Potential distribution of *Artemisia annua* L.under climate change

Hongjun Jiang[1], Ting Liu[2], Shiping Gao[1], Ruijun Wang[1], Ruchun Zhang[1], Henan Meng[1], and Lijuan Li[3]

[1]Institute of Geographical Sciences, Hebei Academy of Sciences
[2]College of Teachers' Education, Hebei Normal Universtiy
[3]Senckenberg Research Institute

March 3, 2021

Abstract

Aim: *Artemisia annua* L. is the one and only original plant used to isolate artemisinin which is a highly effective remedy to fight malaria. Climate change leads to change of distribution and suitable range for many species and *A. annua* is no exception. However, it is not clear that the potential distribution and suitable range change of this unique plant under climate change. Therefore, we present this research to study its change in the future. **Location:** Global. **Methods:** Since the accuracy of species distribution models was affected by occurrence records and environmental variables, 1062 presence records and 7 variables were picked out to build ensemble models with 10 different algorithms by means of biomod2 under current and future climate scenarios. **Results:** At present, except SRE, the AUC values of the rest models were greater than 0.8, and the TSS values were greater than 0.6, the values of ensemble model were 0.968 and 0.826 respectively. Mean temperature of driest quarter was the dominant factor to shape the range of *A. annua* and its optimum interval ranged from 4.8 to 23.3ºC. The high suitable habitats of *A. annua* were mainly located in Eastern Asia, Western Europe, Central Europe. In the future, the high suitable area would decline at 15.55% to 25.87%. **Main conclusions:** Ensemble models showed it performed better than any the single one. At present, the high suitable habitat simulated by ensemble model was in accordance with the actual occurrence records. In the future, the high suitable habitat for *A. annua* would move northeast, and disappear in North America. They would increase with time under each SSP, but sharply decline while comparing with the current one. This study can be used to protect wild resource and guide cultivation for *A. annua*, which would make modest contribution to fight malaria.

Title

Potential distribution of *Artemisia annua* L. under climate change

Running title

Distribution of *Artemisia annua*

Abstract

Aim: *Artemisia annua* L. is the one and only original plant used to isolate artemisinin which is a highly effective remedy to fight malaria. Climate change leads to change of distribution and suitable range for many species and *A. annua* is no exception. However, it is not clear that the potential distribution and suitable range change of this unique plant under climate change. Therefore, we present this research to study its change in the future.

Location: Global.

Methods: Since the accuracy of species distribution models was affected by occurrence records and environmental variables, 1062 presence records and seven variables were picked out to build ensemble models with 10 different algorithms by means of biomod2 under current and future climate scenarios.

Results: At present**,** except SRE, the AUC values of the rest models were greater than 0.8, and the TSS values were greater than 0.6, the values of ensemble model were 0.971 and 0.839 respectively. Mean temperature of driest quarter was the dominant factor to shape the range of *A. annua* and its optimum interval ranged from 4.8 to 23.3ºC. The high suitable habitats of *A. annua* were mainly located in Eastern Asia, Western Europe, Central Europe. In the future, the high suitable area would decline at 15.55% to 25.87%.

Main conclusions: Ensemble models showed it performed better than any the single one. At present, the high suitable habitat simulated by ensemble model was in accordance with the actual occurrence records. In the future, the high suitable habitat for *A. annua* would move northeast, and disappear in North America. They would increase with time under each SSP, but sharply decline while comparing with the current one. This study can be used to protect wild resource and guide cultivation for *A. annua* , which would make modest contribution to fight malaria.

Introduction

Artemisia annua L. also called as annual wormwood, or huang hua hao, or sweet Annie, or sweet wormwood, is an annual herb, belong to Asteraceae(Lin, Humphries, & Gilbert, 2011). Its dry aboveground part is a traditional Chinese medicine called Qinghao(Tang & Eisenbrand, 1992). Qinghaosu (also named artemisinin or atreannuin), known as the frontline treatment of malaria(L. H. Miller & Su, 2011), was isolated only from *A. annua* with none yield in 30 other species of *Artemisia* (Klayman, 1985). Additional,*A. annua* has strong allelopathy(Lydon, Teasdale, & Chen, 1997). During its growth process, some secondary metabolites could be released to affect their surrounding plants(Knudsmark Jessing, Duke, & Cedergreeen, 2014). Therefore, it is an important resource plant not only for medical purpose to cure human diseases, but also for agricultural purpose to exploit natural herbicide.

In 2017, there were still an estimated 219 million malaria cases and 435000 deaths globally(WHO, 2018). Reliable supply of artemisinin which is the core component of ACTs (artemisinin-based combination therapies)- the most effective antimalarial drugs- is an urgent need(Kayani, Kiani, Dilshad, & Mirza, 2018). Besides exploring new synthesis pathway of artemisinin, enlarging cultivated area of *A. annua* would be the best way to ease this urgency while its wild resource is limited. Meanwhile, climate change has shift the ranges of many species(Urban, 2015). So, projecting its potential distribution, predicting its range shift under different climate scenarios and understanding the relationship between its distribution and environmental factors are very necessary. Up to now, there are many researches on isolation(Klayman et al., 1984), pharmacological analysis(Efferth, 2006) and artificial synthesis(Farhi, Kozin, Duchin, & Vainstein, 2013) of artemisinin, climate suitability analysis and relationship between artemisinin content and climate factors. It is rare to see some related report or studies about the range shift of *A. annua* under future climate scenarios. Species distribution models (SDM)(Austin, 2007) and global climate model(Comer, Fenech, & Gough, 2007) make it possible for us to simulate and foresee the range change of species under different future climate scenarios.

With niche theory as its kernel, SDM is the synthesis of geographic information systems (GIS), ecology and statistics(Jane Elith & Franklin, 2013; A. Guisan et al., 2013). Based on species occurrences and environmental variables, SDMs could work out some results using specific algorithms. These results can be explained as species richness, habitat suitability or probability of species presence(J. Elith & Leathwick, 2009). Nowadays, SDM is widely used in many research fields, such as Conservation Biology, Invasion Biology, Biogeography, Lemology, and impact of climate change(Jane Elith & Franklin, 2013; J. Elith & Leathwick, 2009; J. Miller, 2010; Phillips, Anderson, & Schapire, 2006).

Many global climate models are developed to simulate the future climate for the Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC)(Eyring et al., 2016). The past, current and future climate data could be accessed conveniently from the Coupled Model Intercomparison Project (CMIP), and facilitate large scale research on impact of climate change(Balaji et al., 2018).

In this study, by means of SDMs and climate data, we simulated the present and future potential distributions of *A. annua* based on its occurrence records around the world. The following questions were addressed: (1) what is the potential distribution pattern of *A. annua* under present climate? (2) which bioclimatic variable is the dominant variable to shape the range of this species? (3) How do the future distributions shift under different future climate scenarios? (4) What can be inferred from these results for its plant plan and management?

Methods

Species occurrence

In the study, 2130 occurrence records of *A. annua* were collected for modelling its potential distribution. Most of these data were retrieved from several online databases. They were Global Biodiversity Information Facility (GBIF,*http://www.gbif.org/*), Tropicos Database (*http://www.tropicos.org/*), National Specimen Information Infrastructure (NSII,*http://www.nsii.org.cn/*) and Chinese Virtual Herbarium (CVH,*http://www.cvh.ac.cn/*). The others were acquired from field investigation. The detailed addresses of records were georeferenced by means of GPS and Google map. In order to reduce the negative effects caused by high-density collection, SDMtoolbox(Brown, 2014) was introduced to spatially rarefy the occurrence data and remove duplicate records. Each grid cell (˜10 km×10 km) just kept one record. Thus, 1062 presence records remained for building models (Fig. 1).

Environmental variables

Altitude and a set of 19 bioclimatic variables were adopted in our research (Table 1). These bioclimatic variables were derived from the monthly meteorological data(Fick & Hijmans, 2017). They could be clustered into four groups which represented annual trends, seasonality and extreme or limiting environmental factors(van Zonneveld, Castaneda, Scheldeman, van Etten, & Van Damme, 2014). All these variables were at 5 arc-minutes spatial resolution (˜10 km×10 km) and with the range of 180°W to 180°E longitude and 60°S to 84°N latitude.

The altitude data were obtained from CGIAR-CSI (available at*https://srtm.csi.cgiar.org/*). The current global bioclimatic variables data were download from CHELSA (available at http://chelsa-climate.org/). They were the average values for the period 1979-2013(Karger et al., 2017). The future ones were derived from the projections of eight global climate models (GCMs): BCC-CSM2-MR, CNRM-CM6-1, CNRM-ESM2-1, CanESM5, IPSL-CM6A-LR, MIROC-ES2L, MIROC6, MRI-ESM2-0. These GCMs were selected according to their data availability and involvement in relevant model intercomparison projects within the CMIP6(Moseid et al., 2020). The future climate data were 20-year averages for 2021-2040, 241-2060, 2061-2080 and 2081-2100 and for four Shared Socio-economic Pathways (SSPs): 126, 245, 370 and 585. These data were downloaded from WorldClim (available at http://www.worldclim.org). In order to reduce the bias in certain area from sole GCM, multi-model ensemble (MME) was adopted to derived the average values for future climates(Mendlik & Gobiet, 2016; Pierce, Barnett, Santer, & Gleckler, 2009). Since collinearity-correlation among variables would result in accuracy drop of SDMs, total seven variables (shown in bold in Table 1) were picked out from 20 environmental variables according to multiple correlation coefficient ($|R| < 0.6$) and variance inflation factor (VIF < 10) between each other(Naimi & Araújo, 2016).

Modelling procedure

Biomod2 was a R package which was developed for species distribution modelling. It included 10 different species distribution algorithms, such as Artificial Neural Network (ANN), Classification Tree Analysis (CTA), Flexible Discriminant Analysis (FDA), Generalized Additive Model (GAM), Generalized Boosting Model (GBM), Generalized Linear Model (GLM), Multiple Adaptive Regression Splines (MARS), Maximum Entropy (MAXENT), Random Forests (RF) and Surface Range Envelop (SRE)(Wilfried Thuiller, Georges, Engler, & Breiner, 2016). Since absence records requirement of several algorithms above, 10000 pseudo-absences were selected randomly 5 times to follow a common strategy(Antoine Guisan, Thuiller, & Zimmermann, 2017; Merow, Smith, & Silander, 2013). 80% of presence and pseudo-absence data were

used to calibrate the models, and the rest was used for model testing(Antoine Guisan et al., 2017). The model calibrations and evaluations would repeat 10 times. Response curves and relative contributions of each variable involved were calculated. AUC (area under the receiver operating characteristic curve) and TSS (true skill statistics) were employed as performance evaluation criteria for 10 different algorithms. The models whose AUC was greater than 0.90 and TSS was greater than 0.6 were considered to be with good performance in species distribution modelling(Allouche, Tsoar, & Kadmon, 2006; Swets, 1988). By means of biomod2, an ensemble modelling approach were adopted to build the ensemble models for eliminating model bias caused by model selection. Using these ensemble models, the potential distributions under current and future climate scenarios were projected for *A. annua* . And finally, based on these projections, distribution patterns, range sizes and shifts under different climate scenarios were analyzed and compared for *A. annua* .

Results

Model performances

There were 10 algorithms used to build the ensemble models for *A. annua*in the research. The AUC and TSS values of each algorithms and ensemble models were calculated to evaluate their performance under different climate scenarios.

The TSS results were showed in Table 2. The Average TSS value of RF was 0.835. It got the first place in 10 algorithms, followed by GBM, CTA, MAXENT, GAM, ANN, MARS, GLM and FDA. The last one was SRE, its average TSS value was 0.494. The average TSS value for Ensemble was 0.839, it was greater than any other individual algorithm.

The AUC values were showed in Table3. The average AUC value of RF was 0.964. It was the maximum in 10 algorithms, followed by GBM, MAXENT, GAM, MARS, CTA, ANN, FDA and GLM. The last one continued to be SRE, its average AUC value was 0.747. The average AUC value for Ensemble was 0.971, it was greater than any other individual algorithm too.

Contributions of climatic factors

Potential distribution of species strongly depended on environmental factors, especially climatic factors. During model simulating, variable importance and response curves were figured out to indicate the relationship between distributions and environmental variables (Fig. 2a). Each factor played a different part in different algorithms, and showed out different importance. Dominant environmental factors were figured out for each algorithm. BIO9 (mean temperature of driest quarter) performed the best among seven variables in all of 10 algorithms. According its response curves plotted out in Fig. 2b, the proper mean temperature of driest quarter for *A. annua* ranged from 4.8 to 23.3.

Current and future potential distributions

Potential distributions of *A. annua* were projected by ensemble models under current and future climate scenarios. Based on the value of presence probability, the distribution could be categorized into three classes respectively. The area with value ranged from 0-0.2 was low suitable area, the area with value ranged from 0.2-0.5 was medium suitable area, and the area with value ranged from 0.5-1 was high suitable area.

According to the projection of ensemble model under current climate scenario, the high suitable areas for *A. annua* were located in Eastern Asia, Western Europe, Central Europe, eastern part of Northern Europe, western part of Eastern Europe, south and east edge of Australia, eastern coast of North America, northwest part of Great Plains in U.S., southeast coast and Andes Mountains in South America (Fig. 3). The medium suitable areas were located in west of China, Eastern Europe, central and east of U.S. and southeast of South America.

Due to different climate scenarios, the future potential distributions of*A. annua* differed. The outputs showed that, the pixel numbers for high suitable habitats increased with time under each SSP (Fig. 4). The

largest growth occurred under SSP 585, its percentage raised up to 7.18% (average of 2080-2100) from 6.34% (average of 2021-2040).

Although, compared with the current one, the pixel numbers declined sharply under future climate scenarios. The percentages ranged from 15.55% (average of 2081-2100 under SSP 585) to 25.87% (average of 2021-2040 under SSP 126). After analyzing the projections of ensemble models under different climate scenarios and periods, we figured out the gain and loss of high suitable habitats in the future. According to the result, the shift patterns of different periods under SSP 126 was demonstrated below (Fig. 5). Almost all the high suitable habitats which located in North America degraded. The high suitable habitats in Europe presented a trend moving toward east, they were loss in Atlantic coast and gain in East Europe. In Asia, the habitat shift area mainly located around China, most of degraded areas were in northwest China and north of China, major parts of upgraded area spotted along the border of southwest China and south China. Under the other three SSPs, the gain and loss acted a same pattern as the one under SSP 126, which could be checked in supplementary materials (see Appendix S1).

Discussion

In our study, ensemble models were developed for *A. annua* . With good accuracy, potential distributions were simulated and suitable habitats were estimated under present and future climate scenarios. The simulations indicated that the mean temperature of driest quarter was the dominant bioclimatic variable. It also showed that a huge change both on its range and location which would be caused by climate change. This result clearly pointed out that this herb would be dramatically influenced by climate change. Its future conservation and utilization must consider the impact of climate change.

Model performance

Ensemble models were widely adopted in many researches related to species distribution(Hao, Elith, Lahoz-Monfort, & Guillera-Arroita, 2020). Previous studies showed that uncertainty of model result would be brought out by many factors such as species distribution models, dispersal strategies, environmental factors and so on(W. Thuiller, Gueguen, Renaud, Karger, & Zimmermann, 2019). Although there were lack of strong quantitative conclusions about the predictive performance of ensemble models(Hao, Elith, Guillera-Arroita, & Lahoz-Monfort, 2019). There was a popular belief that ensemble model would reduce the uncertainty due to sole algorithm, and it would perform better than sole one(Araujo et al., 2019; Araujo & New, 2007). In our study, except SRE, the AUC values of all the other nine models were greater than 0.8, and the TSS values were greater than 0.6. The values of ensemble model were 0.971 and 0.835 respectively. Based on this, an obvious result could easily draw out that ensemble model outperformed any sole algorithm, which was consistent with the conclusion ahead. This result also showed that the simulation was with high accuracy and could be used to predict the current and future potential distributions of *A. annua* .

Dominant factors

Any organisms cannot live without its biotic and abiotic elements around. Their distribution range was determined with many factors, such as climate, soil, topography, biology, history, human activity and so on. *A. annua* was no exception. In our study, 19 bioclimatic factors and altitude were prepared for distribution modeling. After analyzing their correlations between each other, seven variables were considered to build species distribution models for *A. annua* . Due to their different basic theories and emphases, the same factors behaved diversely. But BIO9 (mean temperature of driest quarter) occupied the leading position in every models. This bioclimatic variable played an important role in shaping the distribution of *A. annua.*

Potential distribution

A. annua is an annual plant. Its native habitat was located in Asia, most likely in China. Now, it grew in every continent except Antarctic (Ferreira, Laughlin, Delabays, & de Magalhaes, 2005). Accompany with development of cultivating techniques and growing demand of raw material for extracting artemisinin, it had been introduced in for experimental cultivation or adopted for breeding in non-native habitats(Ferreira et al., 2005; Woerdenbag et al., 1994). The simulation showed its high suitable habitat was located in

many countries, such as USA, Brazil, Argentina, Australia, Malaysia, Vietnam, Thailand, Japan, Korea, India, Ukraine, Poland, Czech, Slovenia, Hungary, Romania, Austria, Italy, Sweden, German, France, UK, Madagascar, Tanzania, South Africa and so on(Laughlin, Heazlewood, & Beattie, 2002). It was in consistent with the distribution of its occurrence records. Compared with other similar studies, it showed that our simulation basically coincided with others(Qin, Jian, Jie, Chengzhong, & Caixiang, 2018; Wang, Li, Zeng, & Xie, 2015). But it was obvious that the high suitable habitat area in their study was bigger than that in ours(Ding et al., 2020). This may be caused by different quantity of adopted algorithms and occurrence records. Only boosted regression tree (BRT) was used in their research. The uncertainty of single model was unavoidable in that situation.

Shift of suitable habitat in Future

Range of species would shift with climate change, which had been demonstrated by a larger number of observation and simulating studies on species distribution(Bellard, Bertelsmeier, Leadley, Thuiller, & Courchamp, 2012; Urban, 2015). Medicinal plants, as an important class of resource plants, were always the first target of species distribution modeling. According to previous studies, high suitable habitat of some species would shrink, even disappear, others would increase, and still others would remain stable in the future. Kurpis(Kurpis, Serrato-Cruz, & Feria Arroyo, 2019), Zhao(Zhao et al., 2020) and Xie(Xie et al., 2018) found the potential distribution of *Tagetes lucida* Cav. (Asteraceae), *Notopterygium incisum* Ting ex H. T. Chang (Apiaceae) and *Eriocaulon buergerianum* Kornicke. (Eriocaulaceae) would decrease respectively in the future. Asase(Asase & Peterson, 2019) considered that the high suitable habitat of *Alstonia boonei* De Wild (Apocynaceae) would be stable under future climate scenarios. Kunwar(Kunwar et al., 2020) pointed out that there would be more high suitable habitat for *Dactylorhiza hatagirea* (D. Don) Soo (Orchidaceae) in the future. Shen(Shen et al., 2019) indicated that adaptive cultivation regions for *Panax quinquefolius* L. (Araliaceae) would expand by 9.16%–30.97% in 2070s. In our study, the projections revealed that ranges of *A. annua* would change dramatically in the future, its area of high suitable habitats would decline in all mentioned periods and climate scenarios.

Impact of climate change

Global climate is changing but its rate and impacts are localized(IPCC, 2013). Amplitude of the variation differed from area to area under future climate change scenarios. So, shifts of high suitable habitat of *A. annua* would present out different patterns in different regions in the future scenarios. In our study, the ranges of high suitable area of *A. annua* in some areas would expand, some would remain stable, some would shrink, even disappear. In Europe, its high suitable habitat would shrink in the coast area of Atlantic Ocean, and expand in the south of Finland and the northwest of Russia. A good agreement could be observed between our simulation and the one made in other studies. Thuiller (W. Thuiller, Lavorel, Araujo, Sykes, & Prentice, 2005) stated that the total suitable area of forest herb in Europe would move strongly northwards and eastwards under future climate scenarios. In Asia, the expand area scatters in southwest border of China, and shrinks in 40degN in China. This may be caused by the unique topography. There were many mountains. And it should be point out that the high suitable habitat for *A. annua* in North America almost disappeared in the future. This was in accordance with previous studies about extinction risk of North American seed plant(Zhang et al., 2017). They predicted that most species belong in Artemisia would suffer distribution loss in the future.

In conclusion, we simulated the potential distributions of *A. annua* under current and future climate scenarios across the world, and predicted its range shift in different periods and Shared Socio-economic Pathways. In our study, with better performance than sole species distribution model, ensemble model was qualified for simulating the potential distribution of *A. annua* . Among 20 environmental factors, BIO9 (Mean temperature of driest quarter) had the strongest ability of shaping the range of the target species. Its potential distribution under present climate scenario was in accordance with the actual range. It was predicted that the high suitable habitat of *A. annua* would decrease under the climate change in the future. Its high suitable habitats would disappear in North America, and move northeastwards in the Europe. Our simulation and forecast could be used to protect the wild resource of *A. annua* and provide theory guidance

for field selection in its cultivation and management. Finally, we hoped that this study could make modest contribution to fight malaria.

Conflict of interest

The authors have no competing interests to declare.

Data availability statement

All environmental variables used in the manuscript are already publicly accessible, and we have provided the download address in the manuscript; relevant sampling site information can be found in Table S1 in the supporting materials.

Reference

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43* , 1223-1232. doi:10.1111/j.1365-2664.2006.01214.x

Araujo, M. B., Anderson, R. P., Marcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., . . . Naimi, B. (2019). Standards for distribution models in biodiversity assessments. *Science Advances, 5* , eaat4858. doi:10.1126/sciadv.aat4858

Araujo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in ecology & evolution, 22* , 42-47. doi:10.1016/j.tree.2006.09.010

Asase, A., & Peterson, A. T. (2019). Predicted impacts of global climate change on the geographic distribution of an invaluable African medicinal plant resource, *Alstonia boonei* De Wild. *Journal of Applied Research on Medicinal and Aromatic Plants, 14* , 100206. doi:10.1016/j.jarmap.2019.100206

Austin, M. (2007). Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling, 200* , 1-19. doi:10.1016/j.ecolmodel.2006.07.005

Balaji, V., Taylor, K. E., Juckes, M., Lawrence, B. N., Durack, P. J., Lautenschlager, M., . . . Williams, D. (2018). Requirements for a global data infrastructure in support of CMIP6. *Geosci. Model Dev., 11* , 3659-3680. doi:10.5194/gmd-11-3659-2018

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity.*Ecology Letters, 15* , 365-377. doi:10.1111/j.1461-0248.2011.01736.x

Brown, J. L. (2014). SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology & Evolution, 5* , 694–700. doi:10.1111/2041-210X.12200

Comer, A., Fenech, A., & Gough, W. (2007). Selecting a global climate model for understanding future scenarios of climate change. In A. Fenech & J. MacLellan (Eds.), *Linking Climate Models to Policy and Decision-Making* (pp. 133-145). Tornnto,Ontario: Environment Canada.

Ding, F., Ma, T., Hao, M., Wang, Q., Chen, S., Wang, D., . . . Jiang, D. (2020). Mapping Worldwide Environmental Suitability for *Artemisia annua* L. *Sustainability, 12* , 1-10. doi:10.3390/su12041309

Efferth, T. (2006). Molecular pharmacology and pharmacogenomics of artemisinin and its derivatives in cancer cells. *Current drug targets, 7* , 407-421. doi:10.2174/138945006776359412

Elith, J., & Franklin, J. (2013). Species Distribution Modeling. In S. A. Levin (Ed.), *Encyclopedia of Biodiversity (Second Edition)*(pp. 692-705). Waltham: Academic Press.

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time.*Annual Review of Ecology Evolution and Systematics, 40* , 677-697. doi:10.1146/annurev.ecolsys.110308.120159

7

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *geoscientific Model Development, 9* , 1937-1958. doi:10.5194/gmd-9-1937-2016

Farhi, M., Kozin, M., Duchin, S., & Vainstein, A. (2013). Metabolic engineering of plants for artemisinin synthesis. *Biotechnology and Genetic Engineering Reviews, 29* , 135-148. doi:10.1080/02648725.2013.821283

Ferreira, J. F. S., Laughlin, J. C., Delabays, N., & de Magalhaes, P. M. (2005). Cultivation and genetics of *Artemisia annua* L. for increased production of the antimalarial artemisinin. *Plant Genetic Resources, 3* , 206-229. doi:10.1079/PGR200585

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology, 37* . doi:10.1002/joc.5086

Guisan, A., Thuiller, W., & Zimmermann, N. (2017). *Habitat suitability and distribution models: With applications in R* . Cambridge, United Kingdom: Cambridge University Press.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., . . . Buckley, Y. M. (2013). Predicting species distributions for conservation decisions.*Ecology Letters, 16* , 1424-1435. doi:10.1111/ele.12189

Hao, T., Elith, J., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2019). A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and distributions, 25* , 839-852. doi:10.1111/ddi.12892

Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2020). Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models.*Ecography, 43* , 549-558. doi:10.1111/ecog.04890

IPCC. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* . Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Karger, D. N., Conrad, O., Bohner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., . . . Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data, 4* , 170122. doi:10.1038/sdata.2017.122

Kayani, W. K., Kiani, B. H., Dilshad, E., & Mirza, B. (2018). Biotechnological approaches for artemisinin production in Artemisia.*World Journal of Microbiology & Biotechnology, 34* , 54. doi:10.1007/s11274-018-2432-9

Klayman, D. L. (1985). Qinghaosu (artemisinin): an antimalarial drug from China. *Science, 228* , 1049-1055. doi:10.1126/science.3887571

Klayman, D. L., Lin, A. J., Acton, N., Scovill, J. P., Hoch, J. M., Milhous, W. K., . . . Dobek, A. S. (1984). Isolation of artemisinin (qinghaosu) from *Artemisia annua* growing in the United States.*Journal of Natural Products, 47* , 715-717. doi:10.1021/np50034a027

Knudsmark Jessing, K., Duke, S. O., & Cedergreen, N. (2014). Potential ecological roles of artemisinin produced by *Artemisia annua* L.*Journal of Chemical Ecology, 40* , 100-117. doi:10.1007/s10886-014-0384-6

Kunwar, R. M., Rimal, B., Sharma, H. P., Poudel, R. C., Pyakurel, D., Tiwari, A., . . . Bussmann, R. W. (2020). Distribution and habitat modeling of *Dactylorhiza hatagirea* (D. Don) Soo, Paris polyphylla Sm. and Taxus species in Nepal Himalaya. *Journal of Applied Research on Medicinal and Aromatic Plants* , 100274. doi:10.1016/j.jarmap.2020.100274

Kurpis, J., Serrato-Cruz, M. A., & Feria Arroyo, T. P. (2019). Modeling the effects of climate change on the distribution of *Tagetes lucida* Cav. (Asteraceae). *Global Ecology and Conservation, 20* , e00747. doi:10.1016/j.gecco.2019.e00747

8

Laughlin, J. C., Heazlewood, G. N., & Beattie, B. (2002). Cultivation of *Artemisia annua* L. . In C. W. Wright (Ed.), *Artemisia*(pp. 159 -195). London: Taylor & Francis.

Lin, Y., Humphries, C. J., & Gilbert, M. G. (2011). Artemisia. In Z. Y. Wu, P. H. Raven, & D. Y. Hong (Eds.), *Flora of China Volume 20-21 (Asteraceae).* (pp. 676-737). Beijing: Science Press and St. Louis: Missouri Botanical Garden Press.

Lydon, J., Teasdale, J. R., & Chen, P. K. (1997). Allelopathic activity of annual wormwood (*Artemisia annua*) and the role of artemisinin.*Weed Science, 45* , 807-811. doi:10.1007/s12032-013-0675-8

Mendlik, T., & Gobiet, A. (2016). Selecting climate simulations for impact studies based on multivariate patterns of climate change.*Climatic Change, 135* , 381-393. doi:10.1007/s10584-015-1582-0

Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions what it does, and why inputs and settings matter. *Ecography, 36* , 1058-1069. doi:10.1111/j.1600-0587.2013.07872.x

Miller, J. (2010). Species distribution modeling. *Geography Compass, 4* , 490-509. doi:10.1111/j.1749-8198.2010.00351.x

Miller, L. H., & Su, X. (2011). Artemisinin: Discovery from the Chinese Herbal Garden. *Cell, 146* , 855-858. doi:10.1016/j.cell.2011.08.024

Moseid, K., Michael, S., Storelvmo, T., Julsrud, I., Olivie, D., Nabat, P., . . . Gastineau, G. (2020). Bias in CMIP6 models as compared to observed regional dimming and brightening. *Atmospheric Chemistry and Physics, 20* , 16023-16040. doi:10.5194/acp-20-16023-2020

Naimi, B., & Araujo, M. B. (2016). sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography, 39* , 368-375. doi:10.1111/ecog.01881

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190* , 231-259. doi:10.1016/j.ecolmodel.2005.03.026

Pierce, D. W., Barnett, T. P., Santer, B. D., & Gleckler, P. J. (2009). Selecting global climate models for regional climate change studies.*Proceedings of the National Academy of Sciences of the United States of America, 106* , 8441-8446. doi:10.1073/pnas.0900094106

Qin, Z., Jian, W., Jie, G., Chengzhong, S., & Caixiang, X. (2018). Production regionalization of medicinal sweet wormwood *Artemisia annua* . *Journal of Plant Protection* . doi:10.13802/j.cnki.zwbhxb.2018.2017029

Shen, L., Li, X.-W., Meng, X.-X., Wu, J., Tang, H., Huang, L.-F., . . . Chen, S.-L. (2019). Prediction of the globally ecological suitability of*Panax quinquefolius* by the geographic information system for global medicinal plants (GMPGIS). *Chinese Journal of Natural Medicines, 17* , 481-489. doi:10.1016/S1875-5364(19)30069-X

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems.*Science, 240* , 1285-1293. doi:10.1126/science.3287615

Tang, W., & Eisenbrand, G. (1992). *Artemisia annua* l L. In*Chinese drugs of plant origin* (pp. 159-174). Berling, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, Barcelona, Budapest: Springer-Verlag.

Thuiller, W., Georges, D., Engler, R., & Breiner, F. (2016). biomod2: Ensemble Platform for Species Distribution Modeling.

Thuiller, W., Gueguen, M., Renaud, J., Karger, D. N., & Zimmermann, N. E. (2019). Uncertainty in ensembles of global biodiversity scenarios.*nature communications, 10* , 1446. doi:10.1038/s41467-019-09519-w

9

Thuiller, W., Lavorel, S., Araujo, M. B., Sykes, M. T., & Prentice, I. C. (2005). Climate change threats to plant diversity in Europe.*Proceedings of the National Academy of Sciences, 102* , 8245-8250. doi:10.1073/pnas.0409902102

Urban, M. C. (2015). Climate change. Accelerating extinction risk from climate change. *Science, 348* , 571-573. doi:10.1126/science.aaa4984

van Zonneveld, M., Castaneda, N., Scheldeman, X., van Etten, J., & Van Damme, P. (2014). Application of consensus theory to formalize expert evaluations of plant species distribution models. *Applied Vegetation Science, 17* , 528-542. doi:10.1111/avsc.12081

Wang, H., Li, H., Zeng, F. L., & Xie, C. X. (2015). Spatial Distribution and Global Potential Suitability Regions of *Artemisia annua* . *Zhong Yao Cai, 38* , 460-466. doi:10.13863/j.issn1001-4454.2015.03.009

WHO. (2018). *World malaria report 2018* . Geneva: World Health Organization.

Woerdenbag, H. J., Pras, N., Chan, N. G., Bang, B. T., Bos, R., van Uden, W., . . . Lugt, C. B. (1994). Artemisinin, Related Sesquiterpenes, and Essential Oil in *Artemisia annua* During a Vegetation Period in Vietnam. *Planta Medica, 60* , 272-275. doi:10.1055/s-2006-959474

Xie, Y.-F., Yang, L., Deng, R.-Y., Chen, M.-H., Luan, X.-F., Gottardi, E., & Zhang, Z.-X. (2018). Changes in the range of the medicinal herb*Eriocaulon buergerianum* Kornicke. (Eriocaulaceae) under climate change. *Plant Biology, 20* , 771-779. doi:10.1111/plb.12836

Zhang, J., Nielsen, S. E., Chen, Y., Georges, D., Qin, Y., Wang, S. S., . . . Thuiller, W. (2017). Extinction risk of North American seed plants elevated by climate and land-use change. *Journal of Applied Ecology, 54* , 303-312. doi:10.1111/1365-2664.12701

Zhao, Z., Guo, Y., Wei, H., Ran, Q., Liu, J., Zhang, Q., & Gu, W. (2020). Potential distribution of *Notopterygium incisum* Ting ex H. T. Chang and its predicted responses to climate change based on a comprehensive habitat suitability model. *Ecology and Evolution, 10* , 3004-3016. doi:10.1002/ece3.6117

Tables

Table 1 Environmental variables used in modeling potential distribution of *A. annua*

Name	Description
ALT	**Altitude**
BIO1	Annual Mean Temperature
BIO2	**Mean Diurnal Range (Mean of monthly (max temp - min temp))**
BIO3	Isothermality (BIO2/BIO7) ($\times$100)
BIO4	Temperature Seasonality (standard deviation $\times$100)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	**Mean Temperature of Wettest Quarter**
BIO9	**Mean Temperature of Driest Quarter**
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	**Precipitation Seasonality (Coefficient of Variation)**
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	**Precipitation of Warmest Quarter**

Name	Description
BIO19	**Precipitation of Coldest Quarter**

Table 2 TSS of 10 algorithms applied in modeling potential distribution of *A. annua* under current and future climate scenarios. The color from blue to red represented the TSS values got bigger and bigger.

SSP[a]	Periods	SRE[b]	FDA[c]	GLM[d]	MARS[e]	ANN[f]	GAM[g]	MAXENT[h]	CTA[i]	GBM[j]	RF
Current	Current	0.457	0.557	0.572	0.661	0.636	0.638	0.716	0.754	0.772	0.81
SSP126	2021-2040	0.490	0.622	0.647	0.704	0.705	0.720	0.759	0.764	0.792	0.83
SSP126	2041-2060	0.492	0.614	0.642	0.693	0.696	0.704	0.762	0.764	0.793	0.83
SSP126	2061-2080	0.494	0.616	0.645	0.704	0.708	0.711	0.773	0.773	0.798	0.83
SSP126	2081-2100	0.483	0.612	0.650	0.697	0.696	0.709	0.760	0.761	0.787	0.83
SSP245	2021-2040	0.497	0.626	0.651	0.700	0.708	0.720	0.762	0.767	0.796	0.83
SSP245	2041-2060	0.491	0.618	0.646	0.701	0.707	0.723	0.771	0.773	0.795	0.83
SSP245	2061-2080	0.497	0.609	0.645	0.697	0.695	0.708	0.767	0.764	0.791	0.83
SSP245	2081-2100	0.490	0.605	0.642	0.691	0.693	0.706	0.763	0.763	0.786	0.83
SSP370	2021-2040	0.498	0.633	0.652	0.705	0.702	0.725	0.768	0.768	0.794	0.83
SSP370	2041-2060	0.502	0.623	0.653	0.695	0.713	0.714	0.764	0.770	0.791	0.83
SSP370	2061-2080	0.509	0.620	0.649	0.698	0.697	0.715	0.766	0.774	0.790	0.84
SSP370	2081-2100	0.503	0.610	0.646	0.695	0.711	0.709	0.764	0.763	0.786	0.83
SSP585	2021-2040	0.501	0.637	0.657	0.706	0.707	0.731	0.759	0.775	0.794	0.84
SSP585	2041-2060	0.497	0.620	0.649	0.705	0.699	0.722	0.763	0.768	0.796	0.83
SSP585	2061-2080	0.495	0.600	0.640	0.687	0.700	0.701	0.760	0.774	0.790	0.83
SSP585	2081-2100	0.507	0.605	0.632	0.685	0.708	0.713	0.758	0.765	0.789	0.84
Average	**Average**	**0.494**	**0.613**	**0.642**	**0.696**	**0.699**	**0.710**	**0.761**	**0.767**	**0.791**	**0.8**

[a] SSP is the abbreviation of Shared Socio-economic Pathways. [b-k] were the short name for following species distribution models. They are Artificial Neural Network (ANN), Classification Tree Analysis (CTA), Flexible Discriminant Analysis (FDA), Generalized Additive Model (GAM), Generalized Boosting Model (GBM), Generalized Linear Model (GLM), Multiple Adaptive Regression Splines (MARS), Maximum Entropy (MAXENT), Random Forests (RF) and Surface Range Envelop (SRE).

Table 3 AUC of 10 algorithms applied in modeling potential distribution of *A. annua* under current and future climate scenarios. The color from blue to red represented the AUC values got bigger and bigger.

SSP[a]	Periods	SRE[b]	GLM[c]	FDA[d]	ANN[e]	CTA[f]	MARS[g]	GAM[h]	MAXENT[i]	GBM[j]	RF
Current	Current	0.729	0.844	0.840	0.857	0.893	0.886	0.866	0.921	0.942	0.96
SSP126	2021-2040	0.745	0.882	0.889	0.892	0.896	0.910	0.916	0.938	0.950	0.96
SSP126	2041-2060	0.746	0.875	0.885	0.886	0.897	0.906	0.910	0.937	0.947	0.96
SSP126	2061-2080	0.747	0.877	0.885	0.897	0.904	0.910	0.914	0.942	0.950	0.96
SSP126	2081-2100	0.742	0.879	0.883	0.886	0.896	0.908	0.912	0.936	0.948	0.96
SSP245	2021-2040	0.748	0.882	0.889	0.895	0.899	0.908	0.916	0.938	0.949	0.96
SSP245	2041-2060	0.745	0.881	0.885	0.895	0.903	0.907	0.915	0.938	0.949	0.96
SSP245	2061-2080	0.748	0.876	0.883	0.890	0.899	0.906	0.913	0.936	0.947	0.96
SSP245	2081-2100	0.745	0.875	0.880	0.889	0.898	0.905	0.911	0.936	0.946	0.96
SSP370	2021-2040	0.749	0.884	0.891	0.894	0.897	0.910	0.917	0.939	0.949	0.96
SSP370	2041-2060	0.751	0.879	0.886	0.895	0.902	0.905	0.912	0.937	0.948	0.96
SSP370	2061-2080	0.755	0.877	0.885	0.891	0.903	0.905	0.913	0.937	0.948	0.96

SSP[a]	Periods	SRE[b]	GLM[c]	FDA[d]	ANN[e]	CTA[f]	MARS[g]	GAM[h]	MAXENT[i]	GBM[j]	RF
SSP370	2081-2100	0.752	0.876	0.882	0.900	0.896	0.906	0.912	0.935	0.946	0.96
SSP585	2021-2040	0.750	0.885	0.893	0.893	0.901	0.911	0.919	0.937	0.949	0.96
SSP585	2041-2060	0.749	0.880	0.885	0.891	0.903	0.909	0.914	0.935	0.949	0.96
SSP585	2061-2080	0.748	0.874	0.874	0.890	0.898	0.906	0.911	0.936	0.949	0.96
SSP585	2081-2100	0.753	0.870	0.873	0.895	0.897	0.901	0.911	0.934	0.948	0.96
Average	**Average**	**0.747**	**0.876**	**0.882**	**0.890**	**0.899**	**0.906**	**0.911**	**0.936**	**0.948**	**0.9**

[a] SSP is the abbreviation of Shared Socio-economic Pathways. [b-k] were the short name for following species distribution models. They are Artificial Neural Network (ANN), Classification Tree Analysis (CTA), Flexible Discriminant Analysis (FDA), Generalized Additive Model (GAM), Generalized Boosting Model (GBM), Generalized Linear Model (GLM), Multiple Adaptive Regression Splines (MARS), Maximum Entropy (MAXENT), Random Forests (RF) and Surface Range Envelop (SRE).

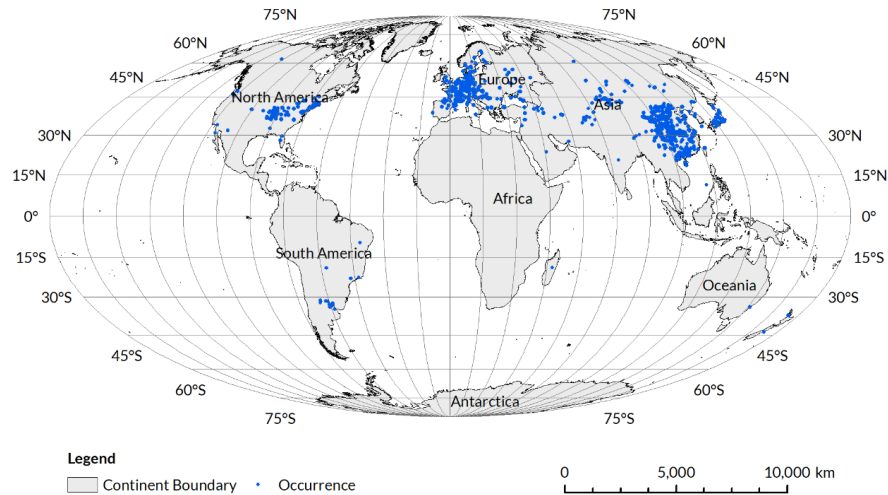**Figure legend and embedded figures**

Fig. 1 Occurrence distribution of *A. annua*

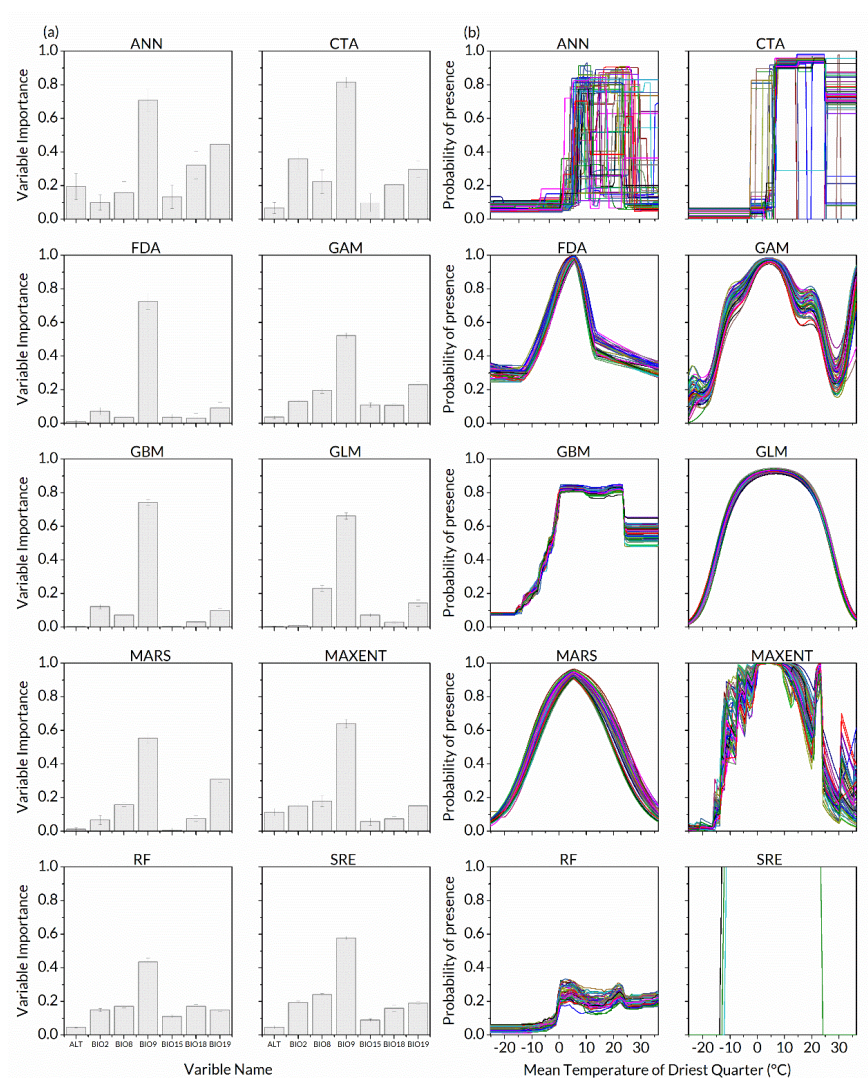(Blue pot represented the location of occurrence)

Fig. 2 (a) Variable importance of seven environmental factors in modeling potential distributions of *A. annua* with 10 different species distribution models. (b) Response curves of BIO9 (mean temperature of driest quarter) for each algorithm in modeling potential distribution of *A. annua.* The 10 species distribution models were Artificial Neural Network (ANN), Classification Tree Analysis (CTA), Flexible Discriminant Analysis (FDA), Generalized Additive Model (GAM), Generalized Boosting Model (GBM), Generalized Linear Model (GLM), Multiple Adaptive Regression Splines (MARS), Maximum Entropy (MAXENT), Random Forests (RF) and Surface Range Envelop (SRE). The seven environmental factors were altitude (ALT), Mean Diurnal Range (Mean of monthly (max temp - min temp)) (BIO2), Mean Temperature of Wettest Quarter (BIO8), Mean Temperature of Driest Quarter (BIO9), Precipitation Seasonality (Coefficient of Variation) (BIO15), Precipitation of Warmest Quarter (BIO18), Precipitation of Coldest Quarter (BIO19).
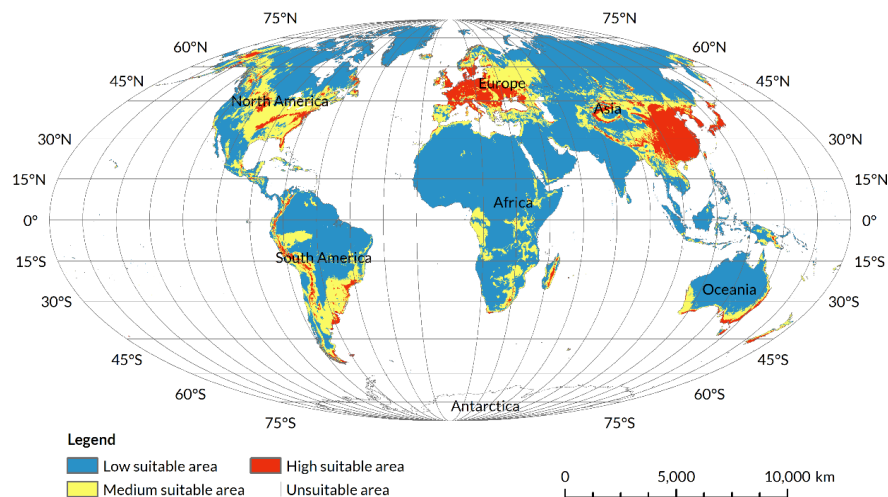
13

Fig. 3 Habitat suitability distribution of *A. annua* projected by ensemble model based on 10 algorithms under current climate scenario.
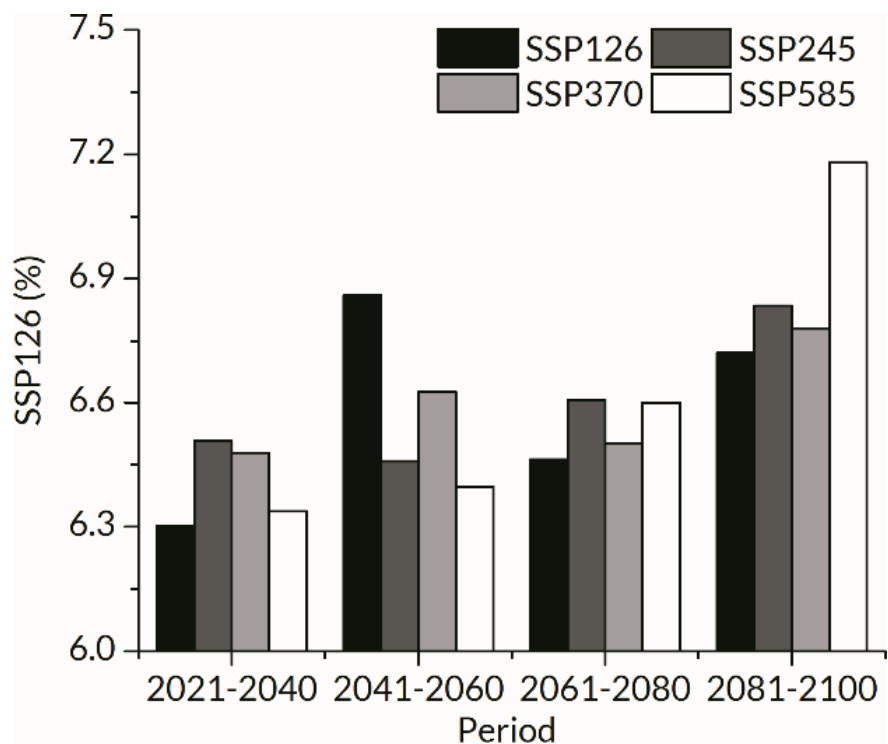

Fig. 4 Proportion of high suitable habitat size for *A. annua* under future climate scenarios. The x axis represented different decades. The y axis represented the proportion of high suitable habitats in total study areas, which was counted by pixel number. SSP is the abbreviation of Shared Socio-economic Pathways.

14

Legend:
- Stable (green)
- Gain (red)
- Loss (blue)
- Unsuitable (grey)
- Non research area

Maps labeled:
- SSP 126 2021-2040
- SSP 126 2041-2060
- SSP 126 2061-2080
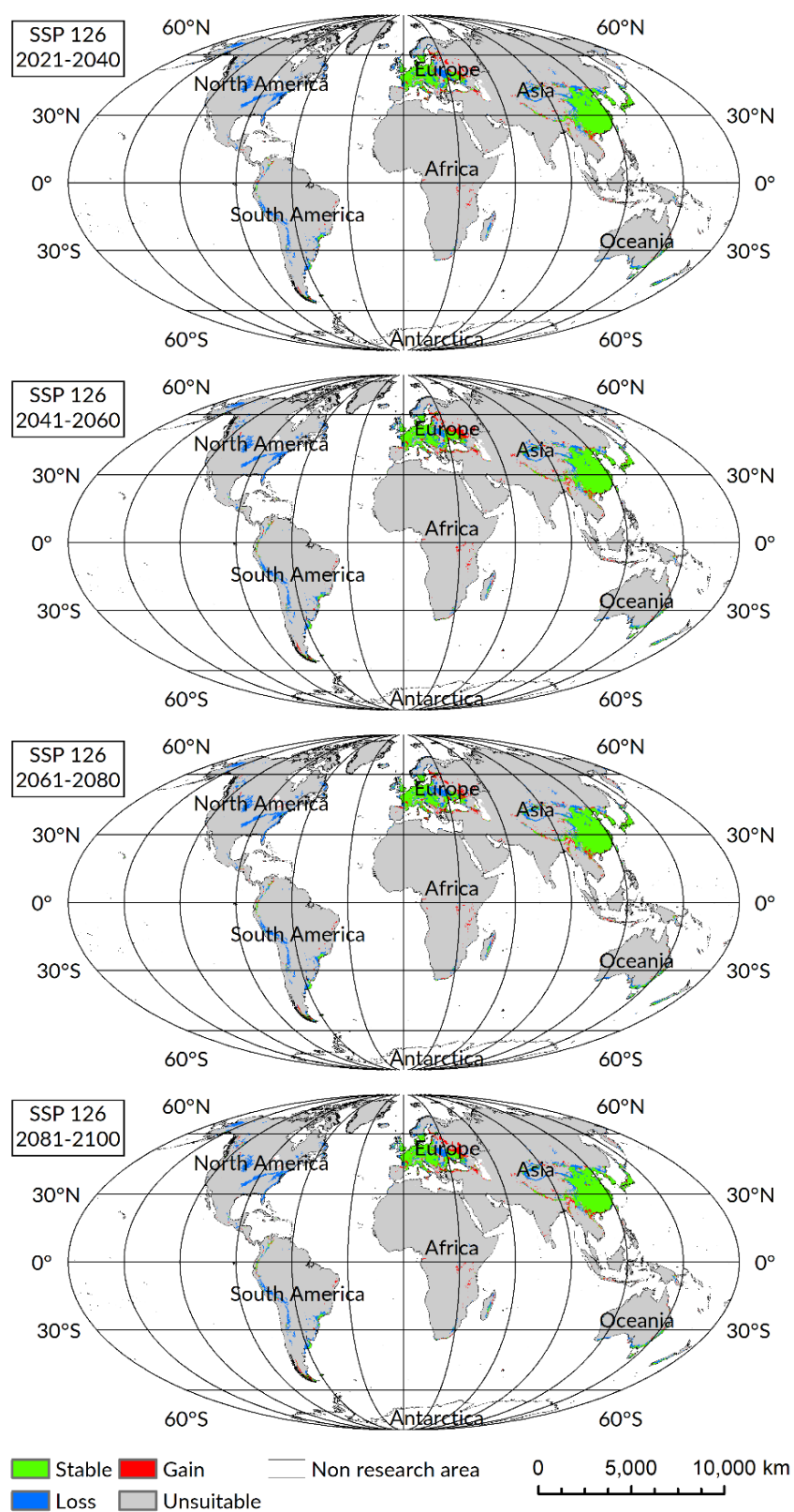- SSP 126 2081-2100

Scale: 0 — 5,000 — 10,000 km

Fig. 5 The gain and loss of high suitable habitats of *A. annua* projected during four periods under SSP 126. SSP is the abbreviation of Shared Socio-economic Pathways.