

Predicting the potential distributions of *Senecio vulgaris* L. in China

Dandan Cheng, Lin Xu

State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences
Lumo Road 388, Wuhan, Hubei, 430074, China

Abstract

Predicting potential distribution for alien plants by species distribution model (SDM, or Ecological Niche Model) using occurrence data and habitat environmental variables plays an important role in management of the invasive risk by an alien plant. Common groundsel (*Senecio vulgaris*, Asteraceae), native in Eurasia and North Africa, has been a cosmopolitan weed in temperate and also listed as one of invasive plants in China. We predict the potential distribution of this species in the world and in China particularly in Maxent (maximum entropy) models by using global occurrence records of *S. vulgaris* and the associated climate variables. The occurrence data were collected from the online databases, Global Biodiversity Information Facility database (GBIF), Chinese Virtual Herbarium database (CVH), and also from field work in China. The climate variables were download from WorldClim (<http://www.worldclim.org>). The occurrence records showed that *S. vulgaris* is present in 16 provinces or regions in north – eastern, south – western, central and north China, and almost not present in south – eastern, north – western China. The mapping of *S. vulgaris* potential distribution is diagonally across China, including the north – eastern, south – western China, and the cool area between the two regions. Analysis of the contribution and importance of climatic factors in the prediction model indicated that *S. vulgaris* adapts to the climate in humid and cool area in China (annual mean temperature ranges 2.4 - 17.5 °C, and annual precipitation ranges 550 - 1500 mm). It is suggested that special attention should be paid to the plain in NE China and Shandong Peninsula, Yungui Plateau, the cool mountain area around Sichuan basin, in western Hubei, southern Shaanxi, Shanxi and around Beijing in order to manage the invasion risk by *S. vulgaris*. The better performance of the model built by using occurrence data in China than that by using the global data in relation the predict outcome in China imply that it is might be better to use regional data than the global data when predict potential distribution for an alien plant with long invasive history in study area.

Keywords:

Maxent, DivGis, ArcGis, Climatic threshold, Species distribution model, Ecological niche model

1. Introduction

Globalization has integrated the widely dispersed human communities into a worldwide economy, providing many benefits through the movement of people and goods, but also leads to the intentional or unintentional transfer of organisms among ecosystems that were previously separate (Perrings et al. 2010). Large numbers of plant species are deliberately transported across borders for the purpose of agriculture, horticulture, medicine, food and fodder, and "hitchhiking" plant species are accidentally introduced with them (Mack & Lonsdale 2001; Reichard & White 2001). Among the species that have been successful in colonizing new areas, some have become pests and were considered as one of the main factors of biodiversity loss, significant economic and health damage (Keller et al. 2011). From a management point of view, it is extremely important to identify area which is not yet invaded but where early warning detection and control programs are essential to implement. Recent studies have developed niche-based models to assess the suitability of a region for a given species and its potential to spread throughout (Jimenez-Valverde et al. 2011; Petterson 2003; Rouget et al. 2004; Thuiller et al. 2005).

Species distribution model (SDM, or Ecological niche model) using occurrence data and habitat environmental variables play important roles in predicting the potential distribution for alien plants (Thalmann et al. 2014; Vicente et al. 2011). During the past decades many kinds of methods developed for SDMs, among which are the genetic algorithm for rule-set prediction (GARP) (Stockwell & Peters 1999), ecological niche factor analysis (ENFA)(Hirzel et al. 2002), and the maximum entropy model (Maxent) (Phillips et al. 2006). The traditional analyses, such as GAM and GLM, require both presence and absence data(Gilberto et al. 2008). However, absence data is usually unavailable and difficult to identify species existence in certain area if it was not observed (Baldwin 2009). In fact, the most data available at atlases, museum and herbarium are presence-only. Therefore, the algorithms such as Bioclim, GARP and Maxent, that only need presence data have been more and more popular(Baldwin 2009; Gilberto et al. 2008), among which Maxent model has been widely used and became one of the most popular SDMs in recent years (Elith et al. 2011; Rodda et al. 2011). Qin et al. (2015) concluded that GARP and Maxent models produced consistent distributional maps and Maxent model results were more conservative.

Common groundsel (*Senecio vulgaris* Asteracea, $2n = 40$) is suggested to be of autotetraploid originating from *Senecio vernalis* ($2n=20$, Kadereit, 1984) and native in Eurasia and North Africa. The cosmopolitan weed *S. vulgaris* var. *vulgaris* is likely to

have originated from the non-weedy *S. vulgaris* ssp. *denticulatus* from which it differs by showing no seed dormancy, by completing its life cycle from germination to seed formation much faster and by lacking ray florets (Comes et al. 1997; Moritz & Kadereit 2001). Groundsel plants can complete its life cycle in as fast as in within 8 weeks (Bolques et al. 2003) and flower at any time only if the environmental is suitable, but is considered to be a summer or winter annual weeds in UK, western Europe, North America, Australia and New Zealand (Frantzen & Hatcher 1997; Müller-Schärer & Frantzen 1996; Paul & Ayres 1987; Robinson et al. 2003; Vitousek et al. 1996). Groundsels have become successful in many crops including forage, mint, berries, ornamentals and vegetables (Figueroa et al. 2007) and shown a high degree of adaptability under different environments and selection pressure (Leiss & Müller-Schärer 2001; Williamson & Fitter 1996).

The ingestion of *S. vulgaris* and other species of this genus has been implicated as a possible cause of liver toxicity in livestock (Wiedenfeld 2011), because the plants of *S. vulgaris* contains high amount of pyrrolizidine alkaloids which cause liver toxicity in livestock and human beings (Borstel et al. 1989). Numerous cases of triazine-resistant *S. vulgaris* have been reported in Europe, North America and South America, and genetic research showed this kind of herbicide-resistance evolved by multiple, independent appearances of mutant alleles followed by spread (Délye et al. 2015; Salzmann et al. 2008).

Senecio vulgaris is listed as one of the alien invasive or naturalized plant species in China (Wu et al. 2010; Xu et al. 2012). *Senecio vulgaris* were reported as new occurrences in Lhasa, Tibet (Yu et al. 2008) and in Nansi Lake, Shandong (Wang 2010), previous studies about *S. vulgaris* in China focused on the PAs and the other bioactive (eg. antibiologic) phytochemicals (Liu et al. 2010; Xie et al. 2010; Xiong et al. 2012; Yang et al. 2011), but little is known about its distribution. Data about the occurrence of *S. vulgaris* as herbarium records or human observation can be obtained from online databank with thanks to the development of biodiversity informatics and the facilitating for data sharing at regional and global scale.

In this study most part of the occurrence record data of *S. vulgaris* were obtained from Global Biodiversity Information Facility (GBIF) and Chinese Virtual Herbarium (CVH). We also collected the occurrence records of *S. vulgaris* by field work in China. After determining the occurrence points with climate parameters, we analyzed these parameters and predicted the potential distribution of this species in China by ecological niche models conducted in Maxent. The aim of this study were to (1) explore the geography pattern of the distribution of *S. vulgaris* in China and (2) identify the regions with high risk of invasion by *S. vulgaris*. There is an urgent need to predict the potential range of *S. vulgaris* in China, since *S. vulgaris* is one of the

noxious weeds. This study also sheds light on developing appropriate strategy methods to predict the invasiveness of alien plant species.

2. Methods

2.1 Collection and collating of occurrence data

2.1.1 Occurrence data collection

Occurrence data of *S. vulgaris* at global level were collected from the Global Biodiversity Information Facility database (GBIF). We collected occurrence records within China Mainland from Chinese Virtual Herbarium database (CVH) and also from GBIF. GBIF and CVH are free and open access biodiversity database that integrates existing worldwide biodiversity data. CVH integrates the herbarium data of national natural museums from 14 institutes in China and color album with identified plant photos, the date and the location (coordination information) when and where the photos were taken.

Checking the occurrence records within China, we found that GBIF and CVH are complementary to each other with some duplication (17 records in both GBIF and CVH). There are more occurrence records of CVH (69 records found only in CVH) than those in GBIF (14 records found only in GBIF). Together with those from our field investigation, we collected 129 occurrence records of *S. vulgaris* in China Mainland, among which half are herbarium and half are observation records (Table 1 and Supplementary Table 1).

2.1.2 Visualization and collating of occurrence data

DIVA-GIS together with ArcGIS were used to map the distribution of *S. vulgaris*. All georeferenced occurrence points were input into the digital map of the world (1:8 000 000) or China (1:1 000 000). We used the province as the spatial unit to illustrate the distribution pattern of *S. vulgaris* in China and calculated the number of occurrence records of *S. vulgaris* by province.

ArcGIS 10.2 (ESRI, Redlands, CA, USA) was used to draw the world abundance map of *S. vulgaris*. First, we selected 53 116 from all of occurrence records, in order to ensure the precision of them. Records of fossil specimen, without specific collection time or places were not used. Then we imported them into ArcGIS 10.2, and show them on the map of the world. We used the point density function in Spatial Analysis Tools to generate a map demonstrating abundance of *S. vulgaris*. Particularly, we selected the occurrence records layer as point features, left blank on population field, set the cell size as 1.5, chose the neighborhood as rectangle, and set the height and width as 1.

Occurrence records provide information of the presence of *S. vulgaris* at a given place

and time, but not give an indication of how long a species was present in an area before it was collected. Hence, in this study, we accepted the date of collection on the specimen label or of the observation as the date when *S. vulgaris* was present in a given location, and we also assume that *S. vulgaris* populations has been established from the record time and would continue to exist there up to now.

Records without coordinates and coordinates less than 2 decimals were deleted, and records from small Pacific and Atlantic islands were also deleted, because assigning coarse resolution coordinates ($0.5^\circ \times 0.5^\circ$) to these records may lead to their corresponding climate data being inaccurate. In addition, we selected the occurrence points by the time scale of 1950-2000. As a consequence, the occurrence records we used to do the prediction are 26 780.

From these 26 780 occurrence records, a total of 1267 grid cells ($0.5^\circ \times 0.5^\circ$) were generated by raster-to-point function in ArcGIS and identified globally as containing *S. vulgaris*. We collated records with a coarse geographic resolution ($0.5^\circ \times 0.5^\circ$) for the reason that there may have been a sampling bias or error at a fine resolution in the GBIF and CVH occurrence records, which could cause the models of lower rather than higher quality. Another consideration was the calculation speed of the computer, identified in many previous studies, which was based on a spatial resolution between $50 \text{ km} \times 50 \text{ km}$ and $200 \text{ km} \times 200 \text{ km}$ (Ballesteros-Mejia et al. 2013; Beck et al. 2013; Jetz et al. 2012).

2.2. Climatic Variables

For the prediction models, we used the climate data from the WorldClim (<http://www.worldclim.org>) (Hijmans et al. 2002), with a resolution of 30 arc seconds (1 km resolution). These climatic variables contain 11 temperature and 8 precipitation metrics, in total 19 bioclimate variables, expressing spatial variations in annual means, seasonality and extreme or limiting climatic factors. The climate variables represent a combination of annual trends, seasonality and extreme environmental conditions. These metrics represent biologically meaningful variables for characterizing species range (Nix 1986).

An excess of climatic variables can cause overfitting, so in this study, we empirically selected 10 of the 19 Bioclim variables (Table 2) to define the climatic niches of the world, which could be used for research on any species, including *S. vulgaris*, at a global scale, and we select six and used the occurrence data in China to predict the potential distribution of *S. vulgaris* in China.

2.3. Model Selection and Evaluation

We used Maxent (version 3.3.3k, <http://www.cs.princeton.edu/~schapire/Maxent/>) as our prediction model. It's a machine learning technique, based on the principle of maximum entropy and is designed by. This approach estimates the most uniform

distribution (maximum entropy) of sampling points compared to background locations given the constraints derived from the data. The algorithms used for the inference of the potential distribution of species are precise, rapid and useful tools in the biogeography (Phillips et al. 2006).

Maxent applies five different feature constraints (linear, quadratic, product, threshold and hinge) to environmental variables, namely “the maximum entropy principle”, to estimate the species distribution probability. We used the Receiver Operating Characteristic (ROC) curve to measure the accuracy of this model (Wang et al. 2007). The accuracy evaluation on test data is useful for assessing if the resulting suitability model tends to overfit the training data; hence, losing its predictive ability. The ROC curve represents the relationship between the percentage of presences correctly predicted (sensitivity) and 1 minus the percentage of the absences correctly predicted (specificity). The area under the curve (AUC) measures the ability of the model to classify correctly a species as present or absent. A rough guide for classifying the model accuracy is: 0.50 - 0.60 = insufficient; 0.60 - 0.70 = poor; 0.70 - 0.80 = average; 0.80 - 0.90 = good; 0.90 - 1 = excellent (Araujo & Guisan 2006)

2.4. Experimental Design and Statistical Analysis

The 1267 grid cells generated by raster-to-point function in ArcGIS were loaded into DivaGIS 7.5 by using a tab delimited (TXT file, *.txt) containing the *S. vulgaris* occurrence data and related coordinates. A Comma Separated Values file (CSV, *.csv) were used in Maxent. Both files contain three fixed fields (columns) corresponding to the following categories: Species, Longitude and Latitude and the columns were listed in this specific order (Scheldeman & van Zonneveld 2010).

We imported the climate data set into DIVA-GIS and exported it into ASCII file. Then we loaded the coordinates of this species occurrence points into Maxent together with the 10 bioclimatic variables.

We calibrated the model by selecting 75% of the points as training data (950 points) and the remaining 25% percent for test data (317 points). The importance of climatic factors was measured by doing a jackknife test and the regularized gain change during the calculation. We run the Maxent model in default settings. The DIVA-GIS was used to transfer the ASCII files to grid format, and generate current distribution map of *S. vulgaris* in the world. Four arbitrary habitat categories were used to describe the probability of presence of *S. vulgaris*: the core area (0.6–1.0), the moderately suitable area (0.4–0.6), the marginal area (0.2–0.4) and the unsuitable area (0–0.2).

Beside the prediction of the global distribution of *S. vulgaris*, we also predict its distribution in China with the selected 6 climatic variables from the 19 ones (Table 2) and the occurrence data in China. And we compared the performance of the two models in relation to their results in China.

3. Results

3.1. Global distribution of *S. vulgaris*

The most abundant area of *S. vulgaris* in North Europe includes: south Norway, southern Sweden, southern Finland, Denmark; in Central or West Europe: Czech, Germany, Austria, Switzerland, Belgium, the Netherlands, northern and central France; in Mediterranean: north Italy, southern Spain, a small part in east Portugal and the coast of Thessaloniki in Greece; SW Iceland, western coast of the USA and central part in Japan. *Senecio vulgaris* also occurs with less abundance in some area, such as NE coast of the USA, SE coast of Canada, central Mexico, SW and NE of China, Japan, SE Australia and the New Zealand (Figure 1).

The AUC values of our model is very high, for training data is 0.928 and test data 0.917, showing a perfect performance (Supplementary Figure 1). The potential suitable areas for *S. vulgaris* are basically parallel with its abundance distribution. Large proportion of its core suitable areas is in Europe. Other areas which may be the core potential distribution of *S. vulgaris* are: western coasts of north and South America (except Central America), SE Australia, New Zealand, and SW China. In China the areas with high probability of *S. vulgaris* presence include: northern Yunnan, western Guizhou, SE Tibet and Western Sichuan (Figure 2).

3.2. Current distribution of *S. vulgaris* in China

The earliest specimen of *S. vulgaris* from China was collected in Guizhou (1921), Taiwan (1930), Liaoning (1932) and Yunnan (1935, Table 3 and Supplementary Table 1). The occurrence records of *S. vulgaris* in China proceeded in different rates in different periods (Figure 3 and Supplementary Table 1). Before 1950, there were only 20 occurrences, however, after, the specimen collection increased steadily, especially from 1990 ~ 2010 (Figure 3). *S. vulgaris* generally occurs in SW and NE China, including 16 provinces. Yunnan province has the most records (35), followed by Taiwan (22), Liaoning (20) and Sichuan (13). The other provinces or regions with more than 5 records are: Hubei, Jilin, Heilongjiang and Chongqing.

For the prediction model in China, the AUC values reach an excellent level, 0.923 for training data and 0.909 for test data (Supplementary Figure 1). Potential distribution of *S. vulgaris* are similar to its distribution based on occurrence records, mainly in SW and NE of China. The core areas are: northern and central Yunnan, western Guizhou, and southern Sichuan, Chongqing, western Hubei, southern Shaanxi, eastern Shanxi, East Shandong, Liaoning, Jilin, Heilongjiang and the central mountain area in Taiwan. The areas with less probability of *S. vulgaris* presence are: southern Yunnan, eastern Guizhou, northern Sichuan, north part of Shaanxi, Shanxi and Hebei, large part of Shandong, and the border area between Liaoning, Jilin, Heilongjiang and Inner

Mongolia (Figure 5).

3.3. Contribution and importance of climatic factors

The climatic variables contributed to modeling the Chinese distribution of *S. vulgaris* are (decline order): Precipitation of Wettest Period (PWP) 42.3 %, Annual Mean Temperature (AMT) 25.0 %, and Mean Temperature of Warmest Quarter (MTWA) 22.8 %, Precipitation of Driest Period (PDP) 8.5 %, Mean Temperature of Coldest Quarter (MTC) 1.3 %, and Annual Precipitation (AP) 0.1 %.

PWP is the most dominate climate factor that determine the distribution of *S. vulgaris* in China. When used in isolation it had the highest gain, which therefore appears to have the most useful information by itself. When it is omitted, it also decreases the gain the most, which therefore appears to have the most information that isn't present in the other variables (Supplementary Figure 2).

The response curves of climatic suitability for the climatic factors showed that the climatic thresholds for the suitable areas of *S. vulgaris* are: PWP 135 - 300 mm, AMT 2.4 - 17.5 °C, MTWA 15 - 24.3 °C, PDP 4 - 28 mm, MTC -13.5 - 12 °C, AP 550 - 1500 mm (Figure 6).

4. Discussion

Senecio vulgaris probably arrived and settled in SW and NE China in the 19 century and dispersed to adjunct area in Central and North China, since the earliest specimen were collected in 1920s, 1930s, and 1940s in Yunnan, Guizhou and Liaoning (Table 3 and Figure 4). However, the lack of occurrence records causes great difficulties to study on the invasive history of *S. vulgaris* in China. Explorations of flora in China have been organized since before the 1900s, first by western missionaries, explorers and scientists, and then later by Chinese botanists (Bretschneider 2011; Chen 1994), which is much later than the flora studies started in wester counties. For the instance of *S. vulgaris*, the earliest collection in Canada was recorded in 1875 in British Columbia (Robinson et al. 2003). Up to today, the number of occurrence records of *S. vulgaris* in China is 129, among which 29 are from our field work after 2005 (Table 1). This number is much lower than that in South Korea and Japan. The number of records in South Korea is 267 and in Japan is 527 according to the data in GBIF.

We regarded that *S. vulgaris* are present in 16 provinces/regions in China according to the occurrence records (Figure 4, Table 4). This adds more details and helps to clarify the some contradicted descriptions about distribution of *S. vulgaris* in the previous studies. Kadereit and Robinson et al. (Kadereit 1984; Robinson et al. 2003) regarded that *S. vulgaris* has been introduced into wide part of Eastern Asia, but only named Hongkong and Tiebet in China, particully. Chen et al (1999) regarded *S. vulgaris* present in NE China (Jilin, Liaoning and NeiMongol/inner Mongolia), SW China

(Yunan, Guizhong, Sichuan, and Xizang/Tibet) and Taiwan (in total, 8 provinces/regions). Li and Xie (2002) declared much wider distribution of *S. vulgaris*, which includes Heilongjiang, Hebei, Shanxi, Xinjiang, Chongqing, Hubei, Shanghai and Hongkong, besides all the parts mentioned by Chen (1999), in total 16 provinces (or regions). Wu et al. (2010) and Xu et al. (2012) presented the distribution of *S. vulgaris* at province or regional level, and they reported different number of provinces. Nine provinces in the former and 14 in the latter study, but no specific province or region names mentioned in either of these two studies. What is common in the previous studies is that the main distribution of *S. vulgaris* is in NE, SW China and some part in northern and central China. It is doubtful that there are stable and thrive populations of *S. vulgaris* present in Shanghai and Hongkong, from where we did not collect any exact specimen, photograph or observation of this species. Moreover, doubts arise to the occurrence records in Xinjiang and western Gansu which are outside of the potential distribution. There is a possibility that these records were caused by identifying *Senecio dubitabilis* (C. Jeffrey et Y. L. Chen) as *S. vulgaris* by mistake. These two species are rather similar to each other morphologically (Chen 1999) and the DNA markers indicated the *S. vulgaris* samples collected from west Gansu in the locations as the specimen recorded were different from *S. vulgaris* of other populations (personal communication).

We prefer the potential distribution predicted by using the occurrence data in China to that by using the global data, because the former one predicted almost all known localities where *S. vulgaris* habits in China and the latter fail to do so (Fig 2 and Fig 5). Theoretically, the global occurrence data would present the most diverse environment which a species can adapt to. In this study, we used 1,267 occurrence points all over the world to predict the global potential distribution of *S. vulgaris*, among which 990 from Europe, 172 from North American, 90 from China and only 15 from Australia and New Zealand. That is to say, the prediction would be determined by the environmental factors of *S. vulgaris* in native range. The ecological niche in the native range, regarded as possible “fulfilled” niche, often used as a constrain for predicting a new area where a species may spread (Peterson et al. 2009). However, *S. vulgaris* has a long invasion history (at least 100 years) and diverse ecological niche in China (from temperate zone to subtropical zone). It is reasonable to consider that this species has reached extremes of its potential niches in China and to predict the Chinese potential distribution with the Chinese localities habituated by *S. vulgaris*. Actually, this strategy has been used in some previous studies. For instance, the potential distribution of *Eupatorium adenophorum* was successfully predicted by using known localities in China (Wang & Wang 2006; Zhu et al. 2007).

The distribution of *S. vulgaris* is diagonally across China (Fig.5), roughly parallel to the so called Heihe (Aihui)–Tengchong Line (or Hu Line), a famous geo-demographic demarcation line in China. This line nearly overlapped the 400 mm isohyetal line, separating China’s semiarid area from semihumid area (Yang et al. 2005). The absence of *S. vulgaris* in west of this line indicates that *S. vulgaris* does not adapted to extreme drought, while the absence of *S. vulgaris* in SE China implying that this species may not survive the hot summer. This agrees the analysis of the contribution and importance of climatic factors in the prediction model (Figure 7).

Although the threshold of the climatic variables of the distribution indicates the ecological niche *S. vulgaris*, further studies with experiment in labs or field work need to be conducted in order to explore the eco – physiological factors limiting its distribution.

The realized distribution of a species is determined jointly by abiotic environmental conditions, biotic factors, and dispersal limitations (Soberón 2007). Considering the frequent economical communication in and the potential distribution of *S. vulgaris* in China, it might be concluded that this species would spread more widely and further into the regions suitable for its inhabitation. We suggest that special attention should be paid to the plain in NE China and Shandong Peninsula, where modern agriculture has developed well. *Senecio vulgaris* tend to invade agriculture area and the frequent use of herbicide may screen the herbicide biotype(Délye et al. 2015). Both of these two trait will bring disadvantage effect on modern agriculture. Moreover, Yungui Plateau, the cooler mountain area around Sichuan basin, in western Hubei, southern Shaanxi, Shanxi and around Beijing could be regarded with invasion risk by *S. vulgaris*. Records of new location as habitat of *S. vulgaris* will help to further identifying its distribution, ecological niche and monitoring its invasion.

Reference

- Araujo MB, and Guisan A. 2006. Six (or so) research priorities for species distribution modelling. *Journal of Biogeography* 33:1677-1688.
- Baldwin RA. 2009. Use of Maximum Entropy Modeling in Wildlife Research. *Entropy* 11:854-866.
- Ballesteros-Mejia L, Kitching I, Jetz W, Nagel P, and Beck J. 2013. Mapping the biodiversity of tropical insects: Species richness and inventory completeness of African sphingid moths. *Global ecology and biogeography* 22:586-595.
- Beck J, Ballesteros-Mejia L, Nagel P, and Kitching I. 2013. Online solutions and the “Wallacean shortfall”: What does GBIF contribute to our knowledge of species’ ranges. *Diversity and Distributions* 19:1043-1050.
- Bolques A, Norcini JG, and Aldrich JH. 2003. Preemergence 5. control of common groundsel (*Senecio vulgaris*). Proc South Nurs Assoc Res Conf. p 333-336.
- Borstel KV, Witte L, and Hartmann T. 1989. Pyrrolizidine alkaloid patterns in populations of *Senecio vulgaris*, *S. vernalis* and their hybrids. *Phytochemistry* 28:1635-1638.
- Bretschneider E. 2011. *History of European Botanical Discoveries in China*. Hamburg: Severus-Verlag.
- Chen J. 1994. the History of plant taxonomy in China. In: China BSo, ed. *the History of Botany in China* Beijing: Science Press.
- Chen Y. 1999. *eFlora of China*: Beijing: Science Press.
- Comes HP, Kadereit JW, Pohl A, and Abbott RJ. 1997. Chloroplast DNA and isozyme evidence on the evolution of *Senecio vulgaris* (Asteraceae). *Plant Systematics & Evolution* 206:375-392.
- Délye C, Causse R, and Michel S. 2015. Genetic basis, evolutionary origin and spread of resistance to herbicides inhibiting acetolactate synthase in common groundsel (*Senecio vulgaris*). *Pest Management Science*.
- Elith J, Phillips S, and Hastie T. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17:43-57.
- Figueroa R, Doohan D, Cardina J, and Harrison K. 2007. Common groundsel (*Senecio vulgaris*) seed longevity and seedling emergence. *Weed science* 55:187-192.
- Frantzen J, and Hatcher P. 1997. A fresh view on the control of the annual plant *Senecio vulgaris*. *Integrated Pest Management Reviews* 2:77-85.
- Gilberto P, Graziano R, and Alessandro F. 2008. Toward improved species niche modelling: *Arnica montana* in the Alps as a case study. *Journal of Applied Ecology* 45:1410-1418.
- Hijmans R, Guarino L, and Rojas E. 2002. DIVA-GIS a geographic information system for the analysis of biodiversity data. Lima, Peru: International Potato Center.
- Hirzel A, Hausser J, and Chessel D. 2002. Ecologicalniche factor analysis: How to compute habitat-suitability maps without absence data. *Ecology* 83:2027-2036.
- Jetz W, McPherson J, and Guralnick R. 2012. Integrating biodiversity distribution knowledge: Toward a global map of life. *Trends in ecology & evolution* 27:151-159.
- Jimenez-Valverde A, Peterson AT, Soberon J, Overton JM, Aragon P, and Lobo JM. 2011. Use of niche models in invasive species risk assessments. *Biological Invasions* 13:2785-2797.

- Kadereit JW. 1984. The origin of *Senecio vulgaris* (Asteraceae). *Plant Systematics and Evolution* 145:135-153.
- Keller RP, Geist J, Jeschke JM, and Kühn I. 2011. Invasive species in Europe: ecology, status, and policy. *Environmental Sciences Europe* 23:23.
- Leiss KA, and Müller-Schärer H. 2001. Adaptation of *Senecio vulgaris* (Asteraceae) to ruderal and agricultural habitats. *Am J Bot* 88:1593-1599.
- Li Z, and Xie Y. 2002. *Invasive Alien Species in China (in Chinese)*. Beijing: China Forestry Publishing House,.
- Liu Y, Zhang Z, and Wang Y. 2010. Chemical constituents in *Senecio vulgaris*. *Chinese traditional and herbal drugs* 41:1608-1612.
- Müller-Schärer H, and Frantzen J. 1996. An emerging system management approach for biological weed control in crops: *Senecio vulgaris* as a research model. *Weed Research* 36:483-491.
- Mack RN, and Lonsdale WM. 2001. Humans as global plant dispersers: Getting more than we bargained for. *Bioscience* 51:95-102.
- Moritz DML, and Kadereit JW. 2001. The genetics of evolutionary change in *Senecio vulgaris* L.: A QTL mapping approach. *Plant Biology* 3:544-552.
- Nix H. 1986. A biogeographic analysis of Australian elapid snakes. In: R. L, editor. Atlas of Elapid Snakes of Australia, Australian Flora and Fauna Series Number 7. Canberra: Australian Government Publishing Service. p 4-15.
- Paul N, and Ayres P. 1987. Effects of rust infection of *Senecio vulgaris* on competition with lettuce. *Weed Research* 27:431-441.
- Perrings C, Burgiel S, Lonsdale M, Mooney H, and Williamson M. 2010. International cooperation in the solution to trade-related invasive species risks. *The Year in Ecology and Conservation Biology* 1195:198-212.
- Peterson AT, Papes M, and Kluza DA. 2009. Predicting the potential invasive distributions of four alien plant species in North America.
- Peterson AT. 2003. Predicting the geography of species' invasions via ecological niche modelling *The Quarterly Review of Biology* 78.
- Phillips, S.J., Anderson RP, and Schapire RE. 2006. Maximum entropy modelling of species geographic distributions. *Ecological Modelling* 190:231-259.
- Qin Z, Zhang J, DiTommaso A, Wang R, and Wu R. 2015. Predicting invasions of *Wedelia trilobata* (L.) Hitchc. with Maxent and GARP models. *Journal of Plant Research* 128:763-775.
- Reichard SH, and White P. 2001. Horticulture as a pathway of invasive plant introductions in the United States. *Bioscience* 51:103-113.
- Robinson DE, O'Donovan JT, Sharma MP, Doohan DJ, and Figueroa R. 2003. The biology of Canadian weeds. 123. *Senecio vulgaris* L. *Canadian Journal of Plant Science* 83:629-644.
- Rodda G, Jarnevich C, and Reed R. 2011. Challenges in identifying sites climatically matched to the native ranges of animal invaders. *Plos One* 6:e14670.
- Rouget M, Richardson DM, Nel JL, Le Maitre DC, Egoh B, and Mgidi T. 2004. Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions* 10:475-484.

- Salzmann D, Handley RJ, and Müller-Schärer H. 2008. Functional significance of triazine-herbicide resistance in defence of *Senecio vulgaris* against a rust fungus. *Basic and Applied Ecology* 9:577-587.
- Scheldeman X, and van Zonneveld M. 2010. Training Manual on Spatial Analysis of Plant Diversity and Distribution. Bioersivity International. Rome, Italy. p 147-179.
- Soberón J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology letters* 10:1115-1123.
- Stockwell D, and Peters D. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13:143-158.
- Thalman DJK, Kikodze D, Khutsishvili M, Kharazishvili D, Guisan A, Broennimann O, and Müller-Schärer H. 2014. Areas of high conservation value in Georgia: present and future threats by invasive alien plants. *Biological Invasions*.
- Thuiller W, Richardson DM, Pysek P, Midgley GF, Hughes GO, and Rouget M. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11:2234-2250.
- Vicente J, Randin CF, Goncalves J, Metzger MJ, Lomba A, Honrado J, and Guisan A. 2011. Where will conflicts between alien and rare species occur after climate and land-use change? A test with a novel combined modelling approach. *Biological Invasions* 13:1209-1227.
- Vitousek PM, DAntonio CM, Loope LL, and Westbrooks R. 1996. Biological invasions as global environmental change. *American Scientist* 84:468-478.
- Wang R, and Wang YZ. 2006. Invasion dynamics and potential spread of the invasive alien plant species *Ageratina adenophora* (Asteraceae) in China. *Diversity and Distributions* 12:397-408.
- Wang Y. 2010. Investigation on Alien Invasive Plants in the Nansihu Lake Wetland of Shandong Province(in Chinese with English abstact). *Chinese Bulletin of Botany* 45:212-219.
- Wang Y, Xie B, Wan F, Xiao Q, and Dai L. 2007. Application of ROC curve analysis in evaluating the performance of alien species' potential distribution models. *Biodiversity Science* 15:365-372.
- Wiedenfeld H. 2011. Plants containing pyrrolizidine alkaloids: toxicity and problems. *Food Additives & Contaminants: Part A* 28:282-292.
- Williamson M, and Fitter A. 1996. The varying success of invaders. *Ecology* 77:1661-1666.
- Wu S-H, Sun H-T, Teng Y-C, Rejmánek M, Chaw S-M, Yang T-YA, and Hsieh C-F. 2010. Patterns of plant invasions in China: Taxonomic, biogeographic, climatic approaches and anthropogenic effects. *Biological Invasions* 12:2179-2206.
- Xie W-D, Li X, and Row K-H. 2010. A New Pyrrolizidine Alkaloid from *Senecio vulgaris*. *Bulletin of the Korean Chemical Society* 31:2715-2716.
- Xiong A, Yang L, Ji L, Wang Z, Yang X, Chen Y, Wang X, Wang C, and Wang Z. 2012. UPLC-MS based metabolomics study on *Senecio scandens* and *S. vulgaris*: an approach for the differentiation of two *Senecio* herbs with similar morphology but different toxicity. *Metabolomics* 8:614-623.

- Xu H, Qiang S, Genovesi P, Ding H, Wu J, Meng L, Han Z, Miao J, Hu B, and Guo J. 2012. An inventory of invasive alien species in China. *NeoBiota* 15:1-26.
- Yang J, Ding Y, Chen R, and Liu J. 2005. Relationship of Asian Monsoon with dry and wet climate boundary and agro-pasturing climate boundary. *Scientia Geographica Sinica* 25:441-447.
- Yang X, Yang L, Xiong A, Li D, and Wang Z. 2011. Authentication of *Senecio scanden* and *S. vulgaris* based on the comprehensive secondary metabolic patterns gained by UPLC–DAD/ESI-MS. *Journal of Pharmaceutical and Biomedical Analysis* 56:165-172.
- Yu S, Li H, and Tu Y. 2008. the new recored geius and species of the flora in Tibet (in Chinese). *The science and tecology of Tibet* 183: 69-71.
- Zhu L, Sun OJ, Sang W, Li Z, and Ma K. 2007. Predicting the spatial distribution of an invasive plant species (*Eupatorium adenophorum*) in China. *Landscape Ecology* 22:1143-1154.

List of the tables

Table 1. Summary of collection of occurrence records of *Senecio vulgaris* in China Mainland

Table 2. Description of 19 climatic factors and corresponding calculated formula

Table 3. Earliest collection of *Senecio vulgaris* in China listed by collected location, year and herbarium.

Table 4 Occurrence records *Senecio vulgaris* in 16 provinces (or region) in China

Supplementary table 1. *Senecio vulgaris*' occurrences records in China

List of the figures

Figure 1. Global spatial abundance of *Senecio vulgaris* (the grid cell is $1.5^{\circ} \times 1.5^{\circ}$). The total number of occurrence records used is 53 116 (53 027 of specimen from Global Biodiversity Information Facility (GBIF) and 129 from Chinese Virtual Herbarium (CVH).

Figure 2. Predication of global potential distribution of *Senecio vulgaris* (the grid cell is $0.5^{\circ} \times 0.5^{\circ}$) by the Maxent model built with the global occurrence records. The standard deviation of each grid square is 0–0.16. The probability of presence of *S. vulgaris*: the core area (0.6–1.0), the moderately suitable area (0.4–0.6), the marginal area (0.2–0.4) and the unsuitable area (0–0.2).

Figure 3. Cumulative occurrence records of *Senecio vulgaris* vs. years in China. The data were summarized according to the decades. Each number represents the number of all the records in and before the particular years.

Figure 4. Distribution of *Senecio vulgaris* in China based on occurrence records. Numbers (No) on the map represents the province where occurrence of *S. vulgaris* were recorded. The No, province names and number of records for every province were in Table 4.

Figure 5. Potential distribution of *Senecio vulgaris* in China (the grid cell is $0.5^{\circ} \times 0.5^{\circ}$) predicted by the Maxent model built with occurrence records within China. The standard deviation of each grid square is 0–0.18. Dots represent occurrence records. The probability of presence of *S. vulgaris*: the core area (0.6–1.0), the moderately suitable area (0.4–0.6), the marginal area (0.2–0.4) and the unsuitable area (0–0.2).

Figure 6. The response curve of climatic suitability for the climatic factors based on Maxent models predicted the distribution of *Senecio vulgaris*; a. Annual mean temperature (AMT); b. Mean temperature of the coldest quarter (MTC); c. Mean temperature of the warmest quarter (MTWA); d. Annual precipitation (AP); e. Precipitation of driest period (PDP); f. Precipitation of wettest period (PWP). The dash lines refer to the core and moderately suitable area with the probability of presence more than 0.4.

Supplementary Figure 1. Accuracy assessment of Maxent model through ROC and AUC

analyses. a. Predication of global potential distribution of *Senecio vulgaris* with the global occurrence records. b. Predication of potential distribution of *S. vulgaris* in China with occurrence records within China.

Supplementary Figure 2. Jackknife test result of variables' contribution. a. Predication of global potential distribution of *Senecio vulgaris* with the global occurrence records. b. Predication of potential distribution of *S. vulgaris* in China with occurrence records within China.

Table 1 Occurrence records of *Senecio vulgaris* in China Mainland

Data resource	Observation	Specimen	Summary
CVH	36	33	69
GBIF	0	14	14
GBIF & CVH	0	17	17
Field work	29	0	29
Total	65	64	129

Table 2. Description of 19 climatic factors and corresponding calculated formula

Variable	Abbreviation	Unit	Formula
Annual Mean Temperature ^{ab}	AMT	°C	
Mean Temperature of Wettest Quarter ^a	MTWE	°C	
Mean Temperature of Driest Quarter ^a	MTD	°C	
Mean Temperature of Warmest Quarter ^{ab}	MTWA	°C	
Mean Temperature of Coldest Quarter ^{ab}	MTC	°C	
Annual Precipitation ^{ab}	AP	mm	
Precipitation of Wettest Period ^{ab}	PWP	mm	
Precipitation of Driest Period ^{ab}	PDP	mm	
Precipitation of Warmest Quarter ^a	PWQ	mm	
Precipitation of Coldest Quarter ^a	PCQ	mm	

Variables superscript with *a* are used only in world model, with *ab* are used in models both world and China.

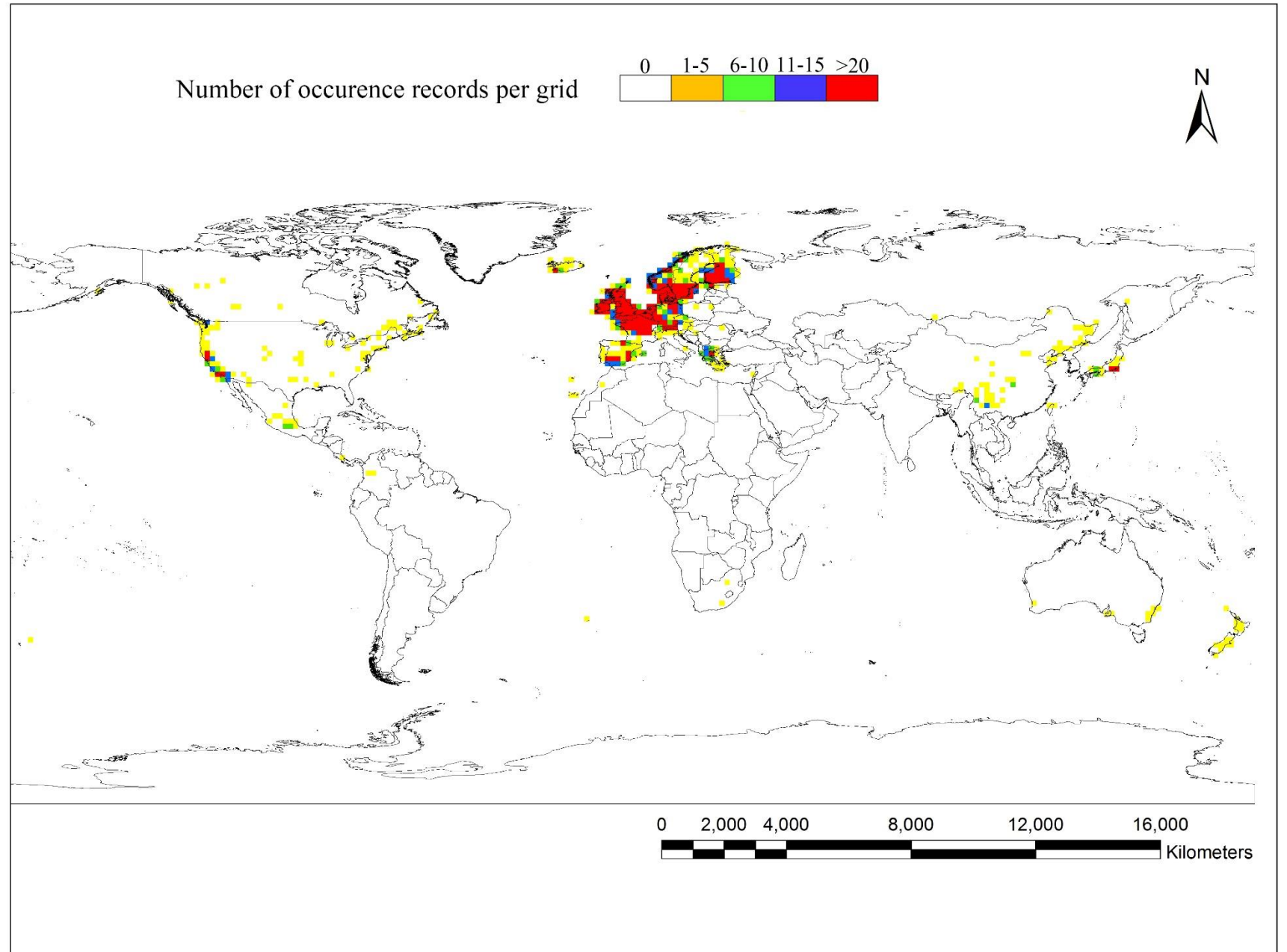
Table 3. Earliest collection of *Senecio vulgaris* in China listed by collected location, year and herbarium. PE: Herbarim, Institute of Botany of the National Academy; TAI: Herbarium of National Taiwan University).

Location	year	herbarium
Weining, Guizhou,	1921	PE
Weining, Guizhou,	1930	PE
Sunning, Taiwan	1930	TAI
Dalian, Liaoning	1932	PE
Kunming, Yunnan	1935	PE
Kunming, Yunnan	1935	Harvard University
Kunming, Yunnan	1941	PE
Kunming, Yunnan	1945	PE

Table 4. Occurrence records *Senecio vulgaris* in 16 provinces (or regions) in China

No.	Province	Number of records
1	Heilongjiang	9
2	Jilin	10
3	Liaoning	20
4	Inner Mongolia	2
5	Shandong	3
6	Hebei	1
7	Shanxi	2
8	Hubei	10
9	Shaanxi	2
10	Sichuan	13
11	Chongqing	7
12	Gansu	3
13	Guizhou	8
14	Yunnan	35
15	Tibet	4
16	Taiwan	22

Figure 1.
Global spatial
abundance of
Senecio
vulgaris



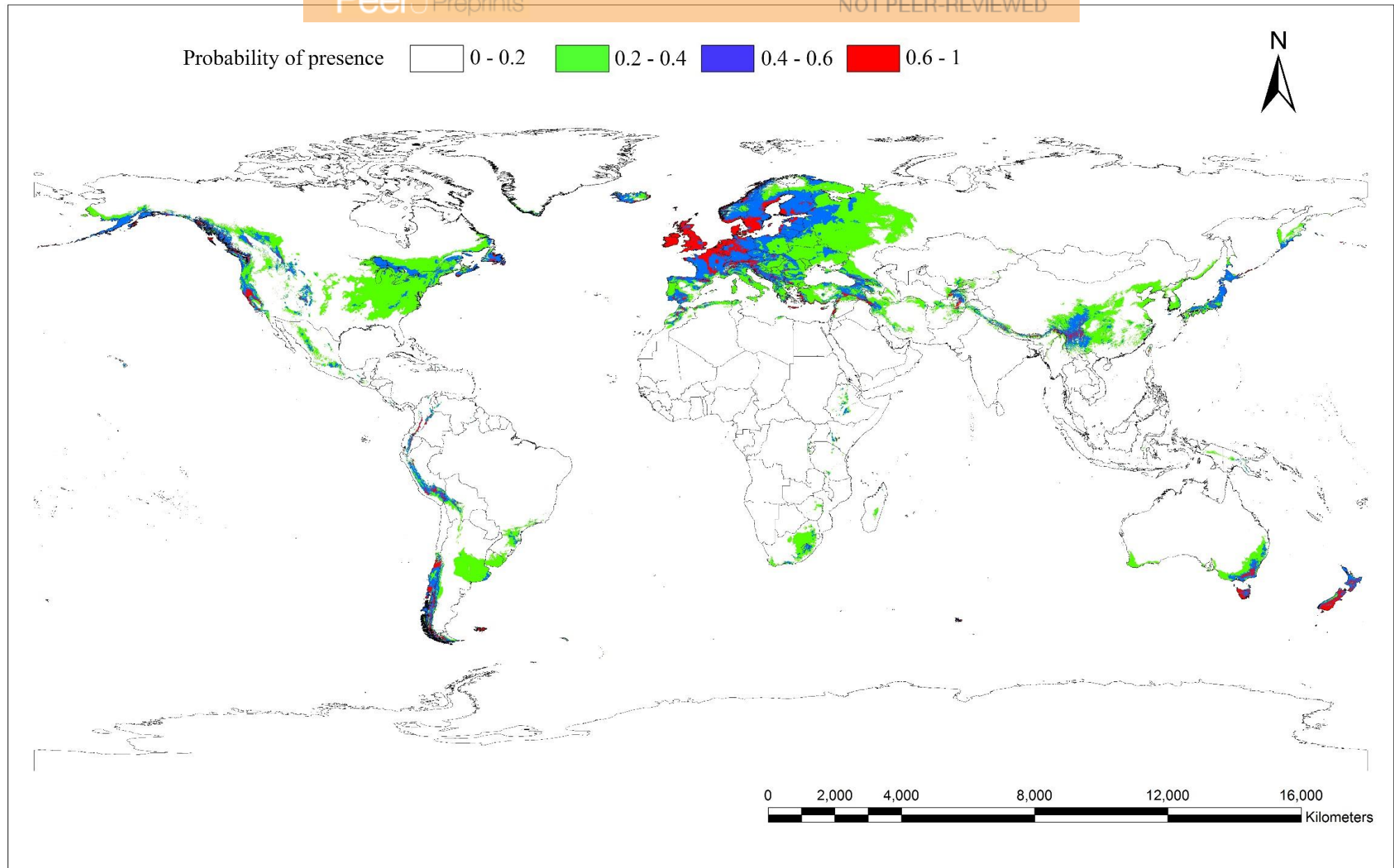


Figure 2. Predication of global potential distribution of *Senecio vulgaris*

Number of records

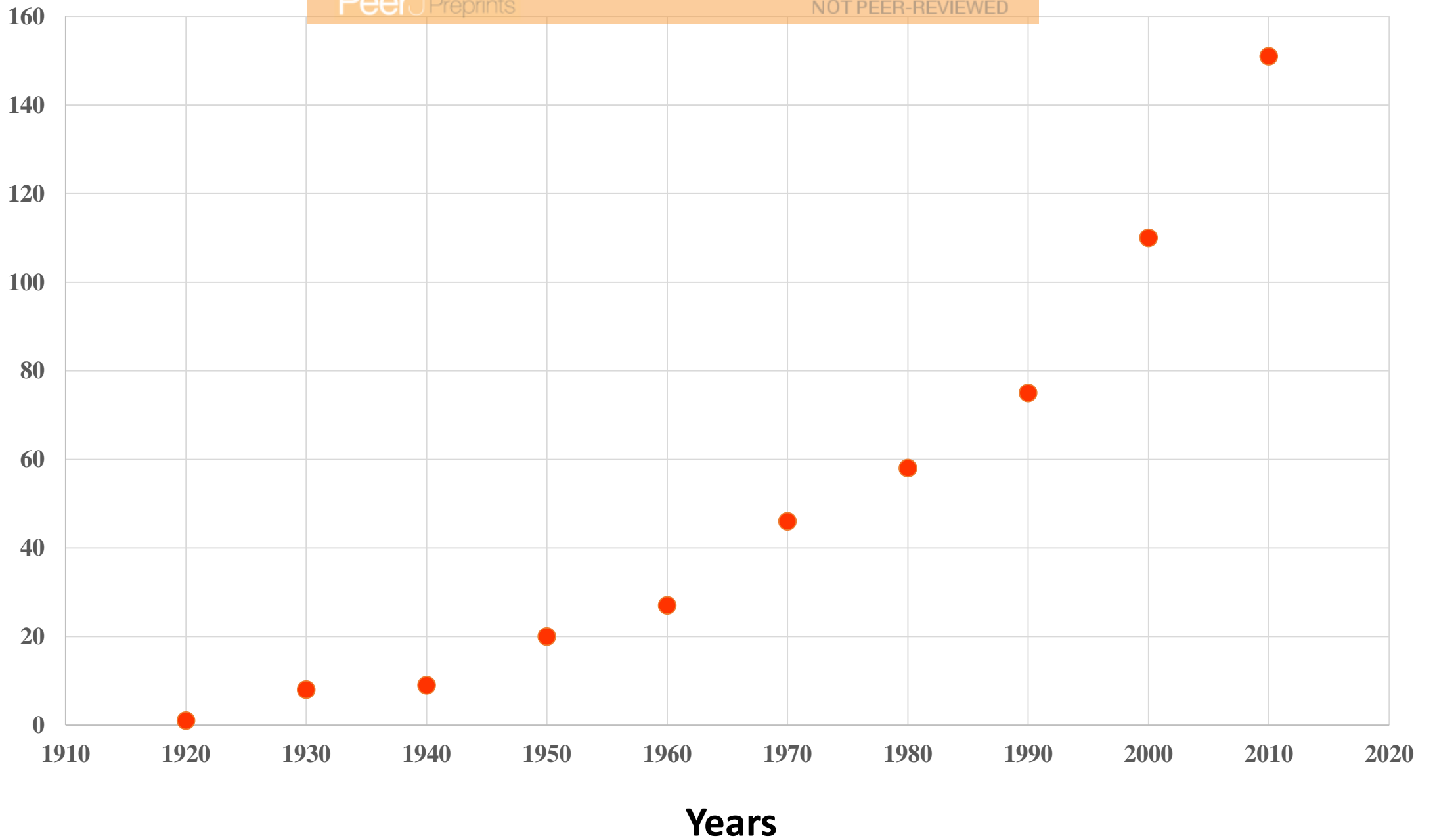


Fig. 3 Cumulative occurrence records of *S. vulgaris*' vs. years in China

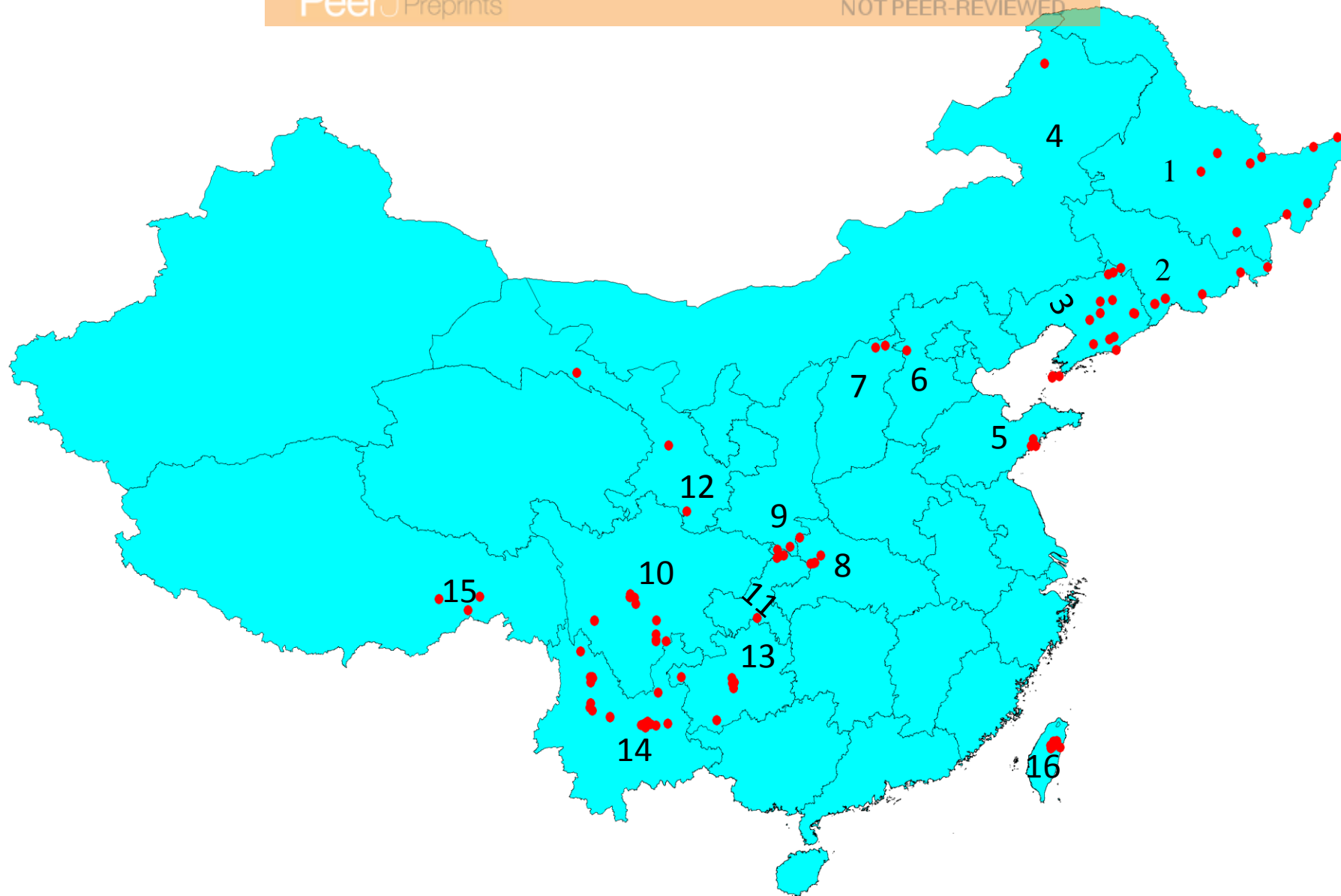


Fig.4 Distribution of *Senecio vulgaris* in China based on occurrence records.

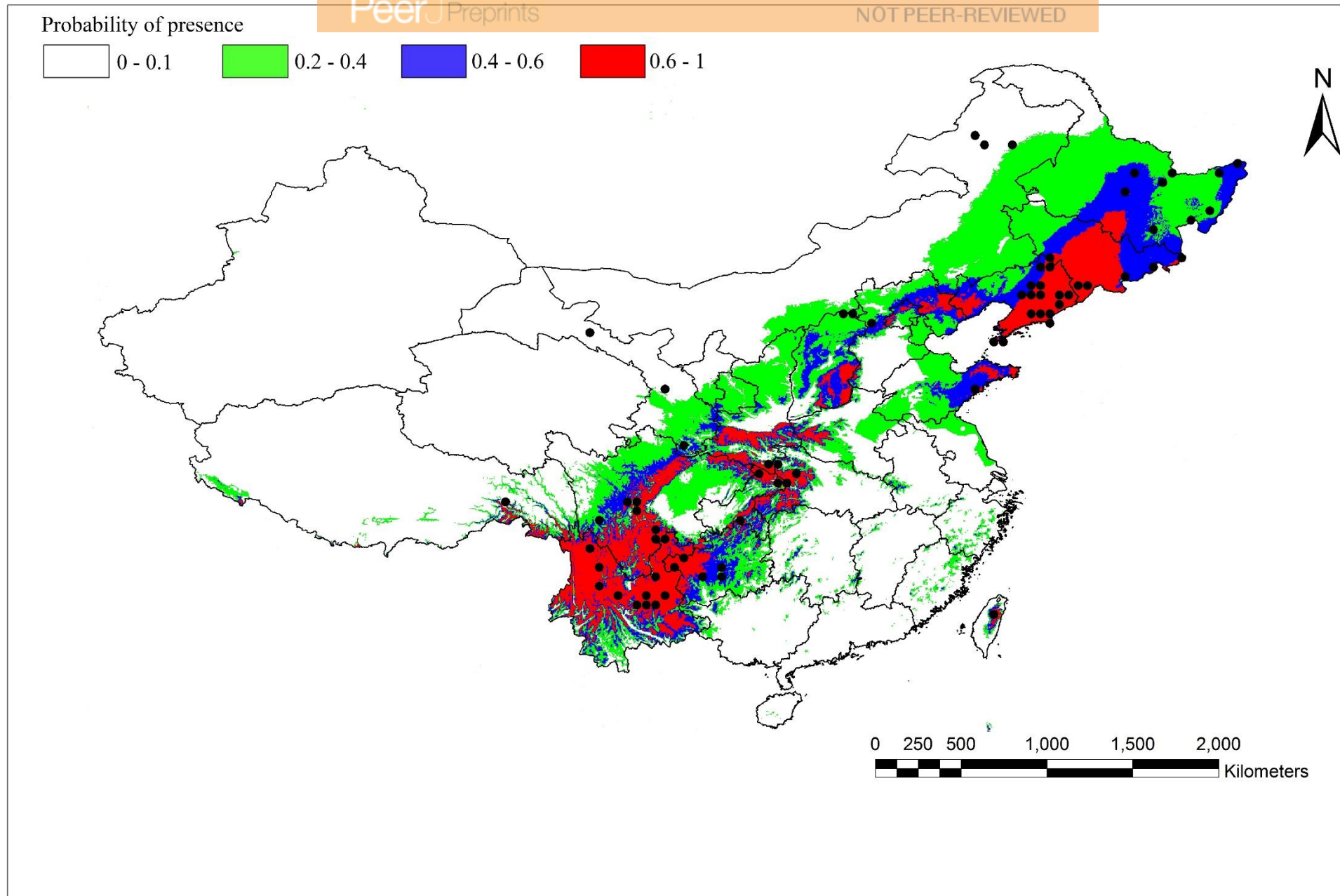
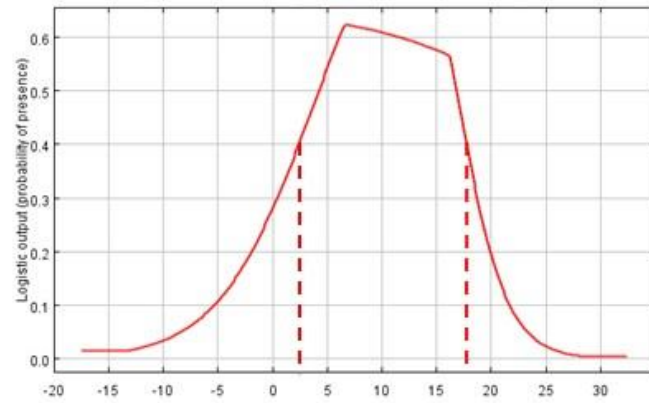
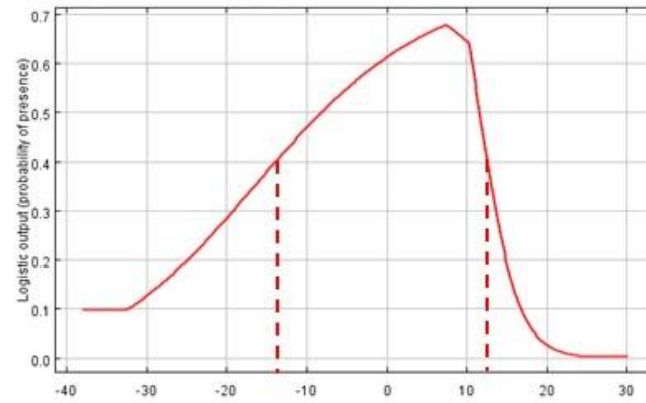


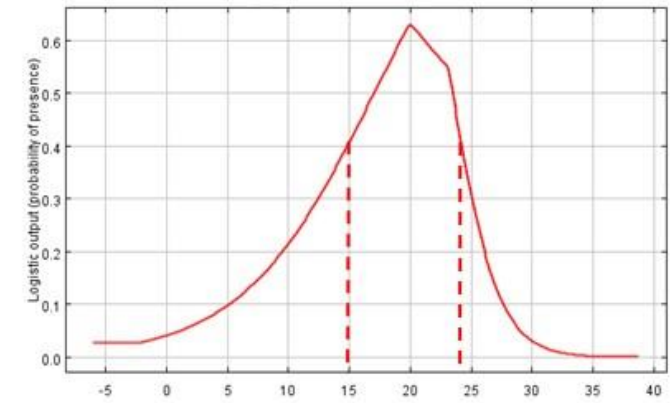
Figure 5. Potential distribution of *Senecio vulgaris* in China



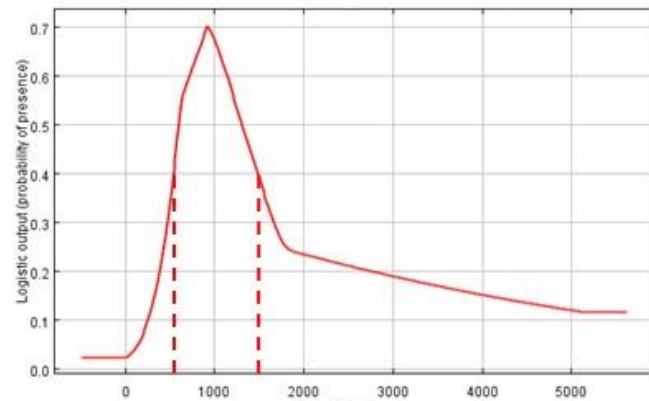
a. annual mean temperature



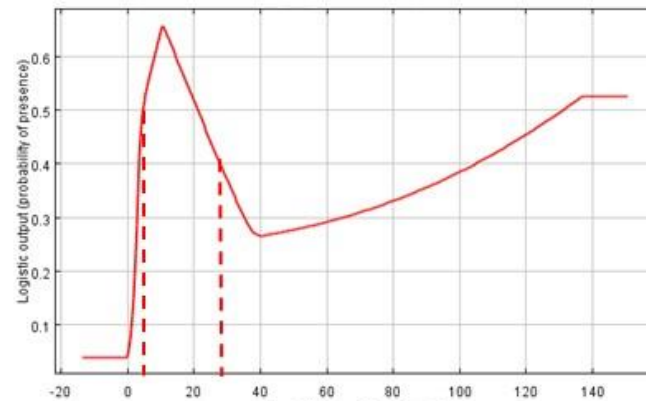
b. mean temperature of coldest quarter



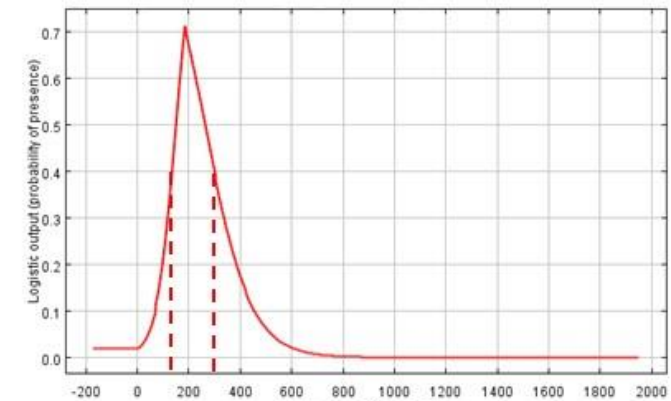
c. mean temperature of warmest quarter



d. annual precipitation

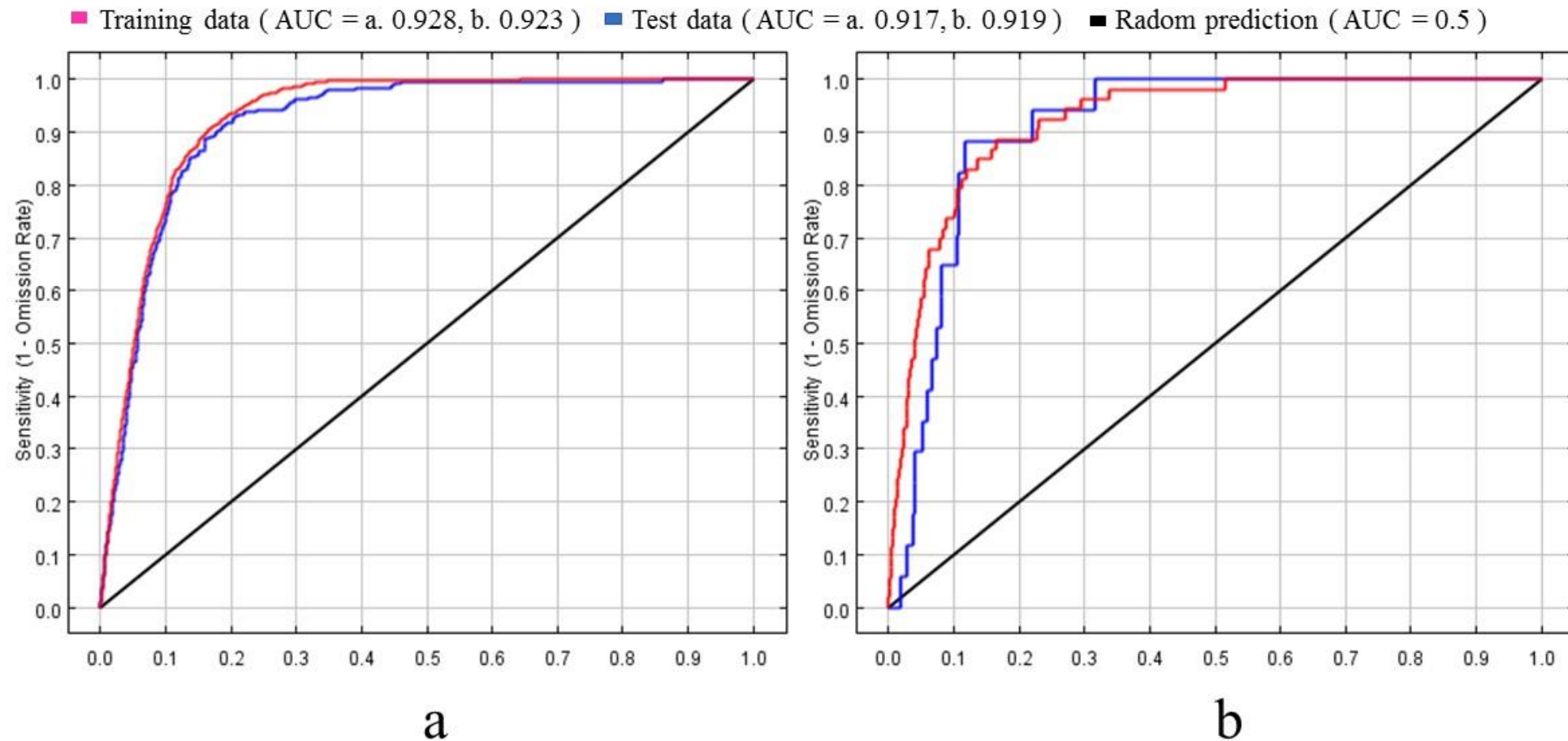


e. precipitation of driest period

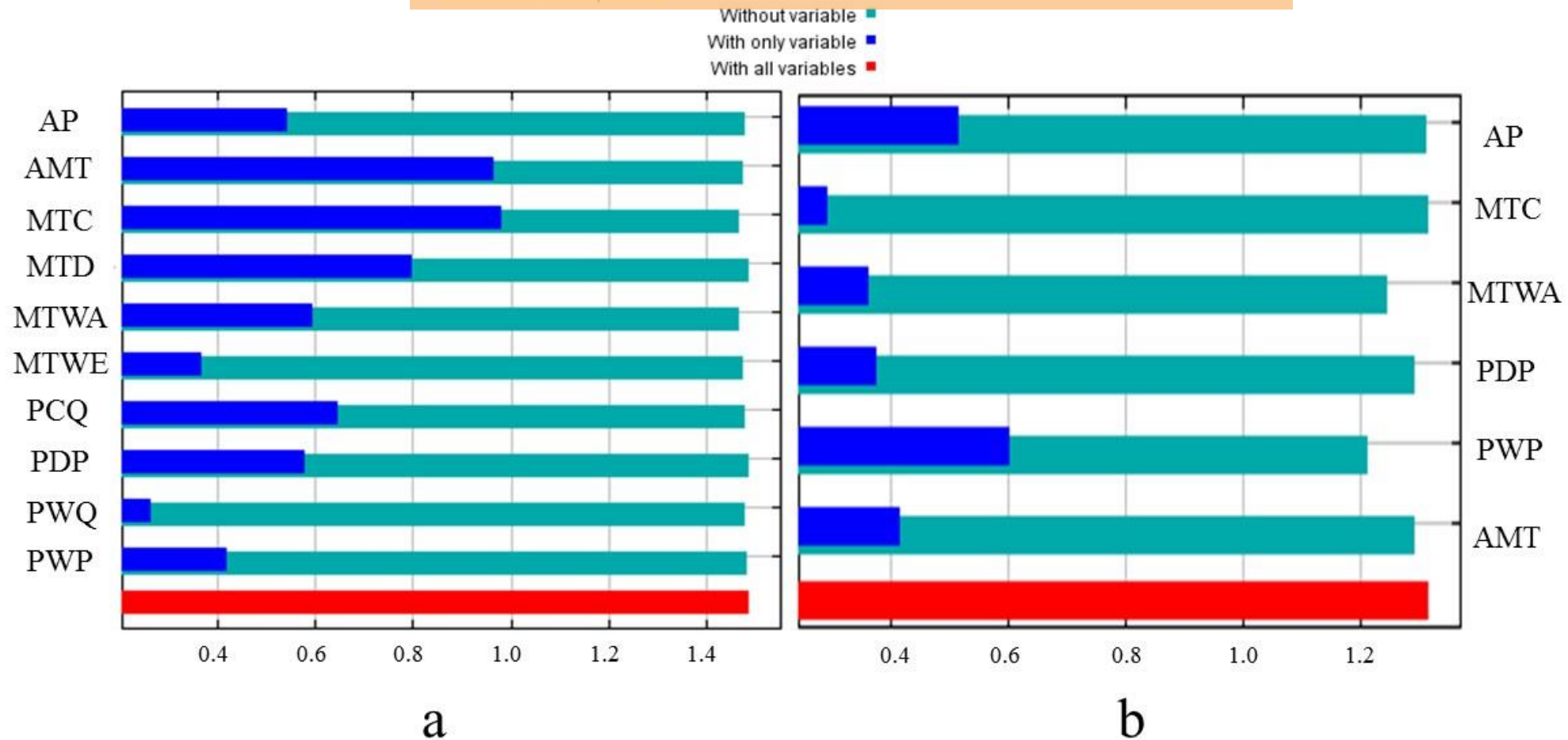


f. precipitation of wettest period

Figure 6. The response curve of climatic suitability for the climatic factors based on Maxent models predicted the distribution of *Senecio vulgaris*



Supplementary Figure 1. Accuracy assessment of Maxent model through ROC and AUC analyses. a. Predication of global potential distribution of *Senecio vulgaris* with the global occurrence records. b. Predication of potential distribution of *S. vulgaris* in China with occurrence records within China.



Supplementary Figure 2. Jackknife test result of variables contribution. a. Predication of global potential distribution of *Senecio vulgaris* with the global occurrence records. b. Predication of potential distribution of *S. vulgaris* in China with occurrence records within China.