



RESEARCH ARTICLE

# Explaining distributional patterns of *Trianthema portulacastrum* and *Ageratum conyzoides* in India under future climatic scenarios using Ensemble modelling approach

Aditi Singh<sup>1</sup>, Yogita Gharde<sup>\*2</sup> and P.K. Singh<sup>2</sup>

Received: 6 December 2023 | Revised: 28 January 2024 | Accepted: 31 January 2024

## ABSTRACT

Weeds are recognized as the most aggressive, troublesome, and competitive elements within croplands. Climate change may affect the geographical distribution of existing weeds or the invasion of weeds in new areas. Therefore, in the current study, modelling was carried out to explore and predict the invasion potential of *Trianthema portulacastrum* and *Ageratum conyzoides* in India under current as well as future climatic conditions. Future climatic scenarios under Representative Concentration Pathways (RCPs) 4.5 and 8.5 for the years 2050 and 2070 were considered and modelling was performed using regression techniques and machine learning approaches with Ensemble technique. Mutually least correlated eight bioclimatic variables along with soil and elevation data were used for the modelling over 375 and 379 occurrence locations of the *T. portulacastrum* and *A. conyzoides*, respectively. True Skill Statistic (TSS) was used to evaluate the models' predictive performance, while AUC of the ROC and Kappa were used to crosscheck their performance. Results revealed that Ensemble model outperformed the individual models for both the species with higher AUC, TSS and Kappa values. Bioclimatic variables such as temperature seasonality, annual mean temperature, and minimum temperature of the coldest month were found to govern the potential distribution of both the species. Results of potential distribution were obtained in four climate suitability classes: not suitable (0-0.2), low (0.2-0.4), moderately (0.4-0.6), and highly suitable (>0.6). Modelling suggests the expansion of suitable ranges of *T. portulacastrum* in India under future climatic scenarios, whilst *A. conyzoides* are expected to predominantly contract in the future.

**Keywords:** Bioclimatic variables, Ensemble modelling, Machine learning, Potential distribution, Species distribution modelling

## INTRODUCTION

Weeds are regarded as the most damaging biotic constraint to agricultural production, in addition to threatening agro-biodiversity and natural water bodies. Weeds compete for primary resources such as light, water, nutrients and space that limit the agricultural production and cause huge yield and economic losses. It is estimated that weeds reduce crop yields by 31.5% (22.7% in winter and 36.5% in summer and rainy seasons) in India (Bhan *et al.* 1999). It is estimated to cause an economic loss of about USD 11 billion in 10 major field crops in India due to weeds (Gharde *et al.* 2018, 2019).

The introduction of alien invasive species poses a serious threat to native biodiversity, ecosystem integrity, and agricultural productivity, resulting in

significant economic consequences (Rai and Singh 2020). Invasive species disrupt trophic balance and reduce species resilience. Furthermore, global climate change increases the risk of alien and invasive plant species while also expanding their range. Previous research has found that many alien and invasive species have characteristics that increase invasion success and allow them to easily adapt to climate change (Priyanka and Joshi 2013; Banerjee *et al.* 2017).

Climate change may be the driver for global range expansions (migration or introduction into new areas), and changes in life cycles of species. Weed migration may result in a difference in the structure and composition of weed communities in natural and managed ecosystems (Ramesh *et al.* 2017). Identifying possible habitats integrating changing climate scenarios is an effective strategy for reducing the spread of invasive plant species. Species distribution models (SDMs) are the most commonly utilised tool for investigating the effects of climate change on distributions (Gharde *et al.* 2023). SDMs anticipate the distributions of given species based on

<sup>1</sup> Department of Agricultural Statistics and Social Science (L.), College of Agriculture, Indra Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, India

<sup>2</sup> ICAR-Directorate of Weed Research, Maharajpur, Adhartal, Jabalpur, Madhya Pradesh 482004, India

\* Corresponding author email: yogita.gharde@icar.gov.in

abiotic factors such as environmental data, making it primarily a tool for predicting the fate of a species in the future in terms of ecology and conservation. These models also aid in understanding existing and potential relationships between species, organisms, environmental conditions, and area richness (Elith *et al.* 2006). SDMs are an evolved version of the predictive habitat distribution models (Guisan and Zimmermann 2000).

Among alien invasive weeds (AIWs), species such as *Trianthema portulacastrum* and *Ageratum conyzoides* have posed serious threats to biodiversity in the country. Both weeds are native to Africa and are considered troublesome weeds in cultivated crops. *T. portulacastrum* also known as horse purslane, is an annual broadleaf weed belonging to the Aizoaceae family and is a widely distributed weed in mustard, corn, pigeon pea, soybean, tomato, potato, onion, cotton, sugarcane, pearl millet, sorghum, maize, direct-seeded rice, summer and rainy season pulses, oilseed crops, fodder crops, vegetables and horticultural crops in India. This weed is said to be a strong competitor, reducing mungbean yield by 50–60% when left untreated (Kaur *et al.* 2017; Aggarwal *et al.* 2017). It is also responsible for significant losses in maize, soybean, and peanut yield. Its extract has an allelopathic effect on soybean seed germination, seedling vigour, and productivity.

*Ageratum conyzoides*, called Billygoat weed, is an annual broad-leaf weed belonging to the family Asteraceae. It reduces yields in major staple crops such as wheat, corn and rice in India (Kohli *et al.* 2006; Batish *et al.* 2006; Kaur *et al.* 2012). It also infiltrates rangeland areas, where it outcompetes native grasses, resulting in a lack of fodder. Its growth strategies, which include rapid growth rates, short life cycles, increased reproductive potential, high competitive abilities, and allelopathy, make it a successful invader of native habitat (Kohli *et al.* 2006; Batish *et al.* 2006; Singh *et al.* 2012).

Considering the negative impact of these two weeds on the performance of different crops in terms of yield and economics, the proposed work aims to study the current geographical distribution of *T. portulacastrum* and *A. conyzoides* and to predict the future expansion/contraction of the species under RCP 4.5 and 8.5 for the years 2050 and 2070 using Ensemble -modelling approach. The study will be helpful to the researchers in identifying the potential areas of invasion for the species and accordingly, plan the strategies for the prevention of these species in newer areas.

## MATERIALS AND METHODS

### Data collection

The study was conducted during the year 2023 at ICAR-Directorate of Weed Research, Jabalpur. For the study, occurrence data for *T. portulacastrum* and *A. conyzoides* were obtained from various sources, including India Biodiversity Portal (<https://indiabiodiversity.org/>); Global Biodiversity Information Facility GBIF (<https://www.gbif.org/>); Centre for Agriculture and Bioscience International CABI (<https://www.cabi.org/>); Flora of Peninsular India, Herbarium JCB, Centre for Ecological Sciences, Indian Institute of Sciences, Bengaluru (<http://flora-peninsula-indica.ces.iisc.ac.in/>) and <https://www.inaturalist.org/>. Occurrence records acquired from Annual Reports of the All India Coordinated Research Project on Weed Management and ICAR-Directorate of Weed Research (DWR) (2010-2022), Weed Atlas, Vol. I and II published in 2008 by the ICAR-DWR were also included. However, over-lapping occurrence points were removed before analysing the data. Altogether 375 occurrence points of *T. portulacastrum* and 379 occurrence points of *A. conyzoides* were considered for the study.

### Climatic variables

Climatic variables (often called bioclimatic data primarily used for ecological applications) namely bio1 through bio19 with a resolution of 30 arc seconds were downloaded from the website [www.worldclim.org](http://www.worldclim.org) for current and two future climate periods, e.g. 2050 (2041-2060 average) and 2070 (2061-2080 average) under Representative Concentration Pathways (RCP) 4.5 and 8.5. As these files were in tiff format, QGIS-v3.22.7 software was used to convert them to ASCII (American Standard Code for Information Interchange) format. Elevation data from the Shuttle Radar Topography Mission (SRTM) and soil layers from [www.worldclim.org](http://www.worldclim.org) and [OpenLandMap.org](http://OpenLandMap.org), respectively, were also considered for the study that influenced the distribution of species. The soil layer's resolution was set to 30 arc seconds and converted to ASCII format before the analysis.

RCP 4.5 denotes the average/moderate GHG emission pathway, whereas RCP 8.5 denotes the greatest GHG concentration pathway (IPCC 2014; Thapa *et al.* 2018). The RCP 4.5 model predicts a steady increase in radiative forcing with projected global mean surface temperatures ranging from 1.4°C to 1.8°C, whereas the RCP 8.5 model assumes

the greatest increase in radiative forcing with projected global mean surface temperatures ranging from 2.0°C to 3.7°C. Future climate data under RCP 4.5 and 8.5 for the years 2050 and 2070 under the model cccma\_canes2 was used for modelling.

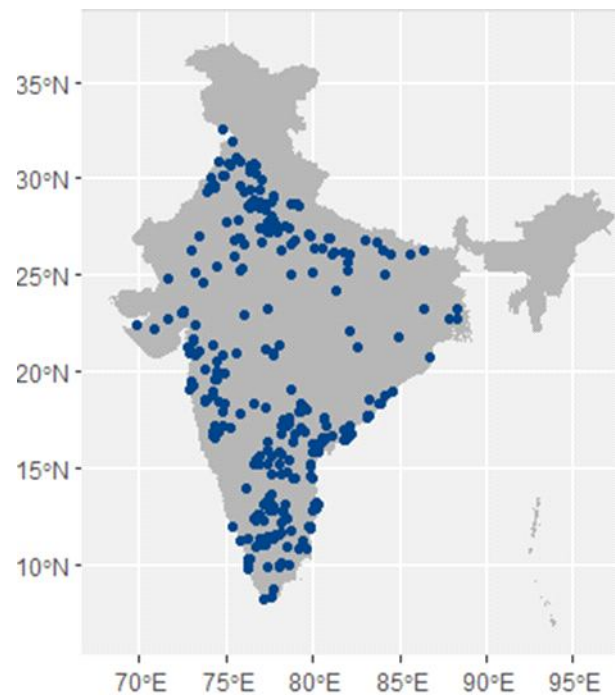
### Data pre-processing

Duplicate records of the occurrence data of *T. portulacastrum* and *A. conyzoides* were omitted using the conditional formatting function available in MS Excel and finally, 375 occurrence points of *T. portulacastrum* and 379 points of *A. conyzoides* were retained for model building and validation for the species distribution modelling. Predictor variables from bio1 through bio19 were tested for multicollinearity using correlogram. Out of 19 bioclimatic variables, strongly associated bioclimatic variables with Pearson correlation coefficient values  $>0.8$  and  $<-0.8$  were eliminated from the analysis, leaving eight variables with the least mutual association across the study area. With this criteria, 8 bioclimatic variables, as well as the elevation layer and soil layer, were retained for model development and validation. Thus, selected climatic variables are bio 1 (Annual Mean Temperature), bio 4 (Temperature Seasonality (standard deviation \*100)), bio 5 (Maximum Temperature of Warmest Month), bio 6 (Minimum Temperature of Coldest Month), bio 9 (Mean Temperature of Driest Quarter), bio 10 (Mean Temperature of Warmest Quarter), bio 12 (Annual Precipitation), bio 14 (Precipitation of Driest Month).

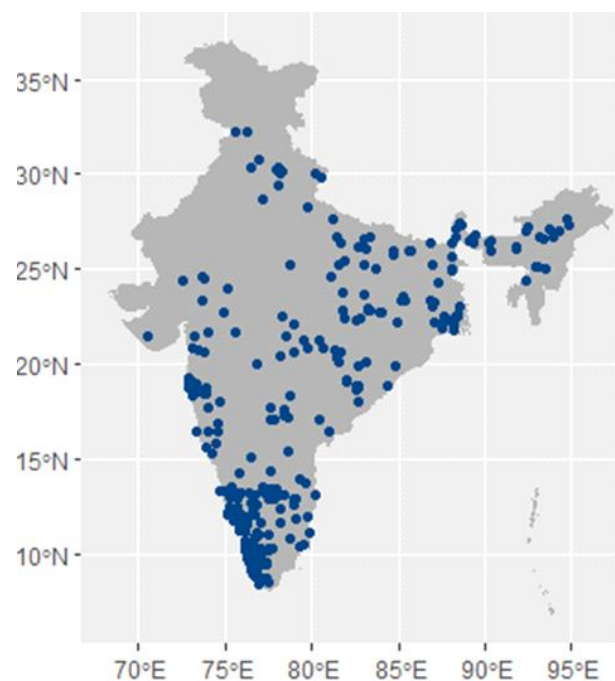
Raster layers of environmental variables were trimmed to get the study region's (India) spatial range (latitudes 8°42' N and 37°6' N, and longitudes 68°7' E and 97°25' E, with a total area of 3,287,263 km<sup>2</sup>). Maps in **Figure 1 and 2** depict the occurrence points of *T. portulacastrum* and *A. conyzoides* in the study region of India, respectively.

### Model fitting, calibration and validation

The geographical distribution of *T. portulacastrum* and *A. conyzoides* was obtained with the BIOMOD2 package available in R. Unlike the single model technique, BIOMOD2 incorporates an Ensemble modelling algorithm and is thought to produce higher accuracy. Ensemble modelling is a process where multiple diverse models are generated to predict an outcome, by using many different modelling algorithms. The Ensemble model then integrates the prediction of each base model and results in one final prediction. The purpose of using Ensemble models is to minimize the generalization error of the prediction.



**Figure 1.** Map depicting the occurrence points of *T. portulacastrum* in the study region



**Figure 2.** Map depicting the occurrence points of *A. conyzoides* in the study region

For generating the Ensemble model, five algorithms from the BIOMOD2 package in R were used. These include two regression approaches (Generalised Linear Model (GLM), Multivariate Adaptive Regression Splines (MARS)) and three machine learning methods (Maximum Entropy (MaxEnt); Artificial Neural Network (ANN) and Random Forest (RF)).

Similar to species presence points, absence points are regarded as valuable data in SDM. SDM algorithms and model assessment procedures consider missing data to be an essential factor. Because there were no actual absence points rather pseudo-absence points were constructed within the model using a random technique. Based on the BIOMOD2 package's guidance, three times as many presence points were generated as pseudo absence points for modelling the weeds *T. portulacastrum* and *A. conyzoides*. The data was divided into two sets: 75% of the data was utilized for the calibration procedure, and the remaining 25% was used as testing data set. A ten-fold cross-validation strategy was used to lower the uncertainty in the response curves and occurrence predictions. The method involves dividing the occurrence data into 10 equal-size groups, or "folds," at random. Models are then created by excluding each fold in turn, with the left-out fold being used for model validation. It uses all of the data for validation as a result. The predicted probability layers were then averaged to get the final model output.

### Model evaluation

There are different ways for assessing the model's predicted performance. For evaluation, many researchers employed Area under Receiver Operating Characteristic (ROC) curve (AUC) and True Skill Statistic (TSS). Both methods can be used alone, but it is best to employ them all for cross-comparisons. Here, TSS was used to evaluate model predictive performance, while AUC of the ROC and Kappa were utilised to crosscheck model predictive performance. Models with  $AUC > 0.7$ ,  $TSS > 0.4$  and  $Kappa > 0.4$  were used for building the Ensemble model.

The area under ROC curve (AUC) indicates the degree or measure of separability, whereas ROC is a probability curve between the True positive rate (TPR) and False positive rate (FPR). It takes values from 0 to 1, where a value of 0 indicates a perfectly inaccurate test and a value of 1 reflects a perfectly accurate test. In general, an AUC of 0.5 suggests no discrimination, 0.7 to 0.8 is considered acceptable, 0.8 to 0.9 is considered excellent, and more than 0.9 is considered outstanding.

True Skill Statistic (TSS) describes the ability of a model to correctly classify presence and background data.  $TSS > 0.8$  excellent;  $0.6 < TSS < 0.8$  good;  $0.4 < TSS < 0.6$  fair;  $0.2 < TSS < 0.4$  poor; and  $TSS < 0.2$  no predictive ability.

The Kappa statistic is frequently used to test interrater reliability. The importance of interrater

reliability lies in the fact that it represents the extent to which the data collected in the study are correct representations of the variables measured. It ranges from -1 to 1. Values  $d = 0$  as indicating no agreement and 0 - 0.20 as none to slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement.

Using TSS scores as a cut-off value, maps depicting the probabilities of suitable and unsuitable areas of current and future distribution of *T. portulacastrum* and *A. conyzoides* was obtained and prepared in four classes of climate suitability i.e., not suitable (0-0.2), low (0.2-0.4), moderately (0.4-0.6) and highly suitable ( $>0.6$ ) categories.

Apart from this, the permutation approach was used to analyse the relative influence of each climate variable on the distribution of selected plant species. To demonstrate the relationship between the probabilities of occurrence for a species with varying values of environmental variables, response curves were also created. One environmental variable is simulated for each plot, with the remaining environmental variables kept constant at their mean.

The biomod range size function in the BIOMOD2 package was used to visualise and measure the range change of the target plant species under future climatic circumstances. The function offers a summary statistic on species range change. "Percentage loss" (i.e., the percentage of currently suitable areas predicted to be lost, calculated as  $[\text{loss}/(\text{loss} + \text{stable})]$ ); "percentage gain" (i.e., the percentage of new habitats predicted to be suitable when compared to the species' current distribution size, calculated as  $[\text{gain}/(\text{gain} + \text{stable})]$ ); and "range change," i.e., the overall output of predictions, calculated as (percentage gain-percentage loss).

## RESULTS AND DISCUSSION

### Variable selection

In **Figure 3**, the correlations for all pairs of variables are represented by the correlogram. Positive correlations are shown in blue, while negative correlations are shown in red. The intensity of the colour is proportional to the correlation coefficient, therefore, the darker the circle, the stronger the correlation (i.e., closer to -1 or 1). The correlation coefficients and related hues are shown in the colour legend on the right side of the correlogram. Strongly connected bioclimatic variables with Pearson correlation coefficient values  $>0.8$  and  $<-0.8$  were removed from the analysis, leaving eight variables with the least mutual association across the study area.

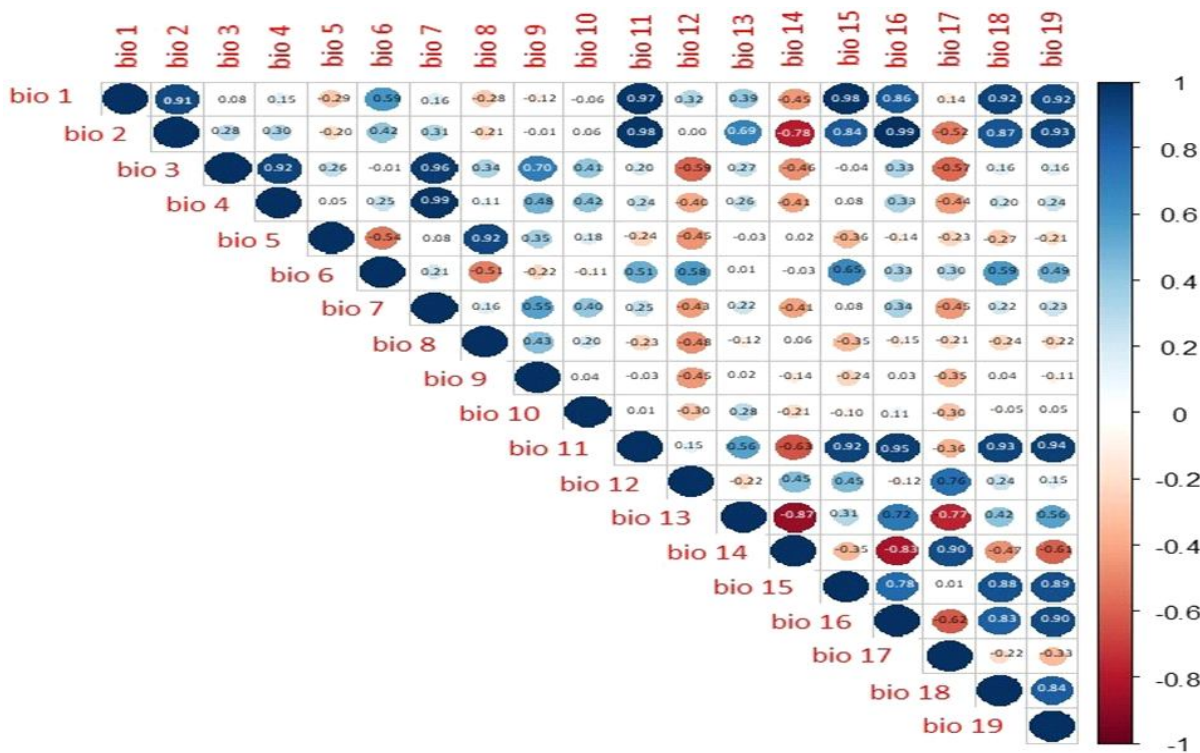


Figure 3. Correlogram- Multicollinearity among environment predictors

The bioclimatic variables that were retained are:

- bio 1 (Annual Mean Temperature)
- bio 4 (Temperature Seasonality (standard deviation \*100))
- bio 5 (Max Temperature of Warmest Month)
- bio 6 (Min Temperature of Coldest Month)
- bio 9 (Mean Temperature of Driest Quarter)
- bio 10 (Mean Temperature of Warmest Quarter)
- bio 12 (Annual Precipitation)
- bio 14 (Precipitation of Driest Month)

Along with eight bioclimatic variables, elevation layer and soil layer were also used for model building and validation.

**Model building and validation for *Trianthema portulacastrum***

Performance score of all models along with Ensemble approach are given in **Table 1** for comparison. From **Table 1**, it is clear that the scores of individual model based on AUC value, ranges from 0.755 (ANN) to 0.849 (RF), whereas TSS value, ranges from 0.403 (RF) to 0.551 (MARS), and KAPPA value from 0.421 (MaxEnt) to 0.510 (GLM). The best performing model based on AUC is RF, whereas the TSS score is high in case of MARS, and the high KAPPA statistic in case of GLM. All the single models outperformed the random models. However, the Ensemble model gave higher AUC value (0.872), TSS value (0.606) and Kappa statistic (0.547) as compared to other models.

**Table 1. Performance score of individual models along with Ensemble model for *Trianthema portulacastrum***

| MODEL    | AUC   | TSS   | KAPPA statistic |
|----------|-------|-------|-----------------|
| MaxEnt   | 0.818 | 0.468 | 0.421           |
| RF       | 0.849 | 0.403 | 0.439           |
| ANN      | 0.755 | 0.496 | 0.422           |
| MARS     | 0.846 | 0.551 | 0.475           |
| GLM      | 0.832 | 0.537 | 0.510           |
| ENSEMBLE | 0.872 | 0.606 | 0.547           |

**Figure 4** shows the response curves of environmental predictors affecting the prediction of Ensemble model of *Trianthema portulacastrum*. As shown in **Table 2**, the three most important variables influencing the model are bio 4 (temperature seasonality), bio 12 (annual precipitation), and bio 6 (minimum temperature of the coldest month). Therefore, they are discussed here. It is clear from the curve that for the growth of *T. portulacastrum*, it needs a temperature of at least 5 °C. There is a higher chance of weed occurrence when temperatures seasonality is above 450. Weed occurrence is quite probable in areas with annual precipitation of approximately 950 mm.

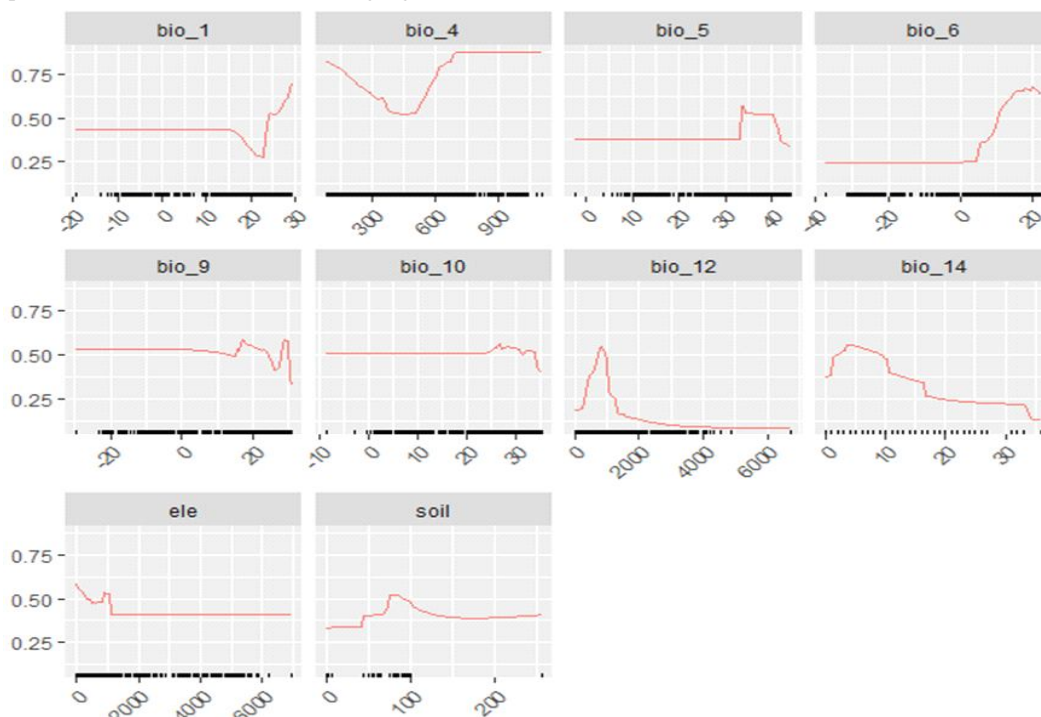
**Current distribution pattern of *T. portulacastrum***

As shown in **Figure. 5**, major parts of Punjab, Haryana, and the north-western part of Uttar Pradesh are classified as highly suitable in the north zone; states such as Tamil Nadu, Karnataka, Telangana, and

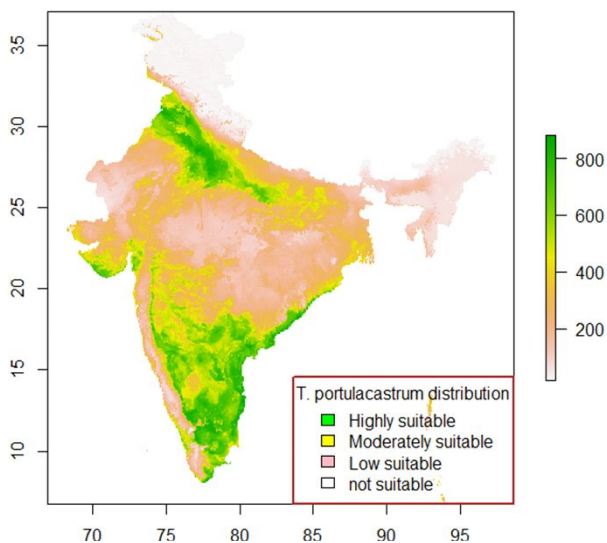
**Table 2. Performance of predictor variables used in species distribution modelling of *Trianthema portulacastrum***

| MODEL     | bio 1         | bio 4         | bio 5         | bio 6         | bio 9  | bio 10        | bio 12        | bio 14        | Elevation     | Soil   |
|-----------|---------------|---------------|---------------|---------------|--------|---------------|---------------|---------------|---------------|--------|
| MaxEnt    | <b>0.4976</b> | 0.1930        | <b>0.3113</b> | 0.0501        | 0.1293 | 0.1377        | <b>0.3502</b> | 0.0647        | 0.1854        | 0.0385 |
| RF        | 0.0493        | <b>0.1288</b> | 0.0596        | <b>0.0787</b> | 0.0622 | 0.0373        | <b>0.1081</b> | 0.0217        | 0.0551        | 0.0102 |
| ANN       | 0.2209        | 0.0680        | 0.1063        | 0.2254        | 0.0436 | 0.1402        | <b>0.7612</b> | <b>0.2387</b> | <b>0.3723</b> | 0.2089 |
| MARS      | 0.0000        | <b>0.5015</b> | 0.0659        | <b>0.7478</b> | 0.1686 | 0.0000        | <b>0.3750</b> | 0.0558        | 0.0517        | 0.0000 |
| GLM       | 0.2732        | <b>0.4995</b> | 0.0000        | <b>0.8409</b> | 0.0000 | <b>0.5926</b> | 0.1539        | 0.0784        | 0.0000        | 0.0140 |
| ENSEM-BLE | 0.0799        | <b>0.2718</b> | 0.0542        | <b>0.3861</b> | 0.0334 | 0.0438        | <b>0.2932</b> | 0.0450        | 0.0312        | 0.0249 |

The most important variables in each model are highlighted in bold



**Figure 4. Response curves of environmental predictors affecting the prediction of Ensemble model for *Trianthema portulacastrum***

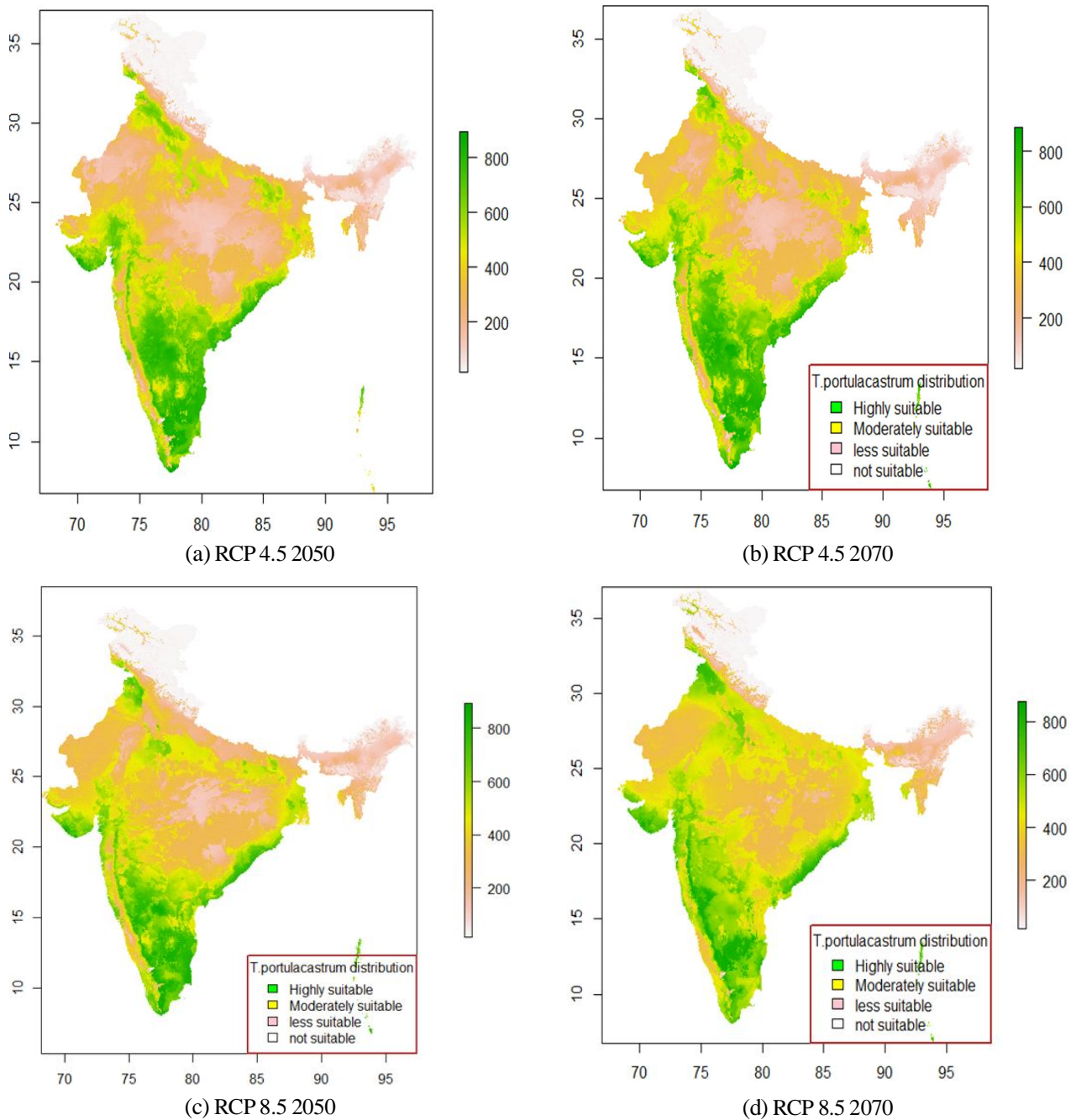


**Figure 5. Predicted distribution of *Trianthema portulacastrum* in India under current climate conditions based on Ensemble model**

Andhra Pradesh are classified as highly suitable in south zone; and, some parts of western Maharashtra and coastal areas of Gujarat are classified as highly suitable in the western zone of India for the growth of *T. portulacastrum*.

**Future distribution pattern of *T. portulacastrum***

From **Figure 6** and **Table 3**, it can be found that *T. portulacastrum* will undergo significant range changes under all the future climatic scenarios with a gain in habitat suitability. Mandal *et al.* (2017) also found that *T. portulacastrum* had high acclimatization capacity and produced more growth under elevated temperature up to ambient+4°C. In the present study, it is observed that invasion hotspots for *T. portulacastrum* will be shifted towards the south and east of the country including a few western parts like Gujarat, in all four future climatic scenarios.



**Figure 6. Predicted suitable regions for *Trianthema portulacastrum* in India under future climatic scenarios based on Ensemble model**

**Table 3. Summary of the range change statistics for the *Trianthema portulacastrum* under different climate change scenarios compared to current climatic conditions**

| Scenario     | Loss % | Gain % | Change % |
|--------------|--------|--------|----------|
| RCP 4.5 2050 | 18.90  | 53.70  | 34.80    |
| RCP 4.5 2070 | 18.32  | 61.22  | 42.90    |
| RCP 8.5 2050 | 22.50  | 50.88  | 28.38    |
| RCP 8.5 2070 | 20.05  | 74.03  | 53.97    |

**Model building and validation for *A. conyzoides***

Performance score of all models along with Ensemble approach are given in **Table 4** for comparison. It is clear that the scores of individual models based on AUC value ranges from 0.774 (ANN) to 0.837 (MARS), whereas TSS value ranges from 0.441 (MaxEnt) to 0.491 (GLM) and Kappa value from 0.445 (MaxEnt) to 0.512 (RF). The best-performing model based on AUC is MARS, whereas the TSS score is high in case of GLM, and Kappa statistic in the case of RF. All of the single models outperformed the

**Table 4. Performance score of individual model along with Ensemble model for *Ageratum conyzoides***

| MODEL          | AUC          | TSS          | Kappa statistic |
|----------------|--------------|--------------|-----------------|
| MaxEnt         | 0.809        | 0.441        | 0.445           |
| RF             | 0.832        | 0.462        | <b>0.512</b>    |
| ANN            | 0.774        | 0.462        | 0.449           |
| MARS           | <b>0.837</b> | 0.483        | 0.510           |
| GLM            | 0.825        | <b>0.491</b> | 0.479           |
| ENSEMBLE MODEL | <b>0.840</b> | <b>0.522</b> | <b>0.525</b>    |

random models. However, Ensemble model gave higher AUC value (0.840), TSS (0.522) and Kappa statistic (0.525) as compared to other models.

**Table 5** displays the importance of predictor variables used in *A. conyzoides* distribution modelling for each model. The model is predominantly influenced by predictor variables such as bio 12 (Annual Precipitation), bio 4 (Temperature Seasonality), and bio 6 (Minimum Temperature of the Coldest Month).

**Figure 7** revealed that the *A. conyzoides* is not a cold region weed, as its occurrence probability is 0 below 0°C. Arora (1999) found that places with temperatures ranging from 20–25 °C are best suited to its growth and development but it also survives well at 15–30°C. This explains its occurrence at higher altitudes (i.e., temperate climates) as well as on the plains (i.e., tropical climates) (Kosaka *et al.* 2010). Response curves show that as temperature seasonality increases, its occurrence probability decreases, indicating it is not tolerant to extreme temperature variability. It indicates that places with less difference between highest and lowest temperature are favourable for this weed. Annual precipitation of 2000–3700 mm has a high likelihood of weed occurrence.

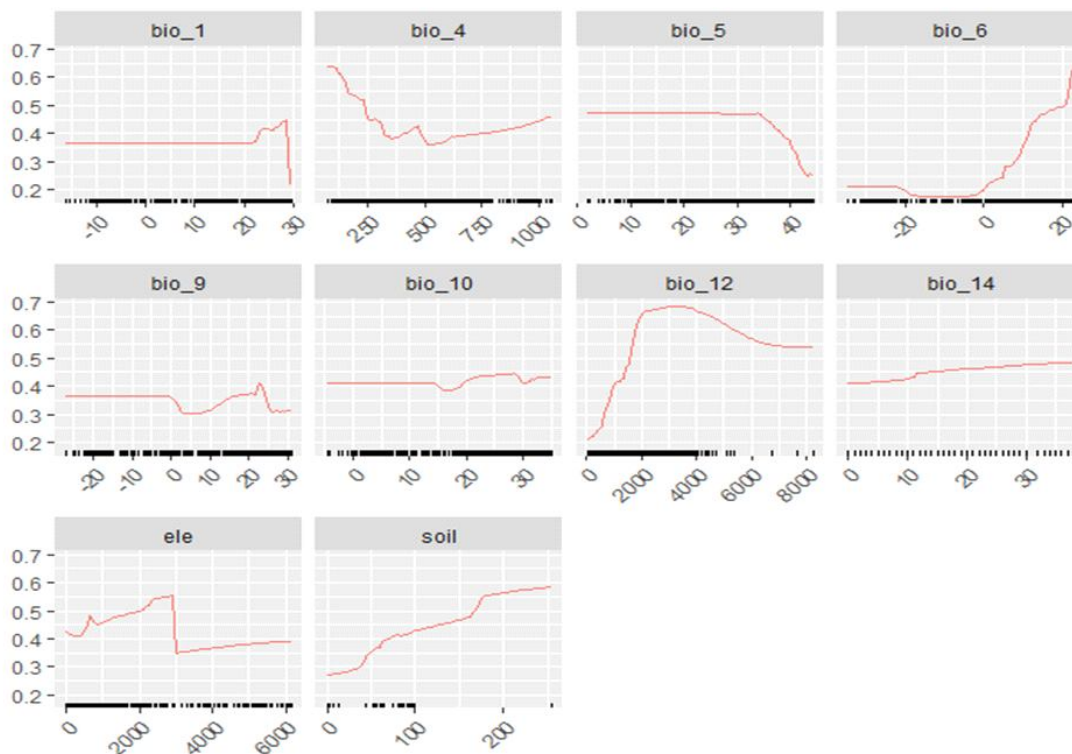
**Current distribution pattern of *A. conyzoides***

According to **Figure 8**, the southern part of Uttarakhand is classified as a moderately suitable area

**Table 5. Performance of predictor variables used in species distribution modelling of *Ageratum conyzoides***

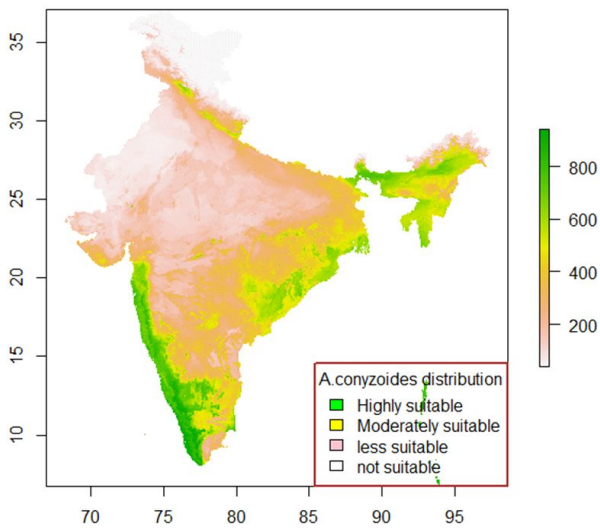
| MODEL    | bio 1  | bio 4         | bio 5         | bio 6         | bio 9         | bio 10 | bio 12        | bio 14 | Elevation | Soil   |
|----------|--------|---------------|---------------|---------------|---------------|--------|---------------|--------|-----------|--------|
| MaxEnt   | 0.0524 | <b>0.3141</b> | 0.0184        | <b>0.1279</b> | 0             | 0      | <b>0.1638</b> | 0.0038 | 0.0069    | 0.0097 |
| RF       | 0.0308 | <b>0.1345</b> | <b>0.0570</b> | 0.0465        | 0.0368        | 0.0407 | <b>0.1123</b> | 0.0167 | 0.0307    | 0.0094 |
| ANN      | 0.0890 | <b>0.5691</b> | <b>0.0934</b> | 0.0424        | 0.0645        | 0.0485 | <b>0.3875</b> | 0.0207 | 0.0854    | 0.0660 |
| MARS     | 0.0508 | <b>0.2687</b> | <b>0.4432</b> | <b>0.9212</b> | 0.0518        | 0      | 0.0813        | 0      | 0.2249    | 0      |
| GLM      | 0      | <b>0.3506</b> | 0             | 0             | <b>0.2723</b> | 0      | <b>0.1749</b> | 0.0084 | 0.0544    | 0.0508 |
| ENSEMBLE | 0.0163 | <b>0.1283</b> | 0.0543        | <b>0.1093</b> | 0.0241        | 0.0050 | <b>0.1632</b> | 0.0034 | 0.0276    | 0.0138 |

The most important variables in each model are highlighted in bold.



**Figure 7. Response curves of environmental predictors affecting the prediction of Ensemble model for *Ageratum conyzoides***



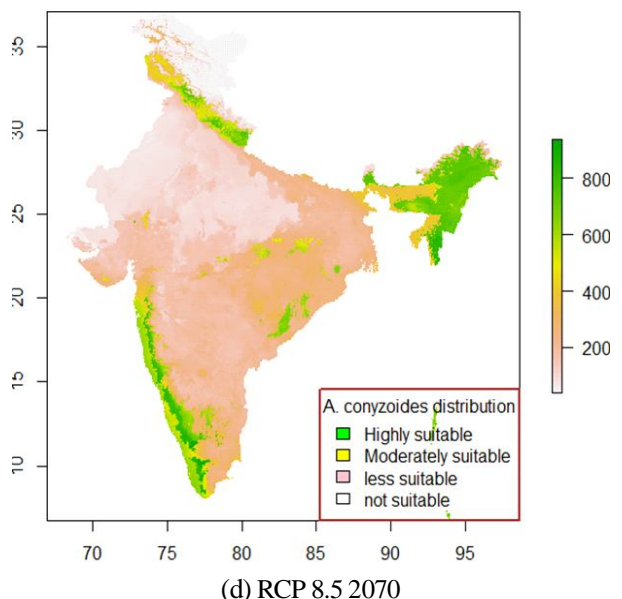
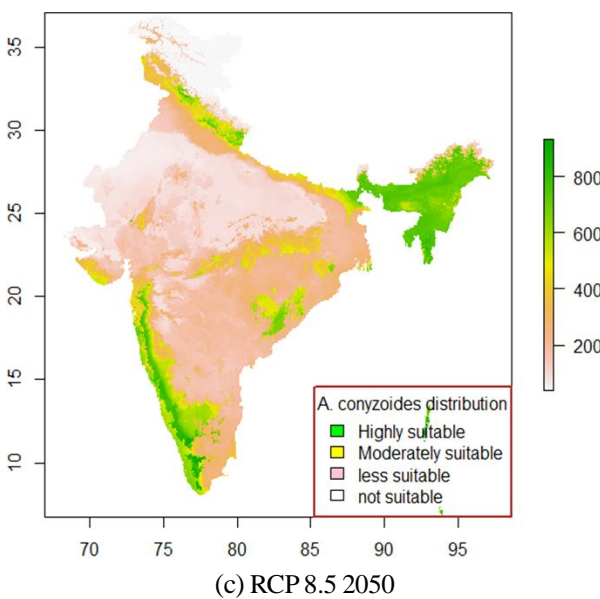
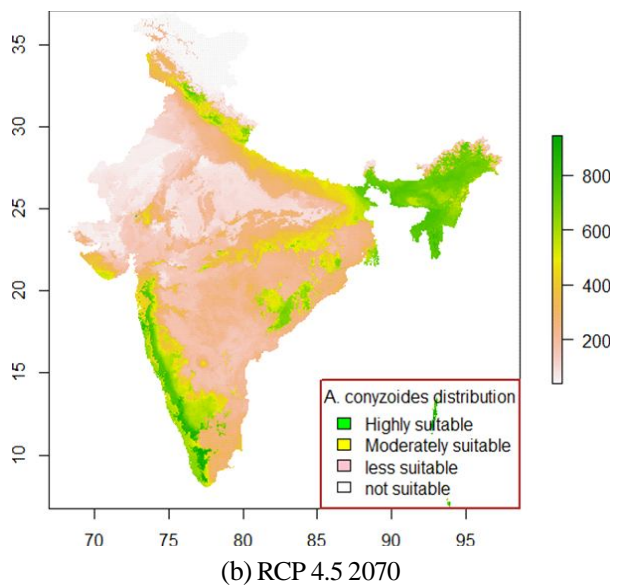
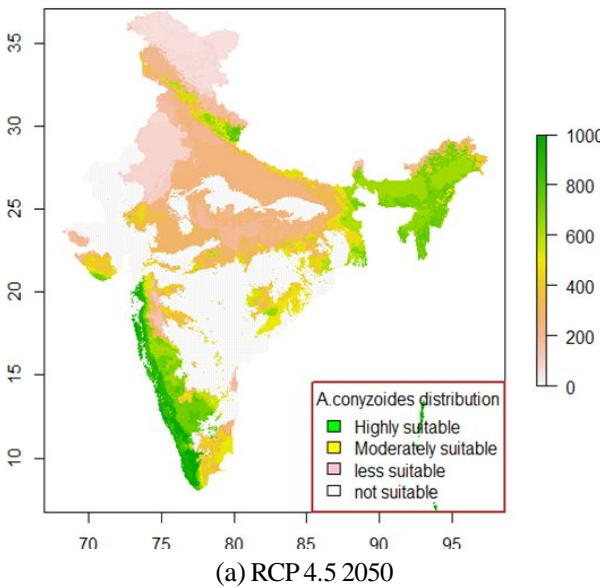


**Figure 8.** Predicted suitable regions of *Ageratum conyzoides* in India under current climate conditions based on Ensemble model

in the north zone; in the south zone, Kerala, coastal Karnataka, and the northern portion of Andhra Pradesh are classified as moderately suitable, and in the eastern zone of India, the coastal regions of Odisha and West Bengal are classified as highly suitable for the growth of *A. conyzoides*. In India’s north-eastern region, all the states, i.e., Assam, Sikkim, Nagaland, Meghalaya, Manipur, Mizoram, and Tripura, come under the moderately to a highly suitable category, except Arunachal Pradesh, which comes under the less suitable category.

**Future distribution pattern of *A. conyzoides***

From **Figure 9** and **Table 6**, it can be seen that *A. conyzoides* will not undergo significant range changes under RCP 4.5 for 2050 and 2070 but will experience significant changes under RCP 8.5 2050 and 2070 predominantly governed by a reduction in habitat



**Figure 9.** Predicted suitable regions *Ageratum conyzoides* in India under future climatic scenarios based on ensemble model

**Table 6. Summary of the range change statistics for the *Ageratum conyzoides* under different climate change scenarios compared to current climatic conditions**

| Scenario     | Loss % | Gain % | Change % |
|--------------|--------|--------|----------|
| RCP 4.5 2050 | 27.211 | 29.096 | 1.885    |
| RCP 4.5 2070 | 31.128 | 28.32  | -2.808   |
| RCP 8.5 2050 | 36.342 | 17.275 | -19.067  |
| RCP 8.5 2070 | 57.397 | 15.668 | -41.73   |

suitability for weeds. It is observed that invasion hotspots for *A. conyzoides* will be completely shifted towards north- eastern side of the country.

### Conclusion

The study highlights the potential impact of climate change on the distribution of *Trianthema portulacastrum* and *Ageratum conyzoides* in India under present and future climatic conditions. The significant contribution of bioclimatic variables such as temperature seasonality, annual mean temperature, and minimum temperature of the coldest month was found to govern the potential distribution of *T. portulacastrum* and *A. conyzoides*. Suitable regions for *T. portulacastrum* are predicted to increase under RCP 4.5 and 8.5 for the years 2050 and 2070, whereas for *Ageratum conyzoides*, suitable regions in India will not undergo significant range change under RCP 4.5 but will decrease under RCP 8.5 for the years 2050 and 2070. This study can aid in the management of weeds in the potentially identified areas as hotspots in climate change, and the findings could be utilised as a preventative strategy to establish early detection and rapid reaction or to develop one if none currently exists. Ensemble modelling was used here and is certainly reliable in estimating species distributions, and it is useful in biodiversity planning and management.

### REFERENCES

- Annual Reports (2010–2022). AICRP-Weed Management, ICAR-Directorate of Weed Research, Jabalpur.
- Arora V. 1999. *Comparative assessment of eco-physiological functions between Ageratum conyzoides L. and Parthenium hysterophorus L.* Ph.D. Thesis, Panjab University, Chandigarh, India.
- Banerjee A, Mukherjee A, Dewanji A. 2017. Potential distribution of *Mikania micrantha* Kunth in India “evidence of climatic niche and biome shifts. *Flora* **234**: 215–223. 10.1016/j.flora.2017.08.001.
- Bhan VM, kumar S, Raghuwanshi MS. 1999. Weed Management in India. *Indian Journal of Plant Protection* **17**: 171–202.
- Centre for Agriculture and Bioscience International (CABI; www.cabi.org)
- Eliith J, Graham C, Anderson R, Dudík M, Ferrier S, Guisan A, Hijmans R, Huettmann F, Leathwick J, Lehmann A, Li J, Lohmann L, Loiselle B, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton J, Peterson A, Zimmermann N. 2006. Novel methods improve prediction of species’ distributions from occurrence data. *Ecography* **29**: 129–151.
- GBIF.org (2 July 2023) GBIF Occurrence Download <https://doi.org/10.15468/dl.ppz7xg>.
- GBIF.org (3 July 2023) GBIF Occurrence Download <https://doi.org/10.15468/dl.hqepwn>.
- Gharde Y, Dubey RP, Singh PK, Mishra JS. 2023. Little seed canary grass (*Phalaris minor* Retz.) a major weed of rice-wheat system in India is predicted to experience range contraction under future climate. *International Journal of Pest Management* DOI **10**:1080/09670874.2023.2199258.
- Gharde Y, Singh PK, Dubey RP, Gupta PK. 2018. Assessment of yield and economic losses in agriculture due to weeds in India. *Crop Protection* **107**: 12–18.
- Gharde Y, Sushilkumar, Sharma AR. 2019. Exploring models to predict the establishment of the leaf-feeding beetle *Zygogramma bicolorata* (Coleoptera: Chrysomelidae) for the management of *Parthenium hysterophorus* (Asteraceae: Heliantheae) in India. *Crop Protection* **122**:10.1016/j.cropro.2019.04.014.
- Guisan A, Zimmermann N. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling. Ecological Modelling* **135**: 147–186.
- IPCC. 2014. Climate change 2014. Synthesis Report.
- Kaur M, Aggarwal NK. 2017. *Trianthema portulacastrum* L. The Noxious weed and its control. *Advances in Plants & Agriculture Research* **6**(3): 213–218.
- Kaur S, Singh H, Batish D, Kohli R. 2012. Phytotoxicity of decomposing below-ground residues of *Ageratum conyzoides*: Nature and dynamics of release of phytotoxins. *Acta Physiologiae Plantarum* **34**: 10.1007/s11738-011-0904-8.
- Kohli RK, Batish DR, Singh HP. 2006. Status, invasiveness and environmental threats of three tropical American invasive weeds (*Parthenium hysterophorus*, *Ageratum conyzoides* L., *Lantana camara* L.) in India. *Biological Invasions* **8**: 1501–1510.
- Kosaka Y, Saikia B, Mingki T, Tag H, Riba T, Ando K. 2010. Roadside Distribution Patterns of Invasive Alien Plants Along an Altitudinal Gradient in Arunachal Himalaya, India. *Source: Mountain Research and Development* **30**: 252–258. 10.1659/MRD-JOURNAL-D-10-00036.1
- Mandal AK, Dheebakaran G, Banik M, Kumar A and Kumar V. 2017. Impact of climate change on adaptation capacity of Horse purslane (*Trianthema portulacastrum*). *The Pharma Innovation Journal* **6**(11): 414–418.
- Priyanka N, Shiju MV, Joshi PK. 2013. A framework for management of *Lantana camara* in India. *Proceedings of the International Academy of Ecology and Environmental Sciences* **3**: 306–323.
- Rai PK and Singh JS. 2020. Invasive alien plant species: Their impact on environment, ecosystem services and human health. *Ecological Indicators* **111**: 106020.
- Ramesh K, Matloob A, Aslam F, Florentine SK, Chauhan BS. 2017. Weeds in a Changing Climate: Vulnerabilities, Consequences, and Implications for Future Weed Management. *Frontiers in Plant Science* **8**: 10.3389/fpls.2017.00095
- Weed Atlas Vol 1 & II (2008). Directorate of Weed Science Research, Jabalpur.