Mapping the Mangrove Forest Restoration Potential and Conservation Gaps in China Based on Random Forest Model

Zhonghua Yu1,2, Wei Li3, Shaowei Zhang4, Buqing Zhong5, Jun Wang5,6, Shi-young Lee2, Jaehyuck Choi2*, and Shulin Deng1,2*

1Doctor, CAS Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement & Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China
2Professor, Department of Landscape Architecture, PaiChai University, Deajeon 35345, Republic of Korea
3Doctor, Huanghuai University, Zhengzhou 450000, China
4Engineer, Shanxi Geomatics center, Ministry of Natural Resources, Xi’an 710054, China
5Doctor, CAS Engineering Laboratory for Vegetation Ecosystem Restoration on Islands and Coastal Zones, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, 510650, China
6Professor, Xiaoliang Research Station for Tropical Coastal Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, China

ABSTRACT

Background and objective: The area of mangroves is gradually decreasing globally, and mangroves are already one of the most threatened ecosystems. Despite net growth in the mangrove areas in China, the restoration potential of mangroves is still insufficient. This study proposed the Random forest model as an excellent data mining method to map the restoration potential based on the predicted probability of mangrove habitat suitability.

Methods: We demonstrated the vital environmental variables influencing habitat suitability. The decisive advantages of RFM were parsimonious (variables selection), cost-effective (using existing open-source data), accurate (training AUC was 0.89, testing AUC was 0.91), highly efficient (fast-training speed); and its results had high explanatory power. Here, we first mapped the conservation gaps using the RFM.

Results: The results showed that temperature was the most important environmental factor influencing the habitat suit-ability of mangroves. The northern limit of suitable areas was around 24°44’ N. The theoretical suitable habitat area for mangrove was 196,566.6 ha (the highly suitable area was 32,551.4 ha, the medium suitable area was 164,015.2 ha). The potential area for mangrove restoration was 176,264 ha (Guangdong with 104215.4 ha, Guangxi with 65957.5 ha).

Conclusion: We proposed 24 sites with conservation gaps for mangrove forests restoration and nine potential sites as examples for the further restoration plan. We took one example site with high restoration potential for further explanation: how the key environmental factors influence the habitat suitability and how to use the information to guide the restoration strategies. RFM can be used as a data mining algorithm for the utmost use of the presence-only ecological data, objectively evaluating the suitability of species distribution, and providing scientifically technical data for species restoration planning.

Keywords: data mining method, mangrove habitat suitability, predicted probability distribution, restoration plan

Introduction

Mangrove forests are the assemblage of trees and shrubs, occurring along tropic or subtropic ocean coastlines, adapted to grow in intertidal environments (Feller et al., 2010). Together, they form a unique and highly dynamic coastal ecosystem and support numerous ecosystem services, including increasing offshore carbon capture and storage,

This work was supported by the National Natural Science Foundation of Guangdong Province (grant number 2018A030313072).

Received: April 13, 2022, Revised: June 29, 2022, Accepted: July 22, 2022

First author: Zhonghua Yu, yuee.zhonghua2000@gmail.com, https://orcid.org/0000-0003-2957-9717
*Corresponding author: Jaehyuck Choi, jhchoi@pcu.ac.kr, https://orcid.org/0000-0002-2268-9924
Shulin Deng, sdldeng@scbg.ac.cn, https://orcid.org/0000-0002-7231-3415

Ⓒ 2022 by the Society for People, Plants, and Environment. This is a Peer-Reviewed Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Mapping the Mangrove Forest Restoration Potential and Conservation Gaps in China Based on Random Forest Model

protecting coastlines, maintaining biodiversity, and improving environment quality (Allard et al., 2020; Duke et al., 2007; Kathiresan, 2012). Globally, mangroves were reported to occupy 18,100,000 ha, but this estimate of worldwide acreage decreased to 13,776,000 ha, then to 8,349,500 ha (Romañach et al., 2018). Despite the repeated demonstration of their social and economic values, more than 35% of mangrove forests worldwide have been destroyed, even reaching 50-80% in some regions (Ben et al., 2017; Giri et al., 2011). The loss is aggregated mainly by anthropogenic interventions and inappropriate urbanization in coastal regions. The potential consequences may lead to faster erosion and discontinuation of many ecosystem services fundamental to coastal communities. In 2015, the National Coastal Shelter Forest System Construction Project Survey suggested that the restored area of mangrove forests reached 34,100 ha, which reversed the trend of mangrove reduction (Institute, 2020). Although China has experienced a net increase in mangrove areas in the past two decades, the mangrove protection and restoration tasks remain arduous.

Today, it is known that better restoration policies are high-priorities to conserving mangrove ecosystems globally. In 2018, The Nature Conservancy (TNC) released the Global Mangrove Restoration Potential Map, in which the restorable Area in East Asia was estimated to cover only 7 km² (Worthington and Spalding, 2018). Although the exact area of these restoration goals is subject to academic discussion, the insufficiency of mapping the potential restoration areas for mangroves becomes an urgent issue. It is necessary to map ecologically suitable areas for mangroves and to predict that they are potentially restorable using remote sensing technology (Hu et al., 2020). Meanwhile, mapping the conservation gaps of the mangroves to assess the protection status can provide a basis for further adjustment of mangrove reserves.

Various factors affecting mangroves' distribution may exist. Species distribution models (SDMs) supported by remote sensing, spatial analysis, and computer modeling are the main models for analyzing the impact of coastal environmental factors on the distribution of mangroves (Buckley et al., 2010; Elith and Leathwick, 2009). SDMs can offer researchers to map ecologically suitable areas by estimating the relationship between mangrove presence data and environmental and/or geographic factors through statistical algorithms (Guisan et al., 2013; Guisan and Zimmermann, 2000; Peterson, 2006).

However, practical applications of SDMs in restoration and conservation planning and ecosystem assessment are highly dependent on very accurate spatial predictions of ecological process and spatial patterns (Elith and Leathwick, 2009; Zellmer et al., 2019). Experts' knowledge and well-structured assumptions can lead to a considerable understanding of the ecological process (Olden et al., 2008). Nevertheless, given the complexity of the ecosystem, it may be challenging to formulate hypotheses or select variables without first knowing whether there is a correlation. These correlations may be highly non-linear, exhibit autocorrelation, be scale-dependent, or function as an interaction with another variable.

The Random forest model (RFM) provides an excellent application of non-parametric analytical methods for identifying these variables, building accurate predictions, and exploring mechanistic relationships identified in the model instead of traditional inferential approaches based on experts' knowledge (Bradter et al., 2013; Olden et al., 2008). The decisive advantage of RFM is identifying and exploring non-intuitive relationships through an iterative process of exploration and discovery, followed by hypothesis generation and testing. Specifically, the flexibility of RFM to handle complex, high-dimensional interactions allows it to discover relationships hidden in traditional parametric analysis and are unlikely to be proposed a priori by a non-omniscient observer.

Simply put, the RFM can be used as a data mining method. It has the following advantages: it is unexcelled in accuracy among current algorithms, while the RFM improves the prediction accuracy, the calculation amount does not increase significantly; it runs efficiently on large databases; it can handle thousands of input variables without variable deletion; it gives estimates of what variables are important in the classification; it generates an internal unbiased estimate of the generalization error as the forest building progresses; it has an effective method for estimating missing data and maintains accuracy when a large proportion of the data are missing; the physical meaning of the prediction results is easy for interpretation, (Lopatin et al., 2016). When
mapping the potential mangrove forests using a random forest model, it is significant to analyze the possible relationship between mangrove suitability and environmental factors (Bradter et al., 2013; Luan et al., 2020; Naidoo et al., 2012).

In this study, we utilized the classification prediction function of the RFM, (1) combined with land use data to map the mangrove forest restoration potential and conservation gaps in mainland China and Taiwan for the first time, (2) evaluated the suitability of mangroves in the five coastal regions, and determined the environmental factors influencing the habitat suitability of mangroves, (3) to inform mangrove restoration plan mainly related to the critical environment variables. RFM can be a parsimonious, cost-effective, accurate and highly efficient data mining method. It can analyze the existing ecological data in-depth, objectively evaluate the suitability of species distribution, and provide scientifical data for species restoration planning.

**Research Methods**

The methodology of the present study comprised four major steps, including (1) generating an occurrence dataset through remote sensing image interpretation and on-site investigation; (2) collecting and processing environmental variables from different open sources; (3) modeling mangrove suitability using Random forest model; and (4) mapping the restoration potential based on the modeling results and land use data (Globeland 30), (5) mapping the conservation gaps of mangroves in mainland China and Taiwan, respectively, (6) proposing the mangrove restoration plan. The step-by-step flow chart is in Fig. 1.

**Study area**

The study area is located between latitude 18°12’-27°32’ N and longitude 108°03’-120°22’ E across 2,510,934.2 ha coastal areas. Since 1990, 30 mangrove nature reserves have been constructed in the study area, including seven national nature reserves (six internationally essential wetlands). The scope of the study area and current mangrove reserves are shown in Fig. 2.

**Occurrence dataset**

Remote sensing imagery of the spatial distribution of mangroves in the study areas was obtained from National Earth System Science Data Center, National Science & Technology Infrastructure of China (http://www.geodata.cn) (2015, resolution 30 m) (Jia et al., 2019; Jia et al., 2018). We used field survey sample points and the google map (resolution 0.5 m) to evaluate the accuracy of the classification results and the overall mangrove forest. Optimized the mangrove distribution data (2015) and evaluated the classification results using landsat8 OLS remote sensing images. Displayed remote sensing images by false-color synthesis method (Yin et al., 2012) and overlayed the classification data. The classification accuracy was > 90%. These processes were performed in ArcGis10.5.

**Field data survey**

We gathered ground reference data to guide the model’s training and analyze the final map's accuracy. The sampling plots were conducted along the coastline with a dimension of 30 m × 30 m. Also, the field data survey was conducted to identify the classification of mangrove types and their distribution circumstances.
Environment variables dataset

The mangroves' distribution is determined by biophysical and bioclimatic variables in the coast area. Based on the current database, we used four groups of environmental variables, including 19 climate factors, seven seawater factors, three topographic data, and ten soil data. Their meanings are shown in Table 1 and Fig. 4. All 39 environmental factors were regarded as ecological variables. (Guisan and Zimmermann, 2000; Shih, 2020).

The sampling process was as follows: (1) Identified the sampling data range (within the coastline extending 2 km inward and 1 km outward). (2) Divided the study areas into 1 km × 1 km raster data, it was believed that the environmental variables in each grid were consistent. (3) Cut and project the environmental variables into raster data and the

Fig. 2. Study area and mangrove occurrence locations and reserves. The study area included approximately 9,515 kilometers of mainland coastline (Guangxi, Guangdong, Fujian, Hong Kong, and Macau) and 3,001 kilometers of island coastline (Hainan, Taiwan), the northernmost city of Fuding, Fujian Province (27°32' N), and the southernmost city of Sanya, Hainan Province (18°12' N).

Fig. 3. The Random Forest modeling process.
Fig. 4. Present condition of various aspects (Mean temperature of the wettest quarter, Mean temperature of the warmest quarter, Precipitation of the warmest quarter, Elevation, Bulk density, Coarse fragments, Mean sea surface temperature of the wettest quarter, and Mean sea surface of the warmest quarter).
mangrove raster data to form an occurrence dataset. Soil parameters were retrieved from https://soilgrids.org/. Temperature and precipitation parameters were achieved from www.worldclim.com with a spatial resolution of 1 km².

| Data type       | Variable | Description                          | Units       |
|-----------------|----------|--------------------------------------|-------------|
| Bioclimatic     | bio_01  | Annual mean temperature              | ℃           |
| Bioclimatic     | bio_02  | Mean diurnal range                   | ℃           |
| Bioclimatic     | bio_03  | Isothermality                         | unitless    |
| Bioclimatic     | bio_04  | Temperature seasonality               | unitless    |
| Bioclimatic     | bio_05  | Max temperature of the warmest month | ℃           |
| Bioclimatic     | bio_06  | Min temperature of the coldest month | ℃           |
| Bioclimatic     | bio_07  | Annual temperature range             | ℃           |
| Bioclimatic     | bio_08  | Mean temperature of the wettest quarter | ℃   |
| Bioclimatic     | bio_09  | Mean temperature of the driest quarter | ℃   |
| Bioclimatic     | bio_10  | Mean temperature of the warmest quarter | ℃   |
| Bioclimatic     | bio_11  | Mean temperature of the coldest quarter | ℃   |
| Bioclimatic     | bio_12  | Annual precipitation                  | mm          |
| Bioclimatic     | bio_13  | Precipitation of the wettest month   | mm          |
| Bioclimatic     | bio_14  | Precipitation of the driest month    | mm          |
| Bioclimatic     | bio_15  | Precipitation seasonality             | unitless    |
| Bioclimatic     | bio_16  | Precipitation of the wettest quarter | mm          |
| Bioclimatic     | bio_17  | Precipitation of the driest quarter  | mm          |
| Bioclimatic     | bio_18  | Precipitation of the warmest quarter | mm          |
| Bioclimatic     | bio_19  | Precipitation of the coldest quarter | mm          |
| Topographic     | CTI      | Compound Topographic Index            | unitless    |
| Topographic     | LDFG     | Local deviation from global          | unitless    |
| Physical soil property | bdod  | Bulk density                         | cg cm⁻³     |
| Physical soil property | sand  | Sand                                 | g kg⁻¹      |
| Physical soil property | silt  | Silt                                 | g kg⁻¹      |
| Physical soil property | clay  | Clay content                         | g kg⁻¹      |
| Physical soil property | cfvo  | Coarse fragments                     | cm⁻³ dm⁻³   |
| Chemical soil property | cec   | Cation exchange capacity             | mmol(c) kg⁻¹ |
| Chemical soil property | nitrogen | Nitrogen                           | cg kg⁻¹     |
| Chemical soil property | phh2o | pH water                             | pH*10       |
| Chemical soil property | soc   | Soil organic carbon                  | dg kg⁻¹     |
| Soil derived property | ocd   | Organic carbon density               | g dm⁻³      |
| Sea surface temperature | SST1  | Mean sea surface temperature of the coldest quarter | ℃ |
| Sea surface temperature | SST2  | Mean sea surface temperature of the wettest quarter | ℃ |
| Sea surface temperature | SST3  | Mean sea surface temperature of the warmest quarter | ℃ |
| Sea surface temperature | SST4  | Mean sea surface temperature of the driest quarter | ℃ |
| Sea surface temperature | SST5  | Annual mean sea surface temperature  | ℃           |
| Sea surface temperature | SST6  | Mean diurnal range of sea surface temperature | ℃ |
| Sea surface temperature | SST7  | Annual range of sea surface temperature | ℃ |
Topographic variables, including elevation, were derived from the ETOPO1 data of the global relief model from NOAA (https://www.ngdc.noaa.gov/mgg/global/global.html). The compound topographic index (CTI) was calculated as (Hu et al., 2020). The sea surface temperature (SST) data was derived from NASA MODIS-Aqua Level-3 (http://oceancolor.gsfc.nasa.gov). All 39 environmental variables were interpolated to the exact 1 km spatial resolution for subsequent model prediction; the total number of these grids was 41,182.

Sampling points were extracted in each grid (1 km × 1 km) if mangroves exist. The value of these sampling points was 1; on the contrary, it was 0. The number of these grids where mangroves exist was 2,991. ArcGIS 10.5 software extracted the environmental variable values corresponding to the mangroves' distribution points and background points.

Random forest modeling process

A random forest model (RFM) was established based on presence-only data to evaluate suitable mangrove habitats. Random forest is a machine learning algorithm based on an ensemble learning theory developed by Breiman in 2001 (Breiman, 2001; Evans et al., 2011; Jordan and Mitchell, 2015; Mazumder et al., 2010). The Random Forest classification prediction accuracy assessment is based on out-of-bag (OOB) data. The mean decrease of Gini can assess the significance of predictors. We imported the Random forest and Raster packages into the R Studio 4.0.4. We took the environmental variable raster data as an input dataset, thereby acquiring the probability of mangrove suitability and the importance of the variables.

Furthermore, we used the real presence points of mangrove samples to analyze the habitat suitability and drew the environmental response curves of each essential variable. The predicted probability of 50% was used as the threshold to obtain a suitable interval for each variable. We used the Seaborn package in Python 3.7 edition to map the environmental response curve. The modeling process is shown in Fig. 3.

Model validation

For RFM, it is reliable to test its performance using the area Under the Curve (AUC) statistics of the Receiver Operator Characteristic (ROC) plot (Bradter et al., 2013; Evans et al., 2011). The AUC of the ROC plot is equivalent to the probability that a test record is correctly differentiated from a random point in the predefined background study area.

Based on the model's outputs, the grids with an occurrence probability of more than 50% were selected as suitable areas (Lopatin et al., 2016). Then, suitable habitat area was calculated in ArcGIS 10.5. The predicted result of RFM probability ranged from 0 to 1 in accordance with the pixel value of each grid. The mangrove suitable area was divided into three groups: suitable high zone ( > 0.7), suitable medium zone (0.5-0.7), the suitable low area (0.1-0.5).

Restoration potential estimation

RFM aims to map the potential restoration locations for mangroves in the study area according to the predicted probability of habitat suitability. Remote sensing images and field data surveys showed that some areas used to be wetlands or marshes. They were transferred into artificial coasts for urban construction and can hardly be transferred to mangroves permanently. In mapping the accurate areas that can be used for mangrove restoration, it is necessary to combine the predicted areas obtained from RFM with land use and land cover parameters (Ficetola et al., 2014; Vessella and Schirone, 2013).

Land use and land cover (LULC) data were derived from Globaland 30 (http://www.globalandcover.com). We introduced the reciprocal of land use and land cover index as a transfer coefficient to revise the restoration potential probability (RPP) (Zhuang and Liu, 1997). The transfer coefficient was assigned to each grid based on the 2020 LULC data of China with a 1-km resolution. (Table 2).

Mapping the potential restoration areas and conservation gaps

We obtained the distribution map of mangrove suitable habitat areas according to the RPP index. Also, we identified suitable habitat areas with medium and high proba-
Mapping the Mangrove Forest Restoration Potential and Conservation Gaps in China Based on Random Forest Model

The calculated restoration potential area is needed to subtract the area of existing mangrove forests reserves. Finally, the existing mangrove reserves and the map of the mangrove’s potential suitable habitat areas were superimposed to recognize the conservation gaps of the mangrove forests. The conservation gap was defined as the sampling sites with a medium and high probability for restoration and less than 20% of existing mangroves under protection in each grid cell (1 × 1 km) (Sabatini et al., 2020).

Results

Performance of the model and the environmental predictors

The AUC value showed that the training accuracy of the RFM performance reached 0.89, and the testing accuracy reached 0.91, indicating that the model results were highly reliable (Fig. 5). The mean decrease Gini illustrated the air temperature, sea surface temperature, precipitation, and topography had a significant contribution to the habitat suitability of mangroves (Fig. 6). In the RFM, response curves can be created to show how predicted probability occurrence depended on the value of each environmental factor. Eight sensitive variables from the different environmental factors were selected to lead the response curves (Fig. 7). Most curves had a single peak shape, displaying that the mangrove niche was relatively concentrated. The variable intervals with the probability of occurrence threshold of

| Type                        | Description                                             | Degree | Coefficient |
|-----------------------------|---------------------------------------------------------|--------|-------------|
| Unused land                 | Unused land, wetland, salt marshes                      | 1      | 1           |
| Forest, grass and water land| Grass, Shrub, Forest, water                             | 1.25   | 0.8         |
| Agriculture land            | Cultivated land, garden land, artificial grassland      | 1.8    | 0.6         |
| Industry land               | Towns, residential areas, industrial and mining land, transportation land | 2.5    | 0.4         |
| Bare tidal flat             | The tide immersion zone between the high tide and low tide and land covered by soil or sand with the vegetation coverage below 5% | 5      | 0.2         |
| Bare rock                   | A surface of rock or grave covering more than 5% of the land | 0.0    | 0.0         |
| Pond                        | Artificially constructed water storage areas blow the annual water level | 1.25   | 0.8         |

Table 2. The land use and land cover degree index and transfer coefficient

*Note.* Higher land use and land cover index represents a lower probability to be converted to mangroves. A higher transfer coefficient indicates a higher possibility for an area to be transferred to mangrove forests from specific land use type. The LULC index was calculated as follows: \[ L_a = \sum_{i=1}^{n} A_i C_i \], where \( L_a \) denotes the LULC index, \( A_i \) denotes LULC classification index of grade \( i \), \( C_i \) denotes the percentage of LULC of grade \( i \).

The RPP was calculated as follows: \[ RPP = P_i \times L_a \], where \( P_i \) denotes the predicted probability of grid \( i \), and \( L_a \) denotes the transfer coefficient of the LULC in grid \( i \).

Fig. 5. ROC curves for prediction distribution of mangroves. Using the area under the curve (AUC) statistics of the receiver operating characteristic (ROC) plot to test RFM performance. The ROC plot is generated by plotting the false positive rate of a model prediction (1-specificity, here given as fractional predicted area) against the true positive rate (model sensitivity, given as 1-rate of false negatives). An AUC value between 0.9 and 1.0 indicates excellent model performance, 0.8–0.9 = good, 0.7–0.8 = average, 0.6–0.7 = poor, and 0.5–0.6 = insufficient.
Field data survey

The field survey collected in situ exhibited that the mangrove forests in five coastal regions are divided into three types: Kandelia candel-Aegiceras corniculatum-Avicennia marina, Rhizophora stylosa-Bruguiera gymnorrhiza and Rhizophora apiculata-Xylocarpus granatum. Among the three types of mangrove forests, Rhizophora stylosa-Bruguiera gymnorrhiza (37.2%) and Rhizophora apiculata-Xylocarpus granatum (5.2%) were relatively less distributed, but mostly (90% of the total) were protected. Kandelia candel-Aegiceras corniculatum-Avicennia marina (71.6%) were mostly distributed and 52.6% protected.

Mangrove potential suitable habitat area and conservation gaps

The total theoretical area of mangroves' suitable habitat throughout the country was 196,566.6 ha. Here we have analyzed mangrove habitat geographical distribution characteristics from regions, latitudes, climate zones, and elevations. And we further investigated the conservation gaps in these regions.

In terms of mangrove suitable habitat distribution among regions, the predicted probability distribution of mangrove habitat suitability suggested that the medium and high suitable areas concentrated along the coastline in Guangdong, Guangxi, and Hainan provinces (Fig. 8). The predicted proportions of mangroves suitable habitat distributed in the five regions (Guangdong, Guangxi, Hainan, Fujian, and Taiwan) were listed in Table 4. The suitable habitat for mangroves in Guangdong and Guangxi provinces was estimated to be 110,666.3 ha and 68,691.2 ha. Comparatively, the proportion of potentially suitable Area in Fujian and Taiwan was relatively low. Our results indicated that the highly suitable areas for mangrove forests in Guangdong province were mainly concentrated along Beibu Bay, Leizhou Peninsula-Yangjiang Coast, and the Pearl River Estuary-Daya Bay. For Guangxi province, the highly suitable habitats were mainly located in Beibu Gulf, including Zhenzhu Bay, Qinjiang Estuary, Qinzhou Bay, Beihai Port, Anpu Port. In Hainan province, the highly suitable habitats were located in Gaolong Bay-Qinglan Port, Dongzhai Port, Chengmai Bay, Mianiao Bay, and a part of Beibu Bay. In Taiwan, there were no high-
Fig. 7. Response curves of eight important environmental factors influencing mangrove distribution (the blue points illustrate the real presence grids of mangroves).
level suitable areas for mangroves restoration. Regarding mangroves habitat distribution along the latitude gradients, 92% of mangroves' suitable habitat ranged from 18°22' N to 24°44' N.

Regarding mangrove suitable habitat distribution within different climate zones, over 80% of the mangrove area was located in the warmest zone, followed by the wettest zone (50%). Therefore, 90% of the mangrove area was located in the warm and wet zones (Fig. 7 (a) to 6 (b)).

Regarding mangrove suitable habitat distribution along the elevation gradients, 90% of mangroves are located in the area with an average elevation of less than 4 m. However, the range of predicted probability of mangrove suitable habitat distribution narrowed to 2 m. The predicted suitability decreased rapidly with the elevation over 2 m (Fig. 7 (d)).

These area statistics in different climate zones and elevation/latitude gradients were not adjusted using the area-adjusted accuracy. No specific confusion matrix was available for each climate zone and elevation/latitude gradient; however, it did show the general suitability distribution of man-

### Table 3. Variables with the top eight Gini index were regarded as major suitable factors

| Type               | Variables                                              | Gini | Optimum suitable range       |
|--------------------|--------------------------------------------------------|------|------------------------------|
| Bioclimatic        | Mean temperature of the wettest quarter                | 16.3 | 27.5-29.0°C                  |
| Bioclimatic        | Mean temperature of the warmest quarter                | 13.7 | 28.1-29.0°C                  |
| Bioclimatic        | Precipitation of the warmest quarter                   | 12.6 | 600-1,100 mm                 |
| Topographic        | Elevation                                             | 14.5 | 0-22 m                       |
| Physical soil property | Bulk density                                      | 14.9 | 100-130 cm$^{-3}$             |
| Physical soil property | Coarse fragments                                   | 9.8  | 50-110 cm$^{-3}$              |
| Sea surface temperature | Mean sea surface temperature of the wettest quarter | 22.3 | 29.2-33.0°C                 |
| Sea surface temperature | Mean sea surface temperature of the warmest quarter | 17.1 | 30.8-34.0°C                 |

### Table 4. Sum of potential suitable areas for mangrove restoration in five regions

| region    | predicted suitable areas (hectares) | Proportion of suitable areas | Total actual mangrove areas (protected proportion %) |
|-----------|-------------------------------------|----------------------------|-----------------------------------------------------|
|           | High > 0.7 | Medium 0.5-0.7 | Low 0.1-0.5 | Total Suitable Area |
| Fujian    | 0.0         | 77.2         | 710,046.8 | 77.2 | 0.4% | 680.9 (94.0%) |
| Guangdong | 18,854.9    | 91,811.4     | 863,436.2 | 110,666.3 | 56.3% | 8,624.2 (74.8%) |
| Guangxi   | 12,110.2    | 56,581.0     | 120,804.7 | 68,691.2 | 34.9% | 6,920.4 (39.5%) |
| Hainan    | 1,586.3     | 13,625.6     | 316,522.4 | 15,211.9 | 7.7%  | 3,666.7 (75.6%) |
| Taiwan    | 0.0         | 1,920.0      | 303,557.6 | 1,920.0 | 1.0%  | 409.9 (80.6%)  |
| Total     | 32,551.4    | 164,015.2    | 2,314,367.6 | 196,566.6 |                      | 20,302.6 |

Restoration potential area 176,264.0
The area of mangroves in the national nature reserves is currently 12,465.6 ha, accounting for 61.4% of the total mangroves, and 38.6% of the mangroves are distributed outside the reserves' boundaries are not strictly protected. In Fujian and Hainan provinces, most existing mangroves were under protection. In Guangxi province, the proportion of protected mangroves is the lowest, accounting for only 39.5% of this region's total area of mangroves. In these regions, only five mangrove wetlands were listed in the List of wetlands of international importance. Due to the limitation of the minimum area of declared national mangrove reserves, most of China's internationally important mangrove reserves are large and medium-sized, lacking small-scale mangrove reserves. Our results indicated that in Guangdong province, the conservation gaps mainly appeared in Leizhou Bay, Yangjiang Port, Zhenhai Bay, Shantou city Coastlines. For Guangxi province, the conservation gaps were mainly appeared in Beibu Gulf, including Zhenzhu Bay, Qinjiang Estuary, Qinzhou Bay, Beihai Port, Anpu Port. In Hainan province, conservation gaps occurred along the northwest coastline (Fig. 9). As for Taiwan, the conservation gaps mainly appeared in Beigang River estuary.

We suggested 24 potential sites for mangroves restoration. We selected nine example sites as restoration opportunities for the further restoration zoning plan (Fig. 10), the sites information was listed in Table 5. Here, we specifically explained one example site, GD-11, how the environmental factors influenced the mangrove forests suitability, and why we chose them as restoration site.

**Mangrove restoration zoning plan**

The components of the restoration concept plan were designed to take advantage of the unique site conditions and location within the selected restoration sites. The restoration concept plan includes three key zones: (1) buffer zone: restoration and agriculture Complex, (2) restoration zone: internal mangrove and marsh, and (3) conservation zone: external mangrove and marsh. The fundamental drivers of successful ecological community assembly at the chosen region are the tidal flows and substrate elevation.

The adjacent areas between restoration areas and agri-
culture and aquacultural areas can be designed as a buffer zone. The ecological development of this zone will build upon the idle aquaculture ponds and establish a restoration and agriculture complex. Further development of the restoration and agriculture complex will be achieved in three primary ways: expanding the area devoted to this system, introducing additional plant species to diversify the brackish marsh, introducing additional sediment in selected locations to create more areas of appropriate elevation, and diversifying site topography. Therefore, in this zone, additional fill will be used to diversify the topography further and create further opportunities to develop mangroves and marsh.

The restoration zone occupied by the internal mudflat and salt marshes has excellent potential for mangrove restoration. The existing surface elevations are too low and the tidal range are too great, not allowing for vegetation establishment (Pfadenhauer and Grootjans, 1999). Therefore, what the planning proposes for this zone is a phased, gradual reintroduction of tidal flows to achieve additional sediment accumulates in these areas.

In the conservation zone, sometimes, the remaining mangrove forests are exposed to open water. This ecological zone will be created by building structures in the open water zone, adjacent to the reserve within the seawalls, to encourage rapid sediment deposition. At current sediment deposition rates, it may take several decades for bottom elevations to increase to a level that would allow natural vegetation colonization. However, by strategic placement of bioengineering structures such as brush boxes and, potentially, the creation of artificial reefs using rock, the rate of sediment deposition can be significantly speeded in selected areas (Ren et al., 2011). Once appropriate surface elevations are achieved, mangrove and marsh species are expected to colonize them naturally, from propagules derived from zones 1 and 2. Techniques pioneered in this zone would have great potential to be useful elsewhere in the other region to restore or create mangrove habitat.

### Restoration application in GD–11 site

Based on RFM prediction results (habitat suitability), mangrove forests conservation gaps analysis, and field survey,

| No. | Location information | Potential area (ha) | Geographical coordinates | Probability of habitat suitability description | Purpose |
|-----|----------------------|----------------------|--------------------------|---------------------------------------------|---------|
| GX-01 | Coastline of Dandou sea, Guangxi province | 1,307.97 | 21.585457394823408, 109.65719677500098 | High (0.85-0.92) | ELEV (-1-7m) estuary, salt marshes, coastal agriculture land |
| GD-11 | Nansha port, Guangzhou, Guangdong province | 3,616.51 | 22.58856172096921, 113.63329034118887 | High (0.73-0.96) | ELEV (0-5m) estuary, wetland, salt marshes, aquaculture ponds |
| GD-07 | Zhenhai bay, Enping, Guangdong province | 1,093.91 | 21.90670253787534, 112.40960517497093 | High (0.89-0.91) | ELEV (2-13m) bay, wetlands, idle aquaculture ponds |
| GD-05 | Jianjiang Estuary, Zhanjiang, Guangdong province | 1,179.32 | 21.29762299504667, 110.62852207941343 | Medium (0.66-0.72) | ELEV (0-7m) estuary, wetlands, salt marshes |
| GD-08 | Huangmao sea, Guangdong province | 6,630.56 | 21.98826935988563, 113.0308967134063 | High (0.72-0.84) | ELEV (0-4m) idle aquaculture ponds |
| GX-07 | Maowei sea, Guangxi province | 1,521.09 | 21.861408164208, 108.5801488959781 | High (0.73-0.92) | ELEV (0-6m) estuary, wetlands, salt marshes |
| HN-01 | Lingao, Hainan province | 156.72 | 19.98725439635757, 109.81608375947049 | Medium (0.65-0.70) | ELEV (0-5m) coastal agriculture land |
| GD-01 | Anpu port, Guangxi province | 3,702.61 | 21.43719665047053, 109.90301367405756 | High (0.75-0.79) | ELEV (0-6m) estuary, wetlands, aquaculture ponds |
| GD-04 | Leizhou bay, Guangdong province | 2,557.33 | 20.91429382743704, 110.17822858288594 | High (0.79-0.94) | ELEV (0-4m) estuary, idle aquaculture ponds |
we found that GD-11 (Nansha Coastal Wetland, 22°36’N and 113°35’E) with high restoration potential and a good match with LULC data 2020. The site was comprised of approximately 1800 ha in the southern portion of Wanqingsha island, Guangdong. A review of historical satellite images showed that the site was in a relatively shallow zone where considerable sediment deposition was occurring. The coastline was accreting rapidly along the southern boundary of the selected site. Past reclamation activities have significantly reduced intertidal mudflat area and mangroves (Fig. 11). Therefore, we recommended that the mangrove forest restoration area should extend to include the suitable and high suitable zones scattered over the southern and southeastern part of this area.

Regional climate factor set important limiting parameters for mangrove habitat suitability. These parameters determine whether a self-sustaining ecological mangroves community can be established in the area. The climate of this region is subtropical oceanic monsoon. The average temperature of the warmest quarter is 29 °C in July to September. The average annual rainfall is 16,356 mm, with the most precipitation occurring from April to October (wettest and warmest quarter). The multi-year average temperature of surface water is 23.5 °C. Water temperature varies according to four distinct seasons; Spring: 24-27 °C, Summer: 28-30 °C, Autumn: 27-28 °C and Winter: 17-18 °C. Five mangrove forests species were reported as planted: *Bruguiera gymnorrhiza*, *Kandelia candel*, *Sonneratia caseolaris*, *Sonnerati. apetala*, and *Aegiceras corniculatum*. Both *Kandelia candel*, and *Aegiceras corniculatum* were reported occurring spontaneously which could better response to the regional environment. Indeed, choosing species with better environmental suitability is a better option, we suggested transplant *Sonnerntia ovata* to this site to rebuilt the ecological restoration systems.

For the further restoration zoning plan, we proposed the following recommendations: first, we designed the areas reclaimed between 1988 and 2000 as the buffer zone because agricultural fields and aquacultural ponds occupy this area. Buffering agricultural and aquacultural activities from sensitive receptors is a sound land use planning practice. Secondly, designed the areas reclaimed between 2000 and 2010 as the conservation zone, because, in the northeastern corner of the site, small areas of mangrove forests (*Kandelia candel*, and *Aegiceras corniculatum*) had been established spontaneously during the last three decades. It suggested that this region was highly suitable for mangroves. Thirdly, designed the areas reclaimed between 2010 and 2020 as the restoration zone, because there are plenty of idle aquacultural ponds with very low economic value.

**Discussion**

This study used the Random forest model (RFM) to map the restoration potential and conservation gaps of mangroves along the coastline in the five coastal regions, identified the suitable area to construct mangrove reserves based on the most significant environmental variables influencing the habitat suitability. Furthermore, the results attained here have important implications for the mangrove restoration plan.

**Evaluation of random forest model**

The biggest advantage of the RFM is to seek parsimony motivated by the model performance and interpretability (Evans et al., 2011). In the study, we were not only interested in describing a process but also in inference and prediction. Variable selection approaches were focused on parsimony. Usually, these approaches are overly aggressive and often result in too few variables to explain a process.
comprehensively (Luan et al., 2020). Random Forest can exhibit a good fit given a single variable if a variable has very high explanatory power. It is vital to seek a parsimonious set of variables that will provide a good fit when predicted to the mangrove forests and adequately represent the complexities of spatial patterns.

We used the presence-only database, the medium-resolution remote sensing images (30 m), as an input dataset to calculate the probability of the mangrove habitat suitability across 2,510,934.2 ha coastal areas. Compared to other SDMs, the RFM exhibited well-fit prediction results. (Hu et al. 2020) used maximum entropy (MaxEnt) and rule-set prediction (GARP) models to map the mangrove distributions, both of the models need to evaluate the importance of the environment variables at first, variables with both contribution and importance rate of ≥ 1 were regarded as major factors. In such screening systems, require selecting key environmental factors based on experts’ knowledge and could lead to the misinterpretation of the environmental variables. We used the RFM model to directly obtain the environment variables’ contributions after training, making the calculation efficient and straightforward.

Furthermore, compared to Hu's results, the RFM prediction results' physical meaning was easier to explain. RFM did not eliminate all four groups of factors (temperature, soil, terrain, surface seawater temperature) when screening. The selected eight important factors had significant contribution rates in the four sample groups. Mean sea surface temperature of the wettest quarter had the highest score in importance ranking (ranking first); that is, it had the biggest impact on the mangrove suitability, while mean temperature of the coldest quarter scored the lowest. It suggested that the temperature of the coldest quarter and the precipitation of the driest quarter might be the extreme restriction factors for mangroves, not suitable factors. Apart from the climatic variables, high elevation also controlled the distribution of mangroves. The elevation (Fig. 6 (d)) ranged between 0 m and 2 m in the major percentage (88.40%) of the region lying within 2-m height from the mean sea level. Among them, soil properties with bulk density ranged from 100-130 cg cm⁻³ (45.5%), and coarse fragments went from 50-110 cm³ dm⁻³ (62.3%) and occupied major landmass.

Continually compelling arguments demonstrated that the superiority of the RFM lies in high robustness to test model sensitivity to sample distribution. Regarding model accuracy validation, the AUC value suggested that the performance of the RFM was accurate and reliable. A big difference in predicted areas between RFM and Hu's results was observed in the present study (Hu et al., 2020). The predicted areas of mangrove suitable habitat in the present study were much less than GARP, but larger than MaxEnt. The moderately and highly suitable habitats predicted by the RFM were more concentrated in the Guangxi, Guangdong, and Hainan coastal areas; the MaxEnt and GARP models tended to over predict the suitability outside the northern limit of the mangroves’ known ecological niche. In previous multi-model studies, the overprediction was a common issue hypothesized to be caused by the factors contributing to the dispersal inability of the species to those false-predicted regions that were not included in the rulemaking process of the model (Elith and Graham, 2009; Ray et al., 2018; Sobek-Swant et al., 2012; Stockman et al., 2006). Hu's study proposed an issue of the underprediction of the MaxEnt, which belied the heterogeneity of the fine-scale habitat (Phillips and Dudík, 2008). As for our study, considering that the habitat was heterogeneous, RFM could handle the higher resolution of the environmental dataset (1 km², soil resolution 250 m²), and provided a better-fitting model to match the actual mangrove distribution.

**Mangroves suitable conditions for restoration**

Our results demonstrated that the probability of mangrove suitability decreases as the latitude increases, which was consistent with the findings of most research (Chakraborty et al., 2019; Estoque et al., 2018; Shih, 2020). The northern limit of the natural distribution of mangroves in China is Fuding County, Fujian Province (27°20' N), the north boundary of artificial planting mangroves is Yueqing County, Zhejiang Province (28°25’ N) (Jia et al., 2014). Our study predicted that the northern limit of mangrove suitable areas was around 24°44’ N. Relatively, Hu's results believed that the predicted northern limit of mangrove suitable areas was about 28°35’ N (Taizhou, Zhejiang province) (Hu et al., 2020). This difference is probably because we have combined different land use and land cover classification data
and directly incorporated land use and land cover data into environmental factors to participate in the calculation. However, there are small patches of mangrove distributed in this area, the physical conditions of the coast can no longer be used for mangrove restoration. Fujian is more susceptible to cold currents, causing large-scale deaths of planted mangroves in such northern extreme areas. Under the RFM results, there were barely any highly suitable areas for mangrove restoration in Fujian. For some mangrove species, the seed production decreases with increasing latitude, while the survival rate of seedlings increases at lower latitudes (Hong et al., 2021).

Among the 39 environmental variables, the temperature had the greatest effect on mangrove suitability, followed by elevation, soil physical properties and precipitation factors. These results were consistent with previous findings on the habitat suitability of mangroves (Hu et al., 2020). The seasonal surface seawater temperature (SST) is the most important factor in the biomineralization of organic matter and biotransformation of minerals (Gonzalez-Acosta et al., 2006; Tahira et al., 2015). It can affect the mangroves' reproduction and root growth seasonally through tidal inundations (Quisthoudt et al., 2012). According to the Gini index of the environmental factors, the high contribution of the mean temperature of the wettest and warmest quarter indicated that the mangroves preferred to grow in the wet and warm areas; the suitable temperature was 27.5 °C to 29 °C. The optimum surface seawater temperature was 29.2 °C to 34.0 °C. The precipitation is highly correlated with tidal currents (Joyce et al., 2018). The warmest quarter response curve of the precipitation exhibited that the higher habitat suitability concentrated from 600 mm to 1,100 mm. We considered areas with less than 600 mm seasonal precipitation or greater than 1,100 mm as unsuitable habitats, too low or too high tidal currents had a negative impact on mangroves. The predicted suitable elevation for mangroves indicated the intertidal areas could be suitable habitat for mangroves along the coastline in the study areas. The contribution of soil's physical properties (bulk density and coarse fragments) informed that organic and well-structured soil could be suitable for mangroves. We used all these findings in the present study for the following restoration planning.

### Mangrove forests restoration potential and conservation gaps

Comparing the data of this study with the area of existing mangrove reserves, among the major regions where mangroves are distributed (Jia et al., 2018), provinces with smaller mangrove areas had a higher proportion of nature reserves. The theoretical predicted potential restoration area in China was 176,264 ha. Respectively, the predicted potential in Guangdong, Guangxi, and Hainan provinces, with large conservation gaps, were estimated to be 104215.4 ha, 65957.5 ha and 12439.8 ha. In Fujian province, the restoration potential for mangroves was relatively low.

Hainan province has a relatively low distribution area of mangroves, with a higher degree of protection, which may be associated with this region's high richness of mangrove species. There are 26 species of true mangrove plants distributed along the northwest coastlines, more than 90% of China's true mangrove plants and 37% of the world's true mangrove plants, including endemic species in China (Someratia x hainanensis) and critically endangered by International Union for Conservation of Nature (IUCN) (Lumnitzera littorea and Xylocarpus granatum), and other national secondary protected species (Nypafruticans, Someratia x guilngai, Rhizophora apiculata, Bruguiera sexangula, etc.) (Li and Lee, 1997). All of the above have been included in the "List of Provincial Key Protected Wild Plants in Hainan Province". Therefore, the proportion of mangroves under protection is higher than in the other regions. On the contrary, Guangxi province has a large ratio of suitable area for mangrove forests but a low degree of protection, which may be related to insufficient scientifically planning of the mangrove reservation. The mangrove species distributed in these regions are mainly Kandelia candel, Aegiceras corniculatum, Avicennia marina, etc. They are not endangered species and are under less protection. Although the mangroves in Guangxi province provide irreplaceable ecological services in reducing waves, promoting siltation, and protecting banks, the restoration strategies in Guangxi are still insufficient.

The analysis of conservation gaps recommended building or expanding the nature reserve to incorporate the mangroves in the main distribution areas. Some of the con-
servation gaps located at the edges of the existing reserves, the relatively scattered and small mangrove patches should be included in the ecological protection red line to achieve strict protection. And the restoration areas should be expanded to larger potential suitable areas.

Yet it is worth noting that, according to the National Coastal Shelter Forest System Construction Project Survey, which suggested that the restored area of mangrove forests reached 34,100 ha, the present data showed that the actual existing mangroves were much lower than that number.

Implications for the mangrove restoration plan

To date, China has developed many mangrove restoration projects with remarkable achievements. We proposed a high adaptable suggestion for a nationwide mangrove restoration plan based on the present study results. The restoration plan in this study focused on the preliminary plans, which spatially organized the proposed site activities instead of detailed design parameters. We selected medium and high suitable areas in the conservation gaps at the planning level. They considered that the suitability of mangroves was affected by temperature, terrain elevation, and soil properties. We proposed mangrove restoration zoning planning to guide mangrove restoration. At the design level, we considered that mangroves were greatly affected by seawater surface temperature and precipitation. We proposed the mangrove ecological community establishment by introducing additional plant species and reintroducing the tidal flows.

The development of a large, ecologically rich mangrove ecosystem reserve provides the opportunity to make an outstanding contribution to improving the local and regional ecology and the possibility to use the mangrove development as a stimulus to catalyze further regional action and coordination with other protected reserves (Romanach et al., 2018; Worthington and Spalding, 2018). Techniques pioneered, and lessons learned in building the mangrove reservations can be used in other places to do additional ecological restoration and creation.

Appropriate restoration sites selection

According to our prediction results, LULC data, and suitable environmental conditions, rather than the bare tidal flats and bare rocks, the land is covered by wetlands, salt marshes, grasses, shrubs, forests, and coastal areas aquaculture ponds could be considered as suitable sites for mangrove restoration. In China, the increased area of mangrove forests has been mainly from the artificial afforestation of bare tidal flats through the planting certain species of mangroves (Avicennia marina, Kandelia candel, Aegiceras corniculatum and Sonneratia apetala). It will lead to problems such as high cost, low survival rate of artificially planted mangroves, and a decrease in biodiversity. The analysis of conservation gaps from the present study suggested that mangrove forests should be restored from existing wetlands, salt marshes. The mangrove reserves should be expanded to include the relatively scattered and small mangrove patches. Our field survey demonstrated that mangrove trees were often found in ecotones with wetlands and salt marshes. Coastal salt marshes and mangroves are predominantly intertidal; they are found in areas at least occasionally inundated by the high tide but not flooded during the low tide.

On the other hand, against the backdrop of coastal over-expansion in mainland China, 12,584.5 ha of the primary mangrove forests were transferred to agriculture and aquaculture ponds between 1980 to 2000 (Institute, 2020). Meanwhile, unfair plans for coastal management and highly intensive aquacultural development proved to be inefficient. Consequently, about 30% of transferred aquaculture ponds are idle (Hu et al., 2020). Therefore, some researchers suggested that mangrove forests should be restored from the idle ponds in the future (Institute, 2020).

A key objective of mangrove restoration is to create natural, self-sustaining ecological communities (Zedler, 2000; Zhang et al., 2017). Mangrove plants prefer to settle in river deltas, estuarine, and lagoons; they rarely settle in stagnant water (Chen et al., 2009). Mangroves can develop where major rivers carrying large sediment loads build marshes into shallow estuaries or out onto the shallow continental shelf where the ocean is fairly quiet (Canales-Delgadillo et al., 2019). The location of the proposed mangrove re-
serves will be better in a shallow estuary with extensive seawater seasonally fluctuating and a high sediment load. In such a system, the elevation of the sediment is related to the fluctuating elevations of the tidal waters, and the elevation will determine the plants that will successfully grow at a given location or not (Hu et al., 2020).

**limitations**

This study has some limitations regarding the reference data for accuracy and GIS data quality, as described below.

**The reference data for accuracy**

The first is to evaluate the protection status of mangroves in nature reserves only from the distribution and area of mangroves and identify the priority areas for protection outside the reserves. A complete evaluation index system should be established regarding distribution, water environment, and other aspects to comprehensively evaluate mangroves' protective effect. Second, due to the limitation of data and technical conditions, it is impossible to distinguish the artificial forest in the mangrove from the natural forest, resulting in the inability to understand the protection status of the natural mangrove deeply.

**The GIS data quality**

In this study, we use up to 39 environmental variables and various GIS data sources to create a database of environmental variables. We do our best to select high-quality data sources, but multi-source data construction's different periods and resolutions may affect overall data quality. This study uses up to 39 environmental variables and various GIS data sources to create a database of environmental variables. We do our best to select high-quality data sources, but multi-source data construction's different periods and resolutions may affect overall data quality. The climate data comes from the average monthly data of Worldclim from 1970 to 2000, which is relatively old. It is the main defect of time climate data. And the resolution of seawater indicators is low, which will have a certain impact on the expression of the mangrove living environment.

**Suggestions for future research**

The next step will be to explore further the classification methods of natural and artificial forests, focusing on assessing the conservation status of natural mangroves and providing a scientific basis for mangrove conservation planning and management.

China's mangrove protection and restoration efforts will be further strengthened in the future process of mangrove ecological restoration. The restoration of mangrove ecosystems is usually complicated. The conservation and management issues and progress will be more ambitious based on the status and threats of China's mangroves. The mangrove conservation, restoration, sustainable utilization, and management projects can be put forward.

**Conclusion**

By testing the Random forest model using the presence dataset, our findings demonstrated that the RFM could be used as a compelling data mining method to assess the suitable habitat distribution of mangroves. The RFM had extraordinary advantages: parsimonious (variables selection), cost-effective (using existing open-source data), accurate (training AUC was 0.89, testing AUC was 0.91), highly efficient (fast-training speed), and its results had high explanatory power. The mangrove restoration potential could be evaluated through the probability of habitat suitability, and mapping the conservation gaps could provide critical information for mangrove restoration projects. The results from the present study could inform the mangrove forests restoration planning. The predicted results indicated that the highly suitable areas for mangrove forests in mainland China were mainly concentrated along the coastline in Guangdong, Guangxi, and Hainan provinces. Our findings highlighted that temperature, elevation, surface sea temperature, and soil physical properties were the main environmental factors influencing mangrove forests' habitat suitability. The northern limit of suitable areas was around 24°44' N. The mangroves preferred to grow in the wet and warm areas; the suitable temperature was 27.5 °C to 29 °C. The optimum surface seawater temperature was 29.2-34.0 °C. The opti-
mum elevation for restoration was 0 to 2 m. The theoretical suitable habitat area for mangrove was 196,566.6 ha. The potential area for mangrove restoration was 176,264 ha (Guangdong with 104215.4 ha, Guangxi with 65957.5 ha). Combining with the LULC data, our results revealed that Guangdong and Guangxi provinces had the largest mangrove restoration potential among the five regions with mangrove distribution. The conservation gaps in these regions were regarded as suitable areas to construct or expand reserves.

Additionally, we proposed 24 conservation gaps for further restoration concept planning and nine sites as examples. The restoration concept planning focused on selecting a suitable location, mangrove restoration zoning plan, and mangrove ecological community establishment. Based on RFM prediction results, we further explained one site located in conservation gaps with high restoration potential in detail for how the key environmental factors influencing the mangrove habitat suitability and how to use the results to guide the restoration planning. In this site, the regional climate factors set important limiting parameters for mangrove habitat suitability. The findings in the present study could provide an adaptive method for mangrove distribution evaluation and significant planning for mangrove restoration projects in China.

References

Allard, S.M., M.T. Costa, A.N. Bulseco, V. Helfer, L.G.E. Wilkins, C. Hasenruck, K. Zengler, M. Zimmer, N. Erazo, J.L. Mazza Rodrigues, N. Duke, V.M.M. Melo, I. Vanwonterghem, H. Junca, H.M. Makonde, D.J. Jimenez, T.C.L. Tavares, M. Fusi, D. Daffonchio, C.M. Duarte, R.S. Peixoto, A.S. Rosado, J.A. Gilbert, and J. Bowman. 2020. Introducing the mangrove microbiome initiative: Identifying microbial research priorities and approaches to better understand, protect, and rehabilitate mangrove ecosystems. MSystems 5(5). https://doi.org/10.1128/mSystems.00658-20

Antizar-Ladislao, B., S.K. Sarkar, P. Anderson, T. Peshkur, B.D. Bhattacharya, M. Chatterjee, and K.K. Satpathy. 2011. Baseline of butyltin contamination in sediments of Sundarban mangrove wetland and adjacent coastal regions, India. Ecotoxicology 20:1975-1983. https://doi.org/10.1007/s10646-011-0739-5

Ben, G., W. David, P. Mike, F. Mark, H. Tim, H. Dominic, T. Nigel, H. Peter, D. Uwe, and P. Tim. 2017. Multivariate extreme value modelling of sea conditions around the coast of England. Maritime Engineering 170(1):3-20. https://doi.org/10.1680/jmaen.2016.16

Bradter, U., W.E. Kunin, J.D. Altringham, T.J. Thom, and T.G. Benton. 2013. Identifying appropriate spatial scales of predictors in species distribution models with the random forest algorithm. Methods in Ecology and Evolution 4(2):167-174. https://doi.org/10.1111/j.2041-210x.2012.00253.x

Breiman, L. 2001. Random forests. Machine Learning 45L5-32. https://doi.org/10.1023/A:1010933404324

Buckley, L.B., M.C. Urban, M.J. Angilletta, L.G. Crozier, L.J. Rissler, and M.W. Sears. 2010. Can mechanism inform species' distribution models? Ecology Letters 13:1041-1054. https://doi.org/10.1111/j.1461-0248.2010.01479.x

Canales-Delgadillo, J.C., R. Perez-Ceballos, M.A. Zaldivar-Jimenez, M. Merino-Ibarra, G. Cardoza, and J.G. Cardoso-Mohedano. 2019. The effect of mangrove restoration on avian assemblages of a coastal lagoon in southern Mexico. PeerJ 7:e7493. https://doi.org/10.7717/peerj.7493

Chakraborty, S., S. Sahoo, D. Majumdar, S. Saha, and S. Roy. 2019. Future mangrove suitability assessment of Andaman to strengthen sustainable development. Journal of Cleaner Production 234:597-614. https://doi.org/10.1016/j.jclepro.2019.06.257

Chen, L., W. Wang, Y. Zhang, and G. Lin. 2009. Recent progresses in mangrove conservation, restoration and research in China. Journal of Plant Ecology 2(2):45-54. https://doi.org/10.1093/jpe/rtp009

Duke, N.C., J.-O. Meynecke, S. Dittmann, A.M. Ellison, K. Anger, U. Berger, S. Cannici, K. Diele, K.C. Ewel, C.D. Field, N. Koedam, S.Y. Lee, C. Marchand, I. Nordhaus, and F. Dahdouh-Guebas. 2007. A world without Mangroves? Science 317:41-42. https://doi.org/10.1126/science.317.5834.41b

Elith, J. and C.H. Graham. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32:66-77. https://doi.org/10.1111/j.1600-0587.2008.05505.x

Elith, J. and J.R. Leathwick. 2009. Species distribution models: Ecological explanation and prediction Across space and time. Annual Review of Ecology, Evolution, and Systematics
Mapping the Mangrove Forest Restoration Potential and Conservation Gaps in China Based on Random Forest Model

Estoque, R.C., S.W. Myint, C. Wang, A. Ishtiaque, T.T. Aung, L. Emerton, M. Ooba, Y. Hijikaka, M.S. Mon, Z. Wang, and C. Fan. 2018. Assessing environmental impacts and change in Myanmar's mangrove ecosystem service value due to deforestation (2000-2014). Global Change Biology 24(11):5391-5410. https://doi.org/10.1111/gcb.14409

Estoque, R.C., S.W. Myint, C. Wang, A. Ishtiaque, T.T. Aung, L. Emerton, M. Ooba, Y. Hijikaka, M.S. Mon, Z. Wang, and C. Fan. 2018. Assessing environmental impacts and change in Myanmar's mangrove ecosystem service value due to deforestation (2000-2014). Global Change Biology 24(11):5391-5410. https://doi.org/10.1111/gcb.14409

Evans, J.S., M.A. Murphy, Z.A. Holden, and S.A. Cushman. 2011. Modeling species distribution and change using random forest. Predictive Species and Habitat Modeling in Landscape Ecology 139-159. https://doi.org/10.1007/978-1-4419-7390-0_8

Feller, I.C., C.E. Lovelock, U. Berger, K.L. McKee, S.B. Joyce, and M.C. Ball. 2010. Biocomplexity in mangrove ecosystems. Annual Review of Marine Science 2:395-417. https://doi.org/10.1146/annurev.marine.010908.163809

Guisan, A., and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 138(2-3):147-186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hong, P., Y. Wen, Y. Xiong, L. Diao, X. Gu, H. Feng, C. Yang, and L. Chen. 2021. Latitudinal gradients and climatic controls on reproduction and dispersal of the non-native mangrove Sonneratia apetala in China. Estuarine, Coastal and Shelf Science 248:106749. https://doi.org/10.1016/j.ecss.2020.106749

Hu, W., Y. Wang, D. Zhang, W. Yu, G. Chen, T. Xie, Z. Liu, Z. Ma, J. Du, B. Chao, G. Lei, and B. Chen. 2020. Mapping the potential of mangrove forest restoration based on species distribution models: A case study in China. Science of Total Environment 748:142321. https://doi.org/10.1016/j.scitotenv.2020.142321

Institute, P. 2020. Research report on mangrove protection and restoration strategy in China.

Jia, M., Z. Wang, L. Li, K. Song, C. Ren, B. Liu, and D. Mao. 2014. Mapping China’s mangroves based on an object-oriented classification of Landsat imagery. Wetlands 34:277-283. https://doi.org/10.1007/s13157-013-0449-2

Jia, M., Z. Wang, C. Wang, D. Mao, and Y. Zhang. 2019. A new vegetation index to detect periodically submerged mangrove forest using single-tide sentinel-2 imagery. Remote Sensing 11(17):2043. https://doi.org/10.3390/rs11172043

Jia, M., Z. Wang, Y. Zhang, D. Mao, and C. Wang. 2018. Monitoring loss and recovery of mangrove forests during 42 years: The achievements of mangrove conservation in China. International Journal of Applied Earth Observation and Geoinformation 73:535-545. https://doi.org/10.1016/j.jag.2018.07.025

Jordan, M.I. and T.M. Mitchell. 2015. Machine learning: Trends, perspectives, and prospects. Science 349(6245):255-260. https://doi.org/10.1126/science.aaa8415

Joyce, J., N.B. Chang, R. Harji, T. Ruppert, and P. Singhofen. 2018. Cascade impact of hurricane movement, storm tidal surge, sea level rise and precipitation variability on flood assessment in a coastal urban watershed. Climate Dynamics 51:383-409. https://doi.org/10.1007/s00382-017-3930-4

Kathiresan, K. 2012. Importance of mangrove ecosystem. International Journal of Marine Science 2(10):70-89

Li, M.S. and S.Y. Lee. 1997. Mangroves of China: a brief
review. Forest Ecology and Management 96(3):241-259. https://doi.org/10.1016/S0378-1127(97)00054-6

Lopatin, J., K. Dolos, H.J. Hernández, M. Galleguillos, and F.E. Fassnacht. 2016. Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. Remote Sensing of Environment 173:200-210. https://doi.org/10.1016/j.rse.2015.11.029

Luan, J., C. Zhang, B. Xu, Y. Xue, and Y. Ren. 2020. The predictive performances of random forest models with limited sample size and different species traits. Fisheries Research 227:105534. https://doi.org/10.1016/j.fishres.2020.105534

Mazumder, R., T. Hastie, and R. Tibshirani. 2010. Spectral regularization algorithms for learning large incomplete matrices. Journal of Machine Learning Research 1(1):2287-2322. http://www.ncbi.nlm.nih.gov/pubmed/21552465

Naidoo, L., M.A. Cho, R. Mathieu, and G. Asner. 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. ISPRS Journal of Photogrammetry and Remote Sensing 69:167-179. https://doi.org/10.1016/j.isprsjprs.2012.03.005

Olden, J.D., J.J. Lawler, and N.L. Poff. 2008. Machine learning methods without tears: A primer for ecologists. The Quarterly Review of Biology 83:171-193. https://doi.org/10.1086/587826

Peterson, A.T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3:59-72. https://doi.org/10.17161/bi.v3i0.29

Pfadenhauer, J. and A. Grootjans. 1999. Wetland restoration in Central Europe: Aims and methods. Applied Vegetation Science 2:95-106. https://doi.org/10.2307/1478886

Phillips, S.J. and M. Dudik. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31:161-175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Quishoudt, K., N. Schmitz, C.F. Randin, F. Dahdouh- Guexes, E.M.R. Robert, and N. Koedam. 2012. Temperature variation among mangrove latitudinal range limits worldwide. Trees 26:1919-1931. https://doi.org/10.1007/s00468-012-0760-1

Ray, D., M.D. Behera, and J. Jacob. 2018. Evaluating ecological niche models: A comparison between maxent and GARP for predicting distribution of hevea brasiensis in India. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences 88:1337-1343. https://doi.org/10.1007/s40011-017-0869-5

Ren, H., X. Wu, T. Ning, G. Huang, J. Wang, S. Jian, and H. Lu. 2011. Wetland changes and mangrove restoration planning in Shenzhen Bay, Southern China. Landscape and Ecological Engineering 7:241-250. https://doi.org/10.1007/s11355-010-0126-z

Romañach, S.S., D.L. DeAngelis, H.L. Koh, Y. Li, S.Y. Teh, R.S. Raja Barizan, and L. Zhai. 2018. Conservation and restoration of mangroves: Global status, perspectives, and prognosis. Ocean and Coastal Management 154:72-82. https://doi.org/10.1016/j.ocecoaman.2018.01.009

Sabatini, F.M., W.S. Keeton, M. Lindner, M. Svoboda, P.J. Verkerk, J. Bauhus, H. Bruelheide, S. Burrascano, N. Debaive, I. Duarte, M. Garbarino, N. rigoriadis, F. Lombardi, M. Mikoláš, P. Meyer, R. Motta, G. Mozgeris, L. Nunes, P. Ódor, M. Panayotov, A. Ruete, B. Simovski, J. Stillhard, J. Svensson, J. Szwagrazyk, O. Tilkkanen, K. Vandekerkhove, R. Volosyanchuk, T. Vrska, T. Zlatanov, and T. Kuemmerle. 2020. Protection gaps and restoration opportunities for primary forests in Europe. Diversity and Distributions 26:1646-1662. https://doi.org/10.1111/ddi.13158

Shih, S.S. 2020. Spatial habitat suitability models of mangroves with kandelia obovata. Forests 11(4):477. https://doi.org/10.3390/f11040477

Sobek-Swant, S., D.A. Kluza, K. Cuddington, and D.B. Lyons. 2012. Potential distribution of emerald ash borer: What can we learn from ecological niche models using Maxent and GARP? Forest Ecology and Management 281:23-31. https://doi.org/10.1016/j.foreco.2012.06.017

Stockman, A.K., D.A. Beamer, and J.E. Bond. 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. Diversity and Distributions 12:81-89. https://doi.org/10.1111/j.1366-9516.2006.00225.x

Tahira, N., B. Nazima, M. Roomina, and I. Noshin. 2015. Effects of siltation, temperature and salinity on mangrove. European Academic Research 11:14172-14179.

Vessella, F. and B. Schirone. 2013. Predicting potential distribution of Quercus suber in Italy based on ecological niche models: Conservation insights and reforestation involvements. Forest Ecology and Management 304:150-161. https://doi.org/10.1016/j.foreco.2013.05.006
Wang, H.T., J.Q. Su, T.L. Zheng, and X.R. Yang. 2014. Impacts of vegetation, tidal process, and depth on the activities, abundances, and community compositions of denitrifiers in mangrove sediment. Applied Microbiology and Biotechnology 98:9375-9387. https://doi.org/10.1007/s00253-014-6017-8

Worthington, T. and M. Spalding. 2018. Mangrove restoration potential: A global map highlighting a critical opportunity. https://doi.org/10.17863/CAM.39153

Yin, T., T.T. Zhang, and X.M. Xiong. 2012. False color composite system of multi-spectral RS images based on IDL. Advanced Materials Research 468:1671-1674. https://doi.org/10.4028/www.scientific.net/AMR.468-471.1671

Zedler, J.B. 2000. Progress in wetland restoration ecology. Trends in Ecology and Evolution 15:402-407. https://doi.org/10.1016/S0169-5347(00)01959-5

Zellmer, A.J., J.T. Claisse, C.M. Williams, S. Schwab, and D.J. Pondella. 2019. Predicting optimal sites for ecosystem restoration using stacked-species distribution modeling. Frontiers in Marine Science 6:3. https://doi.org/10.3389/fmars.2019.00003

Zhang, M., S. Liang, A. Suo, and Z. Sun. 2017. Progress of coastal environment repairing and cleaning engineering research and its prospect. Journal of Marine Environmental Science 36:635-640. https://doi.org/10.13634/j.cnki.mes.2017.04.024

Zhuang, D. and J. Liu. 1997. Study on the model of regional differentiation of land use degree in China. Journal of Natural Resources 12:105-111. https://en.cnki.com.cn/Article_en/CJFDTotal-ZRZX702.001.htm