



OPEN

# Global spatial distribution of *Chromolaena odorata* habitat under climate change: random forest modeling of one of the 100 worst invasive alien species

Pradeep Adhikari<sup>1,6</sup>, Yong Ho Lee<sup>1,2,6</sup>, Anil Poudel<sup>3</sup>, Sun Hee Hong<sup>3</sup>✉ & Yong-Soon Park<sup>4,5</sup>✉

Anthropogenic activities and global climate change increase the risk of *Chromolaena odorata* invasion and habitat expansion. To predict its global distribution and habitat suitability under climate change, a random forest (RF) model was employed. The RF model, utilizing default parameters, analyzed species presence data and background information. The model revealed that the current spatial distribution of *C. odorata* covers 7,892,447 km<sup>2</sup>. Predictions for 2061–2080 indicate expansion of suitable habitat (42.59 and 46.30%), reduction of suitable habitat (12.92 and 12.20%), and preservation of suitable habitat (87.08 and 87.80%) under the SSP (Shared Socio-economic Pathway) 2-4.5 and SSP5-8.5 scenarios, respectively, in comparison to the present distribution. Currently, *C. odorata* is predominantly found in South America, with limited presence in other continents. However, the data suggest that climate change will elevate the global invasion risk of *C. odorata* worldwide, particularly in Oceania, Africa, and Australia. Countries such as Gambia, Guinea-Bissau, and Lesotho, which currently have unsuitable habitats, are predicted to have highly suitable habitats with climate change, supporting the idea that global habitat expansion for *C. odorata* will occur due to climate change. This study indicates that proper management of *C. odorata* is crucial during the early invasion phase.

*Chromolaena odorata* is a perennial shrub in the Asteraceae family, indigenous to the tropical and subtropical regions of Central and South America, encompassing Mexico, the Caribbean, and Brazil<sup>1</sup>. *C. odorata* is a herbaceous to woody plant and can grow up to 2 m in height<sup>1,2</sup>. It reproduces both vegetatively and sexually, with high fecundity and rapid germination<sup>1</sup>. *C. odorata* can grow in various types of soils, ranging from sand dunes to heavy clays, and the seeds can survive in the soil for up to 6 years<sup>3</sup>. *C. odorata* can tolerate acidic to neutral soil and establishes well in disturbed areas, grassland, forest, and fallow areas<sup>1</sup>. Because it has both short- and long-distance dispersal and can spread quickly, *C. odorata* is considered one of the world's most invasive weeds<sup>2,4</sup>.

*C. odorata* is recognized as a very serious weed in all types of tropical perennial crops, e.g., coffee, cocoa, and citrus<sup>5</sup>. *C. odorata* transmits pathogenic fungi<sup>6</sup> and acts as a host for insect pests such as *Zonecerus variegatus*<sup>7</sup>, whose nymphs feed on the foliage of native plants<sup>1</sup>, resulting in economic loss in invaded regions. *C. odorata* was introduced to India in the 1800s as an ornamental plant; it spread to South East Asia by 1920<sup>1</sup>, West Africa by 1937, East Timor by 1975<sup>8</sup>, and Australia by 1994, through contaminated pasture and forest seeds<sup>1</sup>. It has since spread throughout Africa, Oceania, and South and South East Asia as a result of anthropogenic activities, including international trade and tourism, transportation, and land-use change, as well as natural phenomena, such as winds, tides, surface runoff, and the movement of animals<sup>3</sup>.

*C. odorata* can grow in areas with ~1000 mm annual rainfall to >3000 mm annual rainfall and cold temperature from 0 to >35 °C<sup>1</sup>. However, it prefers tropical climates, such as tropical rainforest, tropical monsoon, and

<sup>1</sup>Institute of Humanities and Ecology Consensus Resilience Lab, Hankyong National University, Anseong 17579, Republic of Korea. <sup>2</sup>OJeong Resilience Institute, Korea University, Seoul 02841, Republic of Korea. <sup>3</sup>School of Plant Science and Landscape Architecture, College of Agriculture and Life Sciences, Hankyong National University, Anseong 17579, Republic of Korea. <sup>4</sup>Department of Plant Resources, College of Industrial Sciences, Kongju National University, Yesan 32439, Republic of Korea. <sup>5</sup>Agricultural and Fisheries Life Science Research Institute, College of Industrial Sciences, Kongju National University, Yesan 32439, Republic of Korea. <sup>6</sup>These authors contributed equally: Pradeep Adhikari and Yong Ho Lee. ✉email: shhong@hku.ac.kr; yspark2005@kongju.ac.kr

warm temperate climates<sup>1,9</sup>. Because *C. odorata* can survive in a wide range of environments including regions with high temperature and rainfall, it is likely to become highly abundant under future climate change, although semiarid, Mediterranean, and cold temperate climates are considered safe from invasion<sup>9</sup>. *C. odorata* can grow when days are short and bloom in the winter<sup>10</sup>, enhancing its ability to dominate neighboring flora and establish new habitats<sup>11</sup>. Thus, to prevent invasion and conserve native biodiversity, it is essential to understand how *C. odorata* will colonize, establish, and spread under climate change. Estimating the potential invasion risk of *C. odorata* for different regions and different climate change scenarios will enable us to focus efforts toward solid quarantine, early detection, and eradication of the species at national and international levels. Therefore, ecological niche modeling (ENM) approach has been introduced to predict alien and invasive<sup>12</sup>.

The ENM is widely used in the fields of ecology and conservation. The ENM combines environmental variables and species-occurrence records to predict the spatial distribution of species<sup>13</sup>. More recently, ENM has been extensively used to assess the geographical distribution and invasion risk of alien species in order to develop appropriate conservation strategies<sup>12–15</sup>. Various modeling techniques are used to fit ENMs<sup>16</sup>. Earlier studies mostly used regression-based models<sup>17</sup>, but attention has since turned toward algorithms that use machine learning (ML) techniques<sup>18</sup>, which have several advantages. ML techniques have more flexible fitting functions, can handle different types of data, have options to automatically select variables, and have higher predictive performance than other techniques<sup>18,19</sup>. Among ENMs, random forest (RF) modeling is a novel and popular approach to ecological mapping that has high predictive performance<sup>15</sup>. The RF models can make sound predictions with little parameter tuning, can handle thousands of input and correlated variables; can readily assess the importance of each variable; and are robust, creating several decision trees from the majority vote<sup>20–22</sup>.

Our previous study determined that an RF model produced the most accurate predictions for invasive plants in South Korea among five ENMs tested (a generalized linear model, multivariate adaptive regression splines, an artificial neural network, maximum entropy, and an RF model)<sup>15</sup>. Most of our previous studies focused on invasive species that currently exist in South Korea<sup>23–25</sup>, although one study showed that future climate change will likely result in suitable habitats for a new invasive species (*Parthenium hysterophorus*)<sup>26</sup>. All of the studies concluded that invasive species will pose a significant threat to South Korea in the future because of climate change. Given the qualitative features of RF models, and following the recent use of ENMs to predict invasive species<sup>14,27</sup>, we decided to use RF modeling to predict the spatial distribution of *C. odorata* at the global scale. To our knowledge, no studies have yet attempted to reveal the global invasion risk of *C. odorata*.

We collected global occurrence data for *C. odorata* and designed this study with the following main objectives: (1) to predict the current and future spatial distribution of *C. odorata* across the world using the RF algorithm, (2) to estimate the potential habitat change for *C. odorata* under future climate change, and (3) to classify the *C. odorata* habitat suitability in different countries of the world. The results of this study will help us to understand how the current distribution of *C. odorata* and suitable habitat will change in the future in different countries. From these results, management plans for each country (and the world) can be prepared to manage future *C. odorata* habitat expansion.

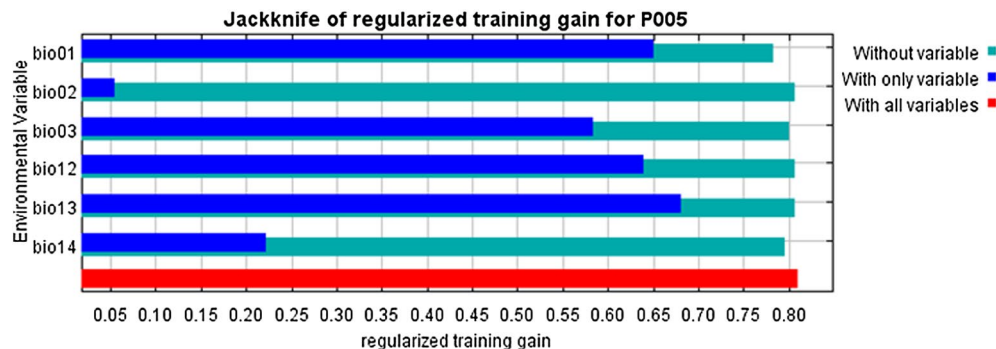
Results

**Selection of bioclimatic variables and their contribution to the model.** We downloaded data for 19 bioclimatic variables from the WorldClim database and performed Spearman's correlation test (Table S1) to select the most important variables for global spatial RF modeling of *C. odorata*. Six bioclimatic variables (annual mean temperature [Bio01], mean diurnal temperature range [Bio2], isothermality [Bio03], annual precipitation [Bio12], precipitation in the wettest month [Bio13], and precipitation in the driest month [Bio 14]; Table 1) were selected on the basis of their weak correlation with each other ( $r < 0.75$ ). These variables were considered to be the most significant factors for predicting the global distribution of *C. odorata*.

To understand the importance of the different variables in modeling the *C. odorata* distribution, we estimated the average contribution of each of the six variables over the historical data (1970–2000) are hereafter referred to as the ‘current climate’ and future (2061–2080) time periods. Among the six variables, Bio13, Bio3, and Bio1 had relatively high contributions to the model, estimated as 69.58, 23.75, and 12.59%, respectively (Table 1). This indicates a prominent role for these three variables in the distribution of *C. odorata*. The other variables had distinctly lower values and thus play more minor roles in the model. Similarly, the relative importance of each variable was assessed by Jackknife test (Fig. 1). The variables Bio13, Bio12, Bio1, and Bio3 show relatively higher importance in the model as compared to other variables (Table 2).

| Code  | Description                        | Unit | Model contribution (%) |
|-------|------------------------------------|------|------------------------|
| Bio1  | Annual mean temperature            | °C   | 12.59                  |
| Bio2  | Mean diurnal temperature range     | °C   | 0.40                   |
| Bio3  | Isothermality (Bio2/Bio7) (* 100)  | %    | 23.75                  |
| Bio12 | Annual precipitation               | mm   | 2.40                   |
| Bio13 | Precipitation in the wettest month | mm   | 69.58                  |
| Bio14 | Precipitation in the driest month  | mm   | 3.25                   |

**Table 1.** Bioclimatic variables selected for modeling of *C. odorata*.



**Figure 1.** Jackknife test showing relative importance of bioclimatic variables used in RF model. The details of variables bio1, bio2, bio3, bio12, bio13, and bio 14 are presented in Table 1.

| Parameter | Excellent | Good      | Fair      | Poor      | Fail  | References    |
|-----------|-----------|-----------|-----------|-----------|-------|---------------|
| AUC       | 0.91–1    | 0.81–0.90 | 0.71–0.80 | 0.61–0.7  | <0.60 | <sup>60</sup> |
| TSS       | 0.81–1    | 0.80–0.61 | 0.41–0.60 | 0.40–0.21 | <0.20 | <sup>61</sup> |
| Kappa     | 0.86–1    | 0.71–0.85 | 0.56–0.70 | 0.41–0.55 | <0.40 | <sup>61</sup> |

**Table 2.** Criteria for model evaluation and validation.

**Evaluation of RF model precision.** The area under the curve (AUC) score, true skill statistic (TSS), and Kappa coefficient were used to evaluate the performance of the model with test data. RF model performance was highest when the rarefied species-occurrence points were used rather than all occurrence points (Table 3). Therefore, we used modeling results obtained from rarefied occurrence points in this study (Fig. 2). The AUC score for the selected RF model was 0.96, indicating excellent model performance in predicting the distribution of *C. odorata* (Table 3). Similarly, the TSS (0.788) and Kappa coefficient (0.685) produced by the model were high. These results provide evidence that the RF model performs well and shows agreement between the observed data and predicted results.

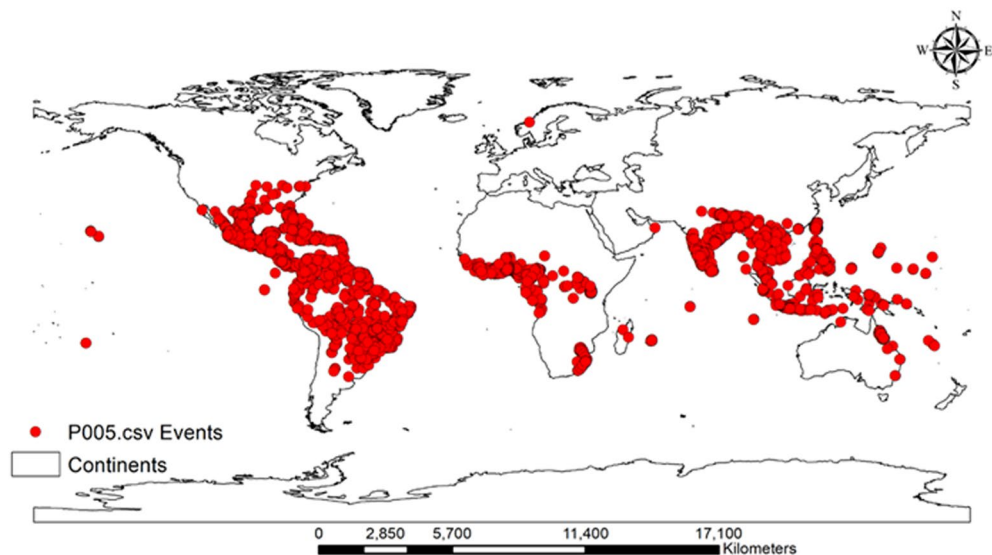
**Spatial distribution of *C. odorata* under the current climate.** RF modeling was performed to assess the current spatial distribution of *C. odorata* (Fig. 3A). *C. odorata* is distributed at latitudes within approximately 30° north and south of the equator, although some suitable habitat areas in South America exist outside of this range. The current spatial distribution of *C. odorata* was estimated in 1,753,877 cells (7,892,447 km<sup>2</sup>) covering 14.34% of the global land surface (Table 4). Under the current climate, *C. odorata* is widespread in all continents except Antarctica. The continent with the highest proportion of land within the estimated spatial distribution of *C. odorata* was South America (76.23%), followed by Africa (30.47%), Australia (21.79%), Oceania (19.95%), Asia (13.15%), North America (6.38%), and Europe (0.43%) (Table S2).

**Potential changes in suitable habitat under future climate scenarios.** We also used RF modeling to predict the future spatial distribution of *C. odorata* (2061–2080) under two climate change scenarios (SSP2-4.5 and SSP5-8.5; Fig. 3B and C). Overall, climate change will increase the worldwide spatial distribution of *C. odorata*: under SSP2-4.5 and SSP5-8.5, the area of suitable habitat for *C. odorata* will increase by 8.38 and 10.86%, respectively. The changes in suitable habitat for *C. odorata* (habitat expansion, habitat loss, and areas with no change) are presented in Fig. 4A (SSP2-4.5) and Fig. 4B (SSP5-8.5).

The change in suitable habitat as a proportion of current potential habitat was estimated and is expressed in Table 4. Future climate change increases the proportion of suitable habitat of *C. odorata* in all continents of the world, but the rate of estimated habitat expansion by 20161–2080 is the highest for Oceania (SSP2-4.5, 275.6% and SSP5-8.5, 326.1%), followed by Africa (SSP2-4.5, 118.7% and SSP5-8.5, 117.4%). Interestingly, under SSP5-8.5,

| Model calibration | Before rarefying | After rarefying |
|-------------------|------------------|-----------------|
| Occurrence points | 39,984           | 4801            |
| AUC               | 0.73             | 0.96            |
| TSS               | 0.64             | 0.76            |
| Kappa             | 0.52             | 0.69            |

**Table 3.** Summary of the random forest model of the worldwide distribution of *C. odorata*.



**Figure 2.** Global occurrence records of *C. odorata* (n = 4801). The red points in the map indicate GIS points for *C. odorata*. Generated using ArcGIS Desktop 10.8 (<https://desktop.arcgis.com>).

even a small portion of Antarctica (0.01%) is predicted to become future potential habitat for *C. odorata* (SSP5-8.5). The rate of estimated habitat loss is highest for Europe (SSP2-4.5, 86.7% and SSP5-8.5, 59.6%) and lowest for South America (SSP2-4.5, 2.7% and SSP5-8.5, 59.6%) for the same future period of time. In Africa, Oceania, and South America, more than 88% of habitat is predicted to remain unchanged, but this value is less than 80% for the other continents. A tropical climate with a hot and humid environment provides a suitable habitat for *C. odorata*; therefore, future climate change favors the retention of current suitable habitats and further habitat expansion.

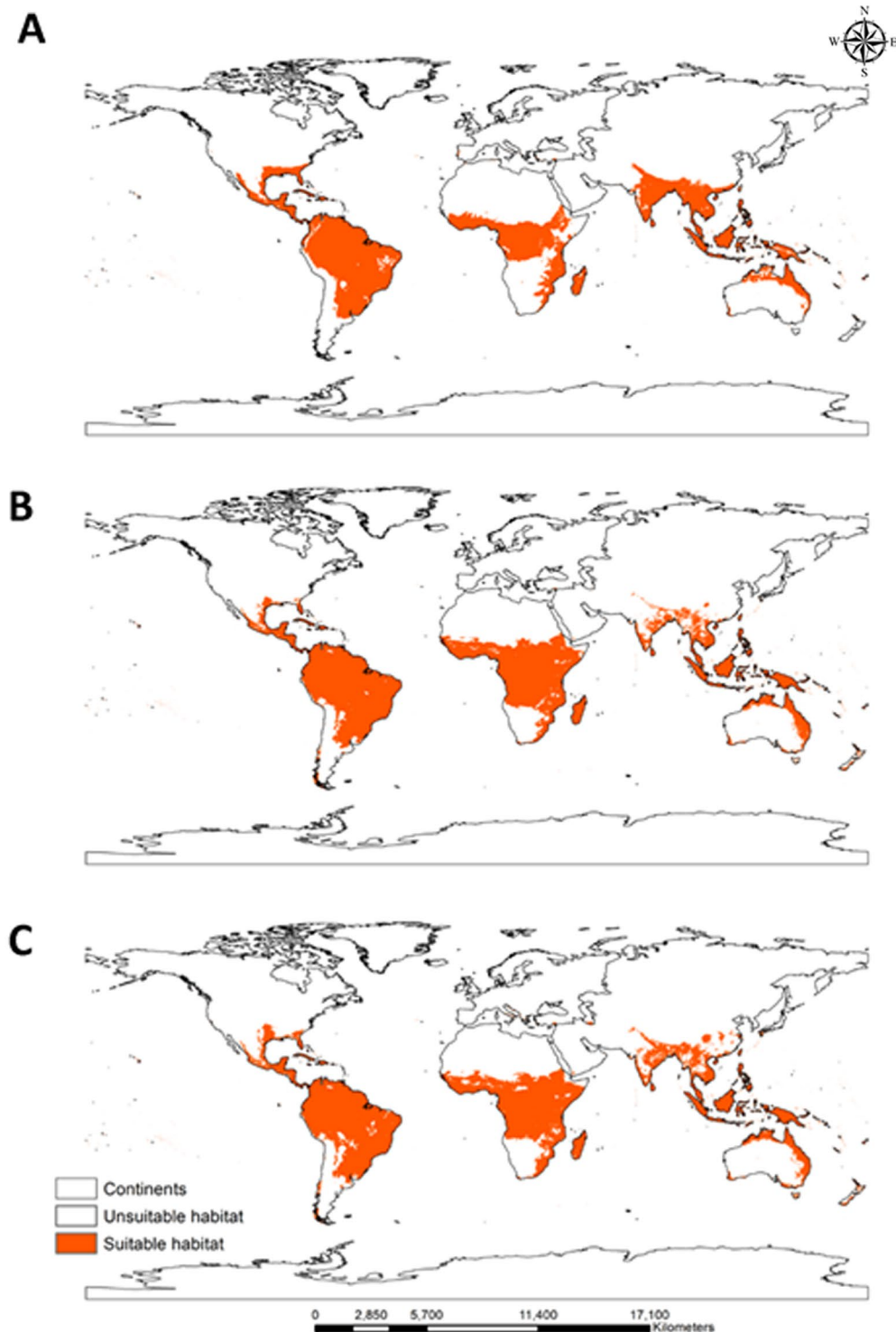
**Classification of habitat suitability in different countries of the world.** The mean habitat suitability of *C. odorata* was estimated for all countries (Table S3). Each country was subsequently classified as being unsuitable or having low, moderate, or high suitability for *C. odorata* under the current climate (Fig. 5A) and the two climate change scenarios (SSP2-4.5, Fig. 5B; SSP5-8.5, Fig. 5C). Under the current climate, 73 countries are classed as having unsuitable habitats (Fig. 6). However, under SSP2-4.5, by 2061–2080, 9 of these 73 countries are predicted to transition to the low-suitability category (Canada, Cape Verde, Djibouti, Japan, Mauritania, Niger, Norway, Yemen), 1 to the moderately suitable category (Senegal), and 3 to the highly suitable category (Gambia, Guinea-Bissau, and Lesotho) (Tables 5 and S4).

Similar results are predicted under SSP5-8.5 but with additional countries, including South Korea, Bosnia and Herzegovina, Israel, and Montenegro, predicted to transition from unsuitable to low suitability (Table 5). Eritrea, New Zealand, Somalia, and Sudan are predicted to transition from low to moderate suitability under SSP5-8.5, and Angola, Burkina Faso, and Zambia are predicted to transition from low to high suitability. The South American and African countries up to 30°S are predicted to have moderate to high suitability, but the Asian and European countries above 30°N latitude are predicted to have unsuitable to low-suitability habitats. These results indicate that several countries in Africa and South America, and some countries in Asia, may be at high risk of invasion by *C. odorata* in the future.

## Discussion

The key findings of this study are as follows. First, the model predictions from rare species-occurrence points are more accurate and less likely to cause overestimation with the RF model (Table 3). Second, the current global spatial distribution of *C. odorata* is primarily concentrated in South America and Africa, with minor regions in Asia and Australia (Fig. 3A). In general, suitable habitats are predominantly located between 30° north and south of the equator. However, future climate change will allow the spread of suitable habitats across all continents, with Oceania experiencing significant habitat expansion (326.04%) relative to its current status. Third, under current climate conditions, 73 countries have habitats that are not suitable for *C. odorata*. Of these countries, 11 (Bosnia and Herzegovina, Canada, Croatia, Djibouti, Israel, Japan, Mauritania, Montenegro, Niger, South Korea, and Yemen) will change to low suitability, 1 (Senegal) will change to moderate suitability, and 3 (Gambia, Guinea-Bissau, and Lesotho) will change to high suitability under the SSP5-8.5 scenario (Table 5). Last, tropical and warm temperate climates are favorable for *C. odorata*; therefore, northern Asia, including Mongolia, Kazakhstan, and Russia, and central and northern Europe are predicted to be safe from invasion of *C. odorata* for the studied time period (2061–2080; Fig. 3B and C).

Species distribution models (SDMs) are commonly used to predict the potential risk of invasion by alien species<sup>28</sup>. The accuracy of these models depends on factors such as the algorithms and variables used<sup>29</sup>, and the quality and quantity of the species-occurrence data<sup>30</sup>. The latter is particularly important for proper modeling. The species occurrence data used in this study was from GBIF. The GBIF data sources may have taxonomic bias

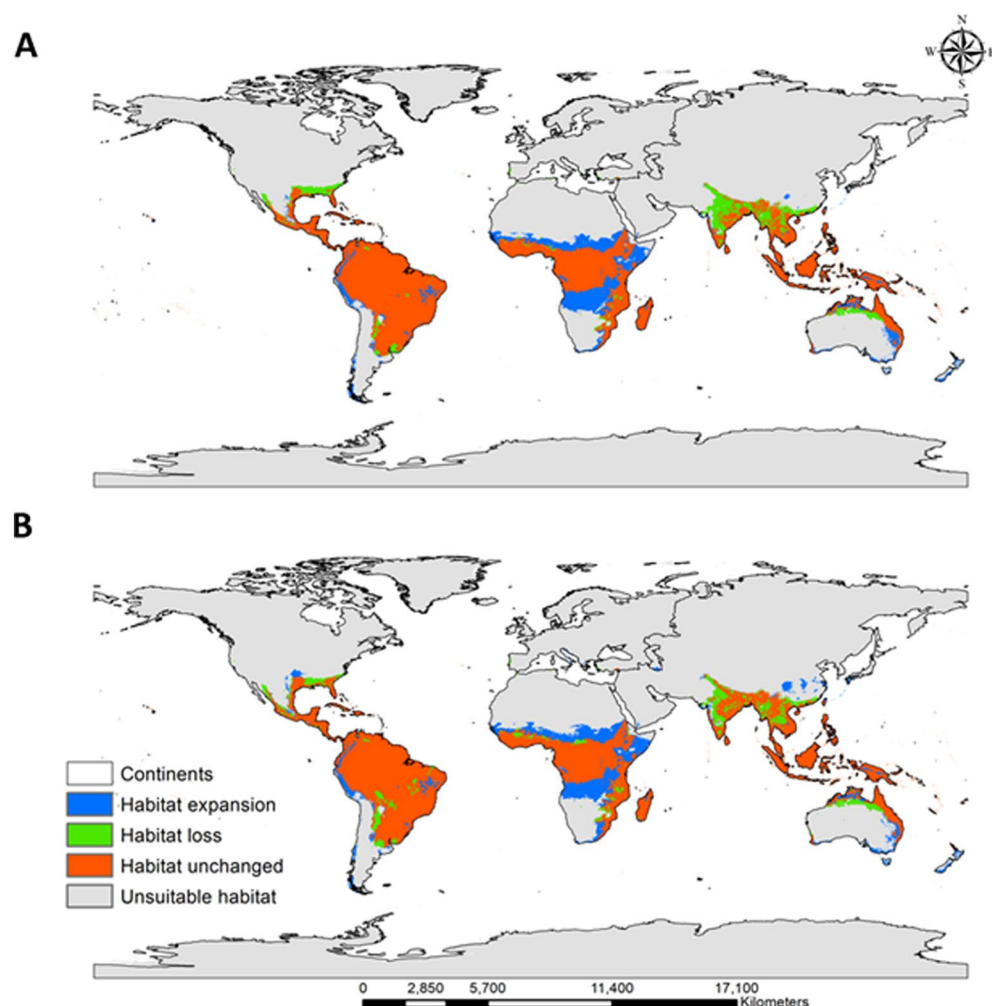


**Figure 3.** Potential global binary distribution of *C. odorata* under the current climate (1970–2000) (A) and under future climate change scenarios SSP2-4.5 (B) and SSP5-8.5 (C) (2061–2080). Orange indicates suitable habitat for *C. odorata*, and white indicates unsuitable habitat. Generated using ArcGIS Desktop 10.8 (<https://desktop.arcgis.com>).

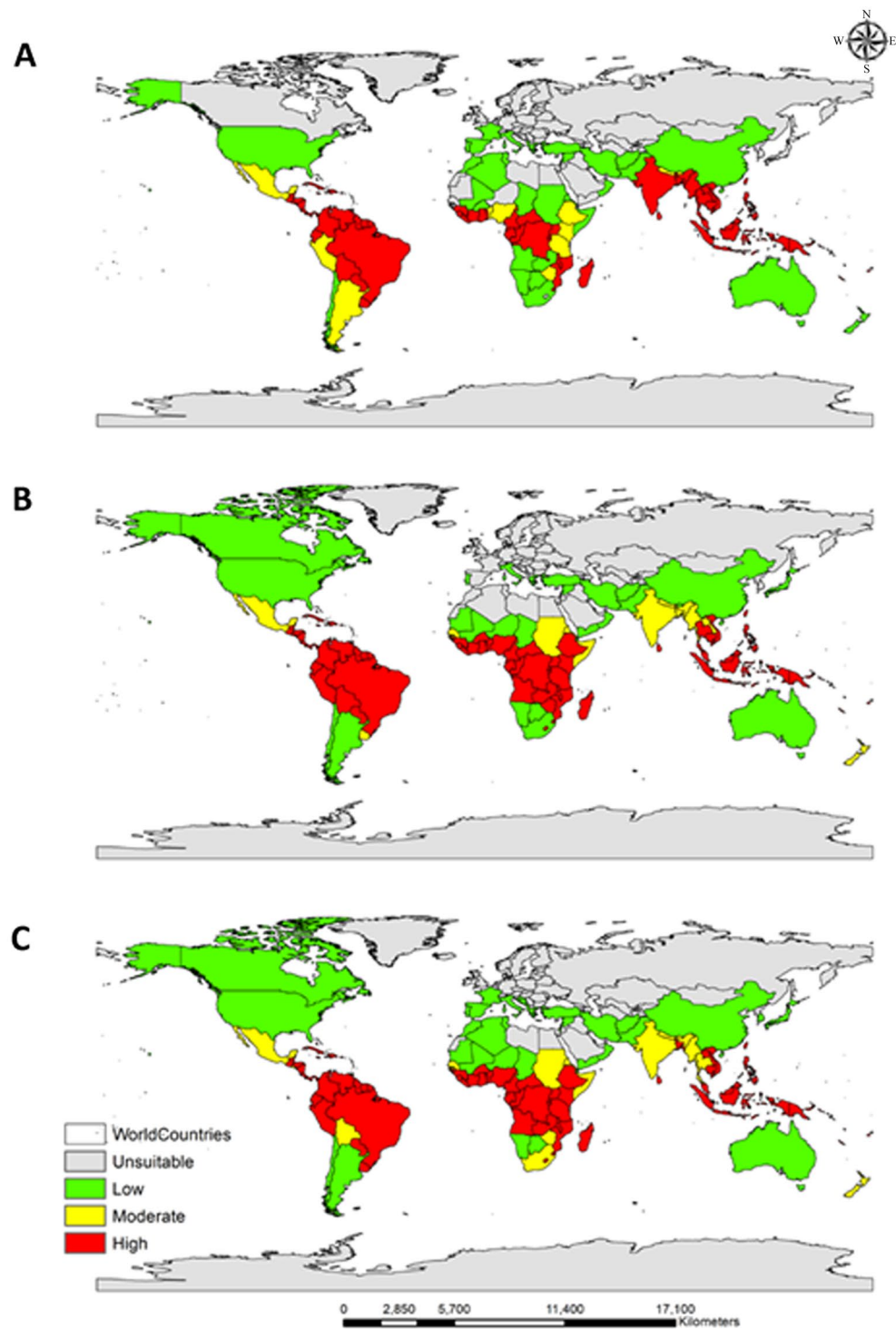
and this bias can arise from a variety of factors, including variation in the ease of identification, the frequency of occurrence, sampling protocol, quality control procedures and the level of interest among the researchers or citizen scientist, and manual georeferencing of herbarium or museum specimens<sup>31</sup>. Therefore, we carefully

| Continent     | Total area (km <sup>2</sup> ) | Current climate (km <sup>2</sup> ) | Expansion (%) |          | Loss (%) |          | Unchanged (%) |          |
|---------------|-------------------------------|------------------------------------|---------------|----------|----------|----------|---------------|----------|
|               |                               |                                    | SSP2-4.5      | SSP5-8.5 | SSP2-4.5 | SSP5-8.5 | SSP2-4.5      | SSP5-8.5 |
| Africa        | 6,629,265                     | 2,020,171.5                        | 118.67        | 117.41   | 2.70     | 6.02     | 97.30         | 93.98    |
| Asia          | 13,730,769                    | 1,805,130                          | 5.73          | 16.74    | 32.02    | 22.76    | 67.98         | 77.24    |
| Australia     | 1,798,479                     | 391,932                            | 85.57         | 81.41    | 28.12    | 29.51    | 71.88         | 70.49    |
| Europe        | 3,959,239.5                   | 17,163                             | 5.98          | 44.21    | 86.65    | 59.60    | 13.35         | 40.40    |
| North America | 9,573,160.5                   | 610,726.5                          | 7.20          | 20.63    | 29.30    | 23.52    | 70.70         | 76.48    |
| Oceania       | 88,200                        | 17,595                             | 275.65        | 326.04   | 11.99    | 2.33     | 88.01         | 97.67    |
| South America | 3,974,377.5                   | 3,029,728.5                        | 14.22         | 14.88    | 2.67     | 5.29     | 97.33         | 94.71    |
| Antarctica    | 15,265,440                    | 0                                  | 230           | 488      |          |          |               |          |
| Total         | 55,018,930.5                  | 7,892,446.5                        | 42.59         | 46.30    | 12.92    | 12.20    | 87.08         | 87.80    |

**Table 4.** Estimated habitat changes (%) from current climate to 2061–2080 for different continents under climate change scenarios SSP2-4.5 and SSP5-8.5 SSP shared socioeconomic pathways. Current climate, climate during 1970–2000.

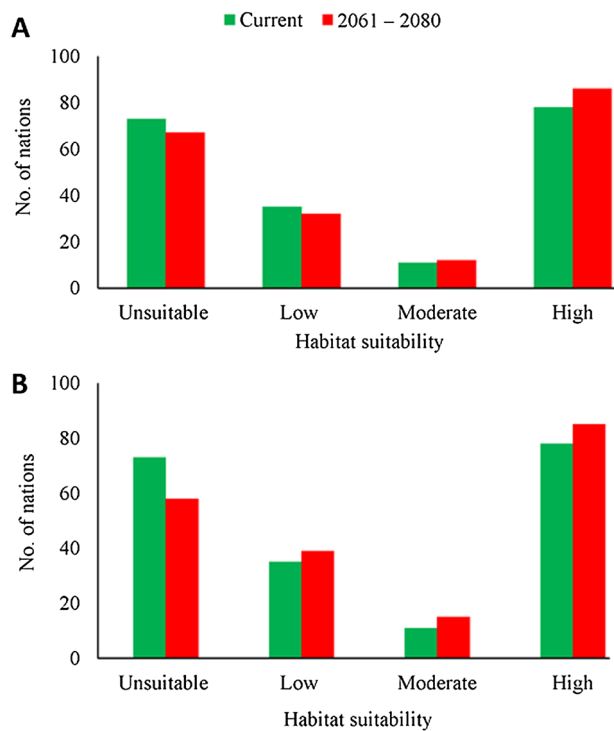


**Figure 4.** Changes in habitat suitability for *C. odorata* under the climate change scenarios SSP2-4.5 (A) and SSP5-8.5 (B) by 2061–2080. Blue indicates habitat expansion, green indicates habitat loss, orange indicates unchanged habitat, and grey indicates unsuitable habitat. Generated using ArcGIS Desktop 10.8 (<https://desktop.arcgis.com>).



**Figure 5.** Mean habitat suitability of *C. odorata* estimated for different countries of the world. (A) Mean habitat suitability under the current climate. (B and C) Mean future habitat suitability under climate change scenarios SSP2-4.5 (B) and SSP5-8.5 (C) by 2061–2080. Gray indicates unsuitable habitat, green indicates habitat with low suitability, yellow indicates moderate suitability, and red indicates high suitability. Generated using ArcGIS Desktop 10.8 (<https://desktop.arcgis.com>).

check the taxonomy, data source, and implemented some quality measures such as filters to remove occurrences with improbable geographic coordinates. During the process we removed the occurrence points present in sea, desert and in arctic regions.



**Figure 6.** Estimation of habitat suitability in different countries of the world under the climate change scenarios SSP2-4.5 (A) and SSP5-8.5 (B). The average habitat suitability of each country was classified as unsuitable (0), low (0.1–0.33), moderate (0.34–0.66), or high (0.67–1.00). Green indicates the number of countries currently in each category; red indicates the predicted number of countries in each category in 2061–2080.

| Change in habitat suitability <sup>a</sup> | SSP2-4.5                                                              | SSP5-8.5                                                                                                            |
|--------------------------------------------|-----------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| Unsuitable to low                          | Canada, Cape Verde, Djibouti, Japan, Mauritania, Niger, Norway, Yemen | Bosnia and Herzegovina, Canada, Croatia, Djibouti, Israel, Japan, Mauritania, Montenegro, Niger, South Korea, Yemen |
| Unsuitable to moderate                     | Senegal                                                               | Senegal                                                                                                             |
| Unsuitable to high                         | Gambia, Guinea-Bissau, Lesotho                                        | Gambia, Guinea-Bissau, Lesotho                                                                                      |
| Low to moderate                            | Eritrea, New Zealand, Somalia, Sudan                                  | Eritrea, New Zealand, Somalia, Sudan                                                                                |
| Low to high                                | Angola, Burkina Faso, Zambia                                          | Angola, Burkina Faso, Zambia                                                                                        |
| Moderate to high                           | Benin, Ethiopia, Nigeria, Peru, Tanzania, Zimbabwe                    | Benin, Ethiopia, Kenya, Nigeria, Peru, Tanzania                                                                     |

**Table 5.** Predicted *C. odorata* habitat transition in different countries from current climate to 2061–2080 under climate change scenarios SSP2-4.5 and SSP5-8.5 <sup>a</sup>The habitat suitability categories were unsuitable (0), low suitability (0.001–0.33), moderate suitability (0.34–0.66), and high suitability (0.67–1.0), and categorization was performed according to the mean habitat-suitability value for each country. This table shows that some countries in the unsuitable, low suitability, and moderate-suitability categories in current climate are predicted to transition to a higher level of habitat suitability by 2061–2080 under climate change scenarios SSP2-4.5 and SSP5-8.5. Details of habitat suitability classification are presented in Table S4.

In this study, we conducted RF modeling both with all available occurrence points and with spatially rarefied occurrence points, to reduce errors and overfitting due to spatial autocorrelation<sup>32</sup>. We compared model performance using the AUC scores, TSS, and Kappa coefficients (Table 3): the model that used the spatially rarefied occurrence points produced the highest scores and was thus considered the most accurate. Our RF model makes accurate predictions and can simulate the spread of *C. odorata*, but there are limitations to this study. The assumption that species niche demand is conservative may not always be true as invasive species can experience niche drift<sup>33</sup>. Additionally, invasive species often move quickly to many places, making it difficult to accurately predict their potential distribution<sup>34</sup>. Using equilibrium data to identify suitable habitat areas would be more effective, but monitoring the population status of invasive species in invaded areas is currently difficult and a low priority. These issues are similar to the limitations reported in studies of *Xanthium italicus*<sup>34</sup>.

The distribution and modeling of invasive species can be influenced by both intrinsic factors, such as dispersal distance and rate of species generation, and extrinsic factors, such as human activities and natural phenomena<sup>35</sup>. According to the Jackknife test performed in this study, two temperature-related variables Bio 1 and Bio 3 and two precipitation-related variables Bio 12 and Bio 13 and were relatively important variables for the habitat

suitability of *C. odorata* (Fig. 1), estimated to contribute 12.59, 23.75, 2.40 and 69.58%, respectively (Table 1). *C. odorata* thrives in areas with high temperatures and precipitation but is unable to tolerate frost<sup>9</sup>. The seeds of *C. odorata* are susceptible to damage at extremely cold temperatures and in excessively dry soil<sup>9</sup>. As a result, it prefers wet-dry seasonal conditions and can grow in various soils up to 1200 m above sea level. Therefore, future climate change is expected to promote the habitat expansion of *C. odorata*'s habitat<sup>36</sup>.

Besides bioclimatic variables, several traits contribute to the invasive nature of *C. odorata*, including its high germination and growth rate, high fecundity<sup>37</sup>, ability to regenerate from roots, and ability to tolerate a wide range of temperatures<sup>10,38</sup>. Environmental factors such as anthropogenic land-use and landcover changes, soil, and roads, may also be crucial factors in determining the distribution of *C. odorata*, but future data on such variables are not yet accessible at a similar resolution. The plant can spread easily via wind, fur, clothes, and machinery and can quickly take over new habitats once introduced, replacing other invasive species<sup>38</sup>. Studies have found that abandoned farms, orchards, urban environments, and roadside ditches make excellent habitats for *C. odorata*, allowing it to rapidly dominate these areas<sup>9,10</sup>.

Our study shows that *C. odorata* is primarily found in areas between 30° north and south of the equator, with humid tropical, subtropical, and warm temperate climates. We also found that the greatest habitat expansion for *C. odorata* is predicted to occur in Africa, specifically in Nigeria, Chad, Sudan, and Somalia. *C. odorata* is also expected to expand in North America, South America, Asia, Australia, and New Zealand, resulting in an increase in its spatial distribution of up to 10.86%. An increase in temperature and precipitation patterns may lead to the future loss of habitat for native species and the transition of habitat from unsuitable for *C. odorata* to highly suitable for *C. odorata*, with its invasiveness aided by its high reproductive capacity, similar to other invasive plants such as *P. hysterophorus*<sup>26</sup>. However, Mediterranean, semiarid, and cold temperate climates are predicted to be unsuitable habitats for *C. odorata* in the future. Most parts of North America, Europe, and northern and central Asia are also predicted to be unsuitable habitats for the plant. These results are consistent with some previous studies<sup>9,10,38</sup>.

The Antarctic Peninsula has experienced the most rapid air-temperature increases in the world over the past 50 years, and the West Antarctic Ice Sheet is projected to lose nearly all of its bulk over several millennia, with sustained warming levels between 2 and 3 °C<sup>39</sup>. These changes could result in the loss of native biodiversity and the establishment of alien and invasive species in Antarctica. Although the northern hemisphere above 30°N latitude is currently unsuitable for *C. odorata*, we found that a small portion of Antarctica (0.01%) could potentially become a suitable habitat for this invasive species in the future because of climate change. Therefore, further research is needed to investigate the potential for *C. odorata* to establish in Antarctica.

*C. odorata* grows rapidly and can negatively impact agriculture, forestry, and grazing animals<sup>40</sup>. Its leaves, particularly the young ones, contain high levels of nitrate, making them toxic and sometimes fatal to grazing cattle and wild herbivores such as roe deer<sup>41</sup>. *C. odorata* is difficult to control once established as it can spread quickly and regrow from seeds and rootstocks, even in hard-to-reach places like cliffs<sup>40</sup>. It also alters soil nutrients and affects nearby vegetation<sup>5</sup> and is named in “100 of the World's Worst Invasive Alien Species”<sup>4</sup>. Despite these negative impacts, practitioners of conventional medicine continue to use it for its various medicinal properties and bioactive compounds, such as fatty acids, flavonoids, saponins, and alkaloids<sup>42</sup>.

*C. odorata* is an invasive plant species that requires a combination of strategies to control its spread and mitigate its negative impacts on the ecosystem, including herbicides, mechanical control, and biological control<sup>43</sup>. Herbicides can be effective but also harm insects and native flora. Alternative control methods, such as mechanical or biological controls, have fewer side effects but can be costly and require more effort. Biological methods of control have been used in many countries with success rates of up to 70%<sup>44</sup>. A strict quarantine system on borders is needed to stop the further expansion of *C. odorata*; collaboration between governments, land-resource managers, and local stakeholders will be required to develop an invasion control plan.

## Conclusions

*C. odorata* can adapt to changing climate conditions such as rising temperatures and shifting precipitation patterns. In this study, an RF model was used to estimate the global spatial distribution of *C. odorata* and to identify areas at high risk of invasion under current and future climate scenarios (SSP2-4.5 and SSP5-8.5). The results revealed that, by 2016–2080, *C. odorata* will retain its current ecological niche and expand its habitat in many countries of Oceania, Africa, and South America, with expansion of up to 326.04% in Oceania. The global distribution of *C. odorata* is expected to increase by up to 10.86% under the SSP5-8.5 climate change scenario. Countries with tropical, subtropical, and temperate climates located between 30° north and south of the equator are at high risk of invasion. Our study identified areas that may become suitable habitats under current and future climate change, which will be useful in developing long-term management strategies for *C. odorata*. For examples, immediate control and management measures such as mechanical control, biological control, and restoration of native species are needed in high-risk countries. These countries should establish the preventive measures such as robust quarantine systems against a potential threat of invasion.

## Materials and methods

**Global species-occurrence records.** Global occurrence records for *C. odorata* (39,984 points) were downloaded mainly from the GBIF ([www.gbif.org](http://www.gbif.org), accessed September 8, 2022), an open access data source. The spatially rarefy occurrence tool in the ArcGIS SDM toolbox v. 2.4 was then used to remove multiple species-occurrence points from the same grid at a spatial scale of 2.5 min (4.5 km<sup>2</sup>) and to select one distinct point per grid<sup>45</sup>. This procedure prevents overfitting and inaccurate inflation of model results due to spatial autocorrelation<sup>32</sup>. Ultimately, there were 4,801 species-occurrence points for *C. odorata* (Fig. 2 and Table S5). In

this study, we used both sets of species-occurrence points in the RF modeling of *C. odorata* to understand model overestimation.

**Selection of bioclimatic variables.** Temperature and precipitation are two of the most significant environmental variables that determine floral diversity and distribution. Thus, the 19 WorldClim bioclimatic variables (derived from temperature and precipitation data) are highly relevant to the ecological and physical tolerance of plants<sup>46</sup>. We therefore downloaded historical data from 1970 to 2000 from WorldClim v2.1 at a spatial resolution of 2.5 min, which is approximately 4.5 km at the equator<sup>47</sup>. Similarly, future bioclimatic variables at 2.5 min resolution from the Coupled Model Intercomparison Project Phase 6 (CMIP6) were also downloaded from the WorldClim data portal<sup>48</sup>. We also used two shared socioeconomic pathways (SSPs; SSP2-4.5 and SSP5-8.5) representing future climate data from 2061 to 2080, developed under the global circulation model (GCM) Max Planck Institute Earth System Model (MPI-ESM1-2-HR)<sup>49</sup>. The SSP scenarios assess changes in energy use and land use as well as the associated uncertainties in greenhouse gas and air pollutant emissions<sup>50</sup>. Among the four SSP scenarios in the WorldClim database, the SSP2-4.5 and SSP5-8.5 depicted an average warming of 3.0 and 5.0 °C, respectively. These scenarios represented intermediate and high radiative forcing, which encompassed greenhouse gas emissions, forest fires, and volcanic eruptions<sup>48</sup>. Consequently, the modeling of *C. odorata* utilized the SSP2-4.5 and SSP5-8.5 scenarios.

According to SSP2-4.5 and SSP5-8.5, the average global surface temperature in 2090–2100 will have risen by 2.4–4.3 °C and 3.8–8.6 °C, respectively, relative to 1880–1900<sup>48</sup>. The SSPs anticipate worldwide socioeconomic development through 2100. The MPI-ESM1-2-HR GCM includes land and ocean carbon cycles, and encompasses the latest ocean biogeochemistry module (the Hamburg Model of the Ocean Carbon Cycle) and land surface scheme (JSBACH)<sup>51</sup> to predict the seasonal climate over decades. Downloading bioclimatic variables from the WorldClim data portal is a popular method of estimating the probable spread of species in response to changing precipitation and temperature. Depending on the particular climate scenarios and associated ecologies, the bioclimatic variables serve to define and forecast the future distribution patterns of species<sup>34</sup>.

We ran a Spearman's correlation test on data from the 19 bioclimatic WorldClim variables (Table S6) using the PROC CORR function of SAS 9.4 (SAS Institute, Inc., Cary, NC, USA), as described before<sup>52,53</sup>. We selected 6 of the 19 bioclimatic variables (Table 1) on the basis of their low correlation with each other ( $r < 0.75$ ; Table S1): annual mean temperature (Bio01), mean diurnal temperature range (Bio2), isothermality (Bio03), annual precipitation (Bio12), precipitation in the wettest month (Bio13), and precipitation in the driest month (Bio14). These six variables were considered the most significant climatic variables for predicting the spatial distribution of *C. odorata*. The relative importance of each variable was accessed using Jackknife test.

**Model development.** The RF model predictions for *C. odorata* in this study were performed with Biomod2 Package 4.1-2, selecting single model RF<sup>54</sup>. The species-occurrence data and bioclimatic data were imported into R. The lack of information about invasive species can make it difficult to determine whether a habitat is suitable because invasive species are likely expanding and have not yet reached equilibrium<sup>55</sup>. The background data (pseudo absent) points of the study area were determined using ArcGIS 10.3, as suggested previously<sup>56,57</sup>. The species-occurrence data were divided into two; 75% of the data was used for model calibration, and 25% was used for model validation<sup>58</sup>. The other model options were run with the default settings, and the model was replicated ten times, as described previously<sup>59</sup>.

**Model evaluation and validation.** The goodness-of-fit of the model was evaluated via three parameters, namely, the AUC score of the receiver operating characteristic (ROC) curves<sup>60</sup>, the TSS<sup>61</sup>, and the Kappa coefficient. The AUC score plots 1 specificity on the abscissa and sensitivity on the ordinate<sup>14</sup>. The AUC is a technique for differentiating between presence and absence that is independent of thresholds, and the score, which assesses the performance of a model, ranges from 0 to 1<sup>62</sup>. The AUC value is unaffected by the size of the dataset (prevalence), but its use is debatable because it gives equal weight to errors of commission and omission and may not assess prediction accuracy reliably<sup>63</sup>. Habitat expansion outside of the species-occurrence range may provide a high AUC score, leading to overfitting, a condition that misleads model evaluation<sup>64</sup>. Therefore, other evaluation parameters (TSS and Kappa coefficient) were also employed to measure the accuracy of the model. The TSS calculates both the specificity and sensitivity of the model [ $\text{TSS} = \text{sensitivity} + (\text{specificity} - 1)$ ]<sup>61</sup>, ranging from  $-1$  to  $+1$ ; estimates both omission and commission errors<sup>61</sup>; and is frequently used as an alternative method for checking model accuracy<sup>61,65</sup>. The Kappa coefficient is the ratio of observation points that the model properly and erroneously predicts. Similar to the TSS, it also ranges from  $-1$  (poor agreement) to  $+1$  (perfect prediction)<sup>61,66</sup>. The greater the scores of these parameters, the higher the accuracy of the model (Table 2)<sup>67</sup>.

**Spatial distribution and change in habitat suitability of *C. odorata* across the world.** The global binary distribution maps of *C. odorata* were obtained from probability distribution maps produced by RF modeling under the current and future climate change scenarios (SSP2-4.5 and SSP5-8.5) for the period 2061–2080 using the threshold TSS defined in the Biomod2 Package<sup>54</sup>. The binary distribution maps represent the suitable and unsuitable habitats for *C. odorata*. We estimated the changes in suitable habitats for *C. odorata* by 2061–2080 by differentiating current suitable and future suitable habitats, and reclassifying the differentiated raster to identify expanded, reduced, and unchanged regions across the world using the Raster v3.6 package in R software (<https://cran.r-project.org/web/packages/raster>, accessed December 12, 2023). The number of cells showing suitable habitat under the current climate and the number of expanded, reduced, and unchanged habitats by 2061–2080 were estimated for the different continents using zonal statistics under the spatial analyst in ArcGIS Desktop 10.8 (Esri, Redlands, CA, USA). Then, the approximate area was estimated (1 grid cell = 4.5 km<sup>2</sup>), and

the proportions of habitat change for the different continents were calculated under the climate change scenarios SSP2-4.5 and SSP5-8.5. Similarly, to understand the risk of invasion in different countries of the world, mean habitat-suitability values for 2061–2080 were estimated for countries under the current and future climate change scenarios, and countries were then classified into four categories: unsuitable habitat (0), low-suitability habitat ( $\leq 0.33$ ), moderate-suitability habitat ( $\leq 0.66$ ), and high-suitability habitat ( $\leq 1$ ).

## Data availability

All data generated or analyzed during this study are included in this article.

Received: 22 February 2023; Accepted: 1 June 2023

Published online: 16 June 2023

## References

1. Pasiecznik, N. *Chromolaena odorata* (Siam weed). *CABI Compendium* <https://doi.org/10.1079/cabicompendium.23248> (2022).
2. Gautier, L. Taxonomy and distribution of a tropical weed: *Chromolaena odorata* (L.) R. King & H. Robinson. *Candollea* **47**, 645–662 (1992).
3. Waterhouse, B. M. & Zeimer, O. On the brink: The status of *Chromolaena odorata* in northern Australia. In *Proceedings of the Fifth International Workshop on Biological Control and Management of Chromolaena odorata*, Pretoria, South Africa: ARC-PPRI (eds. Zachariades, C., Muniappan, R. & Strathie, L. W.) 29–33 (Citeseer, 2002).
4. Lowe, S., Browne, M., Boudjelas, S. & De Poorter, M. *100 of the world's worst invasive alien species: a selection from the global invasive species database*, Vol. 12 (Invasive Species Specialist Group Auckland, 2000).
5. Zahara, M. Description of *Chromolaena odorata* LRM King and H. Robinson as medicinal plant: A Review. In *IOP Conference Series: Materials Science and Engineering* (IOP Publishing).
6. Oritsejafor, J. Weed hosts of *Fusarium oxysporum* f. sp. *elaedis*. *Oleagineux* **41**, 1–7 (1986).
7. Chapman, R., Page, W. & McCaffery, A. Bionomics of the variegated grasshopper (*Zonocerus variegatus*) in West and Central Africa. *Annu. Rev. Entomol.* **31**, 479–505 (1986).
8. McFadyen, R. E. C. In *Proceedings of the Fifth International Workshop on Biological Control and Management of Chromolaena odorata*, Pretoria, South Africa: ARC-PPRI (eds. Zachariades, C., Muniappan, R. & Strathie, L. W.) 13–18 (Citeseer, 2002).
9. Kriticos, D., Yonow, T. & McFadyen, R. The potential distribution of *Chromolaena odorata* (Siam weed) in relation to climate. *Weed Res.* **45**, 246–254 (2005).
10. McFadyen, R. C. & Skarratt, B. Potential distribution of *Chromolaena odorata* (siam weed) in Australia, Africa and Oceania. *Agric. Ecosyst. Environ.* **59**, 89–96 (1996).
11. Wei, H., Xu, J., Quan, G., Zhang, J. & Qin, Z. Invasion effects of *Chromolaena odorata* on soil carbon and nitrogen fractions in a tropical savanna. *Ecosphere* **8**, e01831 (2017).
12. Bradley, B. A., Blumenthal, D. M., Wilcove, D. S. & Ziska, L. H. Predicting plant invasions in an era of global change. *Trends Ecol. Evol.* **25**, 310–318 (2010).
13. Franklin, J. *Mapping Species Distributions: Spatial Inference and Prediction* (Cambridge University Press, 2010).
14. Zhang, X. *et al.* The global potential distribution of invasive plants: *Anredera cordifolia* under climate change and human activity based on random forest models. *Sustainability* **12**, 1491 (2020).
15. Adhikari, P., Lee, Y. H., Adhikari, P., Hong, S. H. & Park, Y.-S. Climate change-induced invasion risk of ecosystem disturbing alien plant species: An evaluation using species distribution modeling. *Front. Ecol. Evol.* 681 (2022).
16. Guisan, A., Thuiller, W. & Zimmermann, N. E. *Habitat Suitability and Distribution Models: With Applications in R* (Cambridge University Press, 2017).
17. Guisan, A., Edwards, T. C. Jr. & Hastie, T. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecol. Model.* **157**, 89–100 (2002).
18. Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151 (2006).
19. Olden, J. D., Lawler, J. J. & Poff, N. L. Machine learning methods without tears: A primer for ecologists. *Q. Rev. Biol.* **83**, 171–193 (2008).
20. Li, X. & Wang, Y. Applying various algorithms for species distribution modelling. *Integr. Zool.* **8**, 124–135 (2013).
21. Probst, P., Wright, M. N. & Boulesteix, A. L. Hyperparameters and tuning strategies for random forest. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **9**, e1301 (2019).
22. Elith, J. 15-Machine learning, random forests, and boosted regression trees. In *Quantitative Analyses in Wildlife Science* 281 (Johns Hopkins University Press, USA, 2019).
23. Adhikari, P. *et al.* Potential impact of climate change on plant invasion in the Republic of Korea. *J. Ecol. Environ.* **43**, 36. <https://doi.org/10.1186/s41610-019-0134-3> (2019).
24. Hong, S. H., Lee, Y. H., Lee, G., Lee, D.-H. & Adhikari, P. Predicting impacts of climate change on northward range expansion of invasive weeds in South Korea. *Plants* **10**, 1604 (2021).
25. Adhikari, P., Kim, B.-J., Hong, S.-H. & Lee, D.-H. Climate change induced habitat expansion of nutria (*Myocastor coypus*) in South Korea. *Sci. Rep.* **12**, 1–12 (2022).
26. Adhikari, P. *et al.* Predicting the impact of climate change on the habitat distribution of *Parthenium hysterophorus* around the world and in South Korea. *Biology* (2022).
27. Ahmed, N., Atzberger, C. & Zewdie, W. Species Distribution Modelling performance and its implication for Sentinel-2-based prediction of invasive *Prosopis juliflora* in lower Awash River basin, Ethiopia. *Ecol. Process.* **10**, 1–16 (2021).
28. Elith, J. Predicting distributions of invasive species. In *Invasive Species: Risk Assessment and Management* (Cambridge University Press, Cambridge, 2017).
29. Kaky, E., Nolan, V., Alatawi, A. & Gilbert, F. A comparison between ensemble and MaxEnt species distribution modelling approaches for conservation: A case study with Egyptian medicinal plants. *Ecol. Inf.* **60**, 101150 (2020).
30. La Marca, W. *et al.* The influence of data source and species distribution modelling method on spatial conservation priorities. *Divers. Distrib.* **25**, 1060–1073 (2019).
31. Guralnick, R. P., Wiczorek, J., Beaman, R., Hijmans, R. J., BioGeomancer Working Group, T. BioGeomancer: Automated georeferencing to map the world's biodiversity data. *PLoS Biol.* **4**, e381 (2006).
32. Boria, R. A., Olson, L. E., Goodman, S. M. & Anderson, R. P. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol. Model.* **275**, 73–77 (2014).
33. Svensson, E. I., Gómez-Llano, M. A., Torres, A. R. & Bensch, H. M. Frequency dependence and ecological drift shape coexistence of species with similar niches. *Am. Nat.* **191**, 691–703 (2018).
34. Zhang, Y., Tang, J., Ren, G., Zhao, K. & Wang, X. Global potential distribution prediction of *Xanthium italicum* based on Maxent model. *Sci. Rep.* **11**, 1–10 (2021).

35. Gao, T., Xu, Q., Liu, Y., Zhao, J. & Shi, J. Predicting the potential geographic distribution of *Sirex nitobei* in China under climate change using maximum entropy model. *Forests* **12**, 151 (2021).
36. Vanderwoude, C., Scanlan, J., Davis, B. & Funkhouser, S. Plan for national delimiting survey for Siam weed. *Natural Resources and Mines Land Protection Services: Queensland Government* 224–300 (2005).
37. Ambika, S. Suppression of plantation crops by Eupatorium weed. *Curr. Sci.* **49**, 874–875 (1980).
38. Muniappan, R., Reddy, G. & Lai, P.-Y. Distribution and biological control of *Chromolaena odorata*. In *Invasive Plants: Ecological and Agricultural Aspects* 223–233 (Springer, 2005).
39. Portner, H., Roberts, D. & Constable, A. IPCC, 2022: Summary for policymakers (2022).
40. Yu, F., Akin-Fajiyi, M., Thapa Magar, K., Ren, J. & Gurevitch, J. A global systematic review of ecological field studies on two major invasive plant species, *Ageratina adenophora* and *Chromolaena odorata*. *Divers. Distrib.* **22**, 1174–1185 (2016).
41. Adhikari, P. *et al.* Seasonal and altitudinal variation in roe deer (*Capreolus pygargus tianschanicus*) diet on Jeju Island, South Korea. *J. Asia Pac. Biodivers.* **9**, 422–428 (2016).
42. Olawale, F., Olofinson, K. & Iwaloye, O. Biological activities of *Chromolaena odorata*: A mechanistic review. *S. Afr. J. Bot.* **144**, 44–57 (2022).
43. Goodall, J. & Erasmus, D. Review of the status and integrated control of the invasive alien weed, *Chromolaena odorata*, in South Africa. *Agric. Ecosyst. Environ.* **56**, 151–164 (1996).
44. Muniappan, R. & Bamba, J. Biological control of *Chromolaena odorata*: Successes and failures. In *Proceedings of the X International Symposium on Biological Control of Weeds* 81–85.
45. Brown, J. L., Bennett, J. R. & French, C. M. SDMtoolbox 2.0: The next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ* **5**, e4095–e4095. <https://doi.org/10.7717/peerj.4095> (2017).
46. Allen, S. K., Plattner, G.-K., Nauels, A., Xia, Y. & Stocker, T. F. Climate Change 2013: The physical science basis. An overview of the working group 1 contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC). In *EGU General Assembly Conference Abstracts* 3544 (2014).
47. Fick, S. E. & Hijmans, R. J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
48. Eyring, V. *et al.* Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
49. Müller, W. A. *et al.* A higher-resolution version of the max planck institute earth system model (MPI-ESM1. 2-HR). *J. Adv. Model. Earth Syst.* **10**, 1383–1413 (2018).
50. Riahi, K. *et al.* The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob. Environ. Change* **42**, 153–168 (2017).
51. Reick, C., Raddatz, T., Brovkin, V. & Gayler, V. Representation of natural and anthropogenic land cover change in MPI-ESM. *J. Adv. Model. Earth Syst.* **5**, 459–482 (2013).
52. Shin, M.-S. *et al.* Prediction of potential species richness of plants adaptable to climate change in the Korean Peninsula. *J. Environ. Impact Assess.* **27**, 562–581. <https://doi.org/10.14249/EIA.2018.27.6.562> (2018).
53. Adhikari, P., Kim, H. W., Shin, M. S., Hong, S. H. & Cho, Y. Potential distribution of the silver striped skipper (*Leptolina unicolor*) and maiden silvergrass (*Miscanthus sinensis*) under climate change in South Korea. *Entomol. Res.* **52**, 483–492 (2022).
54. Thuiller, W. *et al.* Package 'biomod2': Ensemble platform for species distribution modeling, Version 4.1-2 (2022). <https://cran.r-project.org/web/packages/biomod2/index.html>.
55. Jiménez-Valverde, A. *et al.* Use of niche models in invasive species risk assessments. *Biol. Invasions* **13**, 2785–2797. <https://doi.org/10.1007/s10530-011-9963-4> (2011).
56. Barbet-Massin, M., Jiguet, F., Albet, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many?. *Methods Ecol. Evol.* **3**, 327–338 (2012).
57. Jeon, J.-Y., Adhikari, P. & Seo, C. Impact of climate change on potential dispersal of *Paonia obovata* (paeoniaceae), a critically endangered medicinal plant of South Korea. *Ecol. Environ. Conserv.* **26**, S145–S155 (2020).
58. Araujo, M. B. & Guisan, A. Five (or so) challenges for species distribution modelling. *J. Biogeogr.* **33**, 1677–1688 (2006).
59. Adhikari, P. *et al.* Northward range expansion of southern butterflies according to climate change in South Korea. *JCCR* **11**, 643–656. <https://doi.org/10.15531/jkscrc.2020.11.6.643> (2020).
60. Swets, J. A. Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293 (1988).
61. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232 (2006).
62. Thuiller, W., Lavorel, S. & Araújo, M. B. Niche properties and geographical extent as predictors of species sensitivity to climate change. *Glob. Ecol. Biogeogr.* **14**, 347–357. <https://doi.org/10.1111/j.1466-822X.2005.00162.x> (2005).
63. Lobo, J., Jiménez-Valverde, A. & Real, R. AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **17**, 145–151. <https://doi.org/10.1111/j.1466-822X.2007.00358.x> (2008).
64. Lobo, J. M., Jiménez-Valverde, A. & Real, R. AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **17**, 145–151 (2008).
65. Shabani, F., Kumar, L. & Ahmadi, M. Assessing accuracy methods of species distribution models: AUC, specificity, sensitivity and the true skill statistic. *Glob. J. Hum. Soc. Sci. Res.* **18**, 6–18 (2018).
66. Monserud, R. A. & Leemans, R. Comparing global vegetation maps with the Kappa statistic. *Ecol. Model.* **62**, 275–293 (1992).
67. Kim, H. W., Adhikari, P., Chang, M. H. & Seo, C. Potential distribution of amphibians with different habitat characteristics in response to climate change in South Korea. *Animals* **11**, 2185. <https://doi.org/10.3390/ani11082185> (2021).

## Acknowledgements

This research was supported by the Ministry of Environment, Republic of Korea (Grant no. 2018002270001).

## Author contributions

P.A. and Y.H.L. designed the research and methodology. P.A. and A.P. collected the data. P.A. performed modeling, analysis, and prepared the original manuscript. P.A. and Y.-S.P. reviewed and edited the manuscript. S.H.H. performed project administration and funding acquisition. S.H.H. and Y.-S.P. supervised this study. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-36358-z>.

**Correspondence** and requests for materials should be addressed to S.H.H. or Y.-S.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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/>.

© The Author(s) 2023