performance. In general, all the models confirm that Timor, followed by Sumbawa and the Flores Islands, have large, suitable areas for sorghum. It is estimated that up to 99.71% of arid island ecosystems in eastern Indonesia are suitable for sorghum. The geoclimate and machine learning models generated the highest values for AUC in comparison to statistical methods in which the Domain model is 0.962, SVM is 0.903, and both GLM and GAM are 0.894. It is important for plant cultivation planning to consider species distribution modeling and not rely on any single modeling method. The plant cultivation should evaluate the performance of all available models for their crops and area of interest, and select the best representative methods to develop an accurate and representative sorghum crop distribution model.

Keywords: Arid, AUC, machine learning, model, sorghum

INTRODUCTION

The region in eastern Indonesia, which includes the five islands of Lombok, Sumbawa, Sumba, Flores, and Timor Islands, is distinguished by a dry climate and a predominance of desert habitats. Grasslands and shrublands predominate in these tropical habitats. El Nino has extended the dry seasons in this area and risked the supply of crops, particularly rice, when combined with the region's arid environment. 8,400 hectares of rice fields were impacted by the last El Nino, indicating that rice is a crop that is sensitive because of the dry circumstances on Lombok Island (Yasin et al. 2004). Six rivers, water catchment regions, and irrigation systems of Timor Island had water shortfalls as a result of El Nino, leaving insufficient acreage for rice growing.

The circumstances that eastern Indonesia went through between 1980 and 2020—roughly 13 extremely dry seasons—are what caused this susceptibility (Yanti et al. 2022). In eastern Indonesia, 25.46% of all natural disasters have been caused by drought. Low rainfall seasons that last for four months have put agricultural operations in general and rice cultivation in particular at risk. Annual rainfall rates in the region have dropped as low as 1,900 mm, compared to the 2,702 mm national average. One significant abiotic stressor that reduces agricultural production and output is drought. It regularly affects all major crops and occurs in large parts of the planet. Severe droughts severely reduce agricultural yields and quality, and in food-insecure areas, they can cause famine (Akbar et al. 2021). One of the potential alternative crops to replace conventional drought-sensitive rice is sorghum (*Sorghum bicolor* (L.) Moench).

The semi-arid and desert regions are home to about 500 million people who live in impoverished nations and have made sorghum (*S. bicolor*) their main species to be cultivated. Sorghum has a high fiber and protein content and is devoid of gluten (McCann et al. 2015; Impa et al. 2019). Scientific data indicates that sorghum possesses nutrients comparable to those of rice. In addition to providing nutrients for human consumption, sorghum is also collected and used as a feedstock for the production of bioethanol (Mathur et al. 2017).

Sorghum has been harvested recently and is now grown all over the world. Most sorghum is planted in dry, semi-arid areas with limited water supplies. For instance, according to Hadebe et al. (2017), 60% of the land in Sub-Saharan Africa used to grow sorghum is thought to be susceptible to recurrent droughts. Similarly, 80% of sorghum grown in the United States is grown without irrigation, where water is a major limiting factor and yields are significantly lower. Sorghum's advantages were linked to its decreased water needs. In a location with low annual rainfall rates of 300 mm, sorghum is expected to require merely a minimal water input of 350/400 mm/yr (Ruiz-Giralt et al. 2023). In spite of the fact that sorghum resistant to drought, nutrient mobilization and transport as well as nitrogen uptake in the soil were affected by drought stress (Yu et al. 2015; Sarshad et al. 2021).

A technique to simulate and ascertain the possible geographical distributions of sorghum is necessary for its development as an alternative crop. A popular method for simulating a species' possible geographic ranges is to combine statistical, geoclimate, and machine learning models. This model has been frequently used to predict possible ranges of a wide range of creatures, including crop species (Fitzgibbon et al. 2022; Lin et al. 2022; Li et al. 2023), floral species, faunal species (Stephenson et al. 2022), including ticks (Sánchez et al. 2023), and even vegetation (Khanum et al. 2013). At the moment, an increasing number of methods are being used to estimate habitat suitability and species distribution regions. These methods range from general additive/linear models (GAM, GLM) that emphasize statistical methods to Bioclim, Domain, and Biomapper that emphasize geoclimatical methods. In addition, there are Artificial Neural Networks (ANN) that emphasize deep learning techniques and MaxEnt, Random Forest, Support Vector Machines (SVM) that emphasize machine learning-based techniques. Every strategy is different and has pros and cons of its own. According to Marcer et al. (2013), machine learning is seen to be one of the modeling approaches with benefits when compared to other ways. As a result, it is the method that is most commonly used in habitat suitability modeling research. According to Fois et al. (2018), there are several benefits to using machine learning, including the requirement for only species occurrence records, the method's ability to be applied with a limited amount of data, and its capacity to produce prediction models and estimations with high accuracy, high reproducibility, and the capacity to discern the most distinct environmental variables. In an arid region in KwaZulu-Natal, South Africa, MaxEnt informed that sorghum suitability was following the west-to-east trend, with areas in the west being more suitable than the east (Mugiyo et al. 2022). Sorghum has been proposed as an alternative to rice as the main crop in Indonesia (Paesal et al. 2021). The Indonesian government intends to set aside 115,000 hectares in 2023 and an extra 154,000 hectares in 2024 for the cultivation of sorghum in eastern Indonesia. Despite these advancements, nothing is known about the possible range of sorghum, which is mostly found in the dry eastern part of Indonesia. This information is needed in advance to ensure and continue sorghum cultivation. The general purpose of this study is to identify possible sorghum growing locations in dry environments on five specific eastern Indonesian islands. The particular purposes of this study are (i) to predict habitat suitability of drought-tolerant sorghum, (ii) to identify key environmental drivers, and (iii) to generate a suitability map for Eastern Indonesia. The information on possible sorghum growing locations is required by the Indonesian government to support its policy in promoting sorghum cultivation. This research is new in that it uses statistical modeling, geoclimate, and machine learning to attain precision in possible distributions. The outcomes of this study will benefit food safety in the particular arid environment of eastern Indonesia.

MATERIALS AND METHODS

The research was conducted on several islands in eastern Indonesia. The study methodology followed methods developed by Semu et al. (2021), including species occurrence, environmental variables, and model evaluation.

Study area

This research was conducted on several islands located in the eastern Indonesia regions, consisting of Lombok, Sumba, Sumbawa, Flores, and Timor Islands, Indonesia, within the geocoordinates of 8°-10° South latitude and 116°-126° East longitude (Figure 1). These regions are characterized by low rainfall and high temperatures (Table 1), and each island has different areas. The smallest island was Lombok Island, followed by Sumba Island. The most arid islands were Flores, Sumbawa, and the Timor Islands. Those islands have the lowest rainfall, reaching only 5 mm. Among other islands, Sumbawa Island has the hottest temperature, ranging from 32.0 to 36.6°C.

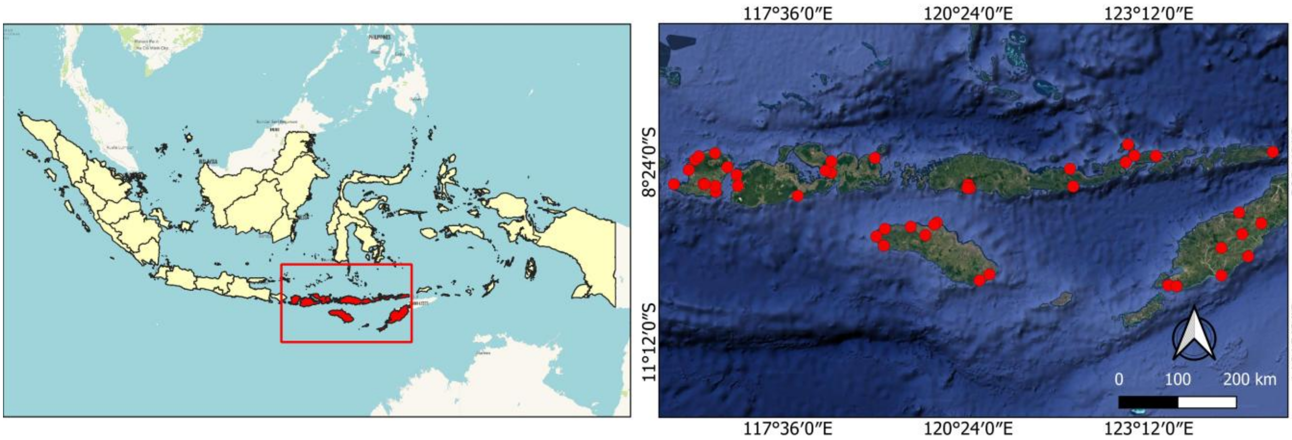

Figure 1. Spatial distribution of presence locations of sorghum (red dots) in Lombok, Sumba, Sumbawa, Flores, and Timor Islands, Indonesia

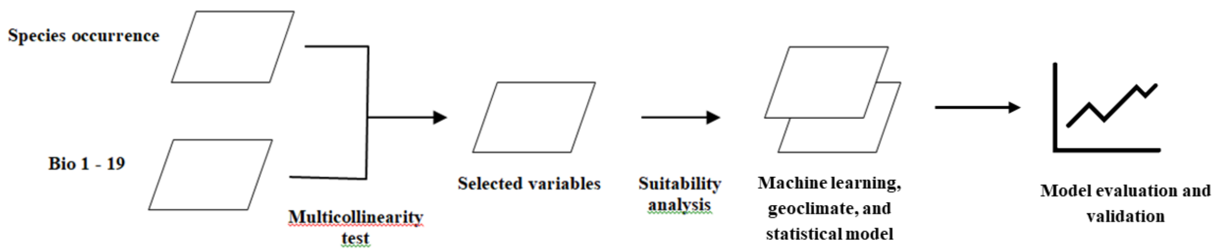

Figure 2. Schematic flowchart showing the potential distribution analysis pipeline

Table 1. Climatic variables in Lombok, Sumba, Sumbawa, Flores, and Timor Islands, Indonesia.

Islands	Area (km^2)	Rainfall (mm)	Temperature (°C)
Lombok	4,739	77-489	23.9 -31.9
Sumba	11,006	15-162	25.0-32
Sumbawa	15,414	10-446	32.0-36.6
Flores	15,531	5-483	16-30
Timor	30,777	10-667	27-31

Because of the dry temperature and arid ecosystems, the ecosystems of those islands are dominated by savanna and grassland ecosystems, which are patchy and fragmented, and range from lowland tropical rainforests to upland tropical forests to sub-alpine vegetation (Sutomo et al. 2021). There are numerous meadows and shrublands on the island (Sapta et al. 2015).

Sorghum occurrence surveys

All regions of a few chosen islands were the subject of field observations and surveys to document the presence of sorghum in real time. The information retrieved from the Herbarium Bogoriense, the Regional Agency for Agriculture and Forestry of the Ministry for Agriculture and Forestry, Indonesia, combined with database developed from literature studies was used to determine the locations for field observations and surveys following Gunawan et al. (2021). The global positioning system (GPS) unit, a Garmin Etrex 30 model, was used to record the geographic locations of *S. bicolor* occurrences (Figure 1) in the field. The data were stored in CSV format and transformed into Microsoft Excel so that habitat suitability modeling could be made. To document the presence of sorghum in real time, field surveys and explorations were carried out from September 2023 to January 2024. From September to October 2023, surveys were conducted in Lombok islands, from October to November 2023 for Sumba and Sumbawa islands, and from November 2023 to January 2024, covering Flores and Timor islands.

Procedures

Environmental predictors and variable selection

Bioclimatic factors (Table 2) from Arshad et al. (2022) and Dong et al. (2023) were used in this investigation for environmental predictors. According to Hijmans et al. (2005), the most frequently used bioclimatic variables that have been extracted from the climate database that is registered in WorldClim (www.worldclim.org, new version 2.0) are Bio 1 through Bio 19 (Figure 2). The pipeline started with performing the multicollinearity test to retrieve the bioclimatic variables with less potential for autocorrelation. The next stage is the habitat suitability modeling using several models. The obtained model was then evaluated. This WorldClim database has been extensively distributed and utilized in several researches on habitat suitability modeling throughout tropical Asia (Rana et al. 2017). Bioclimatic variables were represented at this stage as grids, with a typical spatial resolution of 1 km and a spatial resolution of 30 arc seconds, which in the case of the equator corresponded to around 0.86 km^2 and smaller elsewhere (Fick and Hijmans 2017).

According to Préau et al. (2018), a multicollinearity test together with Pearson's correlation test as parts of variable selection, was used to develop a model with enhanced performance using fewer variables and to prevent collinearity among variables. The model was built on 19 bioclimatic variables, ranging from Bio 1 to Bio 19. In order to build species distribution modeling at a geographical scale, variables exhibiting cross-correlation at high values of r^2 more than 0.8 were removed, whereas variables with r^2 values less than 0.8 were retained for additional study (Figure 3). Multicollinearity in the model indicates that a bioclimatic variable is substantially associated with other variables. Multicollinearity suggests the model has unstable and unreliable estimating capacity, according to As'ary et al. (2023). According to the multicollinearity test, the selected bioclimatic variables to be used were Bio 1, 2, 3, 4, 11, 12, 13, 14, 15, 16, 18, and 19 (Table 2).

Potential distribution analyses

Potential distribution analyses, according to Dolci and Peruzzi (2022), are based on statistical analysis, machine learning, and geoclimate used for crop modelling (Khalil et al. 2021; Ali et al. 2023). Using various species modeling algorithms (Table 3) in the R platform version 3.6.3 package (Table 3) (Mao et al. 2022) will yield estimated suitability maps of *S. bicolor* throughout the Indonesian islands of Lombok, Sumba, Sumbawa, Flores, and Timor. The generated estimated suitability maps were produced using many R packages, including library "sp", library "dismo" (Khan et al. 2022), library "random forest", library "kernlab", library "rgdal" (Bivand 2022), and library "raster" (Lemenkova 2020). Bio 1, 2, 3, 4, 11, 12, 13, 14, 15, 16, 18, and 19 were among the bioclimatic variables that were used as inputs for the model. Since this study only used the data of the presence of sorghum, pseudo-absences, also called background points, are also used and crucial for building accurate species distribution models. The background points are generated from the available environment and used to compare with the observed sorghum presences, helping potential distribution analyses determine the relationship between sorghum presence and environmental variables. Common methods include random sampling from the study area or a biased approach using buffer zones around sorghum presences. The "dismo" R package provides tools for creating background points, including methods for selecting and generating background points.

Table 2. Bioclimatic variables used in this study (Ulak and Paudel 2021)

Variables	Sources	Format	Unit
Annual mean temperature (Bio 1)*	www.worldclim.org	Image data in Raster	°C
Mean diurnal range (Bio 2) * (mean of monthly (max temp-min temp))	www.worldclim.org	Image data in Raster	°C
Isothermality (Bio 3)*	www.worldclim.org	Image data in Raster	%
Temperature seasonality (Bio 4)*	www.worldclim.org	Image data in Raster	°C
Max temperature of warmest month (Bio 5)	www.worldclim.org	Image data in Raster	°C
Min temperature of coldest month (Bio 6)	www.worldclim.org	Image data in Raster	°C
Temperature annual range (Bio 7)	www.worldclim.org	Image data in Raster	°C
Mean temperature of wettest quarter (Bio 8)	www.worldclim.org	Image data in Raster	°C
Mean temperature of driest quarter (Bio 9)	www.worldclim.org	Image data in Raster	°C
Mean temperature of warmest quarter (Bio 10)	www.worldclim.org	Image data in Raster	°C
Mean temperature of coldest quarter (Bio 11) *	www.worldclim.org	Image data in Raster	°C
Annual precipitation (Bio 12) *	www.worldclim.org	Image data in Raster	mm
Precipitation of wettest month (Bio 13) *	www.worldclim.org	Image data in Raster	mm
Precipitation of driest month (Bio 14) *	www.worldclim.org	Image data in Raster	mm
Precipitation seasonality (Bio 15) *	www.worldclim.org	Image data in Raster	coefficient of variation, % variability
Precipitation of wettest quarter (Bio 16) *	www.worldclim.org	Image data in Raster	mm
Precipitation of driest quarter (Bio 17)	www.worldclim.org	Image data in Raster	mm
Precipitation of driest quarter (Bio 18) *	www.worldclim.org	Image data in Raster	mm
Precipitation of coldest quarter (Bio 19) *	www.worldclim.org	Image data in Raster	mm

Note: *: selected variables based on multicollinearity test

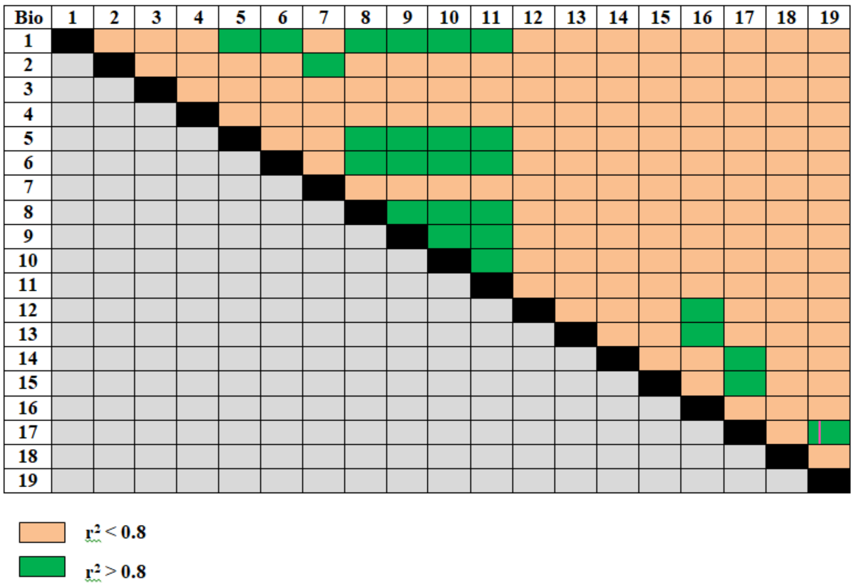

Figure 3. Pearson's correlation analysis matrix of the 19 bioclimatic variables (Bio 1 - Bio 19), and the green squares represent the significant correlation ($r^2 > 0.8$) among variables

Table 3. Evaluated models in this study

Model	Description	Name	Data	R packages
Machine learning	Based on the classification of training data until finally obtaining a better model.	Random Forest	Presence	random forest
		Support Vector Machine	Presence	kernlab
Geoclimate	Based on climate envelope algorithms and calculation of the similarity that exists between candidate pixels with respect to the selected presence records.	Bioclim	Presence	dismo
		Domain	Presence	dismo
Statistical	Based on the median of a response variable and using the Logit Link Function to relate the expected value of the response variable with included predictors.	GAM	Presence	dismo
		GLM	Presence	dismo

The effectiveness of the model was evaluated by looking at the receiving operating curve (AUC) area, and a jackknife test (Promnikorn et al. 2019) was used to analyze the influence and contribution of each bioclimatic variable affecting the appropriateness of *S. bicolor* habitat. AUC values between 0 and 0.5 indicate that the model performs poorly and indicates uninformative data, while values closer to 1.0 indicate that the final model is very informative and regarded as very excellent, according to Zhu et al. (2017). Besides using AUC to measure the model performance, this study also used COR (Calibration of Results) in evaluating models, particularly for classification tasks. COR assesses how well the model's predicted probabilities correspond to the actual probabilities of the event occurring.

Subsequently and according to Hijmans et al. (2012), the machine learning model' output, which estimated the suitability ranges of *S. bicolor*, then imported into GIS for further analysis and mapping display. Consistent with the findings of Wei et al. (2018), the habitat suitability levels on the utilized machine learning model map can be classified into five categories: unsuitable, low, medium, high, and very high suitability, with values ranging from 0.0 to 0.2, 0.2 to 0.4, 0.4 to 0.6, 0.6 to 0.8, and 0.8 to 1.0.

RESULTS AND DISCUSSION

Model evaluation and validation

As indicated by the area under the receiver-operating characteristic (AUC) curve acquired by the accuracy test of the ROC curve analysis method, model assessment and validation ensure the dependability of modeling results. The AUC values fell into multiple value classes and ranged from 0 to 1. The model's execution was inferior to contingency when the AUC value was less than 0.5. The model's performance is deemed moderate when the AUC value is between 0.5 and 0.6; acceptable when it is between 0.6 and 0.7; good when it is between 0.7 and 0.8; very good when it is between 0.8 and 0.9; and excellent when it is between 0.9 and 1. The model is more accurate and descriptive, and has better discrimination when the AUC test value is closer to 1. Based on the result, as can be observed in Figure 4, the AUC of the model is greater than 0.8, among which the Domain model is 0.962, SVM is 0.903, and both GLM and GAM are 0.894.

Response curves of bioclimatic variables

The response curves illustrate the correlations between each bioclimatic variable and the sorghum's degree of occurrence probability and suitability of habitat (Figure 5). The response curves of the areas of sorghum having high suitability were generated for the selected twelve bioclimatic variables. The likelihood of each bioclimatic variable to sorghum illustrate that with the rise in the bioclimatic variable value, the existence probability showed a trend of first increasing rapidly and then decreasing slowly. In this study, the range of climate factors when the probability of suitability is greater than 0.6 was used to represent the bioclimatic characteristics of the sorghum distribution area. The bioclimatic characteristics of the distribution area of sorghum are as follows: Bio1 or the annual mean temperature ranging from 25.0 to 26.0°C, Bio2 or the mean diurnal ranging from 9.0 to 12.0°C, Bio3 or the isothermality ranging from 70 to 80%, Bio4 or the seasonality of temperature ranging from 75.0 to 85.0°C, Bio7 or the annual range of tem-perature ranging from 12.0 to 14.0°C, Bio12 or the annual precipitation ranging from 150 to 250 mm, Bio13 or the precipitation of wettest month ranging from 25 to 35 mm, the precipitation of driest month, or Bio14, is ranging of 1-2 mm, the precipitation seasonality, or Bio15, is ranging of 8-9 mm, the precipitation of wettest quarter, or Bio16, is ranging of 8-10 mm, the precipitation of driest quarter, or Bio18, is ranging of 40-60 mm, and the precipitation of coldest quarter, or Bio19, ranging of is 9-11 mm.

Predictions of suitable habitats for sorghum

The modeling methods differed in predicting distribution areas for all species (Table 4; Figures 6, 7, and 8). For most islands, SVM predicted a larger distribution area. Conversely, Bioclim and Domain predicted less suitable areas of distribution for all islands. All models show consistency since all models confirm that Sumba, Sumbawa, and Flores will have larger suitable areas in comparison to other islands. Under the Bioclim model, Sumba has the smallest suitable area, equal to 485 km^2, and Sumbawa has the largest suitable area, equal to 5,560 km^2. The Domain model has also given similar results, with Lombok having the smallest suitable area at 3,465 km^2 and Sumbawa still has the largest suitable areas equals 15,245 km^2. Other models, including GAM/GLM and SVM, confirm that Timor, followed by Sumbawa and the Flores Islands, have large, suitable areas for sorghum.

Table 4. Distribution area in km^2 predicted by machine learning (SVM), geoclimate (Bioclim, Domain), and statistical (GAM/GLM) methods for the modeling of *Sorghum bicolor* in Lombok, Sumba, Sumbawa, Flores, and Timor Islands, Indonesia

Islands	Area (km^2)	Bioclim		Domain		GAM/GLM		SVM	
		km^2	%	km^2	%	km^2	%	km^2	%
Lombok	4,739	504	10.63	3,465	73.11	2,612	55.11	4,556	96.13
Sumba	11,006	485	4.40	9,556.	86.82	6,495	59.01	10,975	99.71
Sumbawa	15,414	5,560	36.07	15,245	98.90	10,481	67.99	15,331	99.46
Flores	15,531	3,728	24.00	10,333	66.53	5,998	38.61	14,687.	94.56
Timor	30,777	383	1.24	14,482	47.05	17,880	58.09	20,570	66.83

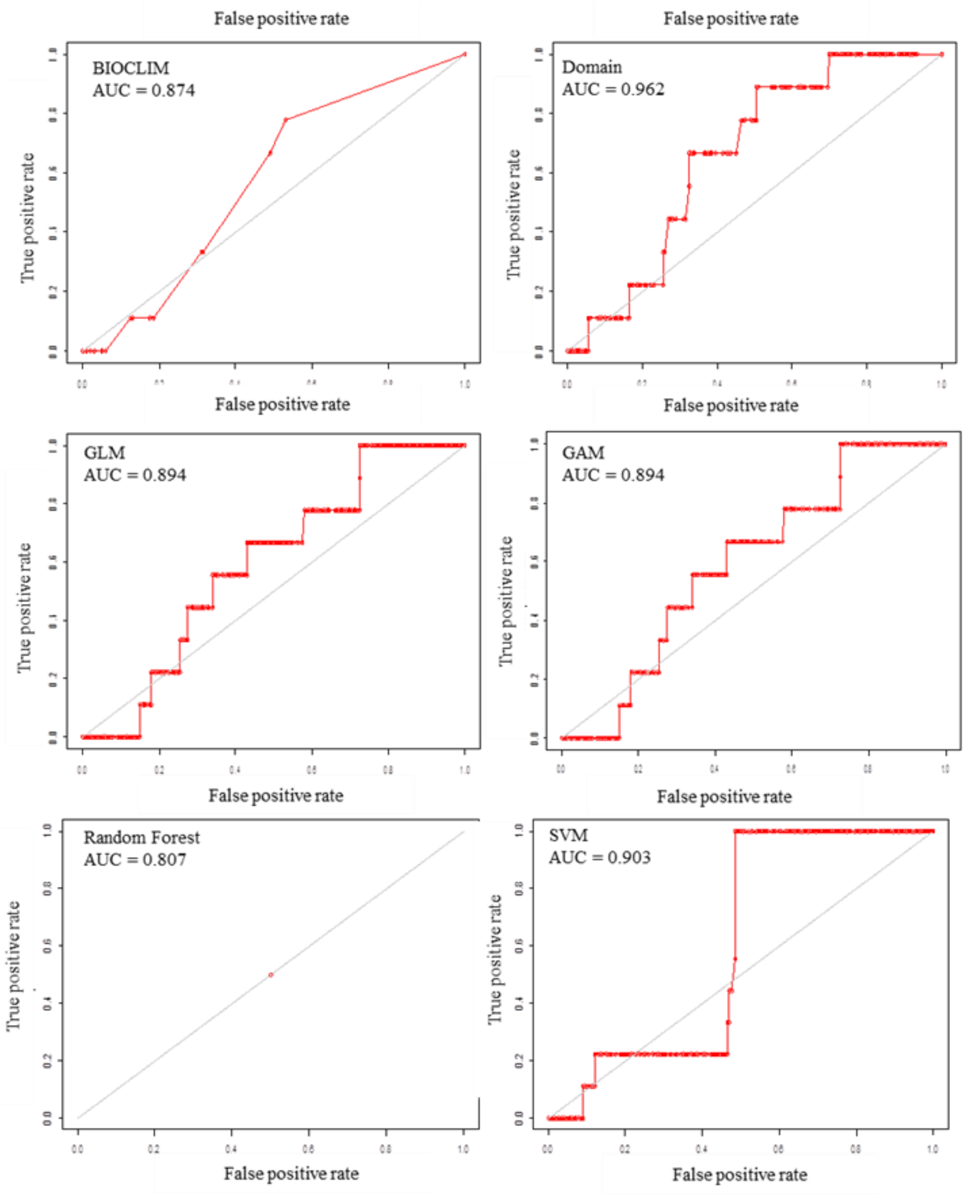

Figure 4. Area under the receiver operating characteristic curves (AUCs) for Bioclim. Domain, GLM, GAM, Random Forest, and SVM

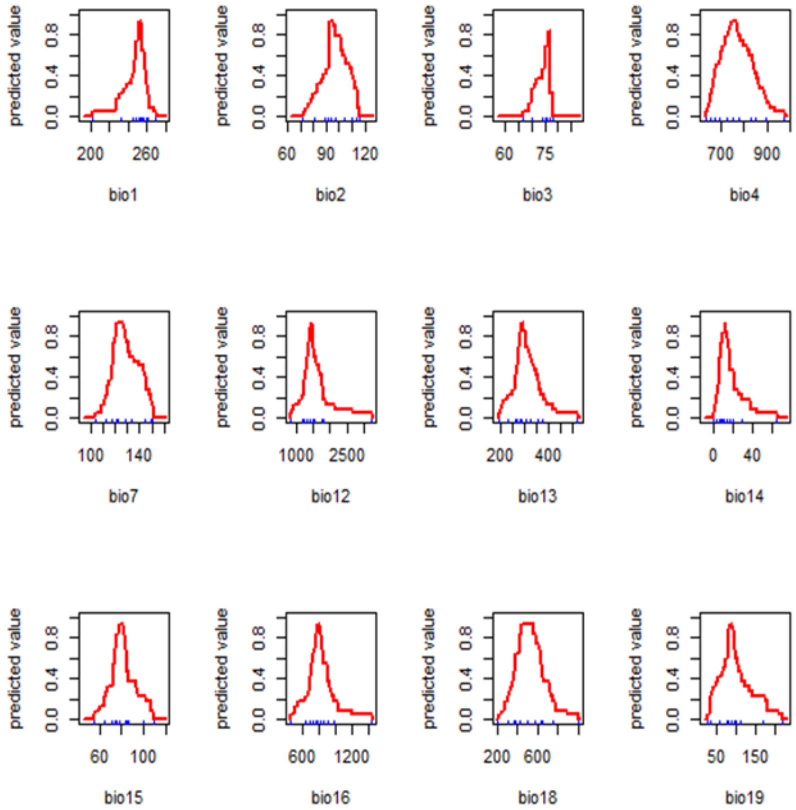

Figure 5. The response curves for Bio 1, 2, 3, 4, 11, 12, 13, 14, 15, 16, 18, and 19 bioclimatic

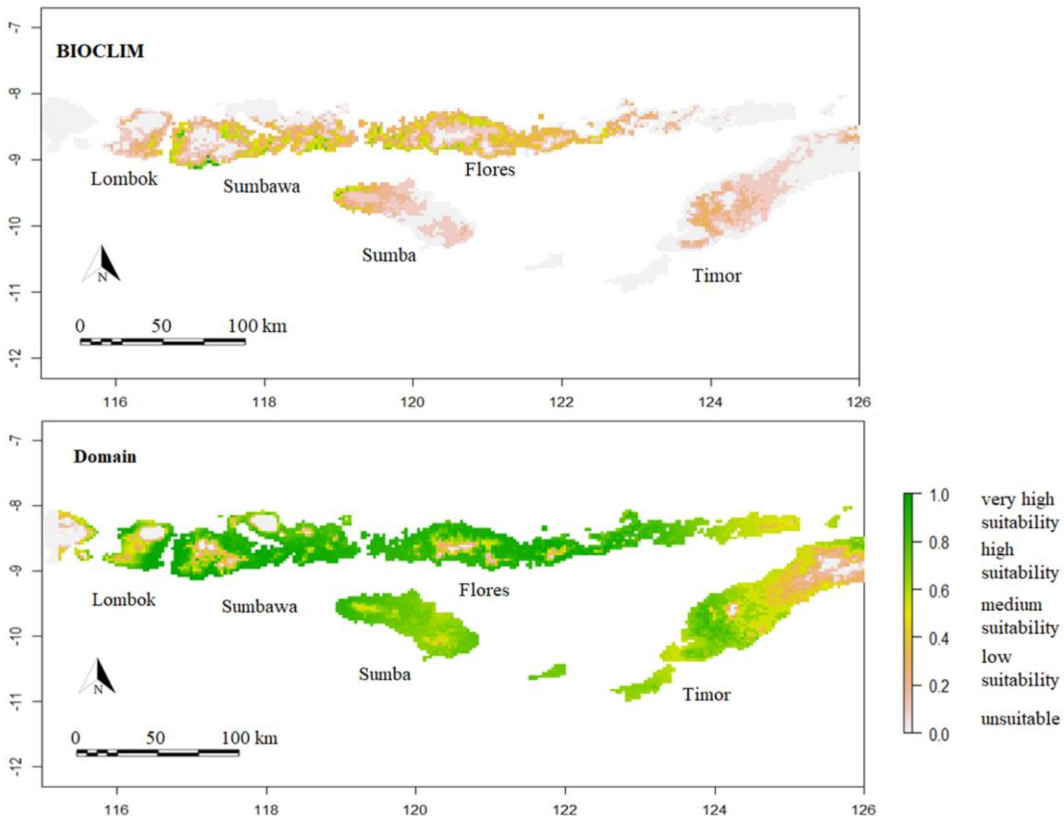

Figure 6. Distribution of suitable area predicted by geoclimate (Bioclim, Domain) for the modeling of *Sorghum bicolor* in Lombok, Sumba, Sumbawa, Flores, and Timor islands, Indonesia. Suitability level included 0.0-0.2: unsuitable, 0.2-0.4: low suitability, 0.4-0.6: medium suitability, 0.6-0.8: high suitability, 0.8-1.0: very high suitability

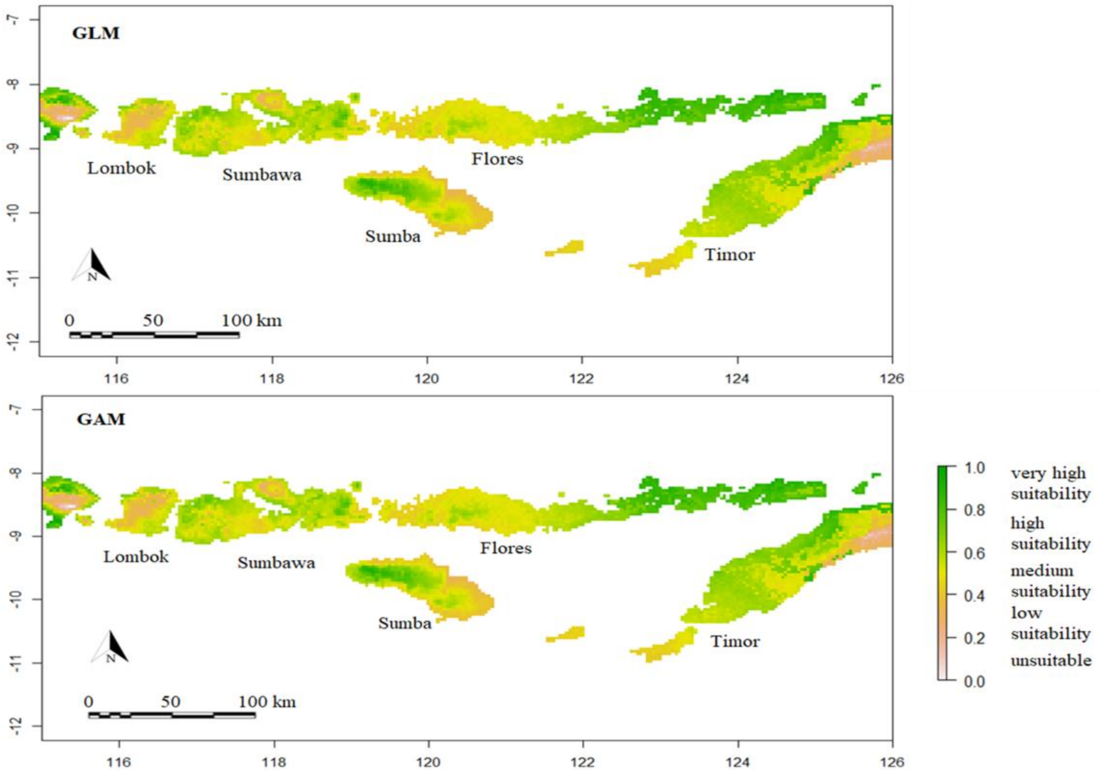

Figure 7. Distribution of suitable area predicted by statistical methods (GLM, GAM) for the modeling of *Sorghum bicolor* in Lombok, Sumba, Sumbawa, Flores, and Timor islands, Indonesia. Suitability level included 0.0-0.2: unsuitable, 0.2-0.4: low suitability, 0.4-0.6: medium suitability, 0.6-0.8: high suitability, 0.8-1.0: very high suitability

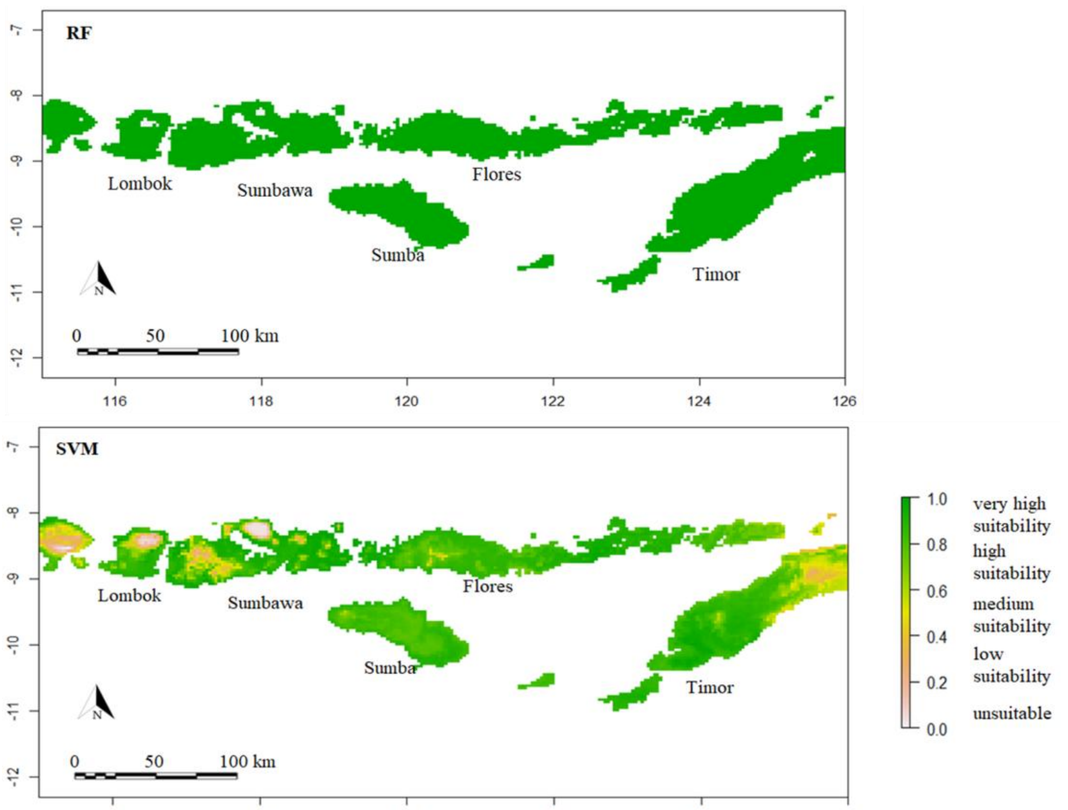

Figure 8. Distribution of suitable area predicted by machine learning (RF, SVM) for the modeling of *Sorghum bicolor* in Lombok, Sumba, Sumbawa, Flores, and Timor islands, Indonesia. Suitability level included 0.0-0.2: unsuitable, 0.2-0.4: low suitability, 0.4-0.6: medium

Table 5. Comparisons of sorghum distribution modeling studies in other locations

Locations	AUC	% high suitable areas	References
KwaZulu-Natal, South Africa	0.93	13.4	Mugiyo et al. (2022)
China	na	na	Niu et al. (2022)
Telangana state, India	na	na	Natarajan et al. (2016)
Kenya, Africa	0.97	-	Kigen et al. (2014)
Lombok, Sumba, Sumbawa, Flores, and Timor Islands, Indonesia	0.807-0.962	1.24 - 99.71	This study

Table 6. AUC: area under the receiver-operated characteristic curve and COR: point biserial correlation coefficient between observed and predicted values of each models

Coeff.	Bioclim	Domain	GLM	GAM	RF	SVM
AUC	0.874	0.962	0.894	0.894	0.807	0.903
COR	0.005	0.112	0.049	0.049	0.024	0.105

Discussion

Certain sorghum species may be comparatively more climate-sensitive in general, and precipitation in particular, since sorghum is a drought-tolerant species. This research forecasted the possible influences of climate variables on the sorghum as a potential crop. This is the first study, especially in the Southeast Asia region, to examine the range extensions of sorghum based on several species distribution models. The species presence data, along with environmental variables, have been wisely chosen and validated to guarantee the accuracy of the model. The AUC was used for sorghum model parameter adjustment and evaluation, and the results showed that the model had a high degree of prediction accuracy. The findings of this study are comparable to those of previous studies (Table 5). Currently, the potential distribution of sorghum has been modeled in China, India, and Africa. At the same time, there is a paucity of this information, mainly in the Southeast Asia region.

The result of this study then contributes significantly to the Indonesian government's policy in promoting sorghum cultivation. The result determines accurately the area suitable for sorghum. The potential areas delineated in this study can assist the small-scale farmer to carefully select the most cultivation areas. The use of modeling to estimate species potential area is in line with previous study (Jeong et al. 2022). Besides that implication, this study also offers a novel method in the form of modeling that can be applied to other species and assist the Indonesian government in shaping cultivation policy, in particular determining potential cultivation zoning for certain species.

Sorghum distributions in arid ecosystems in eastern Indonesia were shown to be impacted by precipitation and to have a negative correlation. At the same time, sorghum can withstand temperature increases. Mugiyo et al. (2022) confirmed that rainfall-related parameters, in this case precipitation, had the greatest influence on potential applicability. Drought is a feature of sorghum, according to Niu et al. (2022). The environmental variables, temperature and precipitation, contributed 86.2% to the model, suggesting that sorghum is heat resistant and avoids places with low temperatures and humidity.

Isothermality and seasonality of temperature factors were observed to have strong contributions adjacent to precipitation variables in this study. This is consistent with the findings of Huang et al. (2021), who found that plant groupings tended to reside in flatter topography areas with more variability of temperature, isothermality, and seasonality of temperature. They also found that the relative influences of isothermality and seasonality of temperature maximized in tropical locations. Sorghum was estimated to be very suitable in flat lowlands with elevations less than 1,000 m on arid islands in eastern Indonesia.

Here we present the comparison of the performance of six modelling methods to estimate the potential distributions of sorghum. Our findings confirm that the preferences of modelling method may affect the determination of suitable habitats for the sorghum (Oppel et al. 2012). Consistent with earlier comparative research, none of the six approaches evaluated produced better estimations in all performance parameters. The Domain and SVM models performed similarly well in our investigation in terms of prediction (Table 6). In spite of the predictions of the six individual models being generally consistent, discrepancies still existed due to the differences in algorithms of each model used, and this is a common phenomenon as observed previously in related studies in Hao et al. (2019) and Liu et al. (2022). The good performances of SVM was in agreement with previous studies (Ghareghan et al. 2020). SVM was recognized as a model with a higher accuracy, and also an inexpensive method for estimating the habitat suitability of a species (Boogar et al. 2019).

In our study, Domain model also shows a good performance. This is contradicted to results by Duan et al. (2014) that stated Domain performance was lower than SVM. The performances of a model were affected by environmental variables and as a result affecting model stability. This study is implemented and limited to small island ecosystems where the data are limited and the environment tends to be homogenous. The domain model is known flexible by Carpenter et al. (1993) and can offer advantages in mapping potential distributions of species, particularly when data is limited and environmental homogeneity is present. Domain can effectively use presence records and biophysical attributes to model potential ranges. Besides that, improved mathematical modeling techniques, machine learning algorithms, and more robust statistical tools will also affect the model performances. The other models that considered have good performances based on AUC values were GAM and GLM. In their study, De et al. (2020) has confirmed higher AUC value of GLM in comparison to RF and SVM. The model's prediction accuracy may be referred to as outstanding based on the ROC results and high AUC values. As a result, this finding can be utilized to determine which models better capture the fitness of sorghum's potential distribution.

Despite this study having succeeded in using and comparing six modeling approaches that are already known to have advantages and have been used widely in the modeling studies, a discrepancy among models is still observed, and this becomes a challenge in this study. In the future, we recommend applying the ensemble approach (Kaky et al. 2020) to gain consensus. Another improvement is regarding the variables used in this study that are limited to bioclimatic variables. Accompanied by the development of modeling approaches, we also recommend including more variables as supplementary data in the model to increase the accuracy of the prediction results and model. Those variables include land use, geological, elevation, and anthropogenic variables. The study is comparing several independent models. Then it is encouraged to compare the individual models with the ensemble approach. This approach has advantages due to its capacity to reduce both bias and variance, particularly in the presence of noise and uncertainty that are common in individual models.

Using machine learning, geoclimate, and statistical methods, the present study estimates comprehensively the spatial distribution of habitat suitability for sorghum in the arid eastern Indonesia ecosystems. The comprehensive result and species distribution model can then generate a knowledge base for future strategies for sorghum crop planners. Modeling methods, such as the one used in this study, can be applied by particular agriculture practitioners and farmers to estimate the areas where the most losses and profits are generated under sorghum cultivation practice, and so maintaining the natural geographic distribution areas of the sorghum under attention needs the greatest consideration and ensure food security, mainly in arid eastern Indonesia ecosystems.

All the models confirm that Timor, followed by Sumbawa and the Flores Islands, have large, suitable areas for sorghum. It is estimated that up to 99.71% of arid island ecosystems in eastern Indonesia were suitable for sorghum. The geoclimate and machine learning model generated the highest values for AUC in comparison to statistical methods. Then geoclimate and machine learning methods can be considered suitable methods to estimate sorghum's potential distribution areas, considering their performances.

The result is considered very useful in contributing to the Indonesian government's zoning policy and climate-resilient agriculture strategies. It provides empirical evidence about suitable areas for sorghum. It is strongly recommended to prioritize Timor, followed by Sumbawa and the Flores Islands, as potential sorghum cultivation zones.

ACKNOWLEDGEMENTS

We are grateful to the students and research staff of School of Environmental Science, Universitas Indonesia that have assisted the data collection and discussion. This study was supported and funded by Direktorat Riset and Pengembangan Universitas Indonesia through PUTI Q2 scheme with grant number of NKB-629/UN2.RST/ HKP.05.00/2022.

REFERENCES

Akbar D, Utami S, Virgianto R. 2021. Analysis of the relationship between meteorological drought and agricultural drought on Lombok Island using Pearson's Correlation. Delta: J Math Educ 9: 133. DOI: 10.31941/delta.v9i1.1275.

Ali S, Makanda T A, Umair M, Ni J. 2023. MaxEnt model strategies to studying current and future potential land suitability dynamics of wheat, soybean and rice cultivation under climatic change scenarios in East Asia. PLoS ONE 18 (12): e0296182. DOI: 10.1371/journal.pone.0296182.

As'ary M, Setiawan Y, Rinaldi D. 2023. Analysis of changes in habitat suitability of the Javan Leopard, 2000-2020. Diversity 15: 529. DOI: 10.3390/d15040529.

Arshad F, Waheed M, Fatima K, Harun N, Iqbal M, Fatima K, Umbreen S. 2022. Predicting the suitable current and future potential distribution of the native endangered tree *Tecomella undulata* (Sm.) Seem. in Pakistan. Sustainability 14: 7215. DOI: 10.3390/SU1412721.

Bivand R. 2022. R packages for analyzing spatial data: a comparative case study with areal data. Geogr Anal 54 (3): 488-518. DOI: 10.1111/GEAN.12319.

Boogar RA, Salehi H, Pourghasemi HR, Blaschke T. 2019. Predicting habitat suitability and conserving *Juniperus* spp. habitat using SVM and Maximum Entropy machine learning techniques. Water 11: 2049. DOI: 10.3390/w11102049.

Carpenter G, Gillison AN, Winter J. 1993. Domain: a flexible modelling procedure for mapping potential distributions of plants and animals. Biodivers Conserv 2 (6): 667-680. DOI: 10.1007/bf00051966.

De K, Ali SZ, Dasgpta N, Uniyal VP, Johnson JA, Hussain SA. 2020. Evaluating performance of four species distribution models using Blue-tailed Green Darner *Anax guttatus* (Insecta: Odonata) as model organism from the Gangetic riparian zone. J Threat Taxa 12 (14): 16962-16970. DOI: 10.11609/jott.6106.12.14.16962-16970.

Dolci D, Peruzzi L. 2022. Assessing the effectiveness of correlative ecological niche model temporal projection through floristic data. Biology 11: 1219. DOI: 10.3390/biology11081219.

Dong H, Zhang N, Shen S, Zhu S, Fan S, Lu Y. 2023. Effects of climate change on the spatial distribution of the threatened species *Rhododendron purdomii* in Qinling-Daba Mountains of Central China: implications for conservation. Sustainability 15 (4): 3181. DOI: 10.3390/su15043181.

Duan RY, Kong XQ, Huang MY, Fan WY, Wang ZG. 2014. The predictive performance and stability of six species distribution models. PloS ONE 9 (11): e112764. DOI: 10.1371/journal.pone.0112764.

Fick SE, Hijmans RJ. 2021. WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. Intl J Climatol 37 (12): 4302-431.

Fitzgibbon A, Pisut D, Fleisher D. 2022. Evaluation of Maximum Entropy (Maxent) Machine Learning Model to assess relationships between climate and corn suitability. Land 11 (9): 1382. DOI: 10.3390/land11091382.

Fois M, Cuena-Lombraña A, Fenu G, Bacchetta G. 2018. Using species distribution models at a local scale to guide poorly known species, review: Methodological issues and future directions. Ecol Model 385: 124-132. DOI: 10.1016/j.ecolmodel.2018.07.018.

Ghareghan F, Ghanbarian G, Pourghasemi HR, Safaeian R. 2020. Prediction of habitat suitability of *Morina persica* L. species using artificial intelligence techniques. Ecol Indic 112 (4). DOI: 10.1016/j.ecolind.2020.106096.

Gunawan, Sulistijorini, Chikmawati T, Sobir. 2021. Predicting suitable areas for *Baccaurea angulatain* Kalimantan, Indonesia Using MaxEnt Modelling. Biodiversitas 22: 2646-2653. DOI: 10.13057/biodiv/d220523.

Hadebe ST, Modi AT, Mabhaudhi T. 2017. Drought tolerance and water use of cereal crops: A focus on sorghum as a food security crop in Sub-Saharan Africa. J Agron Crop Sci 203 (3): 177-191. DOI: 10.1111/jac.12191.

Hao T, Elith J, Guillera-Arroita G, Lahoz-Monfort JJ. 2019. A review of evidence about use and performance of species distribution modeling ensembles like BIOMOD. Divers Distrib 25: 839-852. DOI: 10.1111/ddi.12892.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Intl J Climatol 25: 1965-1978. DOI: 10.1002/JOC.1276.

Hijmans RJ, Guarino L, Mathur P. 2012. Diva-GIS Version 7.5 Manual. Available from: http://diva-gis. org/docs/DIVA-GIS_manual_7. pdf.

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 (3): 697-709. DOI: 10.1111/geb.13251.

Impa SM, Perumal R, Bean SR, Sunoj VSJ, Jagadish SVK. 2019. Water deficit and heat stress induced alterations in grain odeli-chemical characteristics and micronutrient composition in field grown grain sorghum. J Cereal Sci 86: 124-131. DOI: 10.1016/j.jcs.2019.01.013.

Jeong A, Kim M, Lee S. 2024. Analysis of priority conservation areas using habitat quality models and MaxEnt models. Animals 14 (11): 1680. DOI: 10.3390/ani14111680.

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 Inf 60: 101150. DOI: 10.1016/j.ecoinf.2020.101150.

Khalil T, Asad SA, Khubaib N, Baig A, Atif S, Umar M, Kropp JP, Pradhan P, Baig S. 2021. Climate change and potential distribution of potato Solanum tuberosum crop cultivation in Pakistan using Maxent. AIMS Agric Food 6 (2): 663-676. DOI: 10.3934/agrfood.2021039.

Khan AM, Li Q, Saqib Z, Khan N, Habib T, Khalid N, Majeed M, Tariq A. 2022. Maxent modeling and impact of climate change on habitat suitability variations of economically important Chilgoza pine (*Pinus gerardiana* wall.) in South Asia. Forests 13: 715. DOI: 10.3390/F13050715.

Khanum R, Mumtaz A, Kumar S. 2013 Predicting impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. Acta Oecol 49: 23-31. DOI: 10.1016/J.ACTAO.2013.02.007.

Kigen C, Dennis M, Ochieno, Muoma J, Shivoga W, Konje M, Onyando Z, Soi B, Makindi S, Kisoyan P, Mironga J. 2014. Spatial modeling of sorghum *(Sorghum bicolor*) growing areas in Kenyan Arid and Semi-Arid Lands. Elixir Remote Sens 66: 20674-20678.

Lemenkova P. 2020. Using R Packages 'tmap', 'raster' and 'ggmap' for cartographic visualization: An example of DEM-based terrain modeling of Italy, Apennine Peninsula. ZbornikRadova - Geografski Fakultet Univerziteta U Beogradu 68: 99-116. DOI: 10.5937/ZRGFUB2068099L.

Li X, Wu K, Hao S, Yue Z, Ran Z, Ma J. 2023. Mapping cropland suitability in China using optimized MaxEnt model. Field Crops Res 302: 109064. DOI: 10.1016/j.fcr.2023.109064.

Lin Z, Chen C, Liu Y, Liu G, He P, Liao G, Gao W, Cao J, Shao Z. 2022. Simulation of citrus production space based on MaxEnt. Front Environ Sci 10: 993920. DOI: 10.3389/fenvs.2022.993920.

Liu T, Liu H, Tong J, Yang Y. 2022. Habitat suitability of neotenic net-winged beetles (Coleoptera: Lycidae) in China using combined ecological models, with implications for biological conservation. Divers Distrib 28: 2806-2823. DOI: 10.1111/ddi.13545.

Mao M, Chen S, Qian Z, Xu Y. 2022. Using Maxent to predict the potential distribution of the little fire ant (*Wasmannia auropunctata*) in China. Insects 13: 1008. DOI: 10.3390/INSECTS13111008.

Marcer A, Sáe L, Molowny-Horas R, Pons X, Pino J. 2013. Using species distribution modeling to disentangle modeling versus potential distributions for rare species conservation. Biol Cons 166: 221-230. DOI: 10.1016/j.biocon.2013.07.001.

Mathur S, Umakanth AV, Tonapi VA, Sharma R, Sharma MK. 2017. Sweet sorghum as biofuel feedstock: recent advances and available resources. Biotechnol Biofuels 10: 146. DOI: 1010.1186/s13068-017-0834-9.

McCann T, Krause D, Sanguansri. 2015. P. Sorghum—New gluten-free ingredient and applications. Food Aust 67 (6): 24-2 16.

Mugiyo H, Chimonyo VGP, Kunz R, Sibanda M, Nhamo L, Masemola C, Modi AT, Mabhaudhi T. 2022. Mapping the spatial distribution of crop species under climate change using the MaxEnt model: A case of KwaZulu-Natal, South Africa. Clim Serv 28: 100330. DOI: 10.1016/j.cliser.2022.100330.

Natarajan S, Elangovan M, Venkateswaran K, Pandravada SR, Pranusha P, Chakrabarty SK. 2016. Maximum Entropy (Maxent) approach to sorghum landraces distribution modeling. Indian J Plant Genet Resour 29: 16. DOI: 10.5958/0976-1926.2016.00004.8.

Niu K, Zhao L, Zhang Y, Wang Z, Wang Z, Yang H. 2022. Prediction of potential sorghum suitability distribution in China based on Maxent Model. Am J Plant Sci 13: 856-871. DOI: 10.4236/ajps.2022.136057.

Oppel S, Meirinho A, Ramírez I, Gardner B, Connell AFO, Miller PI, Louzao M. 2012. Comparison of five modeling techniques to predict the spatial distribution and abundance of seabirds. Biol Cons 156: 94-104.

Paesal, Syuryawati, Suarni, Aqil M. 2021. Sorghum cultivation of the ratoon system for increased yields in dry land. IOP Conf Ser: Earth Environ Sci 911: 012035. DOI: 10.1088/1755-1315/911/1/012035.6.

Préau C, Trochet A, Bertrand R, Isselin-Nondedeu F. 2018. Modeling potential distributions of three European amphibian species comparing ENFA and Maxent. Herpetol Conserv Biol 13 (1).

Promnikorn K, Jutamanee K, Kraichak E. 2019. Maxent model for predicting potential distribution of *Vitex glabrata* r.br. in Thailand. Agric Nat Resour 53: 44-48.

Rana SK, Rana HK, Ghimire SK, Shrestha KK, Ranjitkar S. 2017. Predicting the impact of climate change on the distribution of two threatened Himalayan medicinal plants of liliaceae in Nepal. J Mt Sci 14 (3): 558-570. DOI: 10.1007/S11629-015-3822-1.

Ruiz-Giralt A, Biagetti S, Madella M, Lancelotti C. 2023. Small-scale farming in drylands: New models for resilient practices of millet and sorghum cultivation. PloS ONE 18 (2). E0268120. DOI: 10.1371/journal.pone.0268120.

Sánchez Pérez M, Feria Arroyo TP, Venegas Barrera CS, Sosa-Gutiérrez C, Torres J, Brown KA, Gordillo Pérez G. 2023. Predicting the impact of climate change on the distribution of *Rhipicephalus sanguineus* in the Americas. Sustainability 15: 4557. DOI: 10.3390/su15054557.

Sapta S, Sulistyantara B, Fatimah I, Faqih A. 2015. Geospatial approach for ecosystem change study of Lombok Island under the influence of climate change. Procedia Environ Sci 24. DOI: 10.1016/j.proenv.2015.03.022.

Sarshad A, Talei D, Torabi M, Rafiei F, Nejatkhah P. 2021. Morphological and biochemical responses of *Sorghum bicolor* (L.) Moench under drought stress. SN Appl Sci 3 (1): 81. DOI: 10.1007/s42452-020-03977-4.

Semu AA, Bekele T, Lulekal E, Cariñanos P, Nemomissa S. 2021. Projected impact of climate change on habitat suitability of a vulnerable endemic *Vachellia negrii* (pic.serm.) Kyal. & Boatwr (Fabaceae) in Ethiopia. Sustainability 13: 11275. DOI: 10.3390/su132011275.

Stephenson K, Wilson B, Taylor M, McLaren K, Veen R, Kunna J, Campbell J. 2022. Modelling climate change impacts on tropical dry forest fauna. Sustainability 14 (8): 4760. DOI: 10.3390/su14084760.

Sutomo van Etten E, Iryadi R. 2021. Short communication: Savanna-forest boundary on Mount Rinjani, Lombok Island, West Nusa Tenggara, Indonesia. Biodiversitas 22: 726-731. DOI: 10.13057/biodiv/d220225.

Ulak S, Paudel P. 2021. Maxent modelling for habitat suitability of vulnerable tree *Dalbergia latifolia* in Nepal. Silva Fenn 55 (4): 17. DOI: 10.14214/sf.10441.

Wei B, Wang R, Hou K, Wang X, Wu W. 2018. Predicting the current and future cultivation regions of *Carthamus tinctorius* using Maxent model under climate change in China. Glob Ecol Conserv 16: E00477. DOI: 10.1016/J.GECCO.2018.E00477.

Yanti LR, Soemeinaboedhy IN, Yasin I, Rahayu R. 2022. The relationship between drought events and The El-Nino Phenomena in The North Lombok Regency. Agrokomplek 1 (3): 285-293.

Yasin I, Mas'hum M, Abawi Y, Hadiahwaty L. 2004. Southern oscillation index for forecasting seasonal rainfall characteristic to determine upland planting strategy in Lombok in Lombok Island. J Argomet 18 (2): 24-26.

Yu SM, Lo SF, Ho THD. 2015. Source-sink communication: regulated by hormone, nutrient, and stress cross-signaling. Trends Plant Sci 20 (12): 844-857. DOI: 10.1016/j.tplants.2015.10.009.

Zhu GP, Gariepy TD, Haye T, Bu WJ. 2017. Patterns of niche filling and expansion across the invaded ranges of Halyomorpha halys in North America and Europe. J Pest Sci 90: 1-13. DOI: 10.1007/s10340-016-0786-z.