Advantage and Disadvantage of Global and Local Climate Datasets on Modeling Species Distribution at Continental and Landscape Scales

Chi-Cheng Liao (hunter_yyl@yahoo.com.tw)
Chinese Culture University https://orcid.org/0000-0002-6409-0506
Chi-Ru CHANG
Chinese Culture University
Yi-Huey CHEN
Chinese Culture University

Research Article

Keywords: Bretschneidera sinensis, Random Forest, microclimate, Taiwan, WorldClim.

Posted Date: October 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-595860/v2

Abstract

Species distribution model based on global and local climate datasets were hypothesized to have advantages on projecting distribution range at continental and landscape scales, respectively. Random Forest (RF) and principle components analysis (PCA) aimed to project potential distribution range and to construct climate space of *Bretschneidera sinensis* in continental East Asia (CEA) and northern Taiwan (NTWN) based on the WorldClim and local climate datasets. Geographical extent of the endangered species at continental scale was available to be projected by RF based on the WorldClim dataset, whereas isolation and fragmentation of natural habitat had not been presented by the projection map in CEA. At landscape scale, projection map of RF in NTWN based on the WorldClim dataset presented gridded distribution far from empirical distribution pattern, while that based on local climate dataset presented a distribution pattern relevant to elevation and topography. PCA had revealed climate differentiation between continental and island populations. Evidently, local climate dataset is essential for identifying ecological adaptation of island population at geographical margin of the endangered species. Meteorological data interpolated and altitudinal adjusted by empirical elevation lapse rate calculated for each watershed had captured climate heterogeneity in mountainous area, whereas it generated huge number of gridded cells that is not available to expand this method to continental region. Global climate dataset has the advantage on modeling geographical extent of plant species at continental scale, while local climate dataset used for modelling species distribution enables conservationists to delineate reliable conservation areas in fragmented natural habitats at landscape scale.

Introduction

The WorldClim dataset with spatial resolution at 30 seconds (approximately 1 km) (Fick & Hijmans 2017) is one of the most common sources of climate dataset widely used as bioclimatic predictors of species distribution models (SDMs) (Karger et al. 2017; Lin et al. 2018; Scheffer et al. 2018; Silva et al. 2019; Synes & Osborne 2011; Xu et al. 2021) to explore contemporary species geographical extent and distribution patterns (Datta et al. 2020; Duque-Lazo et al. 2016; Maria & Udo 2017; Williams et al. 2009), to evaluate ecological characteristics of species (Abdelaal et al. 2019; Datta et al. 2020; Fernández & Hamilton 2015; Lobo et al. 2010; Maria & Udo 2017; Mohapatra et al. 2019), and for conservation applications such as identifying suitable habitats for rare species (Fois et al. 2015; Marmion et al. 2009). However, spatial resolution at 30 seconds was not finer enough to reflect local climate heterogeneity in mountainous areas, since local climate heterogeneity was drastically induced by local ecological forces, such as elevation and topography (Ashcroft et al. 2012; Dobrowski 2011; Lenoir et al. 2017; Meineri & Hylander 2017). Particularly, rare species with locally restrict distribution are often critically affected by certain local elements, such as the occurrences of favorable microsites and microclimates (Elith & Leathwick 2009; Heikkinen et al. 2012). Such local special ecological requirements can be difficult to distinguish in the broader-scale climate dataset (Guisan et al. 2007).

During the last decade, environmental predictors at finer resolution had been applied for SDMs to evaluate habitat differentiation and to predict species distribution patterns. Environmental datasets at fine resolutions had been generated from various data sources, including downscaling of the global climate dataset (Godsoe et al. 2015; Lin et al. 2018; Maria & Udo 2017; Mohapatra et al. 2019; Schorr et al. 2012; Wang et al. 2016; Zhu et al. 2018), interpolating climate data collected from dataloggers (Ashcroft & Gollan 2012; Ashcroft et al. 2012; Dingman et al. 2013; Fridley 2009; Greiser et al. 2018; Vanwalleghem & Meentemeyer 2009) or meteorological stations (Liao & Chen 2021; Meineri & Hylander 2017), interpolation of topographical elements, including elevation, slope and aspect (Bennie et al. 2008; Gogol-Prokurat 2011; Hanberry 2013; Marage et al. 2008; Meineri & Hylander 2017; Williams et al. 2009), or from remote sensing data (Lannuzel et al. 2021). Most recently, several R packages had been developed to model meaningful microclimate data under forest canopy and at open sites by incorporating remote sending data, vegetation structures or temperatures collected from dataloggers (Kearney et al. 2020; Kearney & Porter 2017; Maclean 2020; Maclean et al. 2019). Among these aforementioned methods, climate data from meteorological stations can potentially be used to generate local climate dataset at fine resolution, opening up the opportunity to model local-scale ecological processes over large domains (Meineri & Hylander 2017). This study used meteorological data to generate climate dataset with spatial resolution of 50×50 m^2 for modeling species distribution at landscape scale in mountainous areas. Topographical factors normally used as bioclimatic predictors actually affected on the climate environment but were not directly related to ecological processes constrained species distributions (Tomlinson et al. 2020). Therefore, an alternative method that incorporated topographical factors to interpolate climate dataset was developed to generate high resolution local climate dataset. Before interpolation of climate dataset, study area was divided into several watersheds because this method assigned contrasting effects of monsoon wind to windward and leeward slopes while generating climate dataset. For each watershed, daily data of meteorological stations were interpolated and were subsequently altitudinal adjusted by empirical lapse rates calculated from meteorological stations within the watershed to generate high resolution climate dataset (Liao & Chen 2021). High resolution climate dataset was applied for improving model performance at landscape scale in mountainous areas.

In the 20th century, economic development, urbanization, and population growth led to a highly modification of natural ecosystems, while habitats of plant species were fragmented and currently formed by complex mosaics of rural or urban areas, abandoned grasslands, agricultural landscapes, and secondary forests (Dong et al. 2019; Kier et al. 2009). The IUCN Red List of the Threatened Species (IUCN 2020, http://www.iucnredlist.org) included many plant species that have widely geographical distribution range but they are still in danger of global extinction because of species' locally rare and extremely habitat-restricted as well as habitat fragmented by urban development (Abdelaal et al. 2019; Callmander et al. 2005; Yakimowski & Eckert 2007). Thus, knowledge of contemporary distribution range is necessary for the designation of effective protected areas and conservation of rare or endangered plant species. The WorldClim dataset is advanced in modelling geographical range of endangered species at continental scale but is presumably not available for modelling species distribution at landscape scale.

Thus, the endangered species with locally rare and extremely habitat-restricted was needed to be modelled by using local climate dataset. However, high-resolution local climate dataset had generated huge number of gridded cells within a local geographical extent and, thus, the method interpolated gridded climate dataset at landscape scale is questionable to expand the geographical extent from landscape area to continental region.

SDMs can provide useful predictions for insufficiently surveyed areas and provide guidelines for conservation planning tasks, such as identifying suitable habitats and seeking new populations for rare species (Fois et al. 2015; Mi et al. 2017). Random Forest (RF), one of the SDMs, is a powerful machine learning classifier to provide accurate predictions of species distributions (Boulesteix et al. 2012; Breiman 2001; Liaw & Wiener 2002). RF has become a popular method to identify geographical extent and climate factors of rare or endangered species (Evans & Cushman 2009; Hu et al. 2017; Iturbide et al. 2018; Mi et al. 2017; Mohapatra et al. 2019; Williams et al. 2009). On the other hand, principle components analysis (PCA) was usually used for quantifying climate space and to assess ecological preferences of plant species (Early & Sax 2014; Edwards & Still 2008).

Performance of SDMs have heavily relied on the accuracy of georeferenced species occurrences data and high-resolution environmental dataset that provide discriminatory power regarding presence and absence of species to project maps indicating areas of species distributions. In this study, the WorldClim and local climate datasets were applied for RF and PCA to project potential distribution range and to construct climate space of the endangered plant species, *Bretschneidera sinensis* Hemsl. The endangered plant species is majorly distributed in Southern China and marginal populations of the species extend to Indo China, India and Taiwan (Dong et al. 2019; Hu et al. 2014; Kumar et al. 2017; Li et al. 2016; Wang et al. 2018). *B. sinensis* is a widespread endangered species with locally rare and habitat-restricted. Populations of the species are quiet small and have suffered from isolation and fragmentation of habitats (Dong et al. 2019). Distribution patterns of the endangered plant species in continental region and local area were projected by RF based on the WorldClim and local climate datasets to evaluate geographical extent and locally restricted distribution range of the plant species and to identify whether there is ecological niche diversification between central and marginal populations.

The objectives of this study were to identify the advantage and disadvantage of the global and local climate datasets at continental and landscape scales. Two hypotheses were assessed in this study. (1) The WorldClim dataset was hypothesized to have advantage on projecting geographical distribution range and quantifying climate space of plant species at continental region, whereas it was presumably not appropriate for projecting species distribution at landscape scale. (2) Local climate dataset was presumably available to capture climate heterogeneity induced by elevation and topography in mountainous areas and has the advantage on projecting species distribution range and quantifying climate space at landscape scale. Model predictions based on local climate dataset was expected to capture fine-scale ecological characteristics and to provide valuable geographical information for developing effective conservation management of the endangered species.

Material And Methods

Plant species and collections

Bretschneidera sinensis Hemsl., a deciduous tree species of montotypic genera belongs to the family Akaniaceae, is a relic species of the Tertiary tropical flora mainly occurs in evergreen broad-leaved or mixed evergreen and deciduous forests (Dong et al. 2019; Wang et al. 2018). Habitats of the species are fragmented and isolated in mountainous areas at altitudes of 300-1,700 m above sea level in Southern China (Dong et al. 2019; Li et al. 2016; Wang et al. 2018). The species is listed as endangered on the IUCN Red List of Threatened Species. According to recent studies, most of the populations are seriously threatened (Hu et al. 2014) and the species also suffers from reproductive failure due to low seed productivity and seed germination rates (Qiao et al. 2012). Geographical range of the endangered species extents from Southern China to India, Indo China, Vietnam, and Taiwan (Dong et al. 2019; Hu et al. 2014; Kumar et al. 2017; Li et al. 2016; Wang et al. 2018). Taiwan is a continental island possessed isolated population of *B. sinenis* at geographical margin of the Eurasia continent. In Taiwan, *B. sinensis* is a narrow range species usually scattered in remnant or secondary broadleaved forests only in northern part of the island (Li et al. 2013). Genetic evidence indicated that population of *B. sinensis* from mountainous areas in Taiwan was genetically diversified from continental population and was probably derived from the southeastern Yunnan through postglacial range expansion to its present range (Wang et al. 2018), but little is known about ecological niche differentiation between island and continental populations.

Georeferenced occurrences of the *B. sinensis* in continental East Asia (CEA) was downloaded from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/). A total of 281 data records were downloaded from the GBIF. The data records of *B. sinensis* were spatially verified and duplicated records as well as data records without or with error coordination locations were eliminated from the dataset. The georeferenced data records from botanical gardens or from Taiwan were also eliminated from the dataset. Finally, there were 60 georeferenced data records for modeling distribution of *B. sinensis* in CEA (Fig. 1). Presence data of *B. sinensis* in northern Taiwan (NTWN) was downloaded from the digital herbarium of the Taiwan Forestry Research Institute, National Taiwan University, and Academia Sinica, Taipei. The presence data records from digital herbarium were mapped in ArcInfo software (ESRI, Redlands, California, USA) and records with error or duplicated coordination locations were eliminated. At last, a total of 72 data records was obtained for modeling distribution of *B. sinensis* in NTWN (Fig. 1).

Climate datasets and variables selection

The WorldClim 2 dataset (available at https://www.worldclim.org/) with 30-sec spatial resolution (approximately 1 km in equator) is one of the most commonly used climate datasets in modeling species distribution (Fick & Hijmans 2017). Thirteen climate variables, 9 temperature and 4 precipitation variables, were adopted in this study (Table 1). Projection map in CEA based on the WorldClim dataset was presented to show the geographical extends and distribution patterns of the *B. sinensis* at continental scale.

However, the WorldClim dataset is presumably too coarse for projecting accurate distribution map at landscape scale in NTWN. In order to project accurate current distribution range of *B. sinensis* at landscape scale in NTWN, local climate dataset was constructed based on climate data of local meteorological stations and was used to project current distribution range of the *B. sinensis* in NTWN. Before constructing local climate dataset, NTWN was divided into five watersheds based on the mountain ridges, since climate of NTWN is characterized by monsoon winds and climate heterogeneity was observed between windward and leeward slopes in NTWN. The five watersheds are northeast (NE), northwest (NW), southwest (SW) and southeast (SE) slopes of Yamingshan area and Pingxi area (PX) (Fig. 1). Each of the watershed in NTWN was divided into gridded cells with spatial resolution of 50 × 50 m^2 performed by the "fishnet" function in ArcInfo software. More than 0.4 million gridded cells were generated for five watersheds in NTWN. For each gridded cell, longitude, latitude, elevation, slope, and aspect were obtained from a digital terrain model (DTM). Spatial resolution of the DTM at 20 by 20 meters was generated from the 20 m contour map been built by Department of Geography, Chinese Culture University.

Gridded cells of five watersheds in NTWN were used to extract climate variables independently from the WorldClim and local climate datasets performed by ArcInfo software. The WorldClim dataset was downloaded from the website, while local climate dataset was constructed in this study based on the data of local meteorological stations in NTWN. Local climate dataset was generated from daily data of 30 meteorological stations in NTWN downloaded from the website of the Central Weather Bureau, Ministry of Transportation and Communications, Taiwan (https://e-service.cwb.gov.tw/HistoryDataQuery/index.jsp). The daily data from year 1,980 to 2,000 were used to calculate mean monthly values of climate variables for each meteorological stations. Inverse Distance Weighted (IDW) method performed by ArcInfo software were used to interpolated climate data and to construct climate surfaces. For each watershed, mean monthly climate data from meteorological stations within the watershed were adopted to interpolate and generate raster files of smooth climate surfaces. Also, smooth surface of elevation was constructed for each watershed based on the elevation of meteorological stations. Gridded cells of watersheds were mapped in ArcInfo software and overlapped with the raster files of mean monthly climate data and smooth elevation surface to extract climate data and elevation. The gridded cells with interpolated climate and elevation data were then corrected by empirical elevation lapse rates calculated for each watershed. For each watershed, the elevation lapse rates were calculated by linear regression model based on the elevation and mean monthly climate data of meteorological stations. Subsequently, the interpolated climate data of each gridded cells were adjusted by the following function:

ClimData1 = slope (Elev1 – Elev2) + ClimData2

ClimData1 is the altitudinal adjusted mean monthly data calculated for each gridded cell. Elev1 is the elevation captured from DTM. ClimData2 and Elev2 were from the interpolated mean monthly data and from smooth elevation surface generated in ArcInfo software. The "slope" in the function is the regression slope from the linear regression model. Climate variables for modeling analysis were selected or re-calculated from the altitudinal adjusted mean monthly data. Because of the altitudinal adjustment of mean monthly climate data, elevation is not directly used as a model predictor in this study. An additional model prediction in this study had been performed and indicated that elevation contributed few to the model performance (data not show here). Nonetheless, slope and aspect obtained from DTM were included in the local climate dataset as model predictors. Aspect given by the DTM (0-360°) was normalized between 0 and 10. The normalized function was revised from the previous study (Williams et al. 2009) and the function in this study is "Aspect = |180-X| / 18". Precipitations were rescaled from mm to dm. In order to make the models based on the two climate datasets comparable, the same climate variables independently from the two sources of climate datasets were applied to the RF model.

Random Forest algorithm

RF algorithm implemented by "randomForest" package within the R environment (Breiman 2001; Liaw & Wiener 2002) was applied to project geographical extent of the *B. sinensis* at continental scale and to predict local distribution pattern at landscape scale. A total of 60 and 72 georeferenced data records were available for modeling plant distribution in CEA and NTWN, respectively. The coordination of presence data records was adopted to extract climate variables from the WorldClim and local climate datasets, repectively. The presence data with topographical and climate variables independently derived from the two climate datasets were imported to the RF as two training datasets. False absence points (or background points) were randomly selected within the study range as the same number of presence data. False absence points were combined with the presence data to construct the training dataset of the model and subsequently modelled 100 times with the resampled background points to quantify uncertainties in prediction. The two testing datasets were the gridded cells with climate variables derived from the WorldClim and local climate datasets. RF algorithm categorized gridded cells into presence and absence and gridded cells with 50 times of presences categorized by RF were utilized to project potential distribution range of the endangered plant species. The predictive accuracy of RF was evaluated by AUC (area under operating characteristic curves) and out-of-bag estimates of the error rate (ERR_{OOB}).

Quantification of climate spaces by principle components analysis (PCA)

Further analysis aims to make a comparison between climate spaces of the *B. sinensis* populations in CEA and NTWN constructed by the WorldClim and local climate datasets. Principle components analysis (PCA) was performed to quantify climate spaces of the endangered plant species and to explicitly understand the ecological preferences and climate environments represented by the WorldClim and local climate datasets. In order to construct datasets for quantification of climate spaces, a total of 1.5 million gridded points were generated and randomly scattered in CEA. These 1.5 million random gridded points were imported to ArcInfo software and overlapped with the WorldClim to capture climate data and that were also overlapped with projection map generated by RF in CEA to capture the model results. The areas in CEA for quantification of climate spaces include Mongolia, China, India and Indo China. The 1.5 million random gridded points were subsequently categorized into presence and absence datasets of *B. sinensis* and gridded cells with 50 times of presence were categorized in presence dataset. There were another four datasets for quantification of climate spaces, they were the presence and absence data of *B. sinensis* projected by RF in NTWN based on the WorldClim and local climate datasets. PCA was applied to scaled data for 13 studied climate variables corresponding to the formation of climate spaces of *B. sinensis*. Topographical factors, slope and aspect, were not implemented into PCA, since these factors contributed little to model prediction in NTWN and was not used in the model evaluation of plants in CEA. Among the 13 climate variables implemented to PCA, annual, winter, and summer precipitations were rescaled from mm to dm. The 13 climate variables were thought to provide climate preferences for the plants' distributions and the ecological demands were distilled into three principal components, the first, second, and third axis from a PCA.

Results

Potential distribution range of B. sinensis

Distribution maps of *B. sinensis* in CEA projected by RF based on the construction of presence and absence data extracted climate variables from the WorldClim dataset presented suitable habitats of the plant species in Southern China (Fig. 2a). The AUC (area under operating characteristic curves) value of the model was 0.995. The model had effectively presented potential distribution range of the plant species at continental scale. However, projection map based on the WorldClim dataset presented no isolation and fragmentation of potential habitats at Southern China that is contradict to the proposed empirical distribution pattern of the endangered species. Furthermore, RF had projected gridded square map when the map was zoom in to clearly show the edge of the potential distribution range in a larger scale map (Fig. 2b and 2c). Thus, projection map based on the WorldClim dataset may not reflect the empirical distribution pattern of the plant species and there were uncertainties of the distribution pattern based on the WorldClim dataset at continental scale.

Potential distribution range of the plant species was further examined in NTWN and the map projected by RF based on the WorldClim and local climate datasets had presented completely different distribution patterns at landscape scale (Fig. 3a and 3b). Distribution map of *B. sinensis* based on the WorldClim dataset had presented gridded distribution pattern around the presence data records that was irrelevant to the topography and elevation (Fig. 3a). Gridded distribution map in mountainous areas of NTWN based on the WorldClim was an unrealistic distribution pattern and it was far from empirical distribution pattern of plant species. In contrast, distribution map of *B. sinensis* based on local climate dataset was pertinent to topography and elevation (Fig. 3b), that was more likely to present empirical distribution pattern at landscape scale.

The error rates (ERR_{OOB}) of the two projection results in NTWN presented no significant differences by statistical test (0.9177 for the WorldClim and 0.9103 for local climate datasets). Moreover, the climate variables contributed most to the model performances were different between the two climate datasets (Fig. 4). Diurnal temperature range (Bio2) derived from the WorldClim dataset was the most important factor affecting model performance, while water availability (Bio19) contributed most to the model performance based on the local climate dataset. In addition, topographical factors, slope and aspect, both contributed little to model prediction in NTWN (Fig. 4).

Quantification of climate spaces by PCA

PCA was conducted to identify climate characteristics of the CEA and NTWN and to depict climate space of the plant species. Principle component 1 (PC1) accounted for 64.41% of the variation, while principle component 2 (PC2) and 3 (PC3) 20.86% and 10.12% of the variation, respectively (Fig. 5). PC1 and PC2 were strongly related to Bio12 and Bio3, respectively (Table 2). Apparently, water availability and temperature had evidently played as the important roles for the quantification of climate spaces of *B. sinensis*.

The grey points in the PCA diagram represent climate environments of CEA based on the WorldClim dataset and that of NTWN based on the WorldClim and local climate datasets (Fig. 5). In the PCA diagram, the upper left grey points represent the India and Indo China with humid climates, while the upper right grey points represent northern areas of CEA with relatively drought climates. The dark points clumped at the lower central part of the grey triangle represent the distribution range of *B. sinensis* in southern China projected by RF base on the WorldClim. Interestingly, climate environments in NTWN constructed by the WorldClim and local climate datasets were distant from each other in the PCA diagram. The different climate environments were probably caused by different temperature ranges between the WorldClim and local climate datasets in NTWN. Similarly, distribution of *B. sinensis* in NTWN projected by RF occupies two distinct climate spaces based on the WorldClim and local climate datasets represented by dark circles in the PCA diagram (Fig. 5). It is interesting that the same geographical coordination locations of the plant species presented distinct climate spaces based on the different climate datasets. Furthermore, statistical tests were performed to detect which variable was the most responsible for the different climate spaces in the PCA diagram. Surprisingly, results of analysis of variance (ANOVA) and Tukey's test presented that all the variables have significant differences among the three climate spaces of the plant species (Table 3).

Climate characteristics were drastically different between CEA and NTWN based on the WorldClim and local climate datasets. Histogram had showed wide range of mean diurnal temperature range (Bio2) in CEA based on WorldClim (Fig. 6a), since the CEA had broad latitude range. However, annual precipitation in NTWN based on local climate dataset (Fig. 6i) had extraordinary broad range in contrast to that in CEA and NTWN based on the WorldClim dataset (Fig. 6c and 6f). Wide range of NTWN's climate environment along the PC1 was caused by extraordinary high annual precipitation that was partly a consequence of high winter precipitation at coastal area and lower at inland area in NTWN. Climate preferences of *B. sinensis* are completely different between CEA and NTWN demonstrated that climate niche differentiation exists between island and continental populations of the endangered species. Local climate dataset generated from daily data of meteorological stations was available to reflect climate heterogeneity in NTWN at landscape scale that was not available to be exhibited by the WorldClim dataset.

Discussion

The WorldClim dataset provided global climate surface that had been widely used in many previous SDMs studies as bioclimatic predictors and had accurately projected species distribution range at continental regions (Abdelaal et al. 2019; Datta et al. 2020; Heikkinen et al. 2012; Mi et al. 2017; Mohapatra et al. 2019). In this study, projection of potential distribution range and quantification of climate space were available at continental scale and were evidently the advantage of the WorldClim dataset. Model projections aims to project geographical extent, identify suitable key sites, and find suitable environmental factors based on the presence/absence data and climate variables that can provide scientific information for developing conservation and management strategies at continental scale (Abdelaal et al. 2019; Hu et al. 2017; Lobo et al. 2010; Maria & Udo 2017; Mi et al. 2017; Mohapatra et al. 2019; Williams et al. 2009). Modeling efforts to asses potential distribution range had been made for determining conservation areas of plant species in China (Wan et al. 2017; Xu et al. 2021; Yu et al. 2017; Zhang et al. 2017). However, projection map at landscape scale would be the disadvantage of the WorldClim dataset. Restricted elevation ranges and isolation and fragmentation of suitable habitats of *B. sinensis* had not been presented by the model projection in CEA. In addition, the WorldClim dataset is too coarse and is evidently unavailable to accurately reflect climate heterogeneity induced by elevation and topography (Fig. 7) and gridded distribution map in NTWN projected by RF based on the WorldClim was an unrealistic distribution pattern and it was far from empirical distribution of the plant species. Model predictions based on the global climate dataset had evidently generated bias projection map and provided misleading results of species geographical distribution at landscape scale.

To avoid misleading result caused by the WorldClim dataset, local climate dataset was suggested to apply as bioclimatic predictors of model prediction at landscape scale. Local climate dataset derived from meteorological data and generated by IDW and altitudinal adjustment in this study was advanced in reflecting local climate heterogeneity in NTWN (Fig. 7). In NTWN, winter monsoon from northeast and summer monsoon from southwest carry high moisture atmosphere influenced the NE and SW slopes of Yangmingshan area, respectively. Meteorological stations at NE slope recorded significant low winter temperature as well as extraordinary high winter precipitation and subsequently annual precipitation. Although climate variables were adjusted by the IDW and elevation lapse rates, concentric circles centered at the stations indicated extremely higher values of NE slope in comparison with the values of other slopes. Extremely high climate values had resulted strong color contrast in the three climate maps (Fig. 7b, 7d and 7f) that should not be recognized as artefact data. In addition, climate maps had showed different patterns of temperature and precipitation gradients in NTWN. Mean annual temperature had gently changed along elevation and topography clearly presented in the climate map (Fig. 7b). On the other hand, diurnal temperature range (Fig. 7d), annual precipitation (Fig. 7f), and winter precipitation (Fig. 7h) had showed considerable differences among low and high elevations or between windward and leeward slopes. Gradients of climate variables were not always gentle and sometimes presented extraordinary steep gradient of climate heterogeneity at landscape scale. Climate heterogeneity might have been presented by the co-occurrences of gentle and steep gradients of climate variables. Local climate dataset derived from meteorological stations had generated gentle gradient of mean annual temperature as well as very steep gradients of diurnal temperature range, winter and annual precipitations that should be considered in modeling distribution range of plant species at landscape scale in NTWN.

Because of steep gradients and high variations of climate variables, climate environment of NTWN based on local climate dataset occupies wide range of climate space in PCA diagram. However, local climate dataset generated huge number of gridded cells within a local geographical area and there were more than 0.4 million gridded cells in NTWN that is only 1,041 km^2. Because of huge number of gridded cells, SDMs studies based on local climate dataset are highly difficult to expand geographical range from landscape to continental scales. Model predictions at landscape scale is evidently the advantage of local climate dataset, whereas that at continental scale would be the disadvantage of local climate dataset because of huge number of gridded cells.

Climate environment of NTWN based on the WorldClim dataset occupies a narrow range of climate space in PCA diagram. Narrow climate environment of NTWN in PCA diagram based on the WorldClim was consistent with the smaller standard deviations (SDs) of mean climate data in Table 3. SDs of the mean climate data from the WorldClim dataset in NTWN were almost completely different from and smaller than that from local climate dataset (Table 3). Climate environment based on the WorldClim dataset did not accurately reflect climate heterogeneity of mountainous area at landscape scale. Therefore, projection map based on the WorldClim dataset does not accurately present natural distribution range of plant species at landscape scale.

Previous studies had proposed that RF model with low error rate (ERR_{OOB}) were considered to have more accurate performance. Our study had confirmed that low error rate (ERR_{OOB}) can be achieved even when model did not project accurate distribution range and projection map was evidently far from empirical distribution pattern of the plant species. Low error rate (ERR_{OOB}) did not guarantee an accurate projection map of species distribution.

RF model performance might be affected by bias collections. Bias collection of presence data had commonly been proposed by many previous studies (Datta et al. 2020; El-Gabbas & Dormann 2018; Ferro & Flick 2015; Lannuzel et al. 2021; Tomlinson et al. 2020), since comprehensive collection is a great challenge to be conducted in field survey. Bias collection may lead to bias projection of species distribution. Projection map caused by bias collection at continental scale is unavailable to be applied for conservation management at landscape scale. On the contrary, projection map at landscape scale is available to apply for conservation management, since projection map based on bias collection of presence data can be compensate for the correction of projecting results at landscape scale while designing management strategies or delineating boundaries of conservation area by conservationists.

Over-fit may lead to loss of generality of the model and may have poor performance when this method is expanded to other regions or other species. However, over-fit may not be a deficit of this method, since a previous study had accurately predicted species distribution range at landscape scale based on the same method (Liao & Chen 2021). More studies were expected to identify the generality of model prediction of this method. Despite of, this method generated precise distribution map of the plant species in mountainous area at landscape scale that is evidently close to its empirical distribution range and is very useful for further conservation applications.

Marginal population of continental species had showed striking genetic differences from central populations in neighboring continent (Sexton et al. 2009; Thompson et al. 2005; Wake et al. 2009; Wu et al. 2001), while this study had further identified ecological differentiation between island and continental populations. *B. sinensis* had locally adapted to climate environments in NTWN. Water availability had evidently played as a major climate factor related to the potential distribution range of the *B. sinensis* in NTWN and the plant species was evidently, locally adapted to extraordinary high precipitation in NTWN. Local adaptation of marginal population to novel habitats at geographical margin is akin to niche evolution (Sexton et al. 2009). Climate characteristics of marginal population distinguished from central populations may consequently contribute to the genetic variation. Thus, marginal population on continental islands warrant a high priority in biodiversity conservation (Kier et al. 2009; Weigelt & Kreft 2013), since genetic variation (Li et al. 2016; Wang et al. 2018; Wu et al. 2001) and ecological differentiation both existed between central and marginal populations.

The growing impacts of climate change on plant species calls a request to evaluate geographical extent where species with narrow distribution ranges exist or likely exist in order to enhance their conservation and restoration (Abdelaal et al. 2019). However, designation and effective management of conservation areas is a great challenge on islands because anthropogenic disturbances caused fragmentation and isolation of natural habitats and, most importantly, the complex mosaics of natural and artificial ecosystems (Kier et al. 2009). Conservation and management planning, such as the selection of representative conservation sites, may critically depend on the detailed knowledge of empirical species distribution patterns (Allouche et al. 2006; Kier et al. 2009). Model predictions based on local climate dataset had illustrated projection map accurately showing distribution patterns of plant species along elevation and topography that is useful to delineate conservation area in fragmented habitat at landscape scale. In fact, projection map of *B. sinensis* in NTWN had provided useful information for delineating effective conservation area to protect isolated populations in fragmented habitats.

Conclusion

This study suggested a two-step procedure to identify geographical extent and climate characteristics of central and marginal populations of plant species at continental and landscape scales. Global climate dataset has the advantage on projecting geographical extent and climate characteristics of plant species and is evidently appropriate to evaluate effects of climate change on species distribution range at continental scale. High-resolution local climate dataset had accurately captured climate heterogeneity induced by elevation and topography and is available for SDMs to project accurate distribution map represented species distribution pattern at landscape scale. Accurate distribution map based on local climate dataset offers an advantage in designing conservation areas for plant species in mountainous areas with isolation and fragmentation of natural habitat, since it delineated accurate geographical boundaries of the species along elevation and topography. Model predictions based on local climate dataset was also a powerful tool to correlate species distribution with climate factors at landscape scale that offered detailed geographical and ecological information of plant species and can be confidently applied for assessing the impacts of future climate change on shifting species distribution range at landscape scale.

Declarations

Acknowledgement

The authors deeply appreciate assistant researcher Lin, Huan-Yu in Taiwan Forestry Research Institute for his helping and sharing technique of climate data interpolation. The authors want to thank for Mr. Shih-Chieh Kuo and many students in the Department of Life Science, Chinese Culture University, Taipei, Taiwan for their helps in data collections. The authors appreciate Mr. Kai-Jie Yang in the Institution of Geography, Chinese Culture University, Taipei, Taiwan for his technical support of ArcGIS software. This work was in part supported by Yangmingshan National Park, Construction and Planning Agency Ministry of the Interior, Executive Yuan, Taipei, Taiwan.

Data Availability

The data analyzed in this study was downloaded from Global Biodiversity Information Facility (GBIF) and digital herbarium of the Taiwan Forestry Research Institute, digital herbarium of National Taiwan University, and digital herbarium of the Academia Sinica, Taipei.

Consent to Publish (Ethics)

All the authors agree to publish the manuscript.

Author Contribution

CC Liao conceived and designed the study and analyzed data. CR Chang and YH Cheng discussed with CC Liao and contributed novel ideas of the study. All authors had read and approved the manuscript.

Conflict of Interest

The authors declare no conflict of interest

Funding

This work is partly funded by Yangmingshan National Park, Construction and Planning Agency Ministry of the Interior, Executive Yuan, Taipei, Taiwan.

References

1. Abdelaal M., Fois M., Fenu G. and Bacchetta G. 2019. Using MaxEnt modeling to predict the potential distribution of the endemic plant *Rosa arabica* Crép. in Egypt. Ecological Informatics 50: 68-75.
2. Allouche O., Tsoar A. and Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43: 1223-1232.
3. Ashcroft M.B. and Gollan J.R. 2012. Fine-resolution (25 m) topoclimatic grids of near-surface (5 cm) extreme temperatures and humidities across various habitats in a large (200× 300 km) and diverse region. International journal of climatology 32: 2134-2148.
4. Ashcroft M.B., Gollan J.R., Warton D.I. and Ramp D. 2012. A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. Global Change Biology 18: 1866-1879.
5. Bennie J., Huntley B., Wiltshire A., Hill M.O. and Baxter R. 2008. Slope, aspect and climate: spatially explicit and implicit models of topographic microclimate in chalk grassland. Ecological Modelling 216: 47-59.
6. Boulesteix A.L., Janitza S., Kruppa J. and König I.R. 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2: 493-507.
7. Breiman L. 2001. Random forests. Machine learning 45: 5-32.
8. Callmander M.W., Schatz G.E. and Lowry P.P. 2005. IUCN Red List assessment and the Global Strategy for Plant Conservation: taxonomists must act now. Taxon: 1047-1050.
9. Datta A., Schweiger O. and Kühn I. 2020. Origin of climatic data can determine the transferability of species distribution models. NeoBiota 59: 61.
10. Dingman J.R., Sweet L.C., McCullough I., Davis F.W., Flint A., Franklin J. and Flint L.E. 2013. Cross-scale modeling of surface temperature and tree seedling establishment in mountain landscapes. Ecological Processes 2: 1-15.
11. Dobrowski S.Z. 2011. A climatic basis for microrefugia: the influence of terrain on climate. Global Change Biology 17: 1022-1035.
12. Dong F., Liu X., Zhang Y., Lu R., Zheng Z., Shen X. and Li Y. 2019. Characterization of the complete plastid genome sequence of the endangered species *Bretschneidera sinensis*. Conservation Genetics Resources 11: 287-290.
13. Duque-Lazo J., Van Gils H., Groen T. and Navarro-Cerrillo R. 2016. Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. Ecological Modelling 320: 62-70.
14. Early R. and Sax D.F. 2014. Climatic niche shifts between species' native and naturalized ranges raise concern for ecological forecasts during invasions and climate change. Global Ecology and Biogeography 23: 1356-1365.
15. Edwards E.J. and Still C.J. 2008. Climate, phylogeny and the ecological distribution of C4 grasses. Ecology Letters 11: 266-276.

16. El-Gabbas A. and Dormann C.F. 2018. Wrong, but useful: regional species distribution models may not be improved by range-wide data under biased sampling. Ecology and Evolution 8: 2196-2206.
17. Elith J. and Leathwick J.R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40: 677.
18. Evans J.S. and Cushman S.A. 2009. Gradient modeling of conifer species using random forests. Landscape Ecology 24: 673-683.
19. Fernández M. and Hamilton H. 2015. Ecological niche transferability using invasive species as a case study. Plos One 10: e0119891.
20. Ferro M.L. and Flick A.J. 2015. “Collection Bias” and the Importance of Natural History Collections in Species Habitat Modeling: A Case Study Using *Thoracophorus costalis* Erichson (Coleoptera: Staphylinidae: Osoriinae), with a Critique of GBIF.org. The Coleopterists Bulletin 69: 415-425.
21. Fick S.E. and Hijmans R.J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International journal of climatology 37: 4302-4315.
22. Fois M., Fenu G., Lombrana A.C., Cogoni D. and Bacchetta G. 2015. A practical method to speed up the discovery of unknown populations using Species Distribution Models. Journal for Nature Conservation 24: 42-48.
23. Fridley J.D. 2009. Downscaling climate over complex terrain: high finescale (< 1000 m) spatial variation of near-ground temperatures in a montane forested landscape (Great Smoky Mountains). Journal of Applied Meteorology and Climatology 48: 1033-1049.
24. Godsoe W., Murray R. and Plank M.J. 2015. Information on biotic interactions improves transferability of distribution models. The American Naturalist 185: 281-290.
25. Gogol-Prokurat M. 2011. Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. Ecological Applications 21: 33-47.
26. Greiser C., Meineri E., Luoto M., Ehrlén J. and Hylander K. 2018. Monthly microclimate models in a managed boreal forest landscape. Agricultural and forest meteorology 250: 147-158.
27. Guisan A., Zimmermann N.E., Elith J., Graham C.H., Phillips S. and Peterson A.T. 2007. What matters for predicting the occurrences of trees: techniques, data, or species characteristics? Ecological Monographs 77: 615-630.
28. Hanberry B.B. 2013. Finer grain size increases effects of error and changes influence of environmental predictors on species distribution models. Ecological Informatics 15: 8-13.
29. Heikkinen R.K., Marmion M. and Luoto M. 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography 35: 276-288.
30. Hu X.-G., Wang T., Liu S.-S., Jiao S.-Q., Jia K.-H., Zhou S.-S., Jin Y., Li Y., El-Kassaby Y.A. and Mao J.-F. 2017. Predicting future seed sourcing of *Platycladus orientalis* (L.) for future climates using climate niche models. Forests 8: 471.
31. Hu Z.-y., Lin L., Deng J.-f. and Wang S.-h. 2014. Genetic diversity and differentiation among populations of *Bretschneidera sinensis* (Bretschneideraceae), a narrowly distributed and endemic

species in China, detected by inter-simple sequence repeat (ISSR). Biochemical Systematics and Ecology 56: 104-110.

32. Iturbide M., Bedia J. and Gutiérrez J.M. 2018. Background sampling and transferability of species distribution model ensembles under climate change. Global and Planetary Change 166: 19-29.
33. Karger D.N., Conrad O., Böhner J., Kawohl T., Kreft H., Soria-Auza R.W., Zimmermann N.E., Linder H.P. and Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. Scientific data 4: 170122.
34. Kearney M.R., Gillingham P.K., Bramer I., Duffy J.P. and Maclean I.M. 2020. A method for computing hourly, historical, terrain-corrected microclimate anywhere on Earth. Methods in Ecology and Evolution 11: 38-43.
35. Kearney M.R. and Porter W.P. 2017. NicheMapR–an R package for biophysical modelling: the microclimate model. Ecography 40: 664-674.
36. Kier G., Kreft H., Lee T.M., Jetz W., Ibisch P.L., Nowicki C., Mutke J. and Barthlott W. 2009. A global assessment of endemism and species richness across island and mainland regions. PNAS 106: 9322-9327.
37. Kumar V., Dash S.S., Panday S., Lahiri S., Sinha B.K. and Singh P. 2017. Akaniaceae: A New Family Record for Flora of India and Lectotypification of the Name *Bretschneidera sinensis*. Bulletin of the Botanical Survey of India 59: 8-10.
38. Lannuzel G., Balmot J., Dubos N., Thibault M. and Fogliani B. 2021. High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. BIodiversity and Conservation.
39. Lenoir J., Hattab T. and Pierre G. 2017. Climatic microrefugia under anthropogenic climate change: implications for species redistribution. Ecography 40: 253-266.
40. Li C.F., Chytrý M., Zelený D., Chen M.Y., Chen T.Y., Chiou C.R., Hsia Y.J., Liu H.Y., Yang S.Z. and Yeh C.L. 2013. Classification of Taiwan forest vegetation. Applied Vegetation Science 16: 698-719.
41. Li M., Chen H., Wang Z. and Zhang S. 2016. Isolation and characterization of polymorphic microsatellite markers in the endangered species *Bretschneidera sinensis* Hemsl. Genetics and Molecular Research 15.
42. Liao C.C. and Chen Y.H. 2021. Improving performance of species distribution model in mountainous areas with complex topography. Ecological Research 36: 648-662.
43. Liaw A. and Wiener M. 2002. Classification and regression by randomForest. R news 2: 18-22.
44. Lin H.Y., Hu J.M., Chen T.Y., Hsieh C.F., Wang G. and Wang T. 2018. A dynamic downscaling approach to generate scale-free regional climate data in Taiwan. Taiwania 63: 251-266.
45. Lobo J.M., Jiménez-Valverde A. and Hortal J. 2010. The uncertain nature of absences and their importance in species distribution modelling. Ecography 33: 103-114.
46. Maclean I.M. 2020. Predicting future climate at high spatial and temporal resolution. Global Change Biology 26: 1003-1011.

47. Maclean I.M., Mosedale J.R. and Bennie J.J. 2019. Microclima: An r package for modelling meso- and microclimate. Methods in Ecology and Evolution 10: 280-290.
48. Marage D., Garraud L. and Rameau J.C. 2008. The influence of management history on spatial prediction of Eryngium spinalba, an endangered endemic species. Applied Vegetation Science 11: 139-148.
49. Maria B. and Udo S. 2017. Why input matters: Selection of climate data sets for modelling the potential distribution of a treeline species in the Himalayan region. Ecological Modelling 359: 92-102.
50. Marmion M., Parviainen M., Luoto M., Heikkinen R.K. and Thuiller W. 2009. Evaluation of consensus methods in predictive species distribution modelling. Diversity and Distributions 15: 59-69.
51. Meineri E. and Hylander K. 2017. Fine-grain, large-domain climate models based on climate station and comprehensive topographic information improve microrefugia detection. Ecography 40: 1003-1013.
52. Mi C., Huettmann F., Guo Y., Han X. and Wen L. 2017. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. PeerJ 5: e2849.
53. Mohapatra J., Singh C.P., Hamid M., Verma A., Semwal S.C., Gajmer B., Khuroo A.A., Kumar A., Nautiyal M.C. and Sharma N. 2019. Modelling *Betula utilis* distribution in response to climate-warming scenarios in Hindu-Kush Himalaya using random forest. BIodiversity and Conservation 28: 2295-2317.
54. Qiao Q., Chen H., Xing F., Wang F., Zhong W., Wen X. and Hou X. 2012. Pollination ecology of *Bretschneidera sinensis* (Hemsley), a rare and endangered tree in China. Pakistan Journal of Botany 44: 1897-1903.
55. Scheffer M., Xu C., Hantson S., Holmgren M., Los S.O. and van Nes E.H. 2018. A global climate niche for giant trees. Global Change Biology 24: 2875-2883.
56. Schorr G., Holstein N., Pearman P., Guisan A. and Kadereit J. 2012. Integrating species distribution models (SDMs) and phylogeography for two species of Alpine Primula. Ecology and Evolution 2: 1260-1277.
57. Sexton J.P., McIntyre P.J., Angert A.L. and Rice K.J. 2009. Evolution and ecology of species range limits. Annual Review of Ecology, Evolution and Systematics 40: 415-436.
58. Silva S.M., Peterson A.T., Carneiro L., Burlamaqui T.C.T., Ribas C.C., Sousa-Neves T., Miranda L.S., Fernandes A.M., d'Horta F.M. and Araújo-Silva L.E. 2019. A dynamic continental moisture gradient drove Amazonian bird diversification. Science advances 5: eaat5752.
59. Synes N.W. and Osborne P.E. 2011. Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. Global Ecology and Biogeography 20: 904-914.
60. Thompson J.D., Lavergne S., Affre L., Gaudeul M. and Debussche M. 2005. Ecological differentiation of Mediterranean endemic plants. Taxon 54: 967-976.

61. Tomlinson S., Lewandrowski W., Elliott C.P., Miller B.P. and Turner S.R. 2020. High-resolution distribution modeling of a threatened short-range endemic plant informed by edaphic factors. Ecology and Evolution 10: 763-777.
62. Vanwalleghem T. and Meentemeyer R. 2009. Predicting forest microclimate in heterogeneous landscapes. Ecosystems 12: 1158-1172.
63. Wake D.B., Hadly E.A. and Ackerly D.D. 2009. Biogeography, changing climates, and niche evolution. PNAS 106: 19631-19636.
64. Wan J.-Z., Wang C.-J. and Yu F.-H. 2017. Spatial conservation prioritization for dominant tree species of Chinese forest communities under climate change. Climatic Change 144: 303-316.
65. Wang M.-N., Duan L., Qiao Q., Wang Z.-F., Zimmer E.A., Li Z.-C. and Chen H.-F. 2018. Phylogeography and conservation genetics of the rare and relict *Bretschneidera sinensis* (Akaniaceae). Plos One 13: e0189034.
66. Wang T., Hamann A., Spittlehouse D. and Carroll C. 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. Plos One 11: e0156720.
67. Weigelt P. and Kreft H. 2013. Quantifying island isolation–insights from global patterns of insular plant species richness. Ecography 36: 417-429.
68. Williams J.N., Seo C., Thorne J., Nelson J.K., Erwin S., O'Brien J.M. and Schwartz M.W. 2009. Using species distribution models to predict new occurrences for rare plants. Diversity and Distributions 15: 565-576.
69. Wu J.-E., Huang S., Wang J.-C. and Tong W.-F. 2001. Allozyme variation and the genetic structure of populations of *Trochodendron aralioides*, a monotypic and narrow geographic genus. Journal of Plant Research 114: 45-57.
70. Xu Y., Huang Y., Zhao H., Yang M., Zhuang Y. and Ye X. 2021. Modelling the effects of climate change on the distribution of endangered *Cypripedium japonicum* in China. Forests 12: 429.
71. Yakimowski S.B. and Eckert C.G. 2007. Threatened peripheral populations in context: geographical variation in population frequency and size and sexual reproduction in a clonal woody shrub. Conservation Biology 21: 811-822.
72. Yu F., Skidmore A.K., Wang T., Huang J., Ma K. and Groen T.A. 2017. Rhododendron diversity patterns and priority conservation areas in China. Diversity and Distributions 23: 1143-1156.
73. Zhang M.-G., Slik J.F. and Ma K.-P. 2017. Priority areas for the conservation of perennial plants in China. Biological Conservation 210: 56-63.
74. Zhu Y., Wei W., Li H., Wang B., Yang X. and Liu Y. 2018. Modelling the potential distribution and shifts of three varieties of *Stipa tianschanica* in the eastern Eurasian Steppe under multiple climate change scenarios. Global Ecology and Conservation 16: e00501.

Tables

Table1 Bioclimatic predictors used for modeling potential distribution range of *Bretschneidera sinensis* Hemsl.

Abbreviation	Climatic variable
Temperature	
BIO1	annual mean temperature
BIO2	mean diurnal temperature range [mean of monthly (maximum temperature - minimum temperature)]
BIO3	isothermality (BIO2 / BIO7 · 100)
BIO4	temperature seasonality (standard deviation of monthly temperature)
BIO5	maximum temperature of the warmest month
BIO6	minimum temperature of the coldest month
BIO7	temperature range (BIO6 - BIO5)
BIO10	mean temperature of summer
BIO11	mean temperature of winter
Precipitation	
BIO12	annual precipitation
BIO15	precipitation seasonality (standard deviation of monthly precipitation)
BIO18	precipitation of the summer
BIO19	precipitation of the winter

Table 2 The first three axes of the principle components analysis (PCA) on the correlation matrix of bioclimatic predictors from the WorldClim and local climate datasets.

	PC1	PC2	PC3
Bio1	-0.2834	-0.0862	-0.2512
Bio2	0.0167	0.3151	-0.0590
Bio3	-0.2768	0.8011	-0.2620
Bio4	0.1102	0.0208	0.0254
Bio5	-0.1401	0.0547	-0.2561
Bio6	-0.4534	-0.2858	-0.2248
Bio7	0.3133	0.3406	-0.0313
Bio10	-0.1428	-0.0944	-0.2190
Bio11	-0.4439	-0.0681	-0.2698
Bio12	-0.5077	0.1253	0.7278
Bio15	0.0009	0.0298	-0.0018
Bio18	-0.0873	-0.0316	0.1136
Bio19	-0.1561	0.1388	0.2892

Table 3 Statistical test by analysis of variance (ANOVA) and Tukey's test for comparisons of the bioclimatic predictors from three datasets, WorldClim in continental East Asia (CEA) and WorldClim in northern Taiwan (NTWN), and local climate dataset in NTWN. The values in the table are mean ± SD.

	WorldClim in CEA	WorldClim in NTWN	Local climate in NTWN
Bio1	10.5 ± 11.2[a]	20.5 ± 1.0[b]	21.4 ± 1.4[c]
Bio2	12.1 ± 2.4[a]	5.4 ± 0.3[b]	15.0 ± 1.0[c]
Bio3	36.0 ± 10.9[a]	28.9 ± 0.9[b]	55.2 ± 1.5[c]
Bio4	8.9 ± 4.3[a]	4.8 ± 0.2[b]	4.8 ± 0.2[c]
Bio5	27.7 ± 8.3[a]	30.1 ± 1.2[b]	34.3 ± 1.6[c]
Bio6	- 8.9 ± 16.3[a]	11.4 ± 1.0[b]	7.2 ± 1.8[c]
Bio7	36.6 ± 12.2[a]	18.7 ± 0.7[b]	27.2 ± 1.1[c]
Bio10	20.8 ± 8.1[a]	26.2 ± 1.1[b]	25.9 ± 1.5[c]
Bio11	- 1.1 ± 15.6[a]	14.5 ± 0.9[b]	17.0 ± 1.4[c]
Bio12	744.8 ± 717.9[a]	3199.7 ± 175.6[b]	3570.4 ± 882.0[c]
Bio15	98.1 ± 24.4[a]	24.6 ± 3.5[b]	111.9 ± 33.0[c]
Bio18	303.3 ± 285.7[a]	813.0 ± 80.6[b]	767.9 ± 86.9[c]
Bio19	43.4 ± 143.8[a]	678.6 ± 97.8[b]	1042.4 ± 449.3[c]

Figures

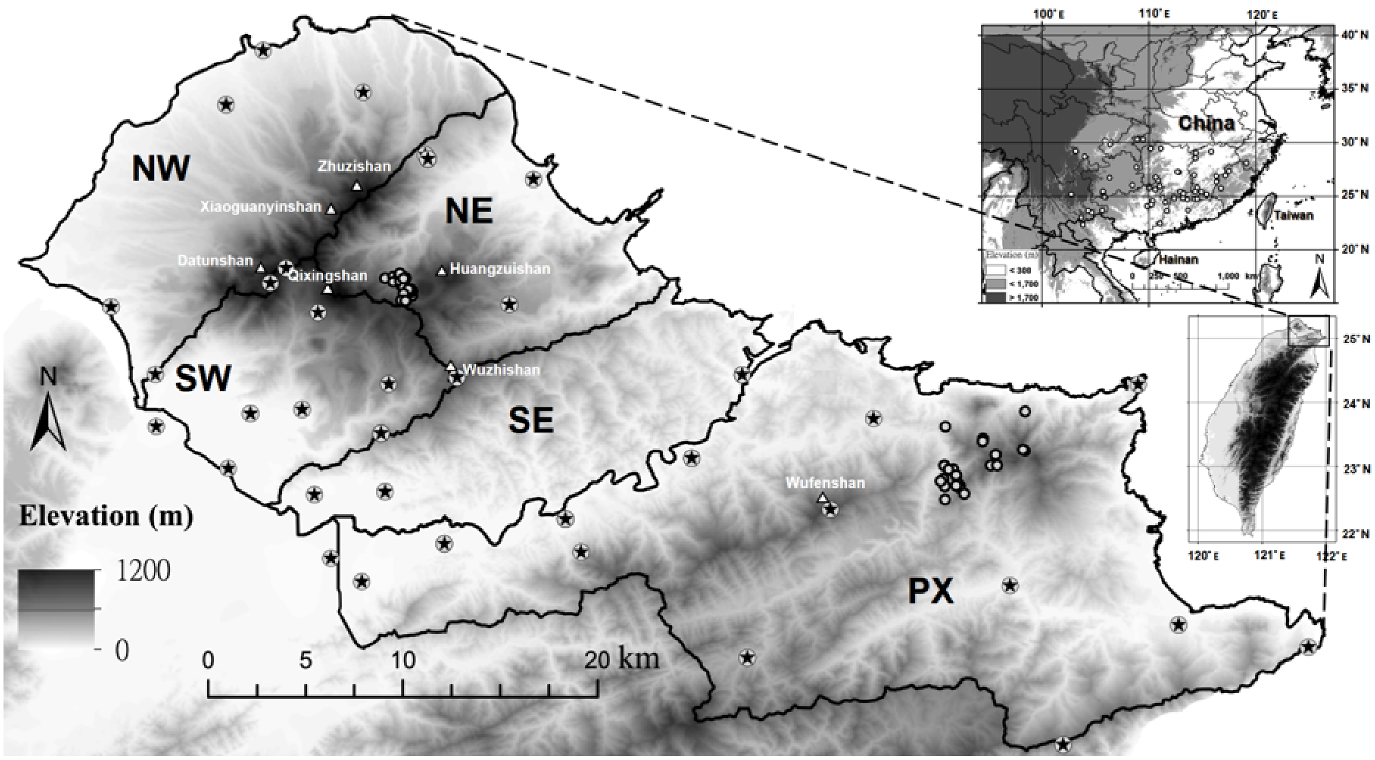

Figure 1

Georeferenced occurrences of Bretschneidera sinensis Hemsl. represented by grey circle with dark outline in the continental East Asia (upper right map) and northern Taiwan (central map). In the map of continental East Asia, the altitudinal range between 300 and 1,700 m was represented by grey color, because it is the altitudinal distribution range of the B. sinensis. In northern Taiwan (central map), five watersheds were delineated for calculating empirical elevation lapse rates of climate data. They are northeast (NE), northwest (NW), southwest (SW) and southeast (SE) slopes of Yamingshan area and Pingxi area (PX). The locations of 30 meteorological stations adopted in this study were represented by dark stars within grey circles. Mountain tops were represented by white triangles.

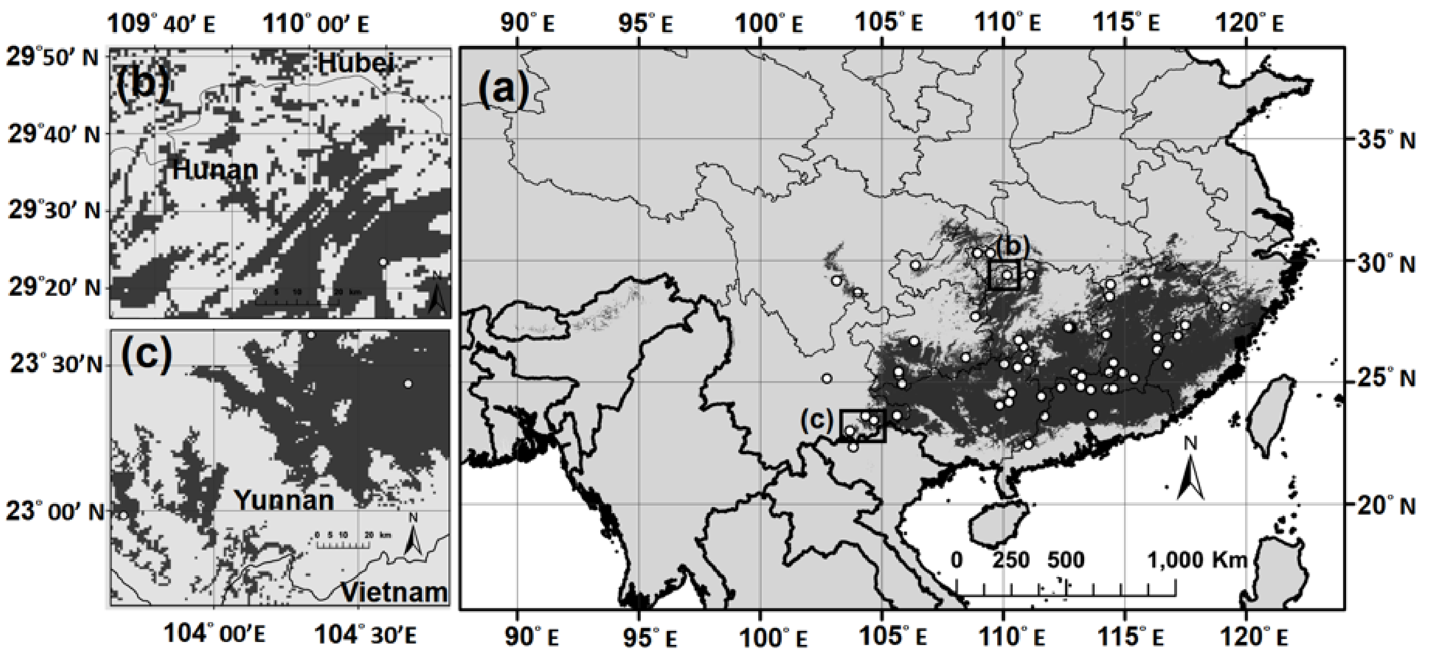

Figure 2

Georeferenced occurrences (open circles) and contemporary potential distribution range of Bretschneidera sinensis (dark grey areas in the maps) in continental East Asia projected by Random Forest based on the WorldClim dataset (a). Distribution pattern of the plant species presents gridded squares when the map was zoom in to show the edge of potential distribution range in larger scale maps (b and c). The two left maps present distribution range at two local areas, Zhangjiajie, Hunan (b) and Wenshan Prefecture, Yunnan (c).

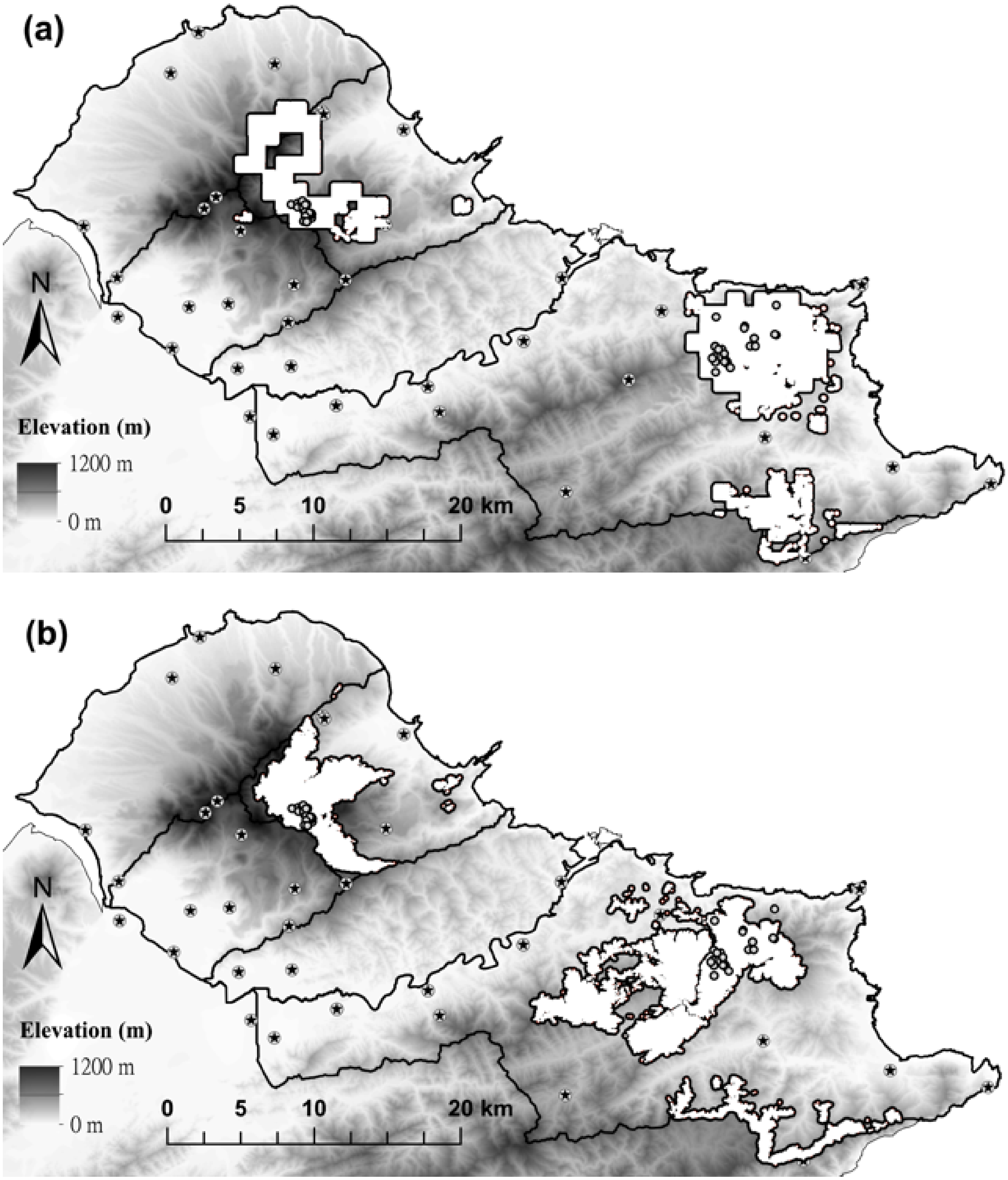

Figure 3

Contemporary potential distribution range of Bretschneidera sinensis in northern Taiwan projected by Random Forest algorithm based on the WorldClim (a) and local climate datasets (b). The symbols of meteorological stations and georeferenced occurrences of B. sinensis are the same as in the Fig. 1. White areas bordered by dark lines are the potential distribution range of the B. sinensis.

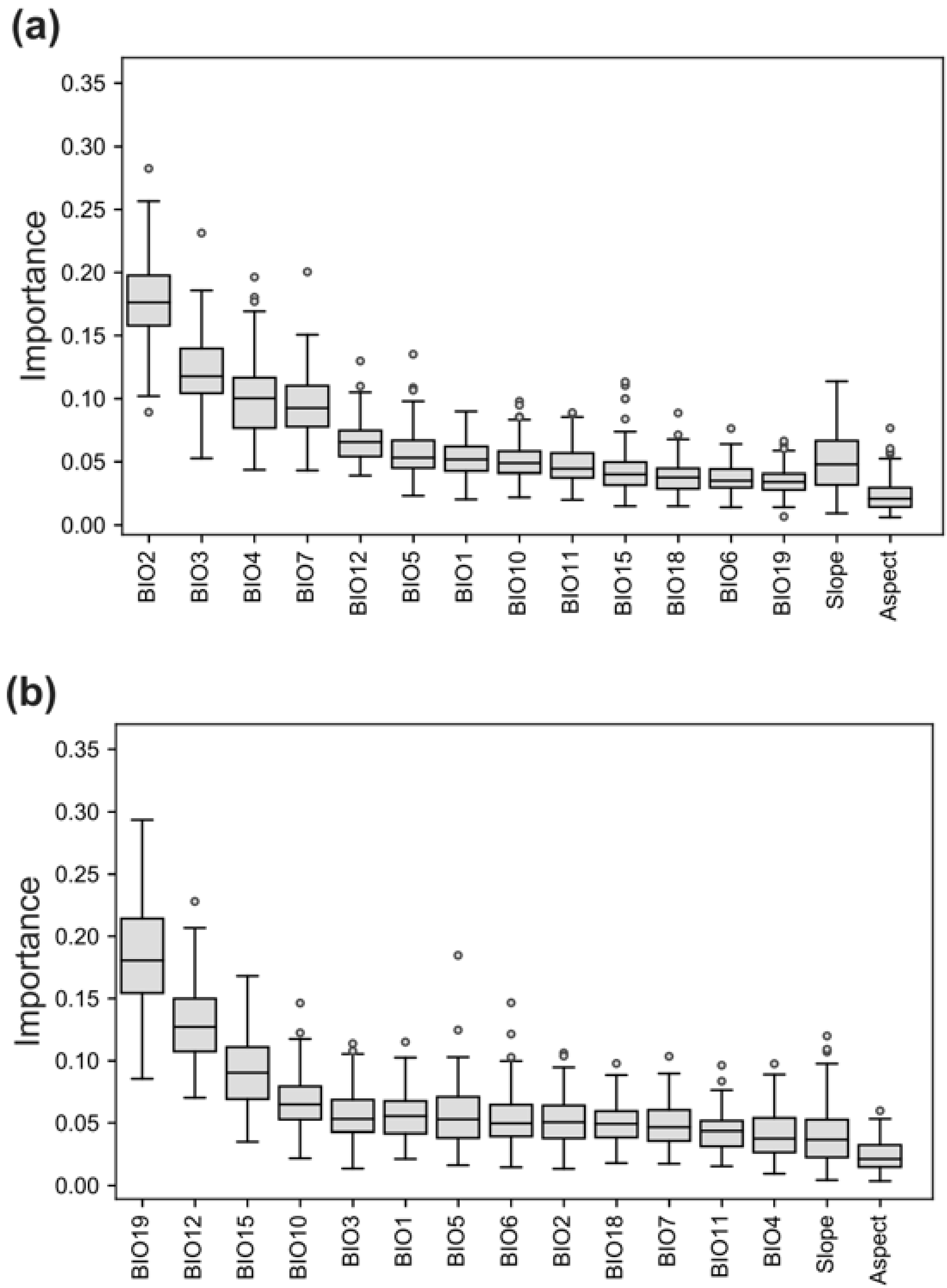

Figure 4

Relative importance of bioclimatic predictors in explaining distribution range of Bretschneidera sinensis in northern Taiwan. Bioclimatic predictors were independently derived from the WorldClim (a) and local climate datasets (b).

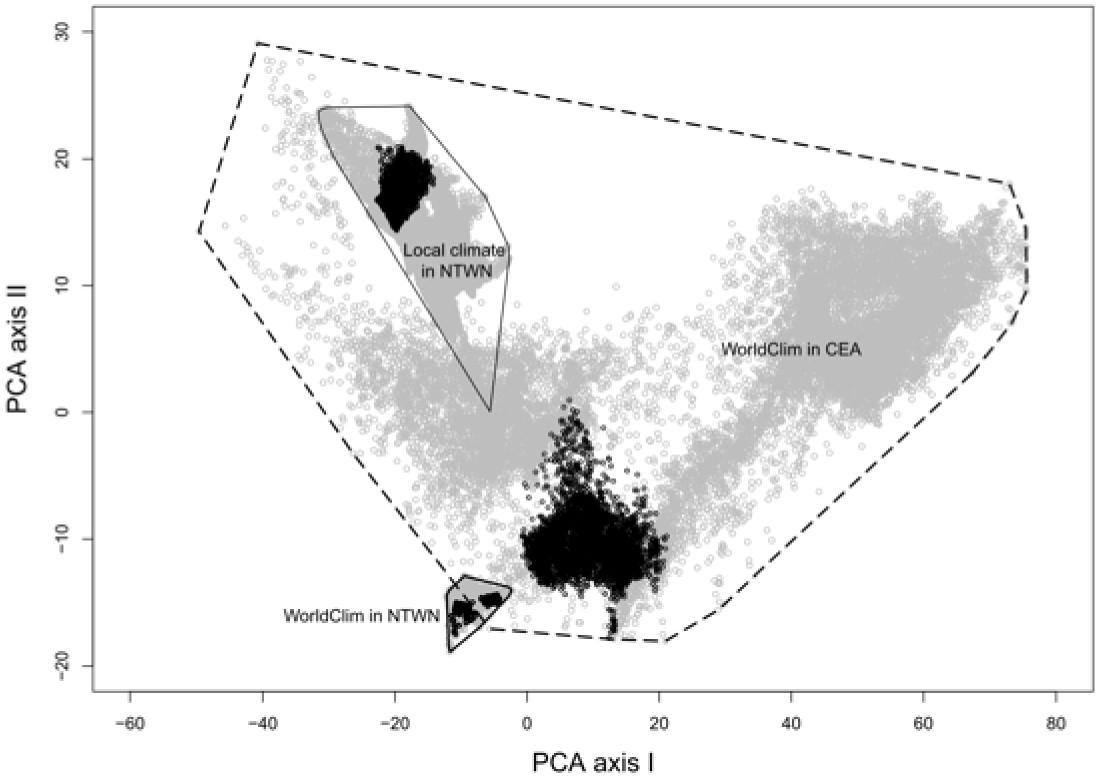

Figure 5

Climate spaces quantified by principle components analysis (PCA). The grey circles enclosed by dashed line, thick solid line, and thin solid line were climate spaces of continental East Asia (CEA) and northern Taiwan (NTWN) based on the WroldClim dataset and NTWN based on local climate datasets, respectively. The three groups of dark circles were climate spaces of B. sinensis based on the WorldClim and local climate datasets in CEA and NTWN.

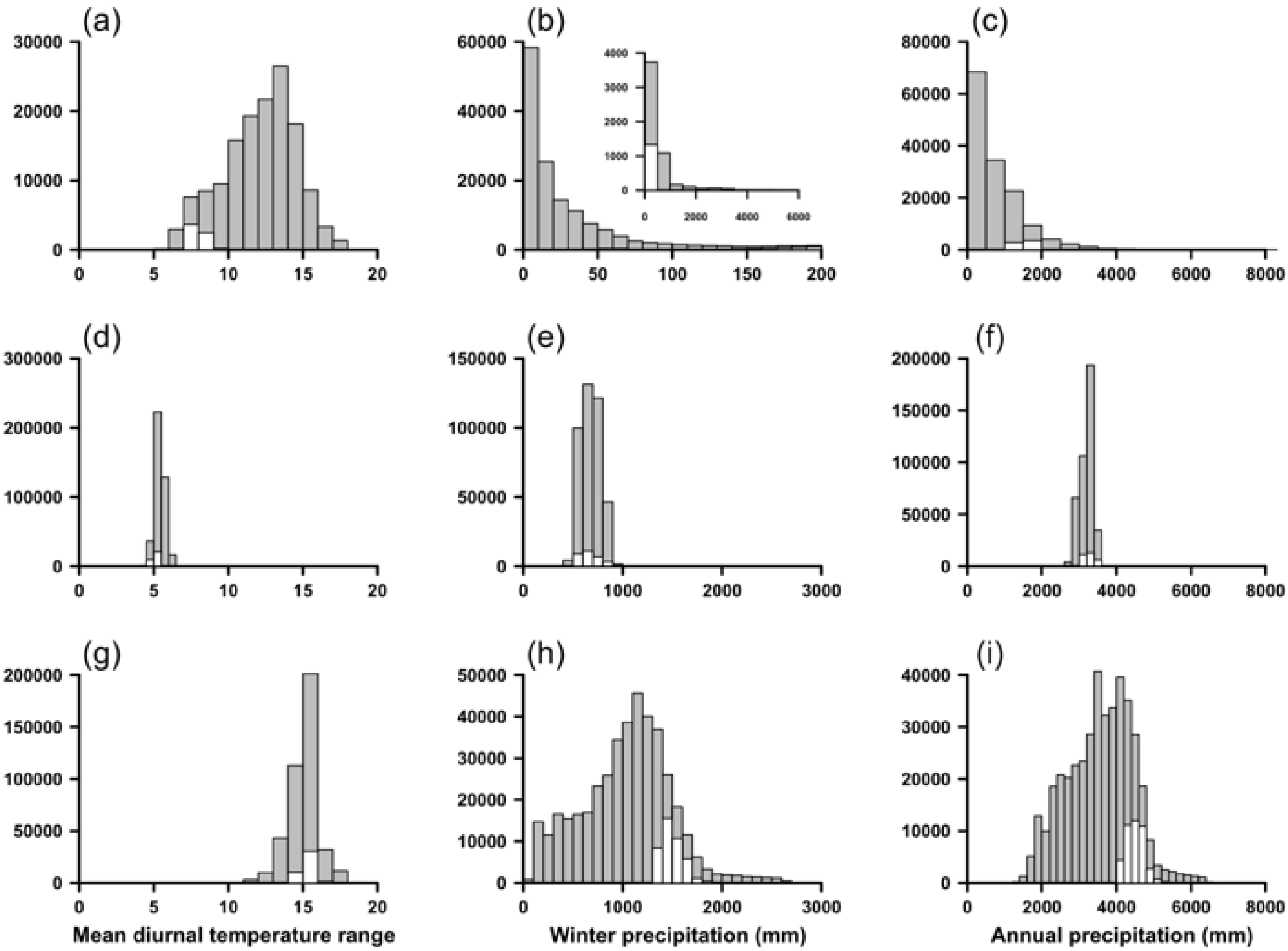

Figure 6

Histogram of three bioclimatic predictors in continental East Asia (CEA) based on the WorldClim (a, b, and c) and northern Taiwan (NTWN) based on the WorldClim (d, e, and f) and NTWN based on local climate datasets (g, h, and i). The three bioclimatic predictors are Bio2, Bio19 and Bio12 from left to right columns. The grey histograms represent the relative frequencies of background cells. The blank histograms within the grey histograms are the cells with presences of Bretschneidera sinensis projected by Random Forest. Winter precipitation based on the WorldClim dataset presented wide range in CEA (b) and gridded cells with winter precipitation higher than 200 mm were showed in the subplot of (b). The x and y scales and intervals were different among histograms.

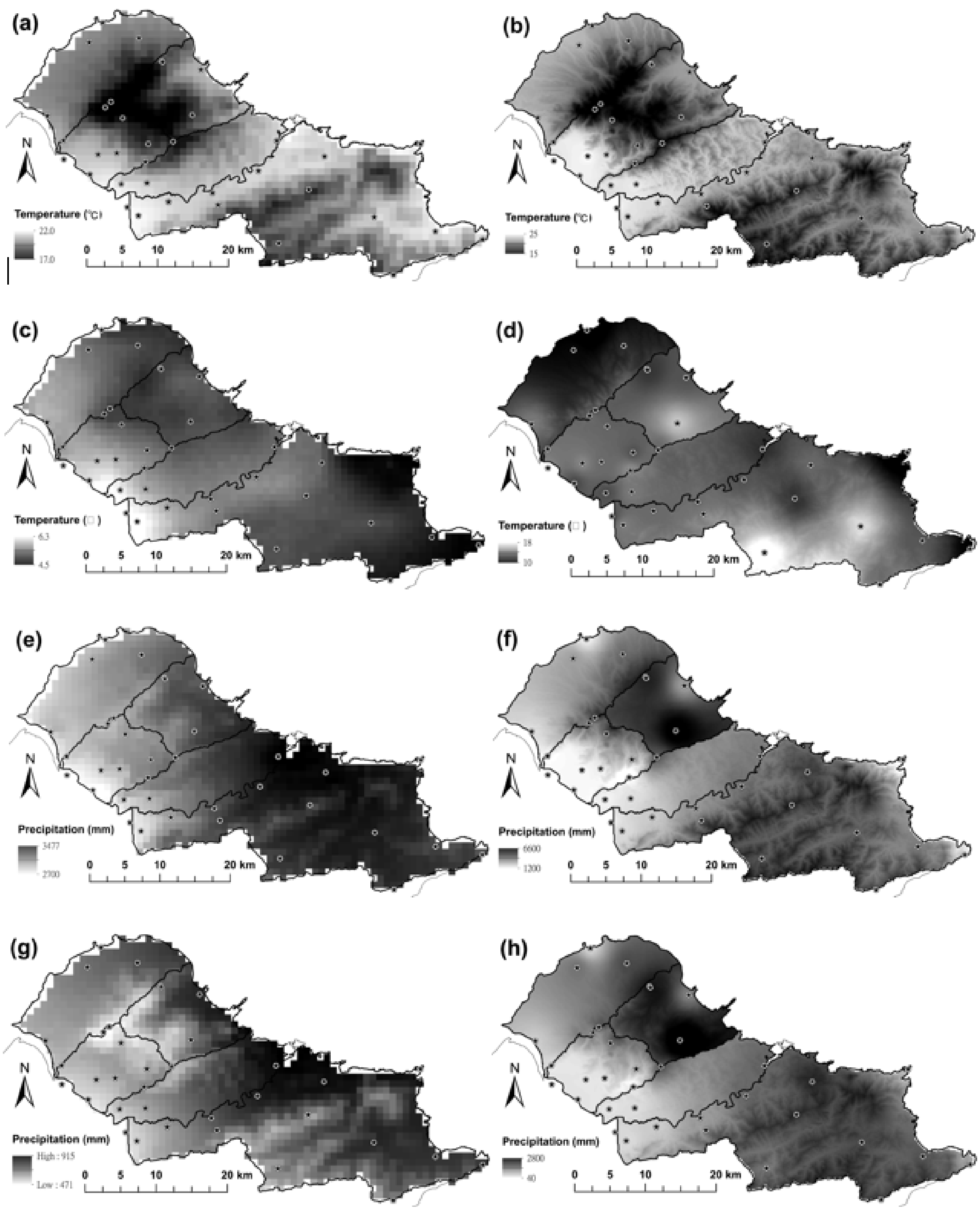

Figure 7

Characteristics of bioclimatic predictors from the WorldClim (left column) and local climate datasets (right column) in northern Taiwan. Bioclimatic predictors from the WorldClim presented significant gridded squares that were not available to reveal climate heterogeneities induced by elevation and topography. Concentric circles centered at the meteorological stations indicated extraordinary high values of mean climate data recorded by the stations that presented very steep gradient of climate variables in the study area and should not be recognized as artefact data. The first row is annual temperature (Bio1) from the WorldClim (a) and local climate dataset (b). The second row (c and d) is diurnal temperature range (Bio2). The third row (e and f) is annual precipitation (Bio12). The fourth row (g and h) is winter precipitation (Bio19). The dark stars within grey circles are meteorological stations adopted in this study.