



# OPEN Prediction of Chinese suitable habitats of *Panax notoginseng* under climate change based on MaxEnt and chemometric methods

Yixin Guo<sup>1,3</sup>, Shiyan Zhang<sup>1,3</sup>, Linghui Ren<sup>1</sup>, Xin Tian<sup>1</sup>, Shicheng Tang<sup>1</sup>, Yisha Xian<sup>1</sup>, Xinjia Wu<sup>1</sup> & Zilong Zhang<sup>1,2</sup>✉

Notoginseng saponin R1; ginsenosides Rg1, Re, Rb1, and Rd; the sum of the five saponins; and underground-part fresh weight (UPFW) of single plants were used as quality evaluation indices for *Panax notoginseng* (Burk.) F. H. Chen (*P. notoginseng*). Comprehensive evaluation of *P. notoginseng* samples from 30 production areas was performed using that MaxEnt model. Spatial pattern changes in suitable *P. notoginseng* habitats were predicted for current and future periods (2050s, 2070s, and 2090s) using SSP126 and SSP585 models. The results revealed that temperature, precipitation, and solar radiation were important environmental variables. Suitable habitats were located mainly in Yunnan, Guizhou, and Sichuan Provinces. The distribution core of *P. notoginseng* is predicted to shift southeast in the future. The saponin content decreased from the southeast to the northwest of Yunnan Province, which was contrary to the UPFW trend. This study provides the necessary information for the protection and sustainable utilization of *P. notoginseng* resources, and a theoretical reference for its application in the quality evaluation of Chinese medicinal products.

**Keywords** *P. notoginseng*, MaxEnt, Quality regionalization, Ecological regionalization, Climate change, Prediction model

*Panax notoginseng* (Burk.) F. H. Chen is a perennial herb of the *Panax* genus (Araliaceae), whose medicinal parts are dry roots and rhizomes<sup>1</sup>. *P. notoginseng* is distributed in Yunnan Province, Guizhou Province, Guangxi Province, Chongqing City, and other regions<sup>2</sup>. It disperses blood stasis, reduces swelling, and relieves pain. Blood circulation promotion is the main function of saponins, including notoginseng saponin R1 and ginsenosides Rg1, Re, Rb1, and Rd<sup>3</sup>. Presently, more than 100 saponins have been described in *P. notoginseng* and are often used to treat hematemesis (vomiting blood), traumatic bleeding, and chest and abdominal (tingling pain)<sup>4</sup>. The medicinal components of *P. notoginseng* are found under the ground. Underground-part fresh weight (UPFW) is related to Notoginseng Radix et Rhizoma yield. Therefore, the saponin content and UPFW reflect the Notoginseng Radix et Rhizoma quality, and yield, respectively. These findings are of importance in evaluating the quality of Notoginseng Radix et Rhizoma.

Since ancient times, medicinal plants distribution in China has changed. Climate change majorly affects on the distribution of medicinal plants. For example, the upslope distribution of many mountain biological communities has changed owing to climate warming, which was also one of the reasons for the changes in *Eucommia ulmoides* distribution<sup>5</sup>. The primary and secondary metabolic pathways of medicinal plants are affected by climate change. The active components in most herbs are derived from secondary metabolic processes. Therefore, studying the environmental impacts of climate change on medicinal plant distribution and quality is conducive to sustainable development.

*P. notoginseng* has continuous cropping obstacles in the planting process, as it needs to grow for at least 3 years to reach the standard composition for medicine, and the land needs to be fallow<sup>6</sup>. Recently, the demand for

<sup>1</sup>School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 102488, China. <sup>2</sup>Meteorological Administration Key Open Laboratory of Transforming Climate Resource to Economy, Chongqing 401147, China. <sup>3</sup>These authors contributed equally: Yixin Guo and Shiyan Zhang. ✉email: zhangzilong76@163.com

Chinese medicinal materials has increased, leading to the short supply of *P. notoginseng* in markets. *P. notoginseng* from different areas have inconsistent product quality, which is not beneficial to standardized quality management of this Chinese medicinal material. Additionally, global warming may affect the distribution, saponin content, and UPFW of *P. notoginseng*.

MaxEnt model can predict potentially suitable habitats for *P. notoginseng* and provide guidance for expanding the planting areas<sup>7,8</sup>. The objective determination of quality differences between the regions is necessary to evaluate the quality of *P. notoginseng* in different regions<sup>9</sup>. To our knowledge, this is the first study to predict the current and future potentially suitable habitats of *P. notoginseng* and analyze changes in its suitable habitats under climate change. The correlation between *P. notoginseng* saponins, UPFW, and environmental variables was analyzed to identify limiting factors and establish a model<sup>10</sup>. The results were then superimposed on an ecological regionalization map to obtain quality distribution patterns. This is beneficial in identifying the distribution of indicators in various regions and has a certain guiding significance for scientific research and production.

## Materials and methods

### *P. notoginseng* distribution data

Using field sampling and collecting data from the Global Biodiversity Information Facility (<https://www.gbif.org/>), Chinese Virtual Herbarium (<http://www.cvh.ac.cn/>), National Specimen Information Infrastructure (<http://www.nsii.org.cn>), the Plant Photo Bank of China (<http://ppbc.iplant.cn>), and previous studies, 201 occurrence records of *P. notoginseng* were collated. These were mainly distributed in southwest China, including Yunnan, Guangxi, Sichuan, and Guizhou Provinces<sup>11,12</sup>. The collected samples were from 30 producing areas in 18 regions distributed across southeastern Yunnan Province (Suppl. Material Table S1). To reduce sampling bias, only one record per grid (1 × 1 km) was selected<sup>13</sup>. Therefore, after removing the distribution points of longitude and latitude duplications and errors, 133 species distribution records that met the conditions were obtained using a buffer analysis method and saved in the CSV format available for the MaxEnt model (Fig. 1A).

### Handling of environmental variables

Current climate data were used as the baseline data (1970–2000), including 19 climate variables and 12-month averages for temperatures, solar radiations, and precipitations<sup>14</sup>. The future periods used for climate data were 2050s, 2070s, and 2090s. These climate data were selected from the Beijing Climate Center-Climate System Model-Medium Resolution (BCC-CSM2-MR) (Coupled Model Intercomparison Project Phase 6, CMIP6), which has two shared socioeconomic pathways (SSP126 and SSP585)<sup>13,15</sup>. These data were obtained from the WorldClim database (<http://www.worldclim.org/>) at a data spatial resolution of 2.5 m (1 × 1 km).

The elevation, slope, and aspect were also obtained from the WorldClim database (<http://www.worldclim.org/>) at the same spatial resolution. Thirty-six soil variables were obtained from the Harmonized World Soil Database (<http://www.fao.org/soil-s-portal/soil-survey/soil-maps-and-database/s/harmonized-world-soil-database-v12/en/>)<sup>16,17</sup>. Therefore, 94 environmental variables (Suppl. Material Table S2) were used in the model analysis for the current period. Terrain, soil, and climate data were not available for future scenarios. ArcGIS was used to extract and transform the data into the format required for the MaxEnt model operation<sup>18</sup>.

Environmental variables with percentage contribution > 0 were filtered out. When the Pearson correlation between two variables was > 0.80, the lower contribution rate was eliminated, and 10 environmental factors were selected to be used in the model analysis.

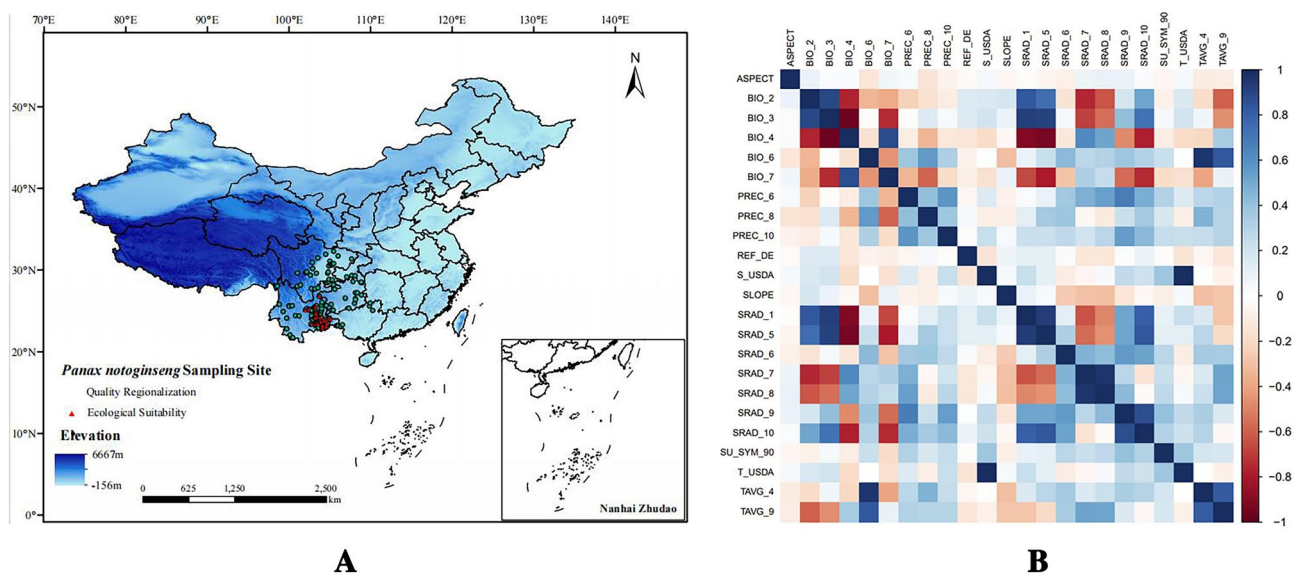


**Figure 1.** Pictures of (A) *Panax notoginseng* distribution in China. (B) Correlation heat map of environment variables.

## Ecological regionalization model establishment

### *Prediction model and parameter optimization process*

Almost 75% *P. notoginseng* distribution points were randomly selected as the training set to establish the prediction model, and the remaining 25% of the distribution points were selected as the test set to verify the model<sup>19</sup>. The jackknife method was selected to measure the importance of each variable, and the response curves of the environmental variables were created and projected onto the future environmental situation. This specific parameter setting was referred to as the optimized value, and the operation was repeated 10 times<sup>20,21</sup>.

In the MaxEnt model default values, the regularization multiplier (RM) was 1. Maxent has five features: linear (L), quadratic (Q), product (P), threshold (T), and hinge fragmentation (H)<sup>22</sup>. Model parameter optimization can avoid overfitting to a certain extent<sup>23</sup>. Delta Akaike information criterion coefficient (DELTA AICc) and omission rate (OR) were used to evaluate the degree of model optimization, using RGui scripts to simulate the different combinations of conditions. DELTA AICc, a standard measure of the goodness-of-fit of a statistical model, is calculated using R software. It can weigh the model complexity and fit the goodness-of-fit of the data, prioritizing the model with the lowest DELTA AICc. OR refers to the percentage of abnormal samples in the test sample that are not correctly classified to the total number of abnormal samples in the test sample; a small OR indicates a high prediction accuracy. The OR provides information on model differences and overfitting and evaluates the data used at a specific threshold. The model parameters were optimal when the OR was < 5% and the DELTA AICc was minimal<sup>24,25</sup>.

### *Classification of suitable habitual levels and changes in core distribution*

The reclassification tool ArcGIS was used to classify the ecological suitability of *P. notoginseng*. Suitable habitats were divided into four levels: unsuitable (0–0.1), poorly suitable (0.1–0.25), moderately suitable (0.25–0.5), and highly suitable (0.5–1)<sup>13</sup>. Four different colors were used to distinguish the habitats of each level.

The SDM Toolbox v2.4 in ArcGIS converted the MaxEnt operation results into binary, converted suitable habitats at different periods into particles, and then analyzed their movement trajectory and distance<sup>26</sup>.

## Quality regionalization analysis

### *Saponin analysis*

For quality analysis, samples from 3-year-old *P. notoginseng* were collected from 30 sampling sites (Fig. 1A). Notoginseng saponin R1 and ginsenosides Rg1, Re, Rb1, and Rd (National Institutes for Food and Drug Control, Beijing, China) were prepared as standard solutions and five test saponin solutions were prepared as reported by Chinese Pharmacopoeia, 2020 edition<sup>1</sup> and Zhang et al.<sup>27</sup>. Five saponin solutions were assessed using an e2695 high-performance liquid chromatograph (Waters Corporation, MA, USA). The chromatographic column was an XBridge C<sub>18</sub> column (250 × 4.6 mm, 5 μm). The mobile phase was acetonitrile–water gradient elution using 0–20 min, 20% acetonitrile; 20–45 min, 20–45% acetonitrile; 45–55 min, 45–55% acetonitrile; 55–60 min, 55% acetonitrile; and 60–70 min, 55–20% acetonitrile. The flow rate, detection wavelength, sample size, and column temperature were 1 mL·min<sup>−1</sup>, 203 nm, 20 μL, and 25 °C, respectively. The sum of the five saponins peaks was the total saponin content<sup>27–30</sup>. Subsequently, the linear relationship and methodology were investigated. The chromatograms of mixed reference and *P. notoginseng* test solutions under these conditions are presented in Suppl. Material Fig. S1.

*P. notoginseng* saponins (PNS) were determined as reported by Qin et al.<sup>31</sup> and Liu<sup>29</sup>. A UV 2800 ultraviolet–visible spectrophotometer (Sunny Hengping Instrument, Shanghai, China) was used to scan 190–700 nm, and a maximum absorption wavelength of 546 nm was selected as the determination wavelength. The linear relationship has previously been investigated<sup>32,33</sup>.

Correlation and stepwise regression analyses were carried out on the data from notoginseng saponin R1; ginsenosides Rg1, Re, Rb1, and Rd; and the sum of five saponins in terms of PNS content and environmental factors. Models for chemical constituents and environmental factors were established. The kriging spatial interpolation method was used to construct the model when no regression equation was established for the environmental variables<sup>34</sup>. The chemical composition distribution map obtained by constructing the model method and the ecological suitability habitat map after reclassification were superimposed to obtain a quality regionalization map<sup>30</sup>.

### *Analysis of UPFW of a single plant*

The UPFW of *P. notoginseng* samples from each sampling site was measured using an electronic scale with 0.01 g accuracy and average values were taken from each sampling site for analysis. The UPFW data and environmental factors were analyzed using stepwise regression, a model for UPFW, and environmental factors was established. The kriging spatial interpolation method was used to construct the model when no regression equation was established for the environmental variables. The UPFW distribution map obtained by constructing the model method and the ecological suitability habitat map after reclassification were superimposed to obtain a quality regionalization map<sup>35</sup>.

## Results

### **Credibility evaluation**

The predictive accuracy of the MaxEnt model was measured using the area under curve (AUC)<sup>36</sup>. The AUC value was generally 0–1, in which 0–0.5 indicate model prediction failure. AUC 0.6–0.7, 0.7–0.8, 0.8–0.9, and > 0.9 indicate poor, general, good, and excellent prediction effect, respectively<sup>37</sup>. In the current period, 1,160 candidate models were generated after parameter optimization, with 193 and 3 models meeting the OR criterion and AICc, respectively. However, only one statistically significant model met both the OR and AICc. The parameters of this model were set as follows: FC = LPT and RM = 2.4. The DELTA AICc of this model was 0 and AUC value

$0.965 \pm 0.005$  (Suppl. Material Fig. S2). For the future period, a total of 1160 candidate models were generated after parameter optimization, with 400 and 1 models meeting the OR criterion and AICc, respectively. However, only one statistically significant model met the OR criterion and AICc. The parameters for this model were set at FC = LP and RM = 0.1. The DELTA AICc of this model was 0 and the AUC value was  $0.9640 \pm 0.008$  (Suppl. Material Fig S3). This result indicates high model fitting effect and reliable experimental results.

### Environmental variable analysis

The environmental variables were filtered as described in Section “Handling of environmental variables”, and 14 environmental variables were obtained for modeling. A heat map of the correlations between these parameters was generated (Fig. 1B and Suppl. Material Table S3). The permutation importance and percentage contribution of the replacement of the 14 environmental variables in the model were compiled (Suppl. Material Table S4). In summary, BIO\_7, SRAD\_6, SRAD\_7, SRAD\_10, and PREC\_8 were the key factors limiting *P. notoginseng* growth and development and their permutation importance and percentage contributions were 90.1 and 89.6, respectively.

According to the jackknife test results, the environmental variables of temperature, precipitation, and solar radiation had greater than 1.1 regularization training and test gain values, with AUC > 0.9 (Fig. 2). These results suggest that environmental variables affect the geographical distribution of medicinal plants.

### Potential suitable habitats and optimal environmental conditions of *P. notoginseng* in the current period

In the current period, the total suitable habitat area for *P. notoginseng* in China is 924,676.79 km<sup>2</sup> (Fig. 3A). Among these, the area of highly suitable habitats was 332,953.01 km<sup>2</sup>, distributed in the southeastern Yunnan Province, southern Guizhou Province and eastern Sichuan Province. The area of moderately suitable habitats was 400,844.21 km<sup>2</sup>, distributed in central and western Yunnan Province, most of Guizhou Province, northwestern Guangxi Province, southeastern Sichuan Province, and Chongqing City. The area of poorly suitable habitats was 190,879.57 km<sup>2</sup>, distributed in the area surrounding moderately suitable habitats. Moderately and poorly suitable habitats surrounded and radiated around highly suitable habitats (Fig. 3B). This is consistent with the results reported by Zhan et al.<sup>13</sup>.

According to the response curve results of environmental variables, *P. notoginseng* growth was the most advantageous with 10–15 °C annual temperature range; 500–600 mm average precipitation in August; and 1500–1600, 18,500–19,500, and 8000–9000 kJ·m<sup>-2</sup> day<sup>-1</sup>, average solar radiation in June, July, and October, respectively.

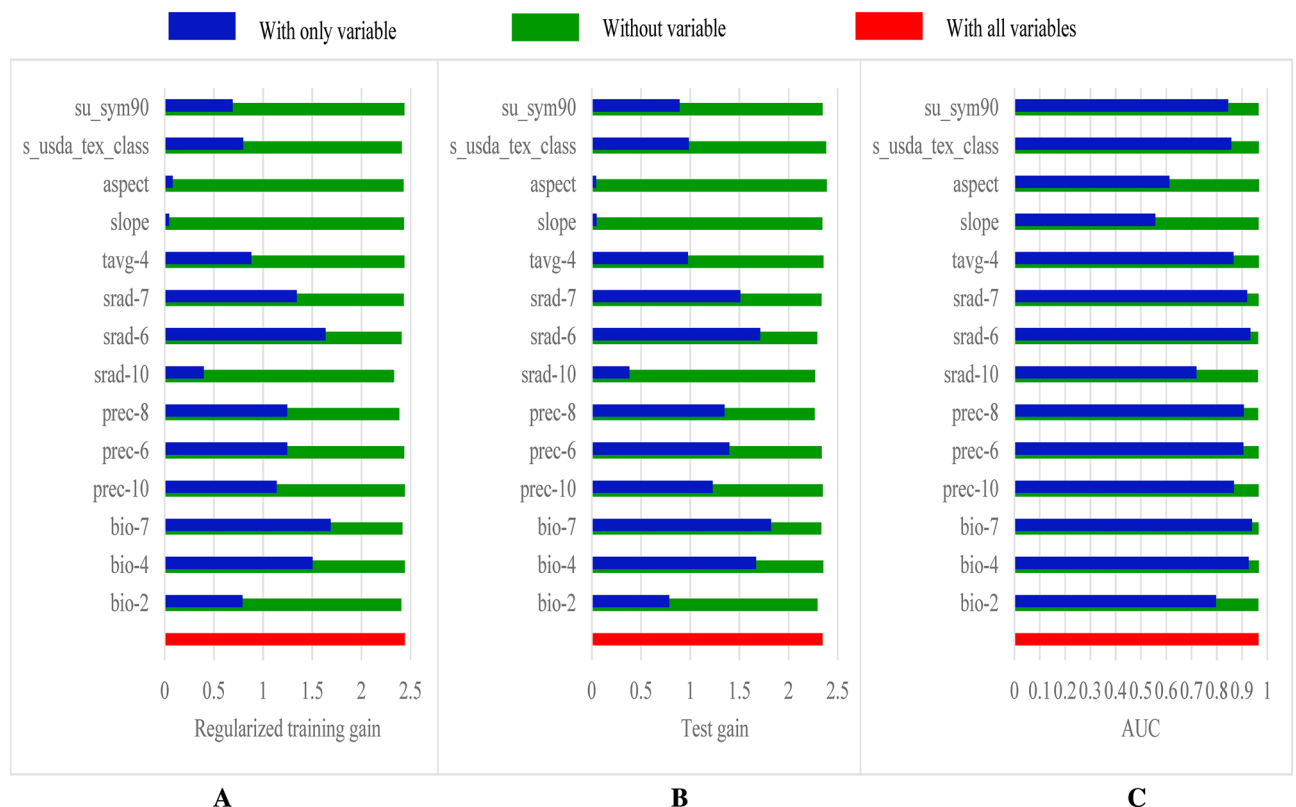


**Figure 2.** Jackknife results of (A) regularized training gain for *P. notoginseng*, (B) test gain, and (C) AUC for *Panax notoginseng*.

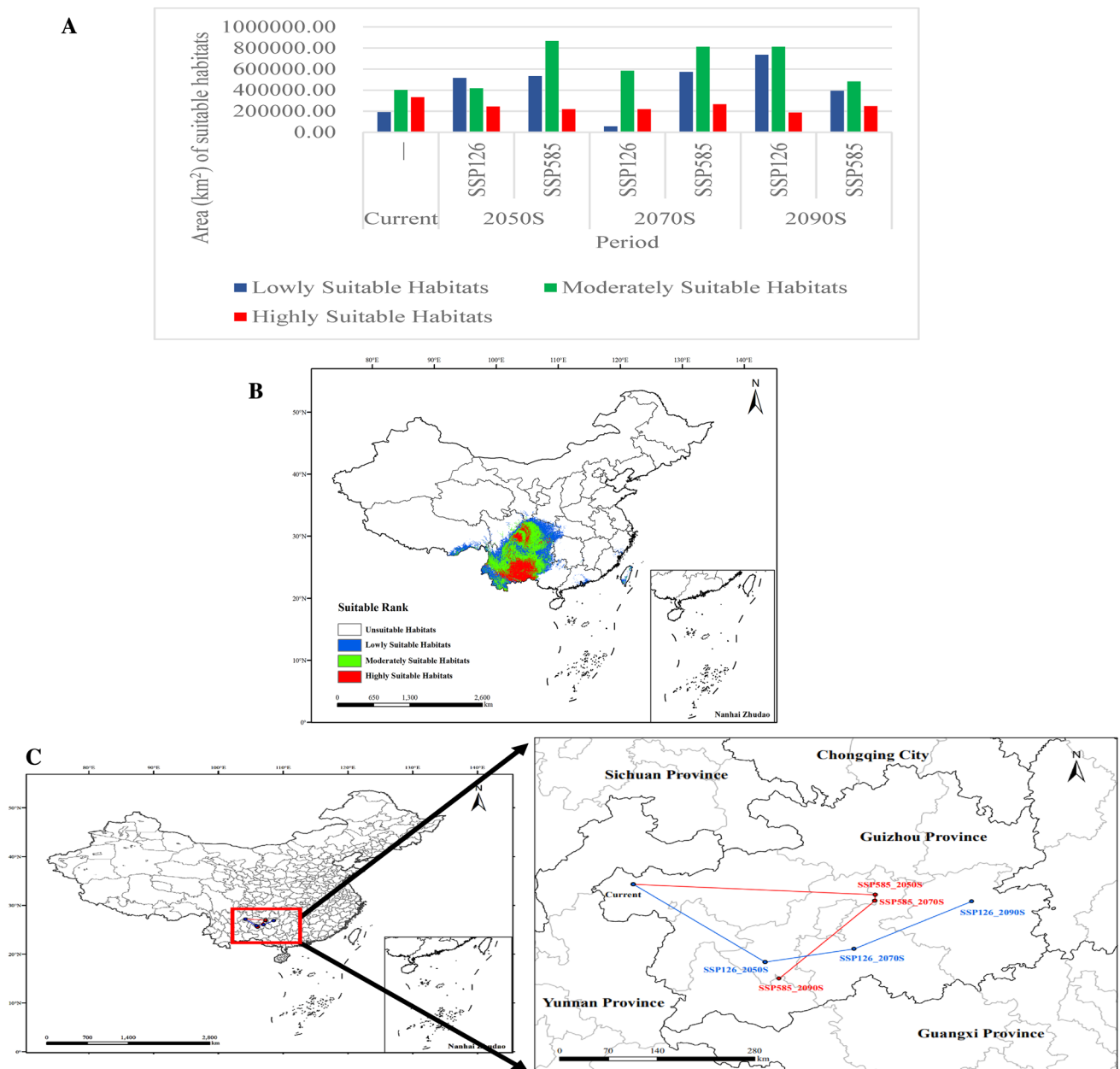


**Figure 3.** Changes in the distribution of *Panax notoginseng*. (A) Area of *P. notoginseng* suitable habitats under different periods and climate scenarios. (B) Ecological regionalization distribution of *P. notoginseng* in current period. (C) Centroid displacement of *P. notoginseng* ecological regionalization.

### Changes of *P. notoginseng* distribution in the future

#### *Changes of P. notoginseng suitable habitats in the future*

Examination of suitable *P. notoginseng* habitat distribution in different future periods and climate scenario models revealed that the trend of the total area of suitable habitats was to first expand, then reduce, and then expand in the SSP126 model (Fig. 4). Among these, the largest suitable habitats appeared in 2070s (1,648,892.29 km<sup>2</sup>), which might be because the increase in area of moderately and highly suitable habitats of *P. notoginseng* in this period. The trend of the total area of suitable habitats first expanded and then reduced in the SSP585 model. The largest suitable habitats appeared in 2090s in the SSP126 model (1,732,954.80 km<sup>2</sup>), which might be owing to relatively balanced proportion of poor, moderate, and highly suitable habitats of *P. notoginseng* in this period. The changes in area for each level of suitable habitat is presented in Fig. 3A. Changes in the distribution, expansion, shrinkage, and stability of ecological regionalization in future periods and climate scenarios were compared with those in the current period (Table 1).

#### *Change of P. notoginseng centroid displacement*

During the present period, the distribution core was in Weining County, Bijie City, Guizhou Province (104.14146° E, 27.157767° N). According to the SSP126 model, by 2050s the distribution core migrated 224.30059 km

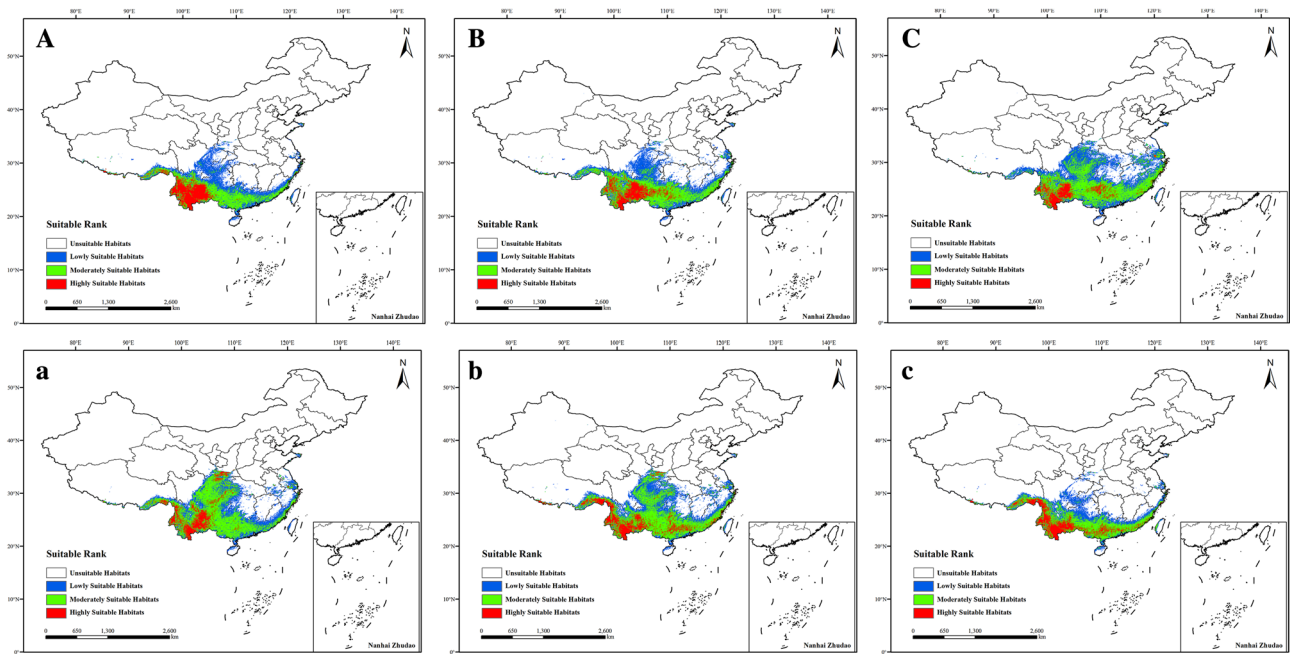


**Figure 4.** Suitable habitats for *Panax notoginseng* under different future climate scenarios. 2050s, 2070s, and 2090s using SSP126 (A–C, respectively); and 2050s, 2070s, and 2090s using SSP585 (a–c, respectively).

|                                  | 2050S      |            | 2070S      |            | 2090S      |            |
|----------------------------------|------------|------------|------------|------------|------------|------------|
|                                  | SSP126     | SSP585     | SSP126     | SSP585     | SSP126     | SSP585     |
| Extented area (km <sup>2</sup> ) | 542,119.71 | 814,712.44 | 688,141.41 | 916,242.34 | 975,496.03 | 591,119.94 |
| Stable area (km <sup>2</sup> )   | 777,640.24 | 982,761.05 | 822,879.14 | 917,581.99 | 953,366.06 | 671,060.89 |
| Shrink area (km <sup>2</sup> )   | 254,713.57 | 49,592.76  | 209,474.67 | 114,771.82 | 78,987.74  | 361,292.92 |

**Table 1.** Stable, shrunken, and increased areas (km<sup>2</sup>) of suitable habitat for *Panax notoginseng* in the future.

southeast to Zhenning County, Anshun City, Guizhou Province (105.840825° E, 25.830653° N). In 2070s, it migrated 117.26312 km northeast to Huishui County, Qiannan Prefecture, Guizhou Province (106.988205° E, 26.051955° N). In 2090s, it continued to migrate 176.03582 km northeast to Zhenyuan County, Qiandongnan Prefecture, Guizhou Province (108.50604° E, 26.866239° N). According to the SSP585 model, by 2050s the distribution core migrated 309.36487 km southeast to Fuquan City, Qiannan Prefecture, Guizhou Province (107.261263° E, 26.976562° N). In the 2070s, it migrated 11.39977 km south to Kaiyang County, Guiyang, and Guizhou (107.255406° E, 26.87411° N). In 2090s, it migrated 192.05186 km southwest to Wangmo County, Qianxinan Prefecture, Guizhou (106.022909° E, 25.545873° N). In the two climate scenarios, the centroid displacement trends were opposite although in the same direction (Fig. 3C).

**Prediction of changes in *P. notoginseng* saponins**

*Saponin content in P. notoginseng*  
Analyses of 300 samples from 30 production areas are presented in Suppl. Material Fig. S4. The concentration of notoginseng saponin R1 content in 179 samples was 0.5–1.5%. The concentration of ginsenoside Rg1 content in 127 samples was 2.5–4.0%. The ginsenoside Re content in 142 samples was <0.5%. The concentration of ginsenoside Rd content in 140 samples was 0.5–1.0%. The sum of five saponin content in samples was 7.5–11.5%. The concentration of PNS content in 122 samples was 6.0–10.5%.

*Construction of a regression model*

The content of notoginseng saponin R1; ginsenosides Rg1, Re, Rb1, and Rd; the sum of five saponins; and PNS from 30 regions were measured, and stepwise regression analysis was performed with environmental variables to obtain the corresponding relationship model.

The relationship model between notoginseng saponin R1 and environmental variables was:

$$Y_1 = 1.709 - 0.073 X_1 + 0.043 X_2,$$
 (1)

where  $Y_1$  is the notoginseng saponin R1 content;  $X_1$  is the average precipitation in October, and  $X_2$  is the average precipitation in June. In the regression equation F-test,  $P=0.012$  ( $<0.05$ ),  $F=1.026$ , and  $R^2=0.280$ , indicating the significant effect of the equation.

The relationship model between ginsenoside Rg1 and environmental variables was:

$$Y_2 = -37.878 + 0.585 X_3 + 0.04 X_4 - 0.353 X_5 - 0.21 X_1 + 0.067 X_2 + 0.026 X_6 - 0.001 X_7 + 0.005 X_8 - 0.002 X_9 - 0.02 X_{10} + 0.14 X_{11}, \quad (2)$$

where  $Y_2$  is the Rg1 content;  $X_1$  is the average precipitation in October;  $X_2$  is the average precipitation in June;  $X_3$  is the mean diurnal range;  $X_4$  is the temperature seasonality;  $X_5$  is the annual temperature range;  $X_6$  is the average precipitation in August;  $X_7$ ,  $X_8$ , and  $X_9$  are the average solar radiation in October, June, and July, respectively;  $X_{10}$  is the average temperature in April; and  $X_{11}$  is the slope. In the regression equation F-test,  $P=0.033$  ( $<0.05$ ),  $F=2.631$ , and  $R^2=0.617$ , indicating the significant effect of the equation.

The relationship model between ginsenoside Rb1 and environmental variables was:

$$Y_3 = -148.751 + 3.214 X_3 + 0.148 X_4 - 1.938 X_5 - 0.299 X_1 + 0.151 X_2 + 0.014 X_6 - 0.002 X_7 + 0.015 X_8 - 0.007 X_9 - 0.639 X_{10} + 0.495 X_{11}, \quad (3)$$

where  $Y_3$  is the Rb1 content;  $X_1$  is the average precipitation in October;  $X_2$  is the average precipitation in June;  $X_3$  is the mean diurnal range;  $X_4$  is the temperature seasonality;  $X_5$  is the annual temperature range;  $X_6$  is the average precipitation in August;  $X_7$ ,  $X_8$ , and  $X_9$  are the average solar radiation in October, June, and July, respectively;  $X_{10}$  is the average temperature in April; and  $X_{11}$  is the slope. In the regression equation F-test,  $P=0.027$  ( $<0.05$ ),  $F=2.770$ , and  $R^2=0.629$ , indicating the significant effect of the equation.

The relationship model between ginsenoside Rd and environmental variables was

$$Y_4 = -5.813 + 0.031 X_4 - 0.456 X_{10} + 0.842 X_{11}, \quad (4)$$

where  $Y_4$  is the Rd content,  $X_4$  is the temperature seasonality,  $X_{10}$  is the average temperature in April, and  $X_{11}$  is the slope. In the regression equation F-test,  $P=0.000$  ( $<0.05$ ),  $F=15.396$ , and  $R^2=0.640$ , indicating the significant effect of the equation.

The relationship model between the sum of five saponins and environmental variables was:

$$Y_5 = 20.736 + 0.002 X_4 - 0.303 X_1 + 0.062 X_6, \quad (5)$$

where  $Y_5$  is the sum of the five saponins,  $X_1$  is the average precipitation in October,  $X_4$  is the temperature seasonality, and  $X_6$  is the average precipitation in August. In the regression equation F-test,  $P=0.034$  ( $<0.05$ ),  $F=3.359$ , and  $R^2=0.279$ , indicating that the effect of the significant equation.

The stepwise regression equation between ginsenoside Re, PNS content, and environmental variables was not significant. Hence, no regression equation was established.

#### Distribution analysis of *P. notoginseng* saponins

Based on the regression equation established between saponin components and environmental variables as well as the kriging spatial interpolation method, ArcGIS was used for spatial calculations to obtain the spatial distribution of notoginseng saponin R and ginsenosides Rg1, Re, Rb1, and Rd; the sum of five saponins; and PNS content (Fig. 5).

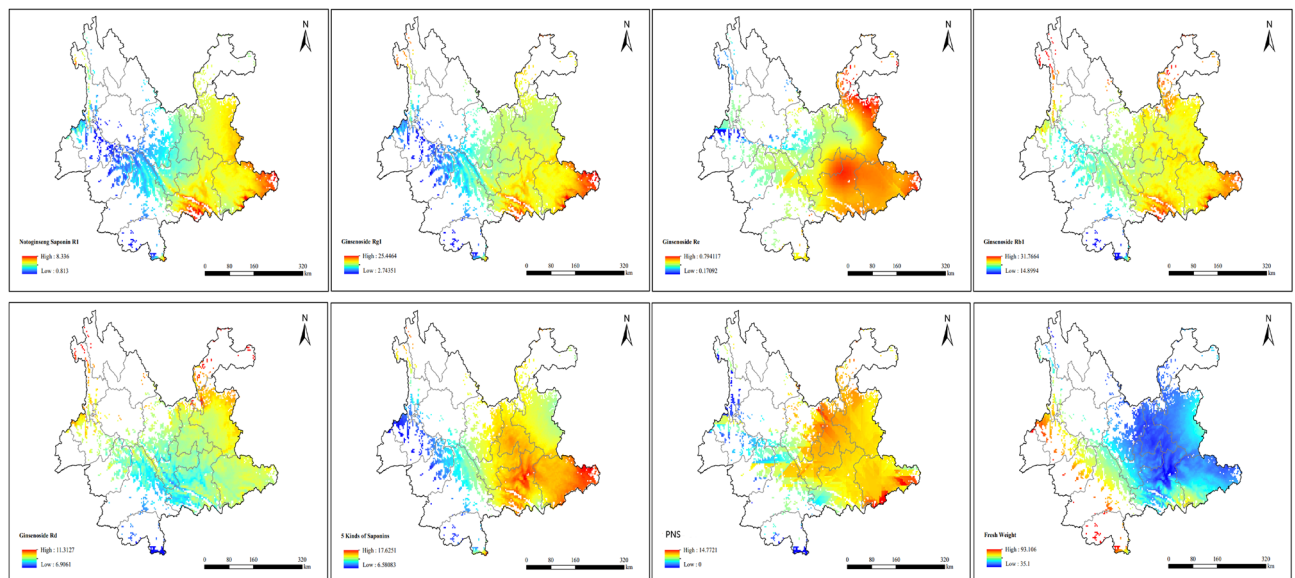


**Figure 5.** Spatial distribution of *Panax notoginseng* qualities.

Regions with high saponin components were found in lower altitudes. Regions with high levels of notoginseng saponin R1 and ginsenosides Rg1 and Rb1 were in the Wenshan and Honghe Prefectures of Yunnan Province, with similar trends of change, decreasing from the southeast to the northwest. The ginsenoside Re content was relatively high in Wenshan Prefecture, Honghe Prefecture, Kunming City, and Qujing City in Yunnan Province. The ginsenoside Rd content was generally low in most areas of Yunnan Province. There were certain differences in the spatial distribution of the sum of five saponins and PNS content, indicating that the sum of five saponins did not fully reflect the overall changes in saponin components in *P. notoginseng*.

### Spatial distribution analysis of UPFW of *P. notoginseng*

Using the measured UPFW from 30 regions and environmental variables for stepwise regression analysis, a corresponding relationship model was obtained. The relationship between fresh weight and environmental variables was as follows:

$$Y_6 = 3.406 + 0.598 X_1, \quad (6)$$

where  $Y_6$  is the fresh weight and  $X_1$  is the average precipitation in October. In the regression equation F-test,  $P = 0.045$  ( $< 0.05$ ),  $F = 4.399$ , and  $R^2 = 0.136$ , indicating the effect of the significant equation.

Based on the regression equation established between the UPFW and environmental variables, ArcGIS was used for spatial calculations and then overlaid with the ecological suitability habitat map to obtain a distribution map of the UPFW of *P. notoginseng* (Fig. 5).

Comparing the UPFW and PNS distribution maps demonstrated an inversely significant relationship between them. In Yunnan Province, the UPFW were lower in areas with low elevations as well as areas with high saponin content. Among them, Maguan, Wenshan, and Pingbian Prefectures in Honghe Prefecture had higher UPFW and saponin content, which could potentially increase *P. notoginseng* cultivation.

## Discussion

### Model precision analysis

Recently, machine learning-based modeling methods have been widely used to predict the ecological distribution of many species<sup>38</sup>. MaxEnt is one of the most widely used and recognized models<sup>39</sup>. It not only analyzes climate, soil, terrain and other environmental factors, but also considers the impact of human activities on the modeling process. When predicting the species distribution on a small scale, human factors such as vegetation cover and the urbanization index can be added to the model to help make the prediction more consistent with the actual situation. In this study, the control unit grating had only one species distribution point, overfitting environmental variables were removed, and the R software was used for parameter optimization. The AUC of the optimized model in both the training and test sets were  $> 0.9$ , which improved model accuracy and reliability.

### Migration of suitable habitats for *P. notoginseng* under climate change conditions

BCC-CSM2-MR has four modes: SSP126, SSP242, SSP370, and SSP585. Among them, there was a significant difference between SSP126 and SSP585, which was representative. Therefore, these two factors were chosen as the focus of this research<sup>40</sup>. The distribution of suitable habitats for *P. notoginseng* was most closely related to annual temperature range, solar radiation, and precipitation. *P. notoginseng* grows in dark and humid environments, which was confirmed from the environmental factors obtained by the model. Under the general trend of global warming, the minimum temperature will increase, solar radiation will enhance, and the future suitable habitats of *P. notoginseng* will thereby be greatly affected and experience change<sup>41</sup>.

The predicted potentially suitable habitat distribution of *P. notoginseng* in this study was different from that reported by Zhan et al.<sup>13</sup>. The number of species distribution data inputs in MaxEnt and the selection of environmental variables had an impact on the modeling effect and result output<sup>42</sup>. The ecological regionalization area in this study was larger than that of the previous study; however, the suitable habitat conditions of the covered areas was consistent. This was because the distribution of *P. notoginseng* in Chongqing City, and the overall collection range of the sampling points might have been more extensive in this study. The distribution of *P. notoginseng* in China is characterized by a subtropical monsoon climate. The rain and heat brought about by this climate are conducive to the growth and accumulation of effective components of *P. notoginseng*<sup>43</sup>. The southeastern part of Yunnan Province and the southern part of Guizhou Province belong to the plateau, that is, the Yungui Plateau. Ocean currents from the Pacific Ocean can directly affect this area, resulting in hot and humid climates, that are the best growth areas for *P. notoginseng*. However, western Yunnan and northern Guizhou are blocked by plateaus and mountains; the influence of ocean currents is greatly reduced, thereby lowering the fitness index of *P. notoginseng*. Although the genuine areas of *P. notoginseng* were Yunnan and Guangxi Provinces, a change in the distribution core occurred in Guizhou Province. This is because the position of the centroid was obtained by a binary operation on suitable habitats rather than on highly suitable habitats<sup>44</sup>. Therefore, the centroid displacement reflected changes in the overall suitable habitats of *P. notoginseng* during different periods, which was also more consistent with reality.

According to ancient Chinese books, *Panax notoginseng* was originally cultivated in Guangxi Province and later moved to Yunnan Province. During the Qing Dynasty, Notoginseng Radix et Rhizoma from Guangxi and Sichuan Provinces were presented to the emperor; however, researchers did not find any records of *P. notoginseng* cultivation in Sichuan Province. Researchers will continue to expand the collection of historical distribution and climate records of *P. notoginseng* in China, which will be of great importance for studying *P. notoginseng* distribution.

### Correlation analysis of saponin content and UPFW with environmental factors

In the quality regionalization analysis of *P. notoginseng* saponins and UPFW, the prediction scope was narrowed to Yunnan Province as the 37 sampling points were concentrated in Yunnan Province. After correlating the quality and environmental variable data of Yunnan Province, the quality regionalization prediction was more representative and accurate. TB Tools v1.115 is used to draw correlation heat maps. The correlation analysis of environmental factors and quality indices revealed that notoginseng saponin R1 was significantly positively correlated with TAVG\_4 (Fig. 6). Ginsenoside Rd positively correlated with temperature seasonality and negatively correlated with SRAD\_10 and TAVG\_4. UPFW was significantly and positively correlated with PREC\_10 and PREC\_8. Since the sampling points were concentrated in eastern Yunnan Province, quality regionalization could only predict changing trends. Owing to sampling point distribution and numbers, *P. notoginseng* quality regionalization of could be further improved. Although we attempted to use multiple models for joint predictions, these shortcomings could not be ignored. In addition, saponins are the main active component of *P. notoginseng* and play an important role in the treatment of trauma and cardiovascular and cerebrovascular diseases. In this study, the content of five saponins changed with the environment, which laid a foundation for the study of *P. notoginseng* quality and human health under the climate change.

In this study, we predicted the potential distribution area and future changes of *P. notoginseng* to a certain extent, which has guiding significance for expanding its cultivation in China. The ecological regionalization was innovatively combined with chemical composition and UPFW analysis, to intuitively present the change trend, which has a certain reference value in clinical medicine and the development of precise Chinese medicine.

### Conclusions

In this study, the MaxEnt model with optimized parameters was used to evaluate and predict the suitable habitat distribution for *P. notoginseng*. Based on the analysis results of the current period, the optimal *P. notoginseng* growth conditions were 10–15 °C annual temperature range; 500–600 mm average precipitation in August; and 1500–1600, 18,500–19,500, and 8000–9000 kJ·m<sup>-2</sup>·day<sup>-1</sup> average solar radiation in June, July, and October, respectively. Highly suitable habitats were mainly distributed in the central and eastern Yunnan Province, south-west Guizhou Province, and Chongqing City. Other suitable habitats radiated around them and were distributed throughout southwestern China. In the future, owing to climate change conditions, *P. notoginseng* distribution core could shift to the southeast. Based on the analyses of saponin composition and single plant UPFW, saponin content decreased from southeastern to northwestern Yunnan Province and was the highest in Honghe and Wenshan Prefectures, Yunnan Province. The UPFW was inversely related to saponin content. The results of this study provide a reference for the planting status and prediction of suitable habitats for *P. notoginseng*. They lay a solid foundation and practical guidance for promoting *P. notoginseng* yield and quality.

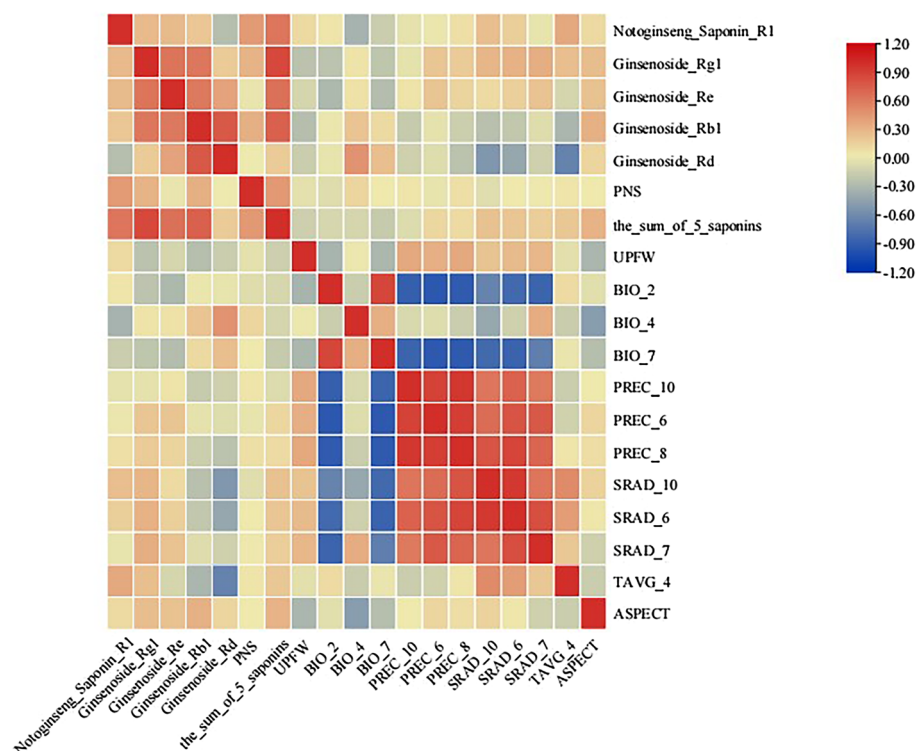


**Figure 6.** Correlation heat map of environment variables and Index of Notoginseng Radix et Rhizoma.

# Data availability

All data generated or analysed during this study are included in this published article. Experimental research and field studies on plants (either cultivated or wild), including the collection of plant material are complying with relevant institutional, national, and international guidelines and legislation. I declare that the plant material used in our study was harvested from farmland. We did not use endangered plant species for the experiments.

Received: 11 September 2023; Accepted: 9 July 2024

Published online: 16 July 2024

# References

1. Committee, N. P. *Pharmacopoeia of Peoples Republic of China* (China Medical Science Press Beijing, 2022).
2. Xu, Y. S., Zeng, J. G., Yu, Z. Y., Xin, W. F. & Duan, M. J. Analysis of Species and Content of Amino Acids in Fibrous Roots of *Panax Notoginseng* From Multiple Producing Areas in Yunnan Province Based On Hplc. *Mol. Plant Breed.* 1–12 (2022). <http://kns.cnki.net/kcms/detail/46.1068.S.20221031.1428.014.html>
3. Huang, Y. D. *et al.* *Panax Notoginseng*: A review on chemical components, chromatographic analysis, *P. Notoginseng* extracts, and pharmacology in recent five years. *China J. Chin. Mater. Med.* **47**, 2584–2596 (2022).
4. Zuo, X. H., Yin, X. N., Li, Y. Q. & Zhao, J. Q. *Panax Notoginseng* saponins reduces inflammation damage in diabetic osteoporosis rats by inhibiting rage/mapk signaling pathway. *Chin. J. Osteoporos.* **29**, 343–348 (2023).
5. Zhang, H. *et al.* Assessing the quality ecology of endemic tree species in china based on machine learning models and uplc methods: The example of *Eucommia ulmoides* oliv. *J. Clean. Prod.* **452**, 142021 (2024).
6. Peng, Z. *Assessment of Suitable Production Areas for Panax Notoginseng in China Based On Saponin Content and Maxent Model* (Zhejiang Sci-Tech University, 2022).
7. Li, J. Y. *et al.* Potential suitable areas of *Symphyotrichum Subulatum* based on maxent under future climate scenarios. *Plant Prot.* **49**, 92–102 (2023).
8. Song, H. F., Dang, X. H., Gao, Y., Meng, Z. J. & Sun, Y. L. Potential distribution of *Pinus Sylvestris* in Inner Mongolia autonomous region based on maxent model. *J. Sichuan Agricult. Univ.* 1–10. <https://doi.org/10.16036/j.issn.1000-2650.202211193> (2023).
9. Li, D. *et al.* Research on distribution and quality suitability division of *Desmodium Styracifolium*. *China J. Chin. Mater. Med.* **42**, 649–656 (2017).
10. Chen, Q. Y. *et al.* Ecology and quality suitability regionalization of *Sabia parviflora*. *Chin. J. Exp. Tradit. Med. Formul.* **27**, 172–180 (2021).
11. Cheng, X. Q., Zhao, W. L., Huang, Q. W., Yi, Z. & Li, Y. H. Textual research on the legal medicinal plants of araliaceae. *Chin. J. Modern Appl. Pharm.* **38**, 1461–1468 (2021).
12. Sun, Q. H. *et al.* Literature study On *Panax Notoginseng*. *Inf. Tradit. Chin. Med.* **34**, 113–117 (2017).
13. Zhan, P. *et al.* Assessment of suitable cultivation region for *Panax Notoginseng* under different climatic conditions using maxent model and high-performance liquid chromatography in China. *Ind. Crop. Prod.* **176**, 114416 (2022).
14. 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).
15. Liu, X. *et al.* Cultural regionalization for coptis chinensis based on 3S technologyplatform I. Study on growth suitability for coptis Chinensis based on ecological factors analysis by maxent and arcgls mode. *China J. Chin. Mater. Med.* **41**, 3186–3193 (2016).
16. FAO, I. & ISRIC, I. Jrc: Harmonized World Soil Database (Version 1.2). Fao, Rome, Italy and Iiasa, Laxenburg, Austria (2012).
17. Peng, L., Anwar, T. & Aidoo, . Prediction of potential distribution of *Calicieaelichens* in Xinjiang by using maxent model. *J. Northeast For. Univ.* **51**, 84–89 (2023).
18. Guo, Y. *et al.* Analysis of the prediction of the suitable distribution of *Polygonatum Kingianum* under different climatic conditions based on the maxent model. *Front. Earth Sci.* **11**, 1111878 (2023).
19. Li, S., Mo, S. H., Hu, X. H. & Deng, T. Prediction of potential suitable areas of endangered plant *Abies Ziyuanensis* based on maxent and arcgls. *Chin. J. Ecol.* **43**, 533 (2023).
20. Aidoo, O. F. *et al.* A machine learning algorithm-based approach (maxent) for predicting invasive potential of *Trioxa Erytreae* on a global scale. *Ecol. Inform.* **71**, 101792 (2022).
21. Yang, H. *et al.* Prediction of potential suitable area of *Cerasus Discoidea* in China based on maxent model. *J. Nanjing For. Univ. Nat. Sci. Ed.* **47**, 131 (2023).
22. Ouyang, Z. Y. *et al.* Prediction of the potential distribution of *Cyclobalanopsis Gilva* in China based on the maxent and arcgls model. *J. Cent. South Univ. For. Technol.* **43**, 19–26 (2023).
23. Ouyang, X. *et al.* Application of an improved maximum entropy sampling method in hull form optimization. *Ocean Eng.* **270**, 112702 (2023).
24. Fang, B., Zhao, Q., Qin, Q. & Yu, J. Prediction of potentially suitable distribution areas for *Prunus Tomentosa* in China based on an optimized maxent model. *Forests.* **13**, 381 (2022).
25. Li, M. *et al.* Predicting the potential suitable area of the invasive ant *Linepithema Humile* in China under future climatic scenarios based on optimized maxent. *Diversity.* **14**, 921 (2022).
26. He, X. *et al.* Prediction of potential distribution of endangered species *Corylus Chinensis* Franch. *Clim. Chang. Context. For. Res.* **35**, 104–114 (2022).
27. Zhang, S. Y. *et al.* Effect of inflorescence removal on saponins in notoginseng radix et rhizome. *Cent. South Pharm.* **20**, 2537–2541 (2022).
28. Dong, Y. *et al.* Determination of five saponins in extract of notoginseng total saponins by quantitative analysis of multi-components by single marker. *Chin. J. Pharm. Anal.* **42**, 518–524 (2022).
29. Liu, G. *Study On Quality Comprehensive Evaluation of Panax Notoginseng From Different Producing Areas* (Beijing University of Chinese Medicine, 2020).
30. Zhang, X. B. & Huang, L. Q. *Chinese Traditional Medicine Regionalization* (Science Press, 2019).
31. Qin, F., Liu, J., Chen, Y. Y. & Zhang, X. H. Study on the transonic extraction and content determination methods of total saponins of *Panax Notoginseng* (Pns). *J. Anhui Agricult. Sci.* **36**, 3062–3063 (2008).
32. Li, J. H. *et al.* Simultaneous determination of seven saponins in steamed *Panax Notoginseng* by quantitative analysis of multi-components by single marker. *Lishizhen Med. Mater. Med. Res.* **28**, 2363–2365 (2017).
33. Li, Y. X. *et al.* Study on quality evaluation of notoginseng radix et rhizoma and its processed products from different producing areas by grey correlation method and topsis method based on entropy weight method. *Chin. Tradit. Herb. Drugs.* **54**, 1252–1259 (2023).
34. Li, Y. F., He, P. & Meng, F. Y. Quality change analysis of *Epimedium Brevicornu* maxim. In Taohe river basin in southern gansu province based on the maxent model. *J. Sichuan Univ. (Nat. Sci. Ed.)* **60**, 156–162 (2023).
35. Yang, F. R. *Eco-Driver Factors Research of Quality Variation for Angelicae Sinensis Radix* (Peking Union Medical College, 2021).

36. Shi, X. *et al.* Prediction of the potentially suitable areas of *Litsea Cubeba* in China based on future climate change using the optimized maxent model. *Ecol. Indic.* **148**, 110093 (2023).
37. Chen, K., Wang, B., Chen, C. & Zhou, G. Maxent modeling to predict the current and future distribution of *Pomatosace Filicula* under climate change scenarios on the Qinghai-Tibet Plateau. *Plants*. **11**, 670 (2022).
38. Cha, Y. *et al.* An interpretable machine learning method for supporting ecosystem management: Application to species distribution models of freshwater macroinvertebrates. *J. Environ. Manag.* **291**, 112719 (2021).
39. Zhao, Z., Xiao, N., Shen, M. & Li, J. Comparison between optimized maxent and random forest modeling in predicting potential distribution: A case study with *Quasipaa Boulengeri* in China. *Sci. Total Environ.* **842**, 156867 (2022).
40. Yao, J. M., Liu, D., Duan, L. C. & Cai, Z. Prediction of potential distribution of *Diaphorina Citri* Kuwayama in China under future climate change scenarios of Cmp6. *Chin. J. Agrometeorol.* **42**, 1031–1041 (2021).
41. Langridge, P., Braun, H., Hulke, B., Ober, E. & Prasanna, B. M. Breeding crops for climate resilience. *Theor. Appl. Genet.* **134**, 1607–1611 (2021).
42. Zhu, G. P. & Qiao, H. J. Effect of the maxent model's complexity on the prediction of species potential distributions. *Biodivers. Sci.* **24**, 1189–1196 (2016).
43. Wei, L. *et al.* Soil water hydraulic redistribution in a subtropical monsoon evergreen forest. *Sci. Total Environ.* **835**, 155437 (2022).
44. Yang, C. P. *et al.* Predicting the future cultivation regions of *Cyrtotrachelus* Bu Queti using maxent model under climate change in China. *J. Sichuan Agricult. Univ.* **38**, 755–763 (2020).

## Acknowledgements

National Natural Science Foundation of China (No.82274090) and Open research project of China Meteorological Administration Key Open Laboratory of Transforming Climate Resource to Economy (No.2024003) supported this work.

## Author contributions

Y.-X.G. and S.-C.T. conceived and designed the experiments, Y.-X.G. and S.-Y.Z. analyzed the data, modified the picture and wrote the paper, L.-H.R., X.T., Y.-S.X. and X.-J.W. participated in the experiments, all authors have read and approved the manuscript for publication. As a grantee, Z.-L.Z. provided technical guidance and financial support for the research.

## 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/s41598-024-67178-4>.

**Correspondence** and requests for materials should be addressed to Z.Z.

**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) 2024