SCIENCE CHINA
Earth Sciences

•RESEARCH PAPER•

November 2018 Vol.61 No.11: 1669–1684
https://doi.org/10.1007/s11430-017-9241-2

Modeling the distribution of *Populus euphratica* in the Heihe River Basin, an inland river basin in an arid region of China

Yanlong GUO[1,2], Xin LI[1,2,3*], Zefang ZHAO[4] & Haiyan WEI[4]

[1] *Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China;*
[2] *University of Chinese Academy of Sciences, Beijing 100049, China;*
[3] *CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing 100101, China;*
[4] *School of Geography and Tourism, Shaanxi Normal University, Xi'an 710119, China*

Received December 15, 2017; revised April 15, 2018; accepted July 3, 2018; published online September 12, 2018

Abstract *Populus euphratica* is a dominant tree species in riparian Tugai forests and forms a natural barrier that maintains the stability of local oases in arid inland river basins. Despite being critical information for local environmental protection and recovery, establishing the specific spatial distribution of *P. euphratica* has rarely been attempted via precise and reliable species distribution models in such areas. In this research, the potential geographic distribution of *P. euphratica* in the Heihe River Basin was simulated with MaxEnt software based on species occurrence data and 29 environmental variables. The result showed that in the Heihe River Basin, 820 km^2 of land primarily distributed along the banks of the lower reaches of the river is a suitable habitat for *P. euphratica*. We built other MaxEnt models based on different environmental variables and another eight models employing different mathematical algorithms based on the same 29 environmental variables to demonstrate the superiority of this method. MaxEnt based on 29 environmental variables performed the best among these models, as it precisely described the essential characteristics of the distribution of *P. euphratica* forest land. This study verified that MaxEnt can serve as an effective tool for species distribution in extremely arid regions with sufficient and reliable environmental variables. The results suggest that there may be a larger area of *P. euphratica* forest distribution in the study area and that ecological conservation and management of *P. euphratica* should prioritize suitable habitat. This research provides valuable insights for the conservation and management of degraded *P. euphratica* riparian forests.

Keywords *Populus euphratica*, MaxEnt, Species distribution models, Model comparison, Inland river basin

Citation: Guo Y, Li X, Zhao Z, Wei H. 2018. Modeling the distribution of *Populus euphratica* in the Heihe River Basin, an inland river basin in an arid region of China. Science China Earth Sciences, 61: 1669–1684, https://doi.org/10.1007/s11430-017-9241-2

1. Introduction

Populus euphratica, a broad-leaved deciduous tree species that is particularly hardy and especially adapted to continental arid climates, usually forms a dominant component of desert riparian ecosystems in arid regions of continental river basins (Li et al., 2016). The Heihe River Basin, located in the middle of the Hexi Corridor in northwest China, is one of the most intensively exploited inland river basins in China, *P. euphratica* forests are mainly distributed in the lower reaches of the Heihe River in extremely arid desert zones. As a natural barrier that maintains the stability of local oases, *P. euphratica* forests are vital to maintaining the ecological balance in such vulnerable environments (Zhu et al., 2009; Li et al., 2016). However, the area of *P. euphratica* forest has decreased due to the unsustainable utilization of water resources and intensification of human activities over the past five decades (Zhao et al., 2005; Cheng et al., 2014).

* Corresponding author (email: lixin@lzb.ac.cn)

An ecological water diversion project (EWDP) was successfully implemented in the past ten years, and more water has been delivered to the lower reaches of the Heihe River. As a result, *P. euphratica* forests have recovered, and desert shrublands have been partially rehabilitated (Keyimu et al., 2018; Cheng et al., 2014). Importantly, *P. euphratica* forests in the lower reaches of the Heihe River must be protected and restored (Peng et al., 2013; Li et al., 2016), yet without knowledge of the habitat requirements of this species, protection measures that include habitat protection cannot be implemented. Specific information on the habitat requirements of *P. euphratica* is essential for effective habitat protection and may also facilitate more effective plant cultivation.

Species distribution models (SDMs) are numerical tools that have been widely used to predict the potential distribution of species across space and time in terrestrial, freshwater and marine environments (Araújo and Luoto, 2007; Elith and Leathwick, 2009; Adhikari et al., 2012; Guo et al., 2016). There are two types of basic data for SDMs: species occurrence data and environmental variables. For species occurrence data, most SDMs use presence-only and presence-absence data (Barbet-Massin et al., 2012; Fei and Yu, 2016). By contrast, for environmental variables, SDMs implemented in previous studies have generally included several factors, such as climate, soil, topography, vegetation, and human activity (Ficetola et al., 2009; Márquez et al., 2011; Lu et al., 2012; Andriamparany et al., 2015). SDMs reveal the environmental requirements of species by correlating the occurrence of the target species with the physical environment (Elith and Leathwick, 2009; Naimi and Araújo, 2016) and then predicting distributions of the target species across landscapes and extrapolating species distribution across space and time. Therefore, reliable and precise species occurrence data are required for the construction of successful SDMs for the prediction of species distribution. However, identifying the most critical environmental variables that control the distribution of a given plant species is a major challenge in SDM research (Anderson, 2013; Pliscoff et al., 2014).

Arid ecosystems are among the most vulnerable ecological systems, especially in arid inland river basins in which the entire ecosystem depends on water from the river and groundwater recharge (Aishan et al., 2013; Cheng et al., 2014; Hu et al., 2015). Changes in the water volume and the location of the main river can promote dramatic changes in the vegetation distribution in such areas, especially the distribution of tree species in riparian ecosystems (Zhu et al., 2009; Aishan et al., 2015; Li et al., 2016). Thus, for research on *P. euphratica* distributions, real-time and accurate species occurrence data are required. In general, the choice of predictor environmental variables is limited by both the availability of information and the scale of the research area. Furthermore, the environmental characteristics of suitable habitats for the target species should be fully considered (Márquez et al., 2011; Lu et al., 2012; Guo et al., 2016). At large scales, such as national and regional, many SDMs consider only climatic factors, such as the core set of 19 bioclimatic variables (Hijmans et al., 2005; Pliscoff et al., 2014; Guo et al., 2016), although they occasionally include soil variables derived from the Harmonized World Soil Database (HWSD) (Rödder and Lötters, 2009). However, as the collection of detailed ecological information for conservation planning in an endorheic river basin in an arid zone was the objective of this study, more reliable and sufficient environmental variables were required to depict the environmental characteristics.

With technological advancements and an expanded understanding of niche theory, numerous statistical methods and software implementations are widely available for describing patterns and performing predictions, including surface range envelope (SRE; usually called BIOCLIM) (Busby, 1991), flexible discriminant analysis (FDA) (Kuemmerlen et al., 2014), multiple adaptive regression splines (MARS) (Zhang et al., 2012), generalized boosting models (GBMs) (Ridgeway, 1999), classification tree analysis (CTA) (Edwards et al., 2006), generalized linear models (GLMs) (Marmion et al., 2009a), artificial neural networks (ANNs) (Heikkinen et al., 2006), random forests (RFs) (Bradter et al., 2013) and maximum entropy (MaxEnt) (Phillips et al., 2006). Among these models, MaxEnt (MaxEnt version 3.3.3, http://biodiversityinformatics.amnh.org/open_source/maxent/) is currently the most popular SDM because it is free, has a user-friendly operational interface, and presents stable and reliable results, even with incomplete data and small sample sizes (Phillips et al., 2006; Merow et al., 2013; Yi et al., 2016). In addition, MaxEnt simply requires presence-only data for a species, and both continuous and categorical data can be used as input for environmental variables.

Currently, the most popular methods for validating the accuracy of SDMs include the threshold-independent receiver-operating characteristic (ROC) analysis, Cohen's kappa coefficient and the true skill statistic (TSS) (Allouche et al., 2006; Bucklin et al., 2015). Previous studies have theorized that with functionally relevant predictors based on well-designed survey data and a suitable model, SDMs can accurately map the distribution of species (Phillips et al., 2006; Thuiller et al., 2009; Merow et al., 2013; Bucklin et al., 2015; Brown et al., 2016). However, because of the complexity of the natural ecological environment and the interference of species that have similar niches to the target species, suitable habitats identified by SDMs may not actually be suitable for the target species. Because of the restrictions of technical conditions, the authenticity of the results remains a challenge for SDMs (Yang et al., 2013), and detailed patched distributions are difficult to simulate. In this research, to more effectively manage and conserve *P. euphratica* resources in a fragile ecological environment, we modeled the distribution of *P. euphratica* at a fine spatial scale that can provide significant landscape details and show true distribution characteristics. Only in this way should the results be authenticated to further guide conservation planning practices.

There have been several studies on spatial distributions of *P. euphratica* in inland river basins in arid regions, especially in the Tarim River Basin and the Amudarya Delta. In the Tarim River Basin, most studies focus on spatial distribution pattern in typical sample areas, and they are based merely on experience or on simple statistical methods to show the relationship between the spatial distribution pattern and environmental factors (Tayierjiang et al., 2011; Keyimu et al., 2018). For example, based on semi-quantitative and qualitative knowledge, researchers built a fuzzy habitat suitability index (HSI) model to evaluate the ecological situation in the northern Amudarya Delta under changing environmental conditions (Rüger et al., 2005). However, the fuzzy HSI depends primarily on expert knowledge rather than on data from sample sites. Therefore, this method cannot address environmental variables for which there is insufficient expert knowledge regarding species suitability. In this study, we used a MaxEnt model to predict the distributions of *P. euphratica* in the Heihe River Basin. MaxEnt is a general-purpose machine-learning method with a simple and precise mathematical formulation; moreover, it is an objective approach using information provided by data from sample sites (Phillips et al., 2006; Merow et al., 2013; Yi et al., 2016). The first objective of this paper is to model the habitats suitable for *P. euphratica* in the Heihe River Basin, identify the habitat requirements, and select the key environmental variables highly correlated with *P. euphratica* distribution. The second objective of this paper is to compare the results of several MaxEnt models based on different environmental variables and those of several commonly used SDMs based on the same 29 environmental variables to explain the influence of environmental variables and model selection on the model results. This study provides examples for modeling species distribution in arid areas and valuable insights for the protection of *P. euphratica* resources.

2. Methods

2.1 Study area and species data

The Heihe River rises in the Qilian Mountain along the northern edge of the Qinghai-Tibetan Plateau, and it is the second-largest inland river in China (Li et al., 2011; Cheng et al., 2014). Flowing through the Zhangye, Jiuquan and Jinta Basins, the Heihe River finally empties into Ejin Banner in Inner Mongolia. The Heihe River Basin covers an area of approximately 143×10^3 km^2 and is located between 97°E and 102°E and 38.7°N and 42.7°N (Figure 1). The entire basin can be generally divided into three parts: the upper reaches (Qilian Mountain, the source of the river), the middle reaches (the main distributive oasis area), and the lower reaches (around the Ejin Oasis). In the lower reaches of the Heihe River Basin, according to land cover data from 2011 (Wang et al., 2014), the riparian forest and grassland areas accounted for less than ten percent of the entire area, and outside of the Ejin Oasis, the primary landscape was desert steppe and Gobi peneplain. *P. euphratica* is the dominant tree species in the riparian forest, and along with *Tamarix ramosissima* Ledeb., the primary shrub species, it forms the primary vegetation matrix in the Ejin Oasis. Moreover, *P. euphratica* is vital to maintaining the ecological balance and preventing desertification. Regardless, since the 1950s, the intensive exploitation of water resources in the middle reaches has reduced the surface water supply in the lower reaches, causing severe degradation of the forest area in the Ejin Oasis. Although implementation of the EWDP has alleviated the severe deterioration of the ecosystems in the lower reaches of the Heihe River Basin, the *P. euphratica* forest areas have not been fully restored.

In this study, occurrence data for *P. euphratica* were automatically derived from remote-sensing data with high spatial resolution. First, data with high spatial resolution (higher than 2 m) were extracted from Google Earth. Second, the extracted images were automatically mosaicked using the application programming interface provided by Google Earth. Third, the characteristics of *P. euphratica* were analyzed as image objects, a set of rules for extracting *P. euphratica* by employing object-oriented methods was constructed, and then an accurate distribution of *P. euphratica* was obtained. Finally, the accuracy was evaluated by using the confusion matrix method. The researchers selected 680 rectangular regions of interest in high-resolution images obtained from Google Earth. Each such region was 3×3 and occupied a total of 6188 pixels. Then, three experts performed artificial visual interpretation based on the Google Earth images for the test area, and we thus obtained true *P. euphratica* distribution data for the test area. For comparison with the model result, the researchers built a confusion matrix. The results of the evaluation indicated that the accuracy was greater than 87% (Wang et al., 2016).

A suitable number of sampling points can improve the accuracy of the model in this type of research, and we randomly selected 200 sampling points as candidate occurrence data. Based on the resolution of the environmental variables and the size of the true distribution area of *P. euphratica*, we deleted certain sampling points to ensure that the points were distributed evenly and that the distance between two sampling points was always more than 4 km. At such a distance,

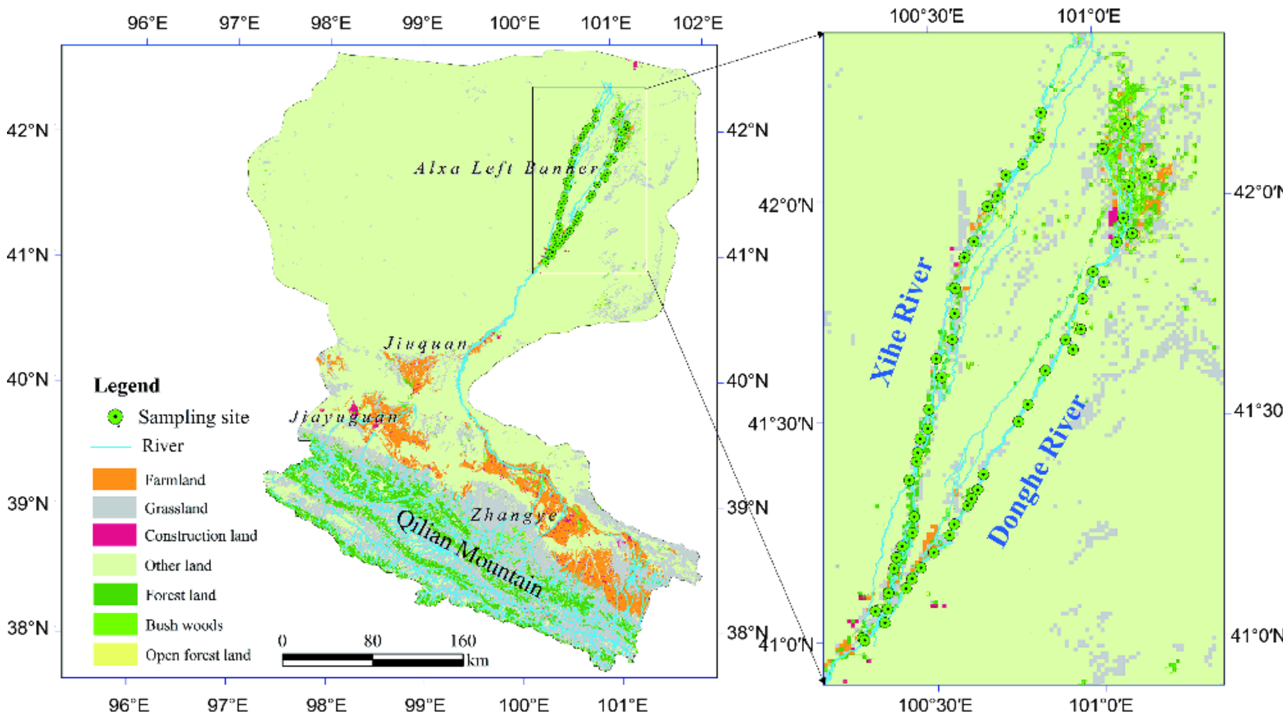


Figure 1 Study area and distribution of sample sites.

the value of environmental data for the two rasters showed a significant difference, and it was conducive to model training. Finally, we obtained 57 locations (occurrence data) of *P. euphratica*, which were used to build the models. The distances between our sampling points and the river channel ranged from 0.1 to 5 km (Figure 1).

2.2 Environmental variables

To determine the geographic distribution of suitable habitat for a target species, a set of environmental characteristics for this species must be defined (Lu et al., 2012; Guo et al., 2016). Hence, we chose four types of environmental factors, including 29 environmental variables, particularly local environmental variables, to simulate the suitable habitat distribution of *P. euphratica* (Table 1).

Topographic variables included elevation, slope and aspect. Elevation variables at a resolution of 1 km×1 km were downloaded from the Data Center for Resources and Environmental Sciences (RESDC) of the Chinese Academy of Sciences (http://www.resdc.cn). The aspect and slope variables were generated from the elevation variables using the spatial analysis function of ArcGIS (ESRI, Redlands, CA, USA). For the climatic variables, the accumulated temperature above 0°C (ATA0), accumulated temperature above 10°C (ATA10), aridity index (AI), and moisture index (MI) were also provided by RESDC at a resolution of 1 km^2. These variables were created by inverse distance weighting for interpolation using meteorological data from 1915 weather stations as independent variables. Corrections were performed using digital elevation model (DEM) data, which resulted in nationwide climatic data for China.

The other four climatic data were the annual precipitation (PRE), mean temperature (TEM), minimum temperature of coldest month (TMIN), and maximum temperature of the warmest month (TMAX), which were derived from Weather Research and Forecasting (WRF) model simulations and remote-sensing data (Pan et al., 2012, 2014, 2015). All of the soil variables were determined by combining soil sample data and other related environmental variables. Based on previous studies, the researchers used boosted regression tree (BRT) and RF models to map different soil properties in the Heihe River Basin (Yang et al., 2014; Yang et al., 2016). Studies have confirmed that this method greatly improves the accuracy and spatial resolution of soil variable data (Liu et al., 2016; Song et al., 2016). Hydrological variables including annual average groundwater table (WD), growing season average groundwater table (GWD), dormant period average groundwater table (DWD), and average evapotranspiration (ET) were derived from an integrated surface water-groundwater model, GSFLOW (Coupled Ground-Water and Surface-Water Flow Model) (Tian et al., 2015). All of the abovementioned data were downloaded from the Environmental and Ecological Science Data Center for West China, National Natural Science Foundation of China (http://westdc.westgis.ac.cn) (Li et al., 2001). In this study, the original data for different types of environmental variables were in different data formats and resolutions. We resampled

Table 1 Evaluation variables for modeling the distribution of suitable habitats for *P. euphratica*

Type	Variable	Range and units	Code
Topographical	Elevation above sea level	From 877 to 5252 m	Elevation
	Slope	From 0° to 33.5°	Slope
	Aspect	From 0° to 359.95°	Aspect
Climatic	Accumulated temperature above 0°C	From 0 to 4070°C	ATA0
	Accumulated temperature above 10°C	From 0 to 3760°C	ATA10
	Aridity index	From 0 to 31.2	AI
	Moisture index	From −58.2 to 21.2	MI
	Annual precipitation	From 13 to 506 mm	PRE
	Mean temperature	From −2.0 to 8.9°C	TEM
	Minimum temperature of coldest month	From −35.3 to −2.0 °C	TMIN
	Maximum temperature of warmest month	From 27.4 to 28.9 °C	TMAX
Soil	Soil type	57 classes	ST
	Wilting moisture	From 16% to 25%	WP
	Field capacity	From 20% to 44%	FC
	Saturated soil moisture	From 43% to 66%	SSM
	Saturated hydraulic conductivity	From 5.6 to 121.4 mm h^{-1}	SHC
	Soil $CaCO_3$ density	From 0 to 321.55 kg m^{-2}	$Caco_3$
	Soil thickness	From 0 to 180 cm	STH
	Soil pH (H_2O)	From 6.7 to 8.86	PH
	Surface soil texture	3 classes	SST
	Soil organic carbon at 0–5 cm depth	From 1.17 to 95.71 g kg^{-1}	SOC_layer1
	Soil organic carbon at 5–15 cm depth	From 1.15 to 83.9 g kg^{-1}	SOC_layer2
	Soil organic carbon at 15–30 cm depth	From 0.7 to 57.4 g kg^{-1}	SOC_layer3
	Soil organic carbon at 30–60 cm depth	From 0.6 to 30.83 g kg^{-1}	SOC_layer4
	Soil organic carbon at 60–100 cm depth	From 0.57 to 5.45 g kg^{-1}	SOC_layer5
Hydrological	Annual average groundwater table	From 0 to 312 m	WD
	Growing season average groundwater table	From 0 to 312 m	GWD
	Dormant period average groundwater table	From 0 to 312 m	DWD
	Average evapotranspiration	From 375 to 1240 mm	ET

all of the variables at a 1 km spatial resolution and converted the data for all environmental variables to ASC files in ArcGIS 9.3.

Land cover data for the Heihe River Basin were used to determine the validity of the model results. These data are available at a resolution of 30 m from the Environmental and Ecological Science Data Center for West China (Ran et al., 2012; Hu et al., 2015). In the lower reaches of the Heihe River Basin, *P. euphratica* is a dominant component of the desert riparian ecosystem; therefore, the distribution of forest land obtained from these data should be consistent with the distribution of *P. euphratica* forest. Accordingly, we extracted the forest land in the lower reaches of Heihe River Basin using ArcGIS 9.3 as actual distributions of *P. euphratica.*

To compare the model results, we used the 19 core bioclimatic variables and soil data (attached list 1) to evaluate different MaxEnt models calibrated with different sets of variables. The bioclimatic variables were downloaded from the WorldClim database, and the resolution was approximately 1 km^2. The soil data were obtained from the HWSD (Ray et al., 2016), http://www.fao.org/soils-portal/), and the resolution was approximately 1 km^2.

2.3 Model evaluation

MaxEnt, one of the most successful SDMs, is based on the maximum entropy principle and Bayesian estimates. The model simulates species distributions based on species occurrence data and environmental variables, and the species distribution results follow the principle that each expected variable value should represent the empirical average (Phillips et al., 2006; Merow et al., 2013) to realize the probability distribution of maximum entropy. In this study, the MaxEnt software platform (MaxEnt version 3.3.3, http://www.cs.princeton.edu/~schapire/maxent/) (Phillips et al., 2006;

Merow et al., 2013) was used to project the distribution of *P. euphratica*. For the modeling process, 70% of the occurrence data were chosen randomly as the training set. We used a threshold-independent ROC analysis to evaluate the model performance. In practice, the area under the ROC curve (AUC) values were used to achieve this function, and those values ranged from 0.5 (random) to 1.0 (perfect discrimination). When the AUC value was above 0.8, the model result was considered satisfactory (Marmion et al., 2009b; Escalante et al., 2013). To determine the relative importance of a single explanatory variable, we used the jackknife test during model development. To further elucidate the environmental characteristics of suitable habitats for *P. euphratica* in the study area, we built different MaxEnt models using single environmental variables and then obtained response curves (Merow et al., 2013; Yuan et al., 2015) that showed how the logistic prediction changed with variations in each environmental variable.

2.4 Model comparison

Successful SDMs require reliable and sufficient environmental variables capable of accurately describing the target environmental characteristics. To compare the effects of using other environmental variables, another two test MaxEnt models were built. Both of these test models included the same 57 locations and the same model settings as the main model; however, one of these two test models was based on the 19 core bioclimatic variables (Table S1, http://earth.scichina.com) and was defined as T1, whereas the other test model used seven bioclimatic variables, three topographic variables and seven soil variables (Table S2) and was defined as T2. For T2 (Figure 2), the seven bioclimatic variables were derived from the results of the correlation test among the 19 bioclimatic variables, and the seven soil variables were derived from the HWSD.

In this research, we also compared the MaxEnt results with those of eight other types of SDMs: SRE, FDA, MARS, GBM, CTA, GLM, ANN and RF. BIOMOD2 (Biodiversity Modeling) (Thuiller et al., 2009) (https://cran.r-project.org/web/packages/biomod2) was used to implement these eight models with the four types of environmental factors, which included 29 environmental variables (Figure 2). During modeling, BIOMOD2 required both presence-only records and presence-absence records, and we randomly generated 112 presence-absence records to ensure that all of the presence-absence records were distributed evenly throughout the Heihe River Basin. Then, we divided the databases that included 57 presence-only records and 112 presence-absence records data into two subsets, the same as the MaxEnt setting; 70% of the presence-only and presence-absence records were randomly chosen as the training set, and 30% of the databases were used for evaluation. The accuracy of the models was evaluated using three widely used evaluation indexes: Cohen's kappa coefficient, the TSS and the AUC.

3. Results

3.1 Distribution of suitable habitats for *P. euphratica* in the Heihe River Basin

According to the model results, the suitable habitat index (SH_i) for *P. euphratica* ranged from 0 to 0.93. For further analysis, the entire study area was classified into three classes of SH_i values: unsuitable habitats presented SH_i values below 0.3, marginally suitable habitats presented SH_i values between 0.3 and 0.5, and suitable habitats presented SH_i values above 0.5. A distribution map of suitable habitats for *P. euphratica* in the Heihe River Basin was drawn (Figure 3), and the habitat suitability grades were calculated. The results showed that the suitable habitat area for *P. euphratica* in the Heihe River Basin was small, only 820 km^2. These areas were mainly located in Ejin Banner in the lower reaches of the Heihe River Basin and distributed within seven kilometers of the banks of the river. The marginally suitable habitat area was also small, 680 km^2, and these areas surrounded the suitable habitats.

In this study, occurrence data for 57 locations and 29 environmental variables were used as inputs, and MaxEnt provided satisfactory results. The average AUC values for 10 replications of the training data set and test data set were 0.994 and 0.989, respectively. To further test the accuracy of the model, we extracted forest information from land cover data from the Heihe River Basin (Figure 4). A comparison of the model results and the forest land data suggested that the model results were consistent with the actual distribution of *P. euphratica* forest land and with the spatial distribution pattern.

3.2 Response of dominant variables to suitability

MaxEnt can calculate the relative contributions of environmental variables to species distribution (Phillips et al., 2006). In this study, the MaxEnt results showed that the saturated soil moisture (SSM), soil organic carbon in the 0–5 cm layer (SOC_layer1), underground water table in the growing season (GWD), soil type (ST) and AI were key variables determining the distribution of *P. euphratica*, with contribution rates of 20.9%, 16.7%, 13.1%, 10.8% and 10.5%, respectively. The slope, mean temperature (TEM), field capacity (FC), and average evapotranspiration (ET) were also important variables and presented relatively high contributions to the distribution of *P. euphratica.*

Based on response curves illustrating how the logistic predictions changed as each environmental variable changed (Figure 5), we calculated the suitable ranges for environ-

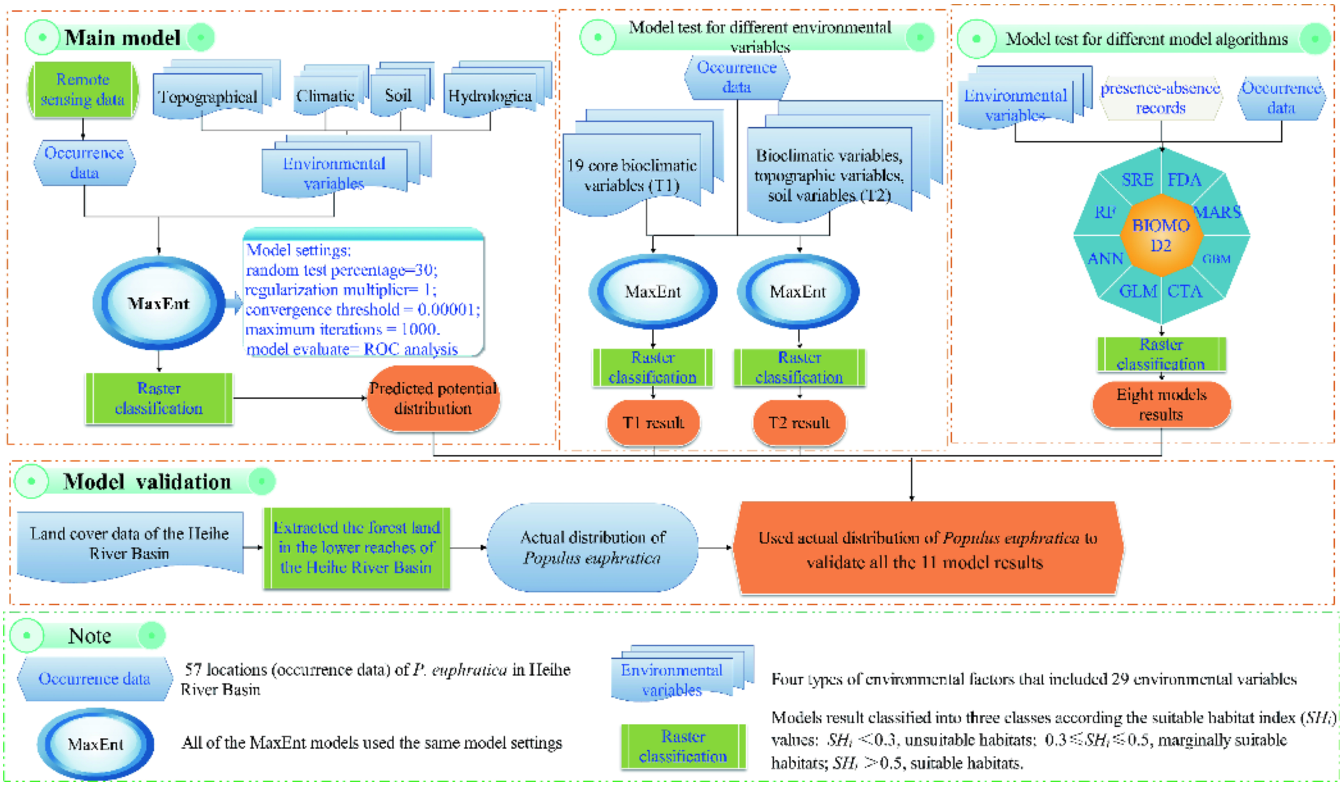


Figure 2 Schematic representation of the methodological approaches.

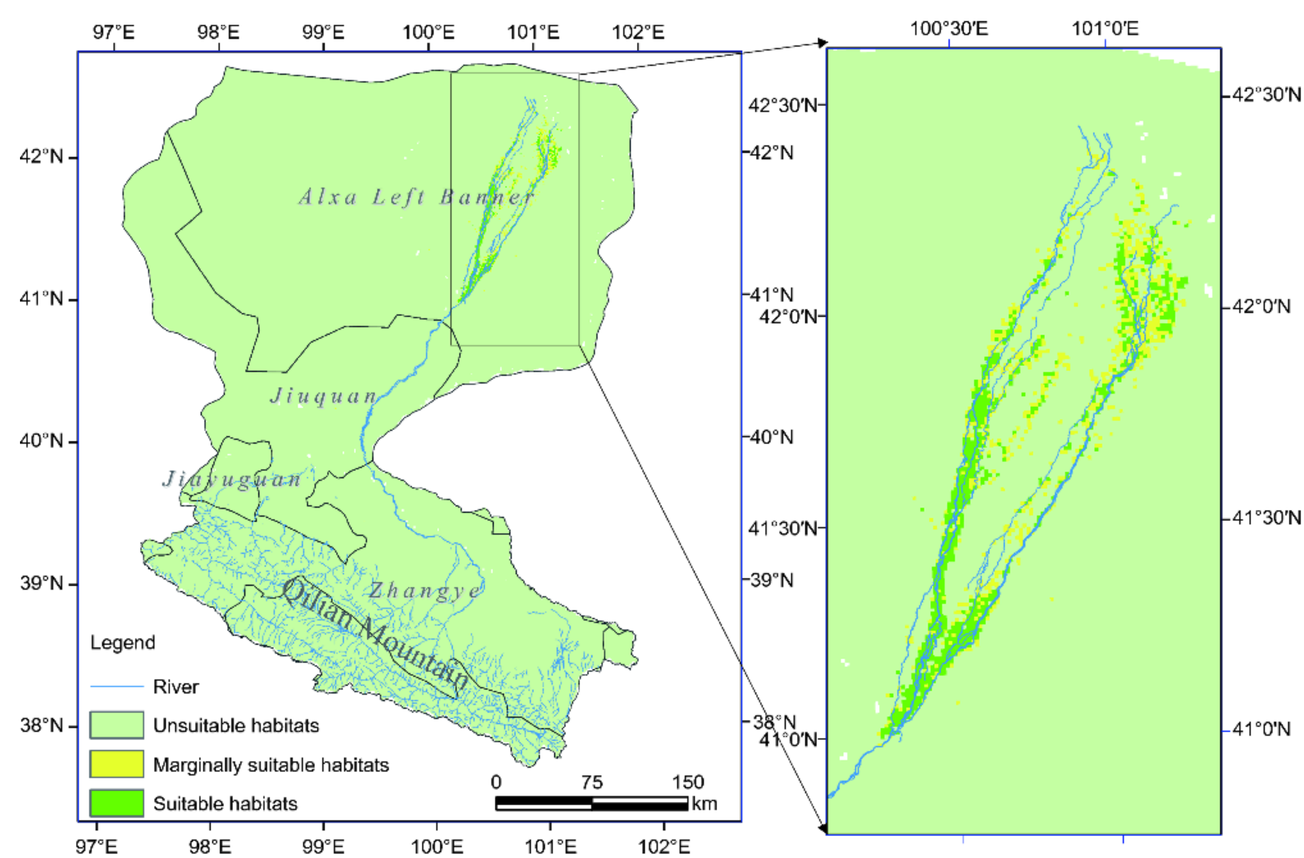


Figure 3 Predicted potential distribution of *P. euphratica* in the Heihe River Basin.

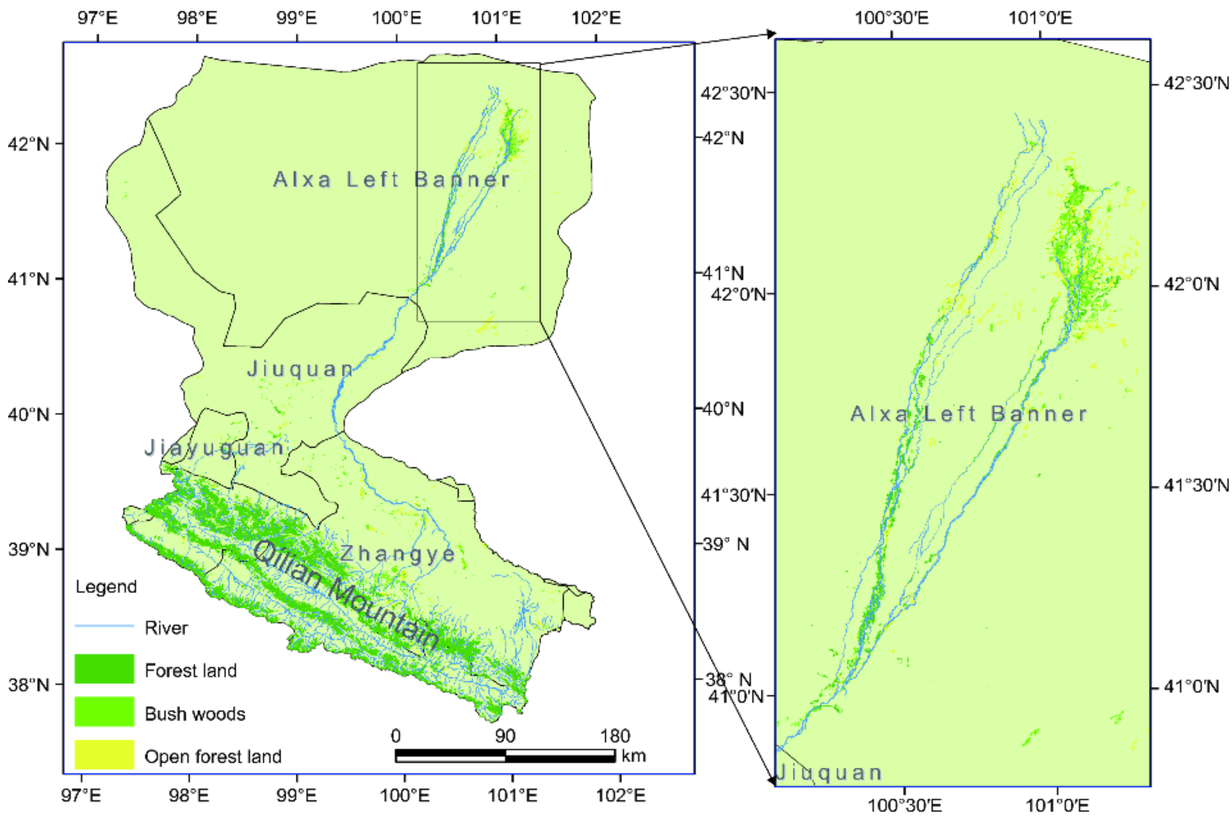


Figure 4 Forest land distribution in the Heihe River Basin.

mental variables within the distribution area of *P. euphratica* (logistic probability of presence >0.3). The observed ranges were 45–47% for SSM; 9–21 g kg^{-1} for SOC_layer1; 0–4 m for GWD; 17–22 for AI; less than 0.3° for slope; 7–8.5°C for TEM; 24–27% for FC; and 1120–1180 mm for ET. In addition, there were four suitable STs for *P. euphratica*.

3.3 Effects of the reliability and resolution of environmental variables for determining the suitable habitat distribution for *P. euphratica* in the Heihe River Basin

The test model results showed that the AUC values for the training and test data sets were 0.997 and 0.997 in T1 (MaxEnt model based on 19 core bioclimatic variables) and 0.994 and 0.949 in T2 (MaxEnt model based on seven bioclimatic variables, three topographic variables and seven soil variables), respectively. Using the same classification standards, we divided the habitat suitability results for *P. euphratica* into three grades (Figure 6). A comparison of the MaxEnt results and the actual distribution of *P. euphratica* forest indicated that the T1 results significantly overvalued the area of suitable habitats for *P. euphratica*, could not provide many landscape details, and did not show the distribution characteristics for suitable habitats for *P. euphratica* along the banks of the river.

The results of the T2 model were better than those of T1. T2 produced a spatial distribution pattern of suitable habitats for *P. euphratica* that was essentially consistent with the actual distribution of *P. euphratica* forest, but this model lost some significant landscape details. For example, only small patches of forest existed along the Donghe River except near Ejin Oasis, but according to the T2 results, suitable habitats for *P. euphratica* included a strip or belt-like shape that was distributed along the Donghe River. The calculations showed that the area of suitable habitats for *P. euphratica* in the T2 results was 1440 km^2, which greatly exceeded the area predicted by the MaxEnt model based on the 29 environmental variables. All of these results indicate that the environmental variables chosen for the T2 model were insufficient for an accurate description of the actual environmental characteristics.

In conclusion, traditional bioclimatic variables and HWSD (Ray et al., 2016) soil variables are not suitable for SDMs of an inland river basin in an arid region because of their low accuracy at the regional scale and their inability to accurately describe the environmental characteristics of these areas.

3.4 Other SDMs for the distribution of suitable habitats for *P. euphratica* in the Heihe River Basin

The BIOMOD2 (Biodiversity Modeling) (Thuiller et al., 2009) results (Table 2) showed that nearly all of the models,

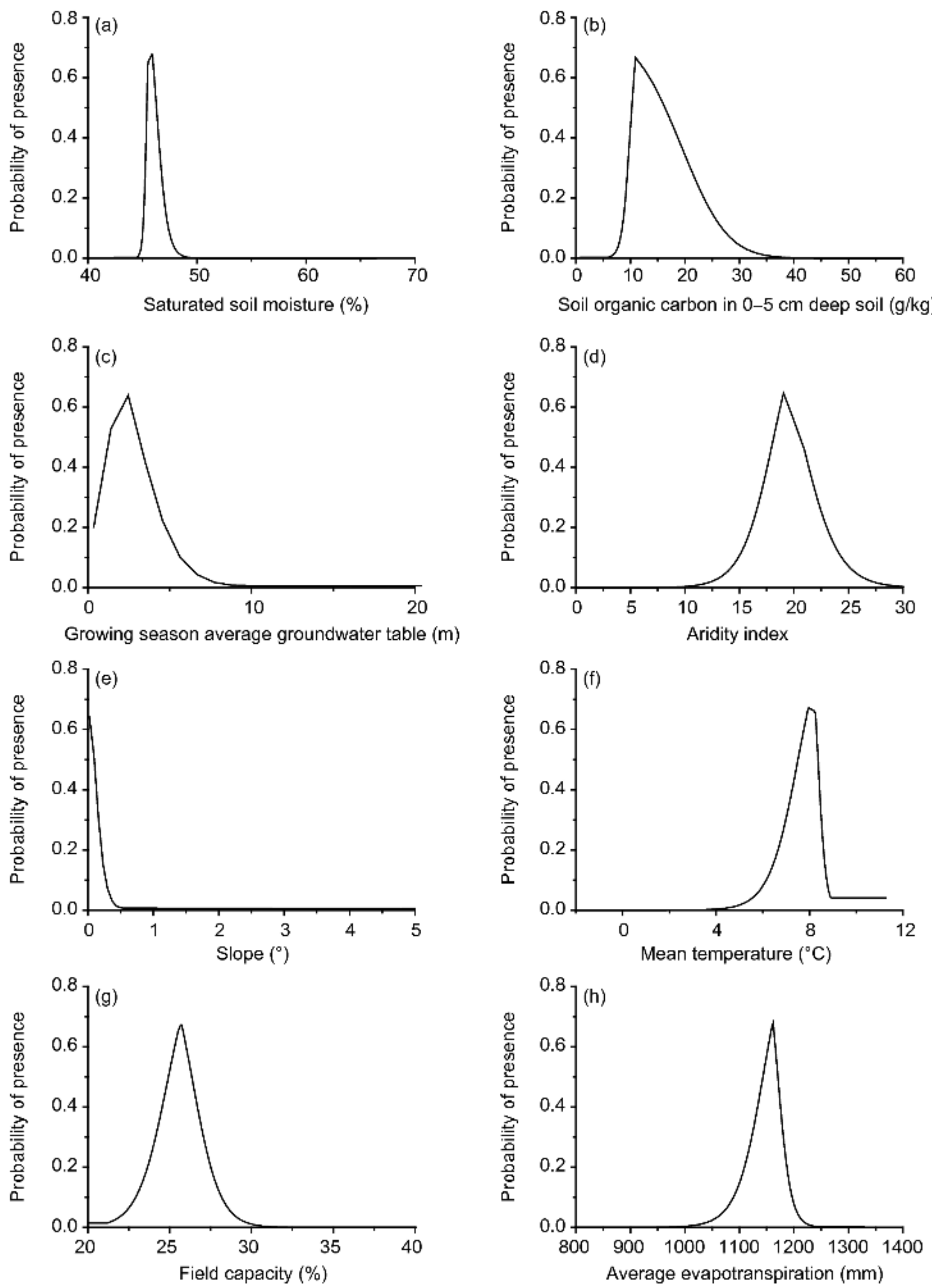


Figure 5 Probability relationships between dominant climate factors and the geographic distribution of *P. euphratica.*

Table 2 Evaluation of accuracy in modeling the distribution of suitable habitats for *P. euphratica*

	GLM	GBM	CTA	ANN	SRE	FDA	MARS	RF	MaxEnt
KAPPA	0.866	0.934	0.718	0.755	0.530	0.931	0.934	0.934	–
TSS	0.866	0.957	0.684	0.826	0.455	0.913	0.957	0.957	–
ROC	0.933	0.996	0.773	0.881	0.727	0.945	0.976	0.996	0.994

with the exception of SRE, passed the evaluation. The statistical accuracy results for these eight models showed that the RF and GBM models were the best. Specifically, the Cohen's kappa coefficient and TSS and AUC values for these two models were significantly higher than those for the other models, and the AUC values for these models were higher than that of the MaxEnt model. The next most accurate models were the MARS, FDA and GLM models. The

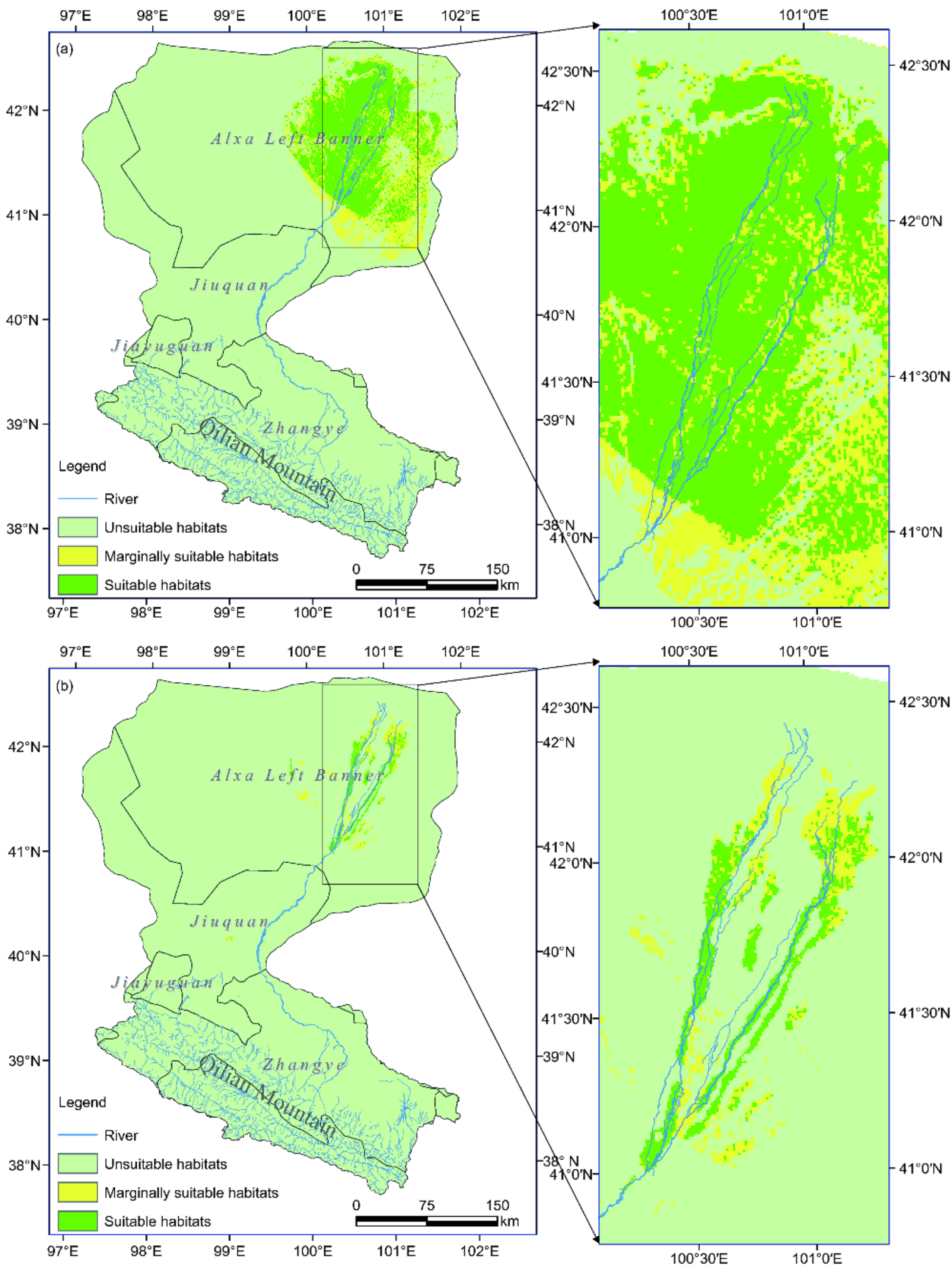


Figure 6 T1 and T2 models for the distribution of suitable habitats for *P. euphratica* in the Heihe River Basin. (a) T1: model based on the core set of 19 bioclimatic variables; (b) T2: model based on 7 bioclimatic variables, 3 topographic variables and 7 soil variables.

AUC values of these three models were greater than 0.9, and the AUC values of the remaining three models (CTA, ANN and SRE) were greater than 0.7. These statistical results indicate that the eight models were successful.

Because these eight models used different mathematical algorithms, the weighting of the environmental variables was also different. Nearly all of the models identified SSM, SOC_layer1, slope, and AI as key factors determining the distribution of *P. euphratica* (Table S3). These results are consistent with the MaxEnt results. To further assess the differences between the results of the different models, the same classification standards were used, and the habitat suitability results for *P. euphratica* from the eight models were classified into three grades, as shown in Figure 7. The figure shows that the spatial distribution patterns of suitable habitats for *P. euphratica* varied widely. A comparison of the MaxEnt results with the actual distribution of *P. euphratica* forest indicates that seven models significantly overestimated the suitable habitat area of *P. euphratica*, whereas the SRE model significantly underestimated the area.

4. Discussion

4.1 Rationality of the model

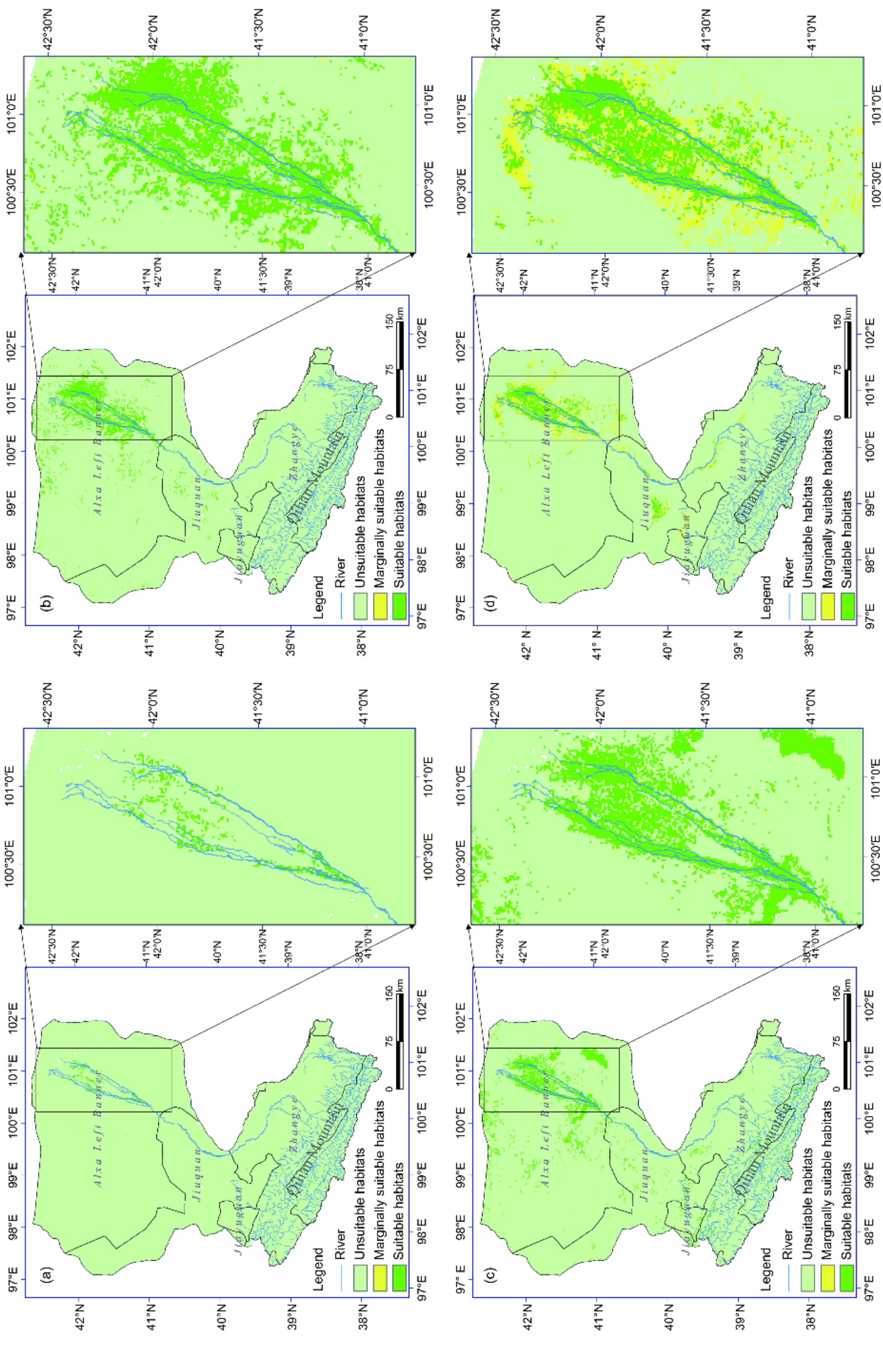
In contrast to previous work, this study selected a particular species adapted to the environment of a continental drainage basin in an arid area. The unique habitat demands of this species increased the difficulty of modeling the species distribution, although they also provided an advantage because the simple vegetation types in the lower reaches of the Heihe River Basin allowed us to accurately obtain the actual distribution of this species using remote-sensing technology; therefore, the realized niche of this species could be accurately described. Moreover, by comparing the actual species distribution with the model results, we were able to intuitively determine the optimal model instead of relying on statistical evaluations.

Successfully predicting the distribution of a species requires the following: accurate and appropriate amounts of species distribution records; reliable and sufficient environmental variables for the study area size and the environmental characteristics of the target species distribution area; and a stable and reliable mathematical algorithm (Elith and Leathwick, 2009; Lu et al., 2012).

For SDMs, species distribution data are usually obtained from herbarium records or published scientific research; however, these data sources are problematic. Some of the species occurrence data from herbariums were collected twenty or even thirty years ago; therefore, the reliability of such data is difficult to guarantee. Moreover, most herbarium and research data lack precise geographic information; therefore, the precision of the model is difficult to guarantee. Field surveys represent another method of obtaining species distribution data; however, this approach presents high labor, material, and financial costs, especially in hostile environments. High-resolution remote-sensing data represent a new method for obtaining species distribution data, and such technology can allow large areas to be observed in a timely and economical manner. Most importantly, because of the relatively simple surface coverage conditions in the study area, we could identify the occurrence of *P. euphratica* at an intuitive level. Hence, in this study, we were able to ensure the accuracy of the species distribution data.

Conditions of appropriate temperature and sufficient water are required for normal plant growth and reproduction (Márquez et al., 2011; Lu et al., 2012; Guo et al., 2016). Thus, the decisive factor for successful species distribution predictions is the use of environmental variables that effectively and reliably describe the environmental characteristics of suitable habitats for a target species. Previous research has shown that the 19 core bioclimatic variables are useful for SDMs, and they are among the most widely used data for that type of model. However, these variables were not sufficient to describe the habitat requirements of *P. euphratica* in this study. In arid zones, such as the lower reaches of the Heihe River Basin, plant growth depends on the level of water that can be absorbed from the environment. Because of the low rainfall and drying temperatures in the study area, plant growth relies mostly on water from rivers and groundwater recharge; therefore, the role of soil moisture and groundwater must be considered. Regarding topographical factors, changes in the condition of the land surface can change the distribution of heat and moisture. Therefore, we chose four types of environmental factors that are important for the habitat of *P. euphratica*: topographical factors, which included elevation, slope and aspect; climatic factors, which primarily included temperature variables; soil factors, comprising ST, soil moisture content, soil organic carbon content, and soil pH; and four hydrological factors, primarily involving groundwater table variables.

MaxEnt was derived from maximum entropy theory, and it has been the most popular tool for species distribution modeling over the past decade. In this study, MaxEnt obtained higher statistical accuracy than the other SDMs using the same model data and strategy. Most importantly, the MaxEnt results were extremely similar to the real spatial distribution pattern of *P. euphratica* forests. This research has shown that on the regional scale, with sufficient and reliable data on environmental variables, the MaxEnt model has excellent capability to model detailed species distributions. Additionally, it demonstrates that MaxEnt has great potential to simulate the patched distributions of species in a fragile or otherwise limited environment. Hence, this method may be used to infer oasis vegetation changes during historical periods by assimilating proxy data (Fang and Li,


(a)
(b)
(c)
(d)
Legend
River
Unsuitable habitats
Marginally suitable habitats
Suitable habitats
Alxa Left Banner
Jiuquan
Jiayuguan
Zhangye
Qilian Mountain
0 75 150 km
97°E
98°E
99°E
100°E
101°E
102°E
38°N
39°N
40°N
41°N
42°N
100°30'E
101°0'E
41°0'N
41°30'N
42°0'N
42°30'N

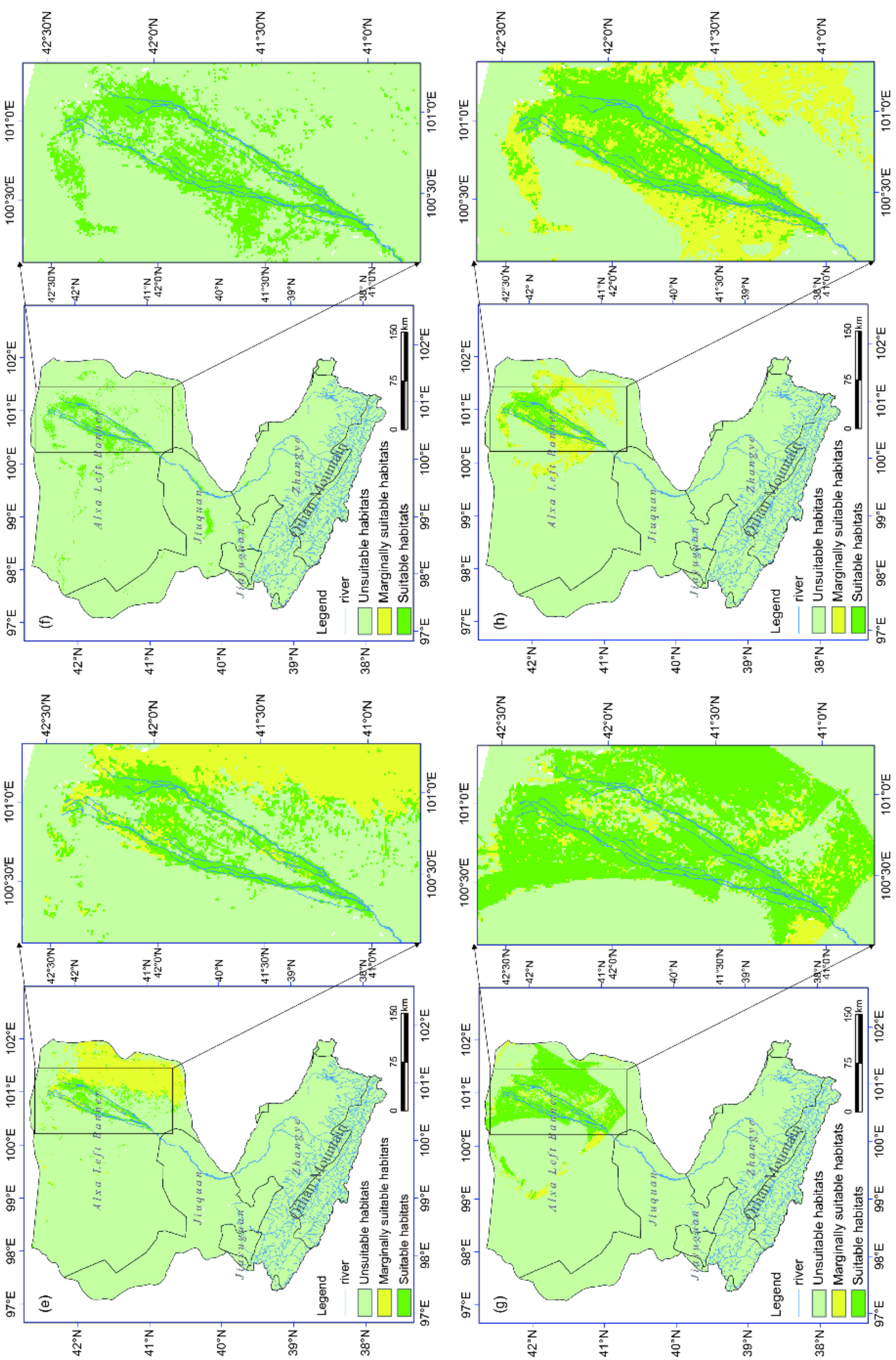


Figure 7 Eight SDMs for the distribution of suitable habitats for *P. euphratica* in the Heihe River Basin. (a) SRE; (b) FDA; (c) MARS; (d) GBM; (e) CTA; (f) GLM; (g) ANN; and (h) RF. SRE: surface range envelope; FDA: flexible discriminant analysis; MARS: multiple adaptive regression splines; GBM: generalized boosting model; CTA: classification tree analysis; GLM: generalized linear model; ANN: artificial neural network; RF: random forest.

2016).

4.2 Model comparison

In this study, we compared nine commonly used SDMs (MaxEnt, SRE, FDA, MARS, GBM, CTA, GLM, ANN, and RF) based on the same input data. In the modeling process, with different mathematical algorithms, these models adopt different response curves and weightings for environmental variables, all of which lead to different model performance. In this case, compared with highly precise land cover data, MaxEnt exhibited superior model performance, with the two main limitations of the other eight models as follows. First, this study used limited occurrence data: considering the small size of the true distribution area and the resolution of the environmental variables, we selected only 57 occurrence data to build the models, which may be insufficient for machine-learning models such as ANN, CTA, and RF (Guo et al., 2016). Second, the environmental variables, especially SSM and underground water table, showed dramatic spatial heterogeneity. This situation makes it challenging to simulate the distribution of *P. euphratica* forests in the study area using regression analysis models such as FDA, MARS, GBM and GLM. With such limited data, these models may be unable to generate complex response curves and determine suitable weighting of environmental variables. Moreover, SRE, which is a traditional climatic envelope model, simply identifies minimum and maximum values for each environmental variable from the occurrence data, and the predicted distribution then includes any site with all variables falling between these minimum and maximum limits (Barbet-Massin et al., 2013). In this case, because of the limited data and dramatic changes in environmental variables, SRE cannot determine the correct range of environmental variables for suitable habitats for *P. euphratica* forests, leading to a significantly undervalued result.

There is another model for simulating suitable habitats for *P. euphratica* forests in the Amudarya Delta. The researchers who developed this model chose geomorphology, groundwater level and flooding regime (i.e., flooding frequency and the timing and duration of the most recent flood) as determining parameters for habitat suitability. Then, based on semi-quantitative and qualitative data, they built a fuzzy HSI model to calculate the HSI value for each evaluation unit (grid cell) in the northern Amudarya Delta (Rüger et al., 2005). This model represents an attempt to automate the evaluation of suitable habitats for *P. euphratica* forests in arid areas, and it is highly meaningful work. Compared with this model, the model used in the present work has the following advantages. First, it includes sufficient and reliable data on environmental variables: because of the lack of environmental data, the HSI model considers only 7 environmental variables, which do not include climatic and soil data. Second, the model in the current study takes an objective approach using only information provided by data. By contrast, the HSI model summarizes the expert's experiences without a precise mathematical model; this type of model makes it difficult to eliminate artificial influences, and it clearly cannot be applied in the absence of sufficient expert knowledge. Third, the current model applies more reasonable weighting of environmental variables. For example, according to the HSI model, groundwater level was the most important variable; this variable alone was nearly completely responsible for determining the distribution of *P. euphratica*. However, in our model, the most important variable was SSM, which is a more powerful index for showing the level of water that can be absorbed from the environment.

4.3 Conservation implications

Species conservation and management is the key to maintaining ecological balance at the regional level, especially in ecologically fragile areas, such as the Heihe River Basin. The results of this study were able to precisely identify the habitat requirements of *P. euphratica*, and these results can be used to purposefully design conservation strategies.

The model calculations indicated that the area of desert riparian forest in the lower reaches of the Heihe River Basin is 430 km^2, which means that under current environmental conditions, *P. euphratica* forest should be distributed more widely. Second, the conservation of resources and the expansion of *P. euphratica* cultivation should prioritized in suitable habitats within seven kilometers from the banks of the river. Third, river flows should be increased in the lower reaches as much as possible to increase SSM levels and raise the underground water table during the growing season.

5. Conclusions

We successfully modeled suitable habitats for *P. euphratica* in the Heihe River Basin, an arid inland river basin region. Moreover, MaxEnt provided satisfactory results based on species distribution data derived from high-resolution remote-sensing data and four types of environmental factors, which were selected on the basis of their environmental characteristics. The AUC value of this model is 0.994, and the model results precisely describe the essential characteristics of the real distribution of *P. euphratica* forest land. Suitable habitats for *P. euphratica* are mainly located in the lower reaches of the Heihe River Basin and distributed along rivers. The potential distribution of *P. euphratica* forest land in the study area is considerable (approximately 390 km^2). These suitable habitats should be prioritized for the management and protection of *P. euphratica* resources.

The present study demonstrates that highly accurate species occurrence data and sufficient and reliable environmental variable data can greatly improve SDM accuracy, especially at the regional scale. Because of the successful prediction of the patched distribution of *P. euphratica* in a fragile arid ecological environment, we can conclude that with the appropriate data inputs, MaxEnt modeling can accurately reflect species distribution characteristics. Furthermore, under suitable conditions, SDMs can predict the realized niches of species.

Acknowledgements *This work was supported by the National Natural Science Foundation of China (Grant No. 91425303), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA20100104) and the 13th Five-year Informatization Plan of Chinese Academy of Sciences (Grant No. XXH13505-06).*

References

Adhikari D, Barik S K, Upadhaya K. 2012. Habitat distribution modelling for reintroduction of *Ilex khasiana* Purk., a critically endangered tree species of northeastern India. Ecol Eng, 40: 37–43

Aishan T, Halik Ü, Cyffka B, Kuba M, Abliz A, Baidourela A. 2013. Monitoring the hydrological and ecological response to water diversion in the lower reaches of the Tarim River, Northwest China. Quat Int, 311: 155–162

Aishan T, Halik Ü, Kurban A, Cyffka B, Kuba M, Betz F, Keyimu M. 2015. Eco-morphological response of floodplain forests (*Populus euphratica* Oliv.) to water diversion in the lower Tarim River, Northwest China. Environ Earth Sci, 73: 533–545

Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). J Appl Ecol, 43: 1223–1232

Anderson R P. 2013. A framework for using niche models to estimate impacts of climate change on species distributions. Ann New York Acad Sci, 1297: 8–28

Andriamparany J N, Brinkmann K, Wiehle M, Jeannoda V, Buerkert A. 2015. Modelling the distribution of four *Dioscorea* species on the Mahafaly Plateau of south-western Madagascar using biotic and abiotic variables. Agric Ecosyst Environ, 212: 38–48

Araújo M B, Luoto M. 2007. The importance of biotic interactions for modelling species distributions under climate change. Glob Ecol Biogeogr, 16: 743–753

Barbet-Massin M, Jiguet F, Albert C H, Thuiller W. 2012. Selecting pseudo-absences for species distribution models: How, where and how many? Methods Ecol Evol, 3: 327–338

Barbet-Massin M, Rome Q, Muller F, Perrard A, Villemant C, Jiguet F. 2013. Climate change increases the risk of invasion by the Yellow-legged hornet. Biol Conserv, 157: 4–10

Bradter U, Kunin W E, Altringham J D, Thom T J, Benton T G. 2013. Identifying appropriate spatial scales of predictors in species distribution models with the random forest algorithm. Methods Ecol Evol, 4: 167–174

Brown C J, O'Connor M I, Poloczanska E S, Schoeman D S, Buckley L B, Burrows M T, Duarte C M, Halpern B S, Pandolfi J M, Parmesan C, Richardson A J. 2016. Ecological and methodological drivers of species' distribution and phenology responses to climate change. Glob Change Biol, 22: 1548–1560

Bucklin D N, Basille M, Benscoter A M, Brandt L A, Mazzotti F J, Romañach S S, Speroterra C, Watling J I. 2015. Comparing species distribution models constructed with different subsets of environmental predictors. Divers Distrib, 21: 23–35

Busby J R. 1991. BIOCLIM—A bioclimate analysis and prediction system. In: Margules C R, Austin M P, eds. Nature Conservation: Cost Effective Biological Surveys and Data Analysis. Victoria: CSIRO. 64–68

Cheng G, Li X, Zhao W, Xu Z, Feng Q, Xiao S, Xiao H. 2014. Integrated study of the water-ecosystem-economy in the Heihe River Basin. Nat Sci Rev, 1: 413–428

Edwards Jr T C, Cutler D R, Zimmermann N E, Geiser L, Moisen G G. 2006. Effects of sample survey design on the accuracy of classification tree models in species distribution models. Ecol Model, 199: 132–141

Elith J, Leathwick J R. 2009. Species distribution models: Ecological explanation and prediction across space and time. Annu Rev Ecol Evol Syst, 40: 677–697

Escalante T, Rodríguez-Tapia G, Linaje M, Illoldi-Rangel P, González-López R. 2013. Identification of areas of endemism from species distribution models: Threshold selection and Nearctic mammals. TIP, 16: 5–17

Fang M, Li X. 2016. Paleoclimate data assimilation: Its motivation, progress and prospects. Sci China Earth Sci, 59: 1817–1826

Fei S, Yu F. 2016. Quality of presence data determines species distribution model performance: A novel index to evaluate data quality. Landsc Ecol, 31: 31–42

Ficetola G F, Thuiller W, Padoa-Schioppa E. 2009. From introduction to the establishment of alien species: Bioclimatic differences between presence and reproduction localities in the slider turtle. Divers Distrib, 15: 108–116

Guo Y, Wei H, Lu C, Gao B, Gu W. 2016. Predictions of potential geographical distribution and quality of *Schisandra sphenanthera* under climate change. PeerJ, 4: e2554

Heikkinen R K, Luoto M, Araújo M B, Virkkala R, Thuiller W, Sykes M T. 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. Prog Phys Geogr, 30: 751–777

Hijmans R J, Cameron S E, Parra J L, Jones P G, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol, 25: 1965–1978

Hu X, Lu L, Li X, Wang J, Lu X. 2015. Ejin oasis land use and vegetation change between 2000 and 2011: The role of the Ecological Water Diversion Project. Energies, 8: 7040–7057

Keyimu M, Halik Ü, Betz F, Dulamsuren C. 2018. Vitality variation and population structure of a riparian forest in the lower reaches of the Tarim River, NW China. J Forest Res, 29: 749–760

Kuemmerlen M, Schmalz B, Guse B, Cai Q, Fohrer N, Jähnig S C. 2014. Integrating catchment properties in small scale species distribution models of stream macroinvertebrates. Ecol Model, 277: 77–86

Li W, Si J, Yu T, Li X. 2016. Response of *Populus euphratica* Oliv. sap flow to environmental variables for a desert riparian forest in the Heihe River Basin, Northwest China. J Arid Land, 8: 591–603

Li X, Lu L, Cheng G, Xiao H. 2001. Quantifying landscape structure of the Heihe River Basin, north-west China using FRAGSTATS. J Arid Environ, 48: 521–535

Li X, Nan Z, Cheng G, Ding Y, Wu L, Wang L, Wang J, Ran Y, Li H, Pan X, Zhu Z. 2011. Toward an improved data stewardship and service for environmental and ecological science data in West China. Int J Digital Earth, 4: 347–359

Liu F, Rossiter D G, Song X D, Zhang G L, Yang R M, Zhao Y G, Li D C, Ju B. 2016. A similarity-based method for three-dimensional prediction of soil organic matter concentration. Geoderma, 263: 254–263

Lu C Y, Gu W, Dai A H, Wei H Y. 2012. Assessing habitat suitability based on geographic information system (GIS) and fuzzy: A case study of *Schisandra sphenanthera* Rehd. et Wils. in Qinling Mountains, China. Ecol Model, 242: 105–115

Marmion M, Luoto M, Heikkinen R K, Thuiller W. 2009a. The performance of state-of-the-art modelling techniques depends on geographical distribution of species. Ecol Model, 220: 3512–3520

Marmion M, Parviainen M, Luoto M, Heikkinen R K, Thuiller W. 2009b. Evaluation of consensus methods in predictive species distribution modelling. Divers Distrib, 15: 59–69

Márquez A L, Real R, Olivero J, Estrada A. 2011. Combining climate with other influential factors for modelling the impact of climate change on species distribution. Clim Change, 108: 135–157

Merow C, Smith M J, Silander Jr J A. 2013. A practical guide to MaxEnt

for modeling species' distributions: What it does, and why inputs and settings matter. Ecography, 36: 1058–1069
Naimi B, Araújo M B. 2016. Sdm: A reproducible and extensible R platform for species distribution modelling. Ecography, 39: 368–375
Pan X, Li X, Cheng G, Li H, He X. 2015. Development and evaluation of a river-basin-scale high spatio-temporal precipitation data set using the WRF model: A case study of the Heihe River Basin. Remote Sens, 7: 9230–9252
Pan X, Li X, Shi X, Han X, Luo L, Wang L. 2012. Dynamic downscaling of near-surface air temperature at the basin scale using WRF—A case study in the Heihe River Basin, China. Front Earth Sci, 6: 314–323
Pan X, Li X, Yang K, He J, Zhang Y, Han X. 2014. Comparison of downscaled precipitation data over a mountainous watershed: A case study in the Heihe River Basin. J Hydrometeorol, 15: 1560–1574
Peng X, Xiao S, Xiao H. 2013. Preliminary dendrochronological studies on *Populus euphratica* in the lower reaches of the Heihe River basin in northwest China. Dendrochronologia, 31: 242–249
Phillips S J, Anderson R P, Schapire R E. 2006. Maximum entropy modeling of species geographic distributions. Ecol Model, 190: 231–259
Pliscoff P, Luebert F, Hilger H H, Guisan A. 2014. Effects of alternative sets of climatic predictors on species distribution models and associated estimates of extinction risk: A test with plants in an arid environment. Ecol Model, 288: 166–177
Ran Y H, Li X, Lu L, Li Z Y. 2012. Large-scale land cover mapping with the integration of multi-source information based on the Dempster–Shafer theory. Int J Geogr Inf Sci, 26: 169–191
Ray D, Behera M D, Jacob J. 2016. Predicting the distribution of rubber trees (*Hevea brasiliensis*) through ecological niche modelling with climate, soil, topography and socioeconomic factors. Ecol Res, 31: 75–91
Ridgeway G. 1999. The state of boosting. Comput Sci Stat, 31: 172–181
Rödder D, Lötters S. 2009. Niche shift versus niche conservatism? Climatic characteristics of the native and invasive ranges of the Mediterranean house gecko (*Hemidactylus turcicus*). Glob Ecol Biogeogr, 18: 674–687
Rüger N, Schlüter M, Matthies M. 2005. A fuzzy habitat suitability index for *Populus euphratica* in the Northern Amudarya delta (Uzbekistan). Ecol Model, 184: 313–328
Song X D, Brus D J, Liu F, Li D C, Zhao Y G, Yang J L, Zhang G L. 2016. Mapping soil organic carbon content by geographically weighted regression: A case study in the Heihe River Basin, China. Geoderma, 261: 11–22
Tayierjiang A, Umut H, Arkin H, Bernd C, Christian O. 2011. Spatial distribution of *Populus euphratica* forests on Argan section in the lower reaches of Tarim River and its influencing factors (in Chinese). J Arid Land Resour Environ, 25: 156–160
Thuiller W, Lafourcade B, Engler R, Araújo M B. 2009. BIOMOD—A platform for ensemble forecasting of species distributions. Ecography, 32: 369–373
Tian Y, Zheng Y, Zheng C, Xiao H, Fan W, Zou S, Wu B, Yao Y, Zhang A, Liu J. 2015. Exploring scale-dependent ecohydrological responses in a large endorheic river basin through integrated surface water-groundwater modeling. Water Resour Res, 51: 4065–4085
Wang J H, Ge C M, Zhao J, Hu X L. 2014. Landuse/Landcover data in the down reaches of the Heihe River Basin in 2011. Cold and Arid Regions Sciences Data Center at Lanzhou (http://westdc.westgis.ac.cn)
Wang Y Y, Zhong B, Shang F J. 2016. An automatic method to extract *populus euphratica* forest in a large area using remote sensing. In: 4th International Conference on Mechanical Materials and Manufacturing Engineering, Wuhan
Yang F, Zhang G L, Yang J L, Li D C, Zhao Y G, Liu F, Yang R M, Yang F. 2014. Organic matter controls of soil water retention in an alpine grassland and its significance for hydrological processes. J Hydrol, 519: 3086–3093
Yang R M, Zhang G L, Liu F, Lu Y Y, Yang F, Yang F, Yang M, Zhao Y G, Li D C. 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. Ecol Indic, 60: 870–878
Yang X Q, Kushwaha S P S, Saran S, Xu J, Roy P S. 2013. Maxent modeling for predicting the potential distribution of medicinal plant, *Justicia adhatoda* L. in Lesser Himalayan foothills. Ecol Eng, 51: 83–87
Yi Y, Cheng X, Yang Z F, Zhang S H. 2016. Maxent modeling for predicting the potential distribution of endangered medicinal plant (*H. riparia Lour*) in Yunnan, China. Ecol Eng, 92: 260–269
Yuan H S, Wei Y L, Wang X G. 2015. Maxent modeling for predicting the potential distribution of Sanghuang, an important group of medicinal fungi in China. Fungal Ecol, 17: 140–145
Zhang X, Nan Z, Wu J, Du E, Wang T, You Y. 2012. Mountain permafrost distribution modeling using multivariate adaptive regression spline (MARS) in the Wenquan area over the Qinghai-Tibet Plateau. Sci Cold Arid Regions, 4: 361–370
Zhao W Z, Chang X L, Li Q Y. 2005. Artificial water diversion effects of HeiheRiver on *Populus euphratica* Oliv. Desert riparian forests in Ejina. Acta Ecol Sin, 25: 1987–1993
Zhu Y, Ren L, Skaggs T H, Lü H, Yu Z, Wu Y, Fang X. 2009. Simulation of *Populus euphratica* root uptake of groundwater in an arid woodland of the Ejina Basin, China. Hydrol Process, 23: 2460–2469

(Responsible editor: Jiancheng SHI)