

# XGB Model: Research on Evaporation Duct Height Prediction Based on XGBoost Algorithm

Wenpeng ZHAO<sup>1,2</sup>, Jincal LI<sup>1</sup>, Jun ZHAO<sup>1</sup>, Dandan ZHAO<sup>3</sup>, Jingze LU<sup>1,2</sup>, Xiang WANG<sup>1</sup>

<sup>1</sup> College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China

<sup>2</sup> College of Computer Science, National University of Defense Technology, Changsha 410073, China

<sup>3</sup> State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry (LAPC), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

zhaowenpeng18@nudt.edu.cn, lijincal@nudt.edu.cn, zhaojun@nudt.edu.cn, zhaodandan@dq.cern.ac.cn, lujingze18@nudt.edu.cn, xiangwangcn@nudt.edu.cn

Submitted September 12, 2019 / Accepted November 4, 2019

**Abstract.** *Evaporation duct is a specific atmospheric structure at sea, which has an important influence on the propagation path of electromagnetic waves (EW). Considering the limit of existing evaporation duct height (EDH) prediction models and aiming at proposing more accurate and stronger generalization ability of EDH models, we applied eXtreme Gradient Boosting (XGBoosting) algorithm to the field of evaporation duct for the first time. And we proposed the new EDH prediction model using XGBoost algorithm (XGB model). Simultaneously, traditional Paulus-Jeske (PJ) model and deep learning Multi-layer Perceptron (MLP) model were introduced into the experiment to make a comparison. In terms of comprehensive performance, XGB model is optimal in all sub-regions and total area. Finally, cross-learning experiments were carried out to test the generalization ability of XGB model. The results show that the generalization ability of XGB model is better than that of MLP model.*

## Keywords

Evaporation duct, machine learning, XGBoost algorithm, XGB model, Paulus-Jeske (PJ) model

## 1. Introduction

Atmospheric duct is a special atmospheric stratification in which the refractivity of the lower troposphere decreases sharply with the increase of altitude. This stratification makes the curvature radius of radio rays less than the curvature radius of the earth. As long as the frequency and angle are appropriate, the energy of EW will be refracted repeatedly in the atmosphere stratification, which restricts the propagation of EW beyond in this narrow region, thus forming over-the-horizon propagation [1], [2]. Previous researches show that evaporation duct is easily formed in the air-sea boundary environment and is widely distributed in the surrounding sea areas of China. Espe-

cially in the South China Sea, the occurrence probability of available evaporation duct is about 80% [3]. The occurrence probability in some specific sea areas is as high as 90% or even exists permanently [1]. Regarding to variations of the height of evaporation duct, multiple researches reported the global annual average EDH is about 13 meters with the maximum of 40 m [1]. Generally, EDH is close to that of radio information system equipment such as ship radar and communication at sea. The trapping effect of evaporation duct can change the propagation path and attenuation characteristics of EW, which enables the equipment to detect beyond-the-horizon but may also cause adverse effects such as detection and communication dead zone [4], [5]. Therefore, accurate prediction of EDH is of great practical significance for detection and communication at sea and has always been the research focus of atmospheric duct [6], [7].

At present, the main measurement methods of EDH include direct measurement, radar clutter inversion, satellite signal inversion, theoretical model prediction and machine learning prediction. Because of the low EDH, the general direct measurement method is to observe the vertical distribution of meteorological elements by the meteorological gradiometer and calculate the refractivity profile by empirical formula [8]. Using the microwave refractivity meter to measure the refractivity profile is also a direct measurement method [9]. However, there are some shortcomings of direct measurement method, such as high cost, low spatial and temporal resolution of measured data and poor confidentiality [10]. Radar clutter inversion method inverts EDH using radar backscattering signals that actively transmit and receive EW from the sea surface, which has the problems of poor concealment and low accuracy [11], [12]. Satellite signal inversion method is a method of inverting evaporation duct by using the power distribution of satellite reflected signals on the sea surface, but the technology is not mature enough with a low precision [13].

Theoretical model prediction method is the main method to obtain the vertical distribution of the corrected

refractivity and EDH, which is generally based on the Monin-Obukhov similarity theory of atmospheric boundary layer [14]. The refractivity profiles of evaporation duct are obtained and EDH is determined according to the meteorological factors (mainly sea surface temperature, air temperature, humidity, pressure, wind speed at specific altitudes). A large number of relevant researchers have been engaged in the research of evaporation duct model [14]. Many prediction models have been proposed, including Liu-Katsaros-Businger (LKB) model [15], Paulus-Jeske (PJ) model [16], Naval Research Laboratory (NRL) model [17], [18], Musson-Gauthier-Bruth (MGB) model [19], Babin-Young-Carton (BYC) model [20], Naval-Postgraduate-School (NPS) model [21] and CORE 3.0 algorithm model [22], as well as domestic pseudo-refractivity model [23], Universal Evaporation Duct (UED) model [24] and non-iterative flux algorithm evaporation duct prediction model [25]. Due to the complexity of the local geographic and hydrometeorological environment, the practicability of the theoretical prediction model at sea based on the Monin-Obkhov similarity theory of homogeneous and steady atmospheric boundary layer usually has some limitations and large prediction errors. PJ model is the most widely used and successful evaporation duct prediction model in the 20th century, which has been applied to the Integrated-Refractive-Effect Prediction System (IREPS) of the U.S. Navy since 1978 and is still in use [26], [27]. At present, many scholars have carried out research on the adaptability of PJ model in Chinese sea area with good results [28–30]. Therefore, we choose PJ model for subsequent contrast experiment analysis, which is valuable for comparison.

Machine learning prediction method is a method that computer trains EDH prediction model with hydrometeorological observation data on the sea surface to predict new observation data. Machine learning is a general term for data analysis algorithms, which enables computers to quickly mine the correlation and implicit rules of large amounts of data without explicit programming, and to do a precise regression prediction or classification [31], [32]. Currently, machine learning has been widely used in many research fields, including artificial intelligence, financial industry, network security and so on, which has a significant impact on social productivity and economy. For atmospheric duct, machine learning prediction method is a new technology for EDH prediction with high efficiency and accuracy, which possesses great application value and achieved some results [33–37]. However, these researchers mainly use support vector machine (SVM) and feedforward neural network. SVM has some problems, such as difficulty to implement large-scale samples, sensitivity to parameter and kernel function selection. Feedforward neural network also has some problems, such as large sample dependence and weak generalization ability. However, the methods still have limitations in application.

In summary, more accurate and stronger generalization ability of evaporation duct models need to be proposed. Based on the data of hydrometeorological and EDH

measured during a voyage, a pure data-driven prediction model of XGBoost evaporation duct was proposed in this paper by using decision tree algorithm XGBoost (XGB model) for the first time. In consideration of performance comparison of MLP and XGBoost have been applied to a land cover urban classification [38], MLP model proposed by Zhu X. Y. et al. [36], [37] and traditional PJ model are introduced in the contrast experiment. The predicted results based on XGB model, MLP model and PJ model were compared with the measured EDH of the voyage. A comprehensive performance evaluation experiment was carried out to evaluate the comprehensive performance of XGB model. In order to test the generalization ability of XGB model, cross-learning comparative experiments were also conducted.

This paper is organized as follows: Section 2 introduces the existing EDH prediction models and methods; Section 3 establishes XGB model; Section 4 introduces the experimental and results analysis; and Section 5 presents the conclusion.

## 2. Introduction of Existing Evaporation Duct Prediction Models and Methods

### 2.1 PJ Model

PJ model is the most widely used EDH prediction model developed by Paulus on the basis of Jeske in the 20th century [16], which has been coupled into the integrated refraction effects prediction system [26]. The input parameters of the model are sea surface temperature and atmospheric temperature, relative humidity, wind speed and assumed atmospheric pressure of 1000 hPa.

The PJ model replaces atmospheric refractivity  $N$ , atmospheric temperature  $T$  and vapor pressure  $e$  with atmospheric potential refractivity  $N_p$ , potential temperature  $\theta$  and potential vapor pressure  $e_p$  respectively. However, Paulus assumes that they are equal. The expression of  $N_p$  is as follows:

$$N_p \approx N = \frac{77.6}{T} \left( 1000 + \frac{4810e}{T} \right). \quad (1)$$

The critical potential refractivity gradient is determined by  $\partial N_p / \partial z$ .  $\partial N_p / \partial z$  is [41]:

$$\frac{\partial N_p}{\partial z} = \frac{S_{N_p} \Phi}{\rho_a \kappa u_* (z + z_0)} \quad (2)$$

where  $S_{N_p}$  is vertical flux of  $N_p$ .  $\Phi$  is dimensionless gradient defined as a stable function. For stable atmosphere,  $\Phi = 1 + \alpha(z/L)$ ,  $\alpha$  is Businger-Dyer constant of 5.2,  $L$  is Monin-Obkhov length. For unstable atmospheric conditions,  $\Phi^4 - 4\alpha(z/L)\Phi^3 = 1$ ,  $\alpha$  is of 4.5.  $\rho_a$  is atmospheric

density.  $\kappa$  is Kalman constant of 0.4,  $u_*$  is friction velocity,  $z$  is the height of 6 meters and  $z_0$  is rough length.

Paulus [16] derived an empirical relationship from many field experiments, and the EDH can be calculated by solving (2). The further details of PJ model can be found in [16].

## 2.2 Evaporation Duct Prediction Model for Multilayer Perceptron

This model is a pure data driven evaporation duct prediction model, MLP model, which is built by Zhu Xiaoyu et al. [38], [39] based on multi-layer perception mechanism. The following is a brief introduction.

Multilayer Perceptron (MLP), also known as Deep Feedforward Network, is a typical neural network. Its goal is to approximate a function  $f(x)$ . For regression problems,  $y = f(x)$  maps the input  $x$  to the function value of the function  $f(x)$  to be fitted. Figure 1 is the simplest MLP network structure, which includes three basic network types: Input, Hidden layer and Output. MLP contains the most quintessential part of the neural network theory. It is widely used in pattern recognition, function approximation, regression and classification, which is the most basic neural network and also one of the most widely used neural networks at present.

The essence of MLP is a non-linear function mapping from input vector to output, which is very similar to the design idea of traditional evaporation duct prediction model. The input format of MLP model based on this idea is similar to that of PJ model, and the input variables are sea surface temperature, atmospheric temperature, relative humidity, wind speed and air pressure. Therefore, the input layer contains five neurons corresponding to these five observation elements. Since it is essentially a regression model, the output layer contains one neuron. According to the specific situation of the experimental data, through a large number of computer experiments, the hyperparameters are finally determined as follows: hidden layers are 4, the number of neurons in each layer from shallow to deep: 10, 15, 10, 5. Batch size is 128 and epoch is 100. However, the hyperparameters will be fine-tuned according to the

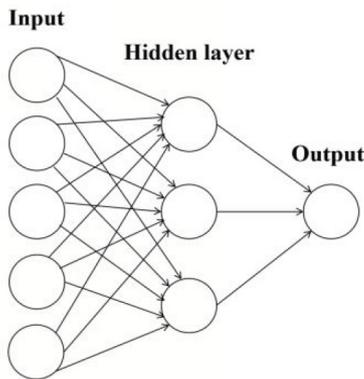


Fig. 1. Structural sketch of MLP neural network.

amount of data. Finally, the MLP model is obtained, which has been trained on the training set with adjusted hyperparameters.

## 3. Establishment of XGBoost Evaporation Duct Prediction Model

### 3.1 Introduction of XGBoost Algorithm

XGBoost is short for eXtreme Gradient Boosting package, which is designed and optimized to be efficient, flexible and portable. The package includes efficient linear model solver and tree learning algorithm. It supports various objective functions, including ranking, classification and regression [39]. It is an efficient and scalable implementation and optimization of gradient boosting framework [40]. Regularized model is used to control the complexity of the model, which makes the learning model simpler and avoids overfitting. XGBoost supports parallelization. It implements parallelization when selecting the best splitting points for enumeration, which makes the training speed very fast. The tree building is stopped ahead of time when the prediction results are good, so that the training speed is accelerated. Furthermore, it supports setting sample weight embodied in the firstly derivative  $g$  and the secondly derivative  $h$ . We can pay more attention to some samples by adjusting the weight. At present, XGBoost is one of the most successful machine learning algorithms.

The traditional gradient boosting frameworks are expressed as follows:

$$F_m(x) = \sum_{i=1}^m \beta_i f_i(x) \quad (3)$$

where  $m$  is the number of weak learners,  $\beta$  is the coefficient,  $f$  is the weak learner and  $F$  is the general model.

Our ultimate goal is to get an excellent general total model, which minimizes the loss function as much as possible. That is:

$$F_m \arg \min \sum_{i=1}^n L(y_i, F_m(x_i)) = \beta_m \arg \min \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + \beta_m f_m(x_i)). \quad (4)$$

It is impossible to solve it simultaneously since  $F$  is eventually weighted by multiple weak learners. Therefore, gradient lifting uses a greedy algorithm. At the beginning, model  $F$  is a constant function, and only one weak learner and its coefficients are solved on time to improve the performance of  $F$  step by step. Gradient boosting minimizes the loss function as quickly as possible by making the new term equal to the negative gradient of the loss function in each step.

The new item can be written in the following form:

$$\beta_m f_m(x_i) = \gamma \frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)} \quad (5)$$

where  $\gamma$  is called step size, which contains minus sign of negative gradient and  $\beta$  parameter. The part on the left of the formula that removes  $\gamma$  is generally called pseudo-residual (also known as gradient):

$$R_{im} = \gamma \frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)}. \quad (6)$$

According to the above formula, the previous step model of  $F$  knows that for each training sample. We can always find its pseudo residual  $R$ . In this way, as long as a weak learner is assumed, we can train the current weak learner  $f_m$  according to  $x$  and  $y$  values of the training samples. Finally, the trained weak learner is drag-in formula (4) to minimize the loss function, and  $\gamma$  can be obtained.

The final general model is:

$$F_m(x) = F_{m-1}(x) + \beta_m f_i(x). \quad (7)$$

However, Xgboost improves the regularization learning target  $L$  on the basis of traditional gradient boosting framework.  $L$  is:

$$L = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k). \quad (8)$$

The first term  $l$  is a differentiable convex loss function, which measures the difference between the predicted value and the target value.  $\hat{y}_i = \sum_{k=1}^K f_k(x_i)$ . The second term  $\Omega$  is the sum of the complexity of all regression trees.  $\Omega(f) = \gamma T + \lambda \|w\|^2/2$ , where  $\gamma T$  is total number of leaf nodes,  $\gamma$  is the difficulty of node segmentation,  $T$  is the number of leaf nodes;  $\lambda \|w\|^2/2$  is the regularization term L2,  $\lambda$  is the regularization coefficient.

The optimal leaf node fraction  $w$  of XGBoost and the objective function  $Obj$  formula are respectively:

$$w_j^* = -\frac{G_j}{H_j + \lambda},$$

$$Obj = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (9)$$

where  $G_j = \sum_{i \in I_j} g_i$ ,  $H_j = \sum_{i \in I_j} h_j$ ,  $g_i$  and  $h_i$  represent the first derivative and the second derivative of the prediction error to the current model respectively.

Different machine learning models are suitable for different types of tasks. Deep neural network can capture high-dimensional and large data such as image, voice and text well by modeling space-time position. XGBoost based on tree model can deal with table data well, which is less demanding on sample size and feature data types. And the application scope of it is wide. Meanwhile, XGBoost has some characteristics that deep neural network does not have, including model interpretability, invariance of input

data and ease to adjust parameters. In theory, the application effect of XGBoost algorithm will be better due to the low-dimensional tabular observed data of evaporation duct and small amount of data.

### 3.2 Building XGB Model

The essence of traditional evaporation duct prediction model is the function of hydrometeorological factors (such as sea surface temperature, air temperature, relative humidity, wind speed, air pressure, etc.). Taking PJ model as an example, Paulus and Jeske obtained the empirical relationship of similar variables in local boundary layer through large-scale field observation experiments. Combining Monin-Obukhov similarity theory with the formation mechanism of evaporation duct itself, the complex functional relationship between EDH and basic hydrometeorological factors was finally obtained

$$EDH = f(SST, T_a, RH, U).$$

$SST$  is the sea surface temperature,  $T_a$  is the temperature,  $RH$  is the relative humidity,  $U$  is the wind speed.

XGBoost is good at regression of low-dimensional tabular data, which is essentially a function mapping from input vector to output and is consistent with the traditional evaporation duct prediction model. The prediction model obtained by XGBoost is entirely based on measured data, which is not subject to basic physical assumptions such as atmospheric boundary layer similarity theory and physical backgrounds. Based on the above theory, we propose a pure data driven evaporation duct prediction model based on XGBoost.

Firstly, a series of pretreatments of the original observation data are carried out, and the processed sample pairs  $(x_i, y_i)_{i=1}^n$  are obtained. The original sample is used as input sample feature for  $x_i = (SST_i, T_i, RH_i, U_i, P_i)$  ( $SST_i, T_i, RH_i, U_i$ , and  $P_i$  represent the sea surface temperature, atmospheric temperature, relative humidity, wind speed and pressure of the  $i$ th sample respectively). The measured  $EDH$   $y_i = EDH_{Obs_i}$  is used as sample label.  $n$  is the number of total samples. 70% of  $n$  is the training set and 30% of  $n$  is the test set. Determining the key parameters of XGBoost is an engineering work. Different application areas and data attributes will have different key parameters configuration. According to the specific conditions of the experimental data and some prior knowledge, the parameters are finally determined as follows through a large number of computer experiments: learning rate (learning\_rate) = 0.01, number of weak learners ( $n_{estimators}$ ) = 1000, maximum depth (max\_depth) = 4, minimum leaf node sample weight (min\_child\_weight) = 5, L2 regularization term weight coefficient (reg\_lambda) = 4. The key parameters will be fine-tuned according to the amount of data. Finally, the XGBoost trained on the training set with adjusted parameters is called XGB model. The flow chart of the model is shown in Fig. 2.

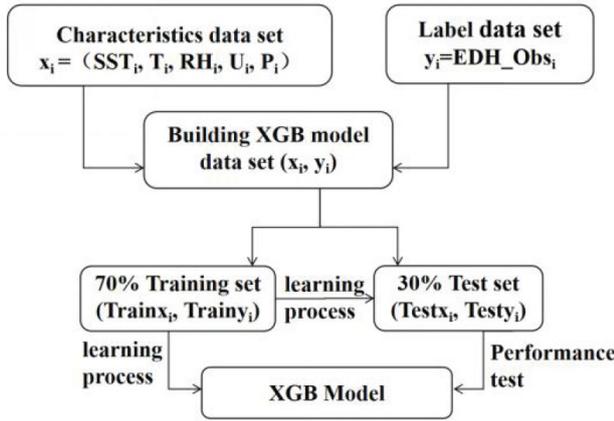


Fig. 2. XGB model construction diagram.

## 4. Experiments and Results Analysis

### 4.1 Introduction to Experiment

(1) Observation experimental areas: mainly in the South China Sea, the equator and the Indian Ocean. Due to the large spatial and temporal locality of evaporation duct, the whole experimental area is further subdivided into five sub-regions and total area (all regions). South China Sea (SCS, including SCS Go, SCS Back), Equator Region (ER), South of Sri Lanka (SSL), North of the Bay of Bengal (NBB) and Total Area are shown in Fig. 3.

Observation experimental data: The hydrometeorological observation elements (including sea surface temperature and atmospheric temperature, humidity, wind speed, pressure at 10 m altitude) and EDH were collected in this experiment. After a series of pretreatments, missing data sample pairs and singular values were eliminated, and

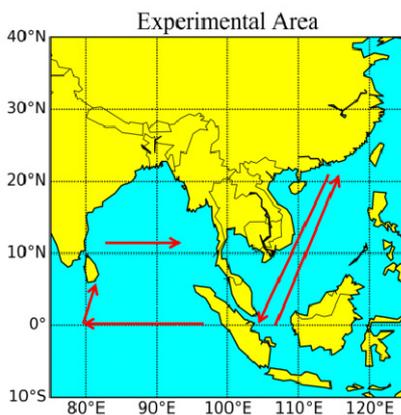


Fig. 3. Experimental area diagram.

finally, about 10,000 sample pairs were included. Considering the spatial and temporal characteristics of the evaporation duct, we shuffled the spatial and temporal order of the observed data of each section of the voyage, make the data attributes uniform, and make the experiment more reasonable and scientific.

(2) Machine learning experimental training set: According to the sub-regions, the training set of XGB model and MLP model is constructed as shown in Tab. 1, as detailed in Sec. 2.2 and Sec. 3.2. (Total area training set and the test set are the sum of the five sub-regionals training sets and the test sets, respectively.)

(3) Experimental performance evaluation indicators: Root Mean Square Error (RMSE) and Square of Correlation Coefficient (SCC) are used to evaluate the performance of models. RMSE represents the accuracy of the prediction model in estimating the height of evaporation duct. The less the value is, the higher the prediction accuracy would be. SCC represents the goodness of fit between the predicted results and the observed data. The numerical range of SCC is usually (0,1). The closer it is to 1, the better the goodness of fit would be with the higher interpreted degree of dependent variables by independent variables.

The expression is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(x_i) - y_i)^2}$$

$$SCC = \frac{\left( n \sum_{i=1}^n f(x_i) y_i - \sum_{i=1}^n f(x_i) \sum_{i=1}^n y_i \right)^2}{\left( n \sum_{i=1}^n f(x_i)^2 - \left( \sum_{i=1}^n f(x_i) \right)^2 \right) \left( n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right)} \quad (10)$$

### 4.2 Comprehensive Performance Evaluation Experiment

In order to evaluate the comprehensive performance of the three evaporation duct prediction models, the predicted results of the three models in the all regions are compared with the measured values. The experimental results are shown in Fig. 4, Fig. 5 and Tab. 2. According to Fig. 4a-4f, XGB model and MLP model are closer to the measured values than PJ model in all regions, and both fit the observation data better than PJ model. The prediction results of XGB model and MLP model are more accurate than PJ model, which indicates that machine learning algorithm can be applied to the optimization research of EDH prediction. Comparing XGB model and MLP model further,

Model	Type of experiment	Sample characteristics $x_i$	Sample label $y_i$
XGB mode	comprehensive performance evaluation	observed data elements 70% data of sub-regions	measured EDH 70% data of sub-regions
MLP model	contrastive experiment of cross-learning	observed data elements 100% data of sub-regions	measured EDH 100% data of sub-regions

Tab. 1. Training set of machine learning prediction models.

it can be found that both models are very close to measured values, and the prediction accuracy and goodness of fit are both good. However, on account of the comparison between the two models it is not intuitive enough, it is impossible to determine who has the better prediction result according to Fig. 4. Therefore, we compare and analyze the bias between the predicted results and the measured values of the two models. According to Fig. 5a–5f, we can see that in all regions, the frequency and the bias of XGB model and MLP model are less than that of PJ model. The prediction accuracy of the two machine learning algorithm models is indeed higher. Comparing XGB model with MLP model further, we can find that the bias and frequency of XGB model is less than that of MLP model in all regions, with the maximum of bias generally less than that of MLP model. It fully shows that the prediction accuracy of XGB model is higher than that of MLP model. From Fig. 5d and 5e, it shows XGB model can achieve ultra-accurate prediction with almost no bias in NBB and SCS Back. In conclusion, compared with PJ model and MLP model, XGB model has higher prediction accuracy and better goodness of fit.

Quantitative comparison results based on RMSE and SCC are shown in Tab. 2. For RMSE firstly: XGB model and MLP model are much less than that of PJ model in all regions with quite low values. This shows that the accuracy of evaporation duct prediction model can be improved by combining machine learning algorithm with measured data. Further comparisons show that the RMSE of XGB model is significantly less than that of MLP model in all regions. In NBB and SCS Back regions, the corresponding RMSE values are as low as 0.32 and 0.35, respectively. The prediction accuracy of XGB model is indeed higher. For SCC secondly: the SCC reaction patterns in all regions were completely consistent with RMSE. In all regions, PJ model has the lowest SCC. XGB model with the highest SCC is obviously higher than MLP model, which means that XGB model has the best goodness of fit.

Considering the accuracy and goodness of fit synthetically, XGB model and MLP model based on machine learning have significantly been improved compared to PJ model, and the prediction results of both models are better. The comprehensive performance of XGB model is the best in all regions with RMSE always at about 0.4 and SCC higher than 0.90. In Total Area, where the learning data is the larger, the prediction accuracy and goodness of fit of XGB model are still excellent whose RMSE is as low as 0.54 and SCC is as high as 0.97. Even in SCS Back with less learning data, XGB model performs equally well, and the corresponding RMSE is as low as 0.35 and SCC is as high as 0.99. It indicates that the XGB model is better than MLP model in learning with stronger, more stable and higher prediction ability. It is little affected by the amount of data and the regional adaptability is better. In conclusion, compared with MLP model, XGB model based on machine learning of XGBoost algorithm has better prediction accuracy and goodness of fit with stable and compre-

hensive performance. In terms of total area, RMSE is reduced by 57.1% and SCC is increased by 11.5%, which shows that machine learning tree boosting algorithm is more suitable for the prediction of EDH than feedforward deep neural network in the field of evaporation duct. Compared with the PJ model, the prediction accuracy and goodness of fit of XGB model improved significantly, in terms of the total area, the corresponding RMSE decreased by 92% and the SCC increased by 110%.

### 4.3 Contrastive Experiment of Cross-learning

In cross-learning experiment, all observation data of one sub-region are used as training set, and all data of other sub-regions are used as test set. That is, learning one sub-region to predict the other sub-regions. The experimental results in previous section show that the XGB model has the best comprehensive performance for EDH prediction in all regions. The cross-learning experiment will base XGB model and MLP model to test the generalization ability of XGB model and explore the spatial correlation of evaporation duct among different regions. The prediction results of other regions based on SCS Go, ER, SSL, NBB and SCS Back (where "Observe" (black solid line) denotes the measured EDH, MLP (yellow dotted line) denotes the predicted results of MLP model, and XGB (red dotted line) denotes the predicted results of XGB model) are shown in Fig. 6–10. RMSE and SCC comparisons represent cross-learning prediction results of XGB model and MLP model in Tab. 3 (where "Go" means SCS Go and "Back" means SCS Back). Based on the analysis of the Fig. 6–10 and Tab. 3, the following conclusions are drawn:

(1) Except for the special case of SCS that Go predicted by ER, the cross-learning prediction accuracy and goodness of fit of XGB model are better than MLP model. That is, the XGB model's comprehensive ability of cross-learning prediction is better than MLP's. In terms of RMSE and SCC, it can be seen that the XGB model is equal to the MLP model only in the special case of learning ER mentioned above. While RMSE of XGB model is absolutely less than that of MLP model in the other cases, accompanied by the larger SCC value. This fully implies that XGB model has stronger generalization ability, better stability and universality than MLP model, which is suitable for the prediction of EDH. For the special case of XGB model (Go predicted by ER), we think, in ER, the upwelling motion of atmospheric convection in the equatorial boundary layer is so stronger. Compared with other regions, the meteorological and oceanic environment of ER is more complex and possesses strong local characteristics of equatorial evaporation duct, which leads to the poor accuracy of model learning ER in predicting EDH in SCS Go.

(2) The best prediction regions of the two models (XGB and MLP) are the same. Learning SCS Go, the best prediction area is SCS Back, followed by SSL. Correspondingly, learning SCS Back, the best prediction area is SCS Go, followed by SSL. For learning SSL, the best prediction

KPI	Model	SCS Go	ER	SSL	NBB	SCS Back	Total Area
RMSE	PJ	4.30	5.80	6.90	7.38	8.23	6.60
	MLP	1.22	1.40	1.00	1.11	1.33	1.26
	XGB	0.65	0.57	0.50	0.32	0.35	0.54
SSC	PJ	0.51	0.40	0.46	0.36	0.51	0.46
	MLP	0.82	0.82	0.87	0.88	0.84	0.87
	XGB	0.94	0.97	0.96	0.99	0.99	0.97

Tab. 2. Comparison of performance indicators for different sub-regions and total area.

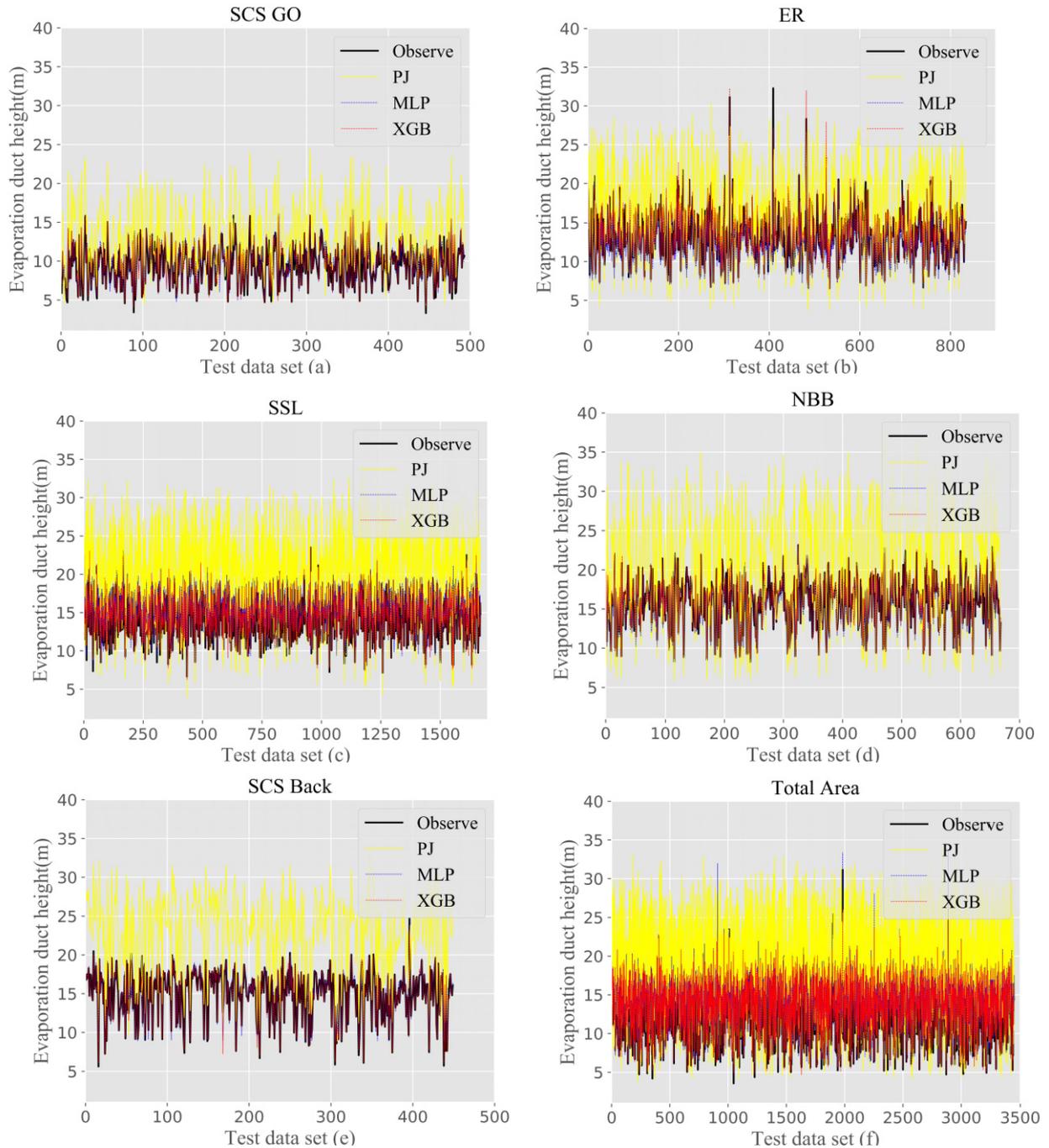
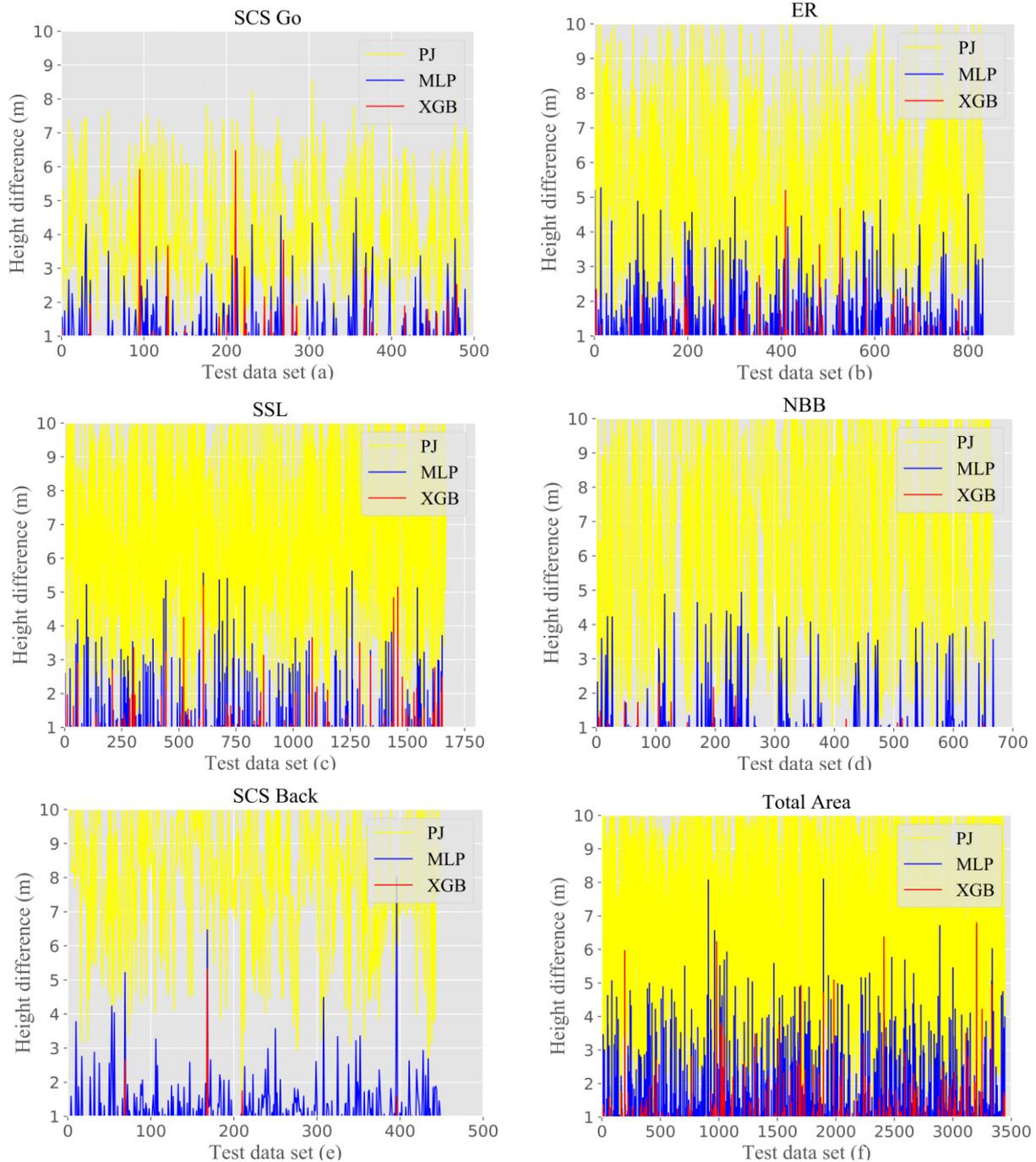


Fig. 4. Comparison of predictive curves for different sub-regions and total area ("Observe" (black solid line) denotes the measured height of evaporation duct, "PJ" (yellow solid line) denotes the predicted result of PJ model, "MLP" (blue dotted line) denotes the predicted result of MLP model, "XGB" (red dotted line) denotes the predicted result of XGB model).



**Fig. 5.** Comparison of the difference between predictive curves for different sub-regions and total area and the measured values ("PJ" (yellow solid line) denotes the difference value of PJ model, "MLP" (blue solid line) denotes the difference value of MLP model, "XGB" (red solid line) denotes difference value of XGB model).

area is SCS Back, followed by SCS Go. The similarity between the meteorology and oceanic environments of SCS and SSL results in similar spatial correlation characteristics and relatively high accuracy of mutual prediction between the two geographically distant regions. Learning ER, the best prediction area is NBB and correspondingly, learning NBB, the best prediction area is ER. ER and NBB are geographically close, with the latitudinal spans of them less than 10 degrees. Thus, their meteorological and ma-

rine environments are comparatively similar, and the relative prediction accuracy is relatively high. The worst prediction region of SCS Go and SCS back both is NBB with the small spatial correlation of evaporation duct, which is due to the large difference of meteorological and marine environment between the South China Sea and the Indian Ocean. The above conclusions show that the method is consistent with the theory and that evaporation duct has strong spatial correlation.

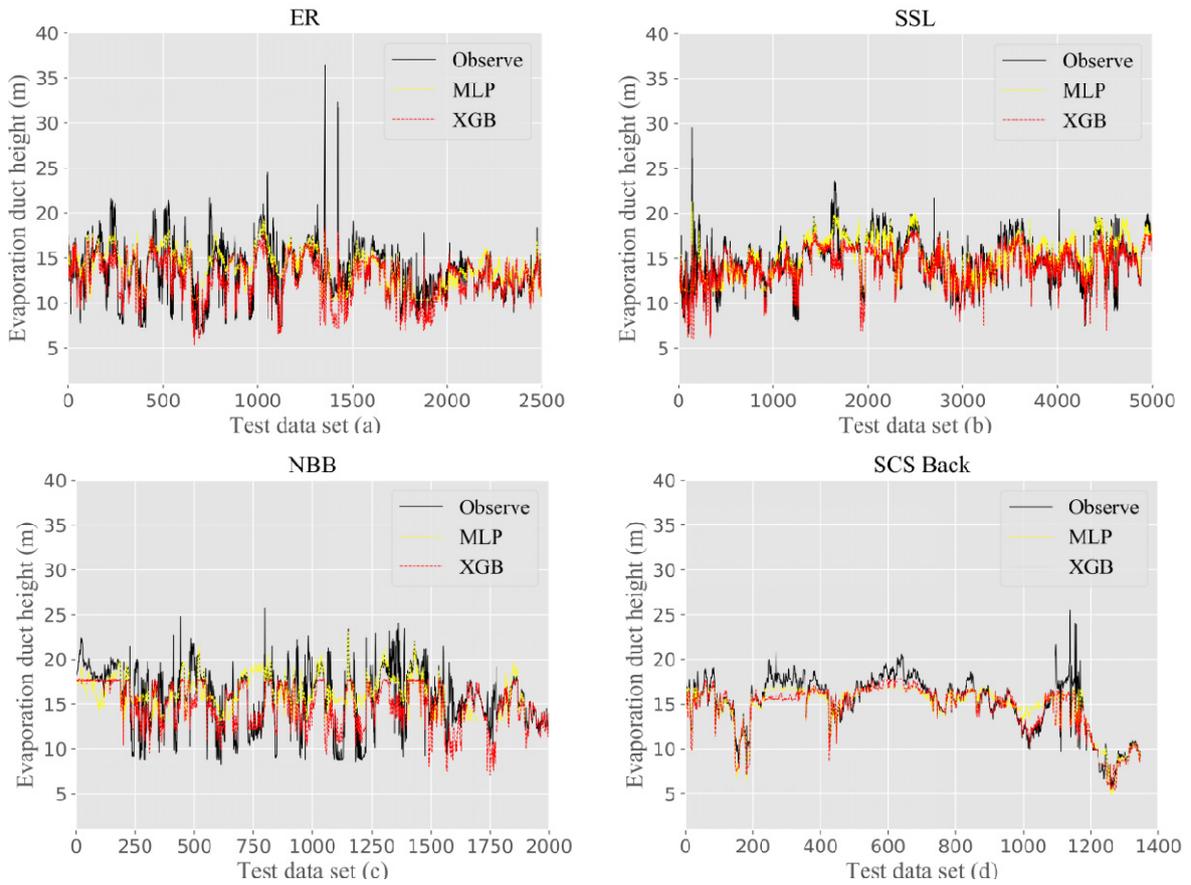


Fig. 6. Prediction of the other areas based on SCS Go data. Prediction of (a) ER, (b) SSL, (c) NBB and (d) SCS Back.

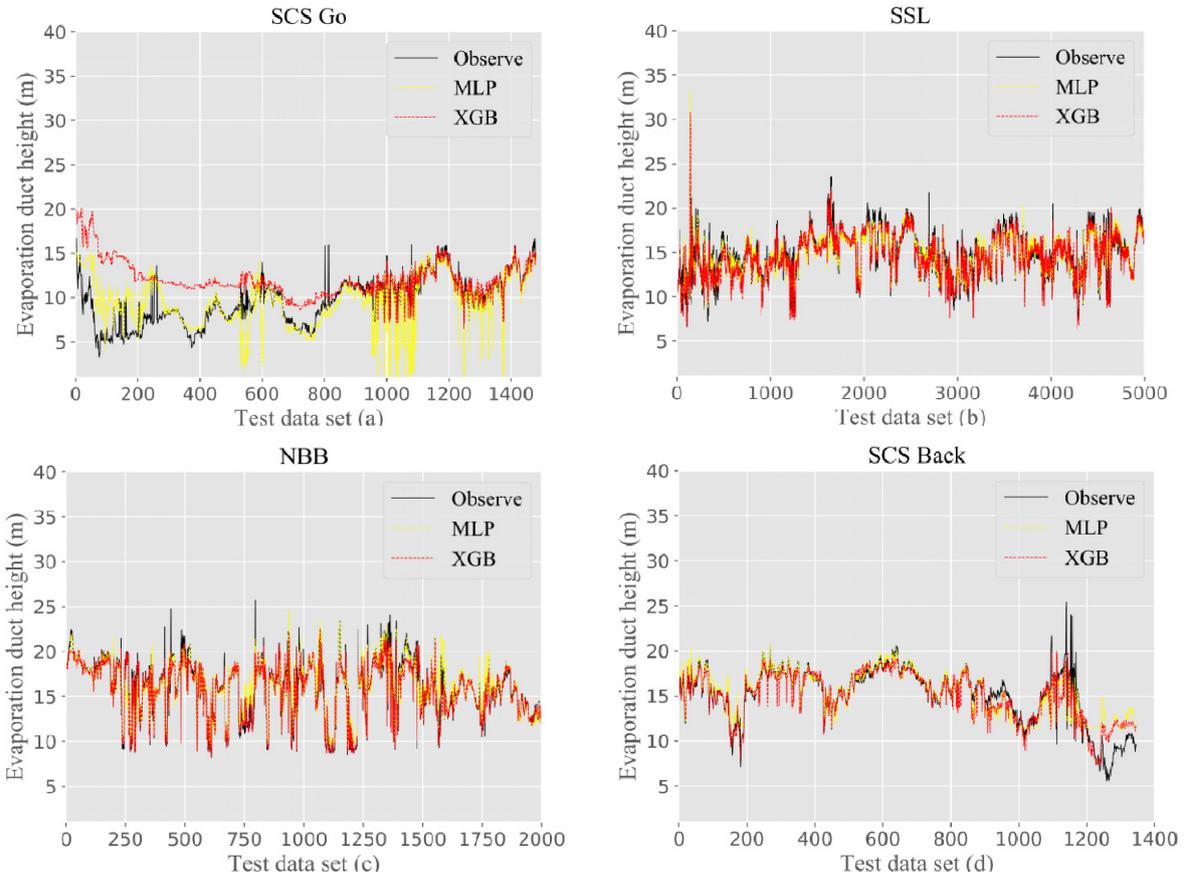


Fig. 7. Prediction of the other areas based on ER data. Prediction of (a) SCS Go, (b) SSL, (c) NBB and (d) SCS Back.

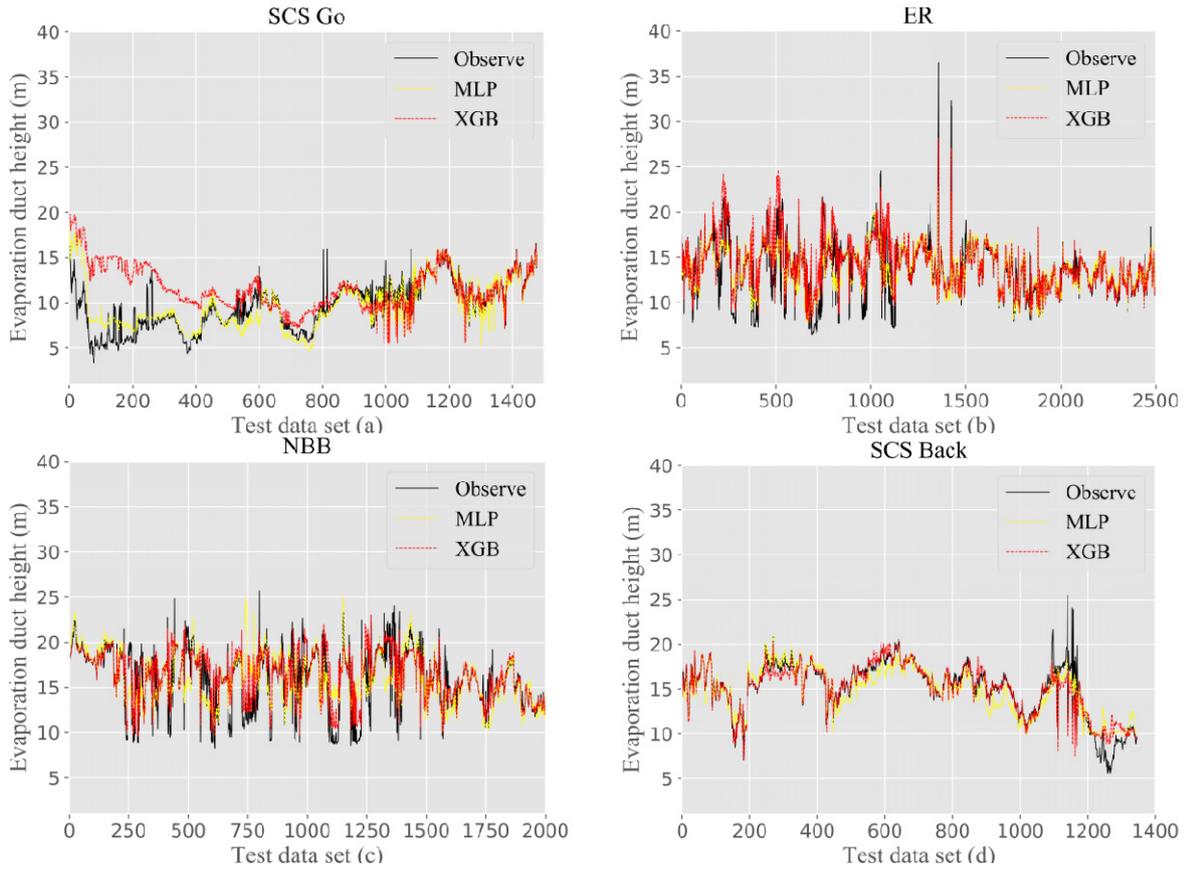


Fig. 8. Prediction of the other areas based on SSL data. Prediction of (a) SCS Go, (b) ER, (c) NBB and (d) SCS Back.

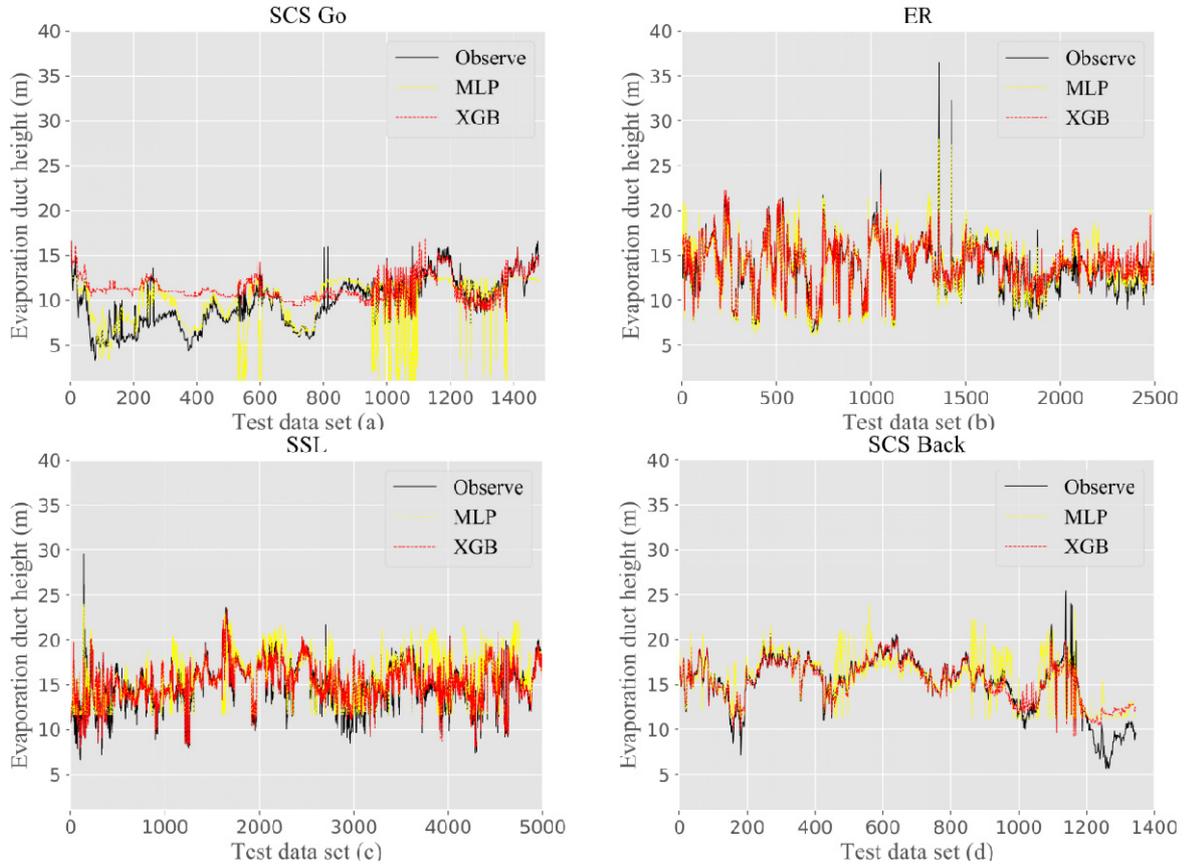


Fig. 9. Prediction of the other areas based on NBB data. Prediction of (a) SCS Go, (b) ER, (c) SSL (d) SCS Back.

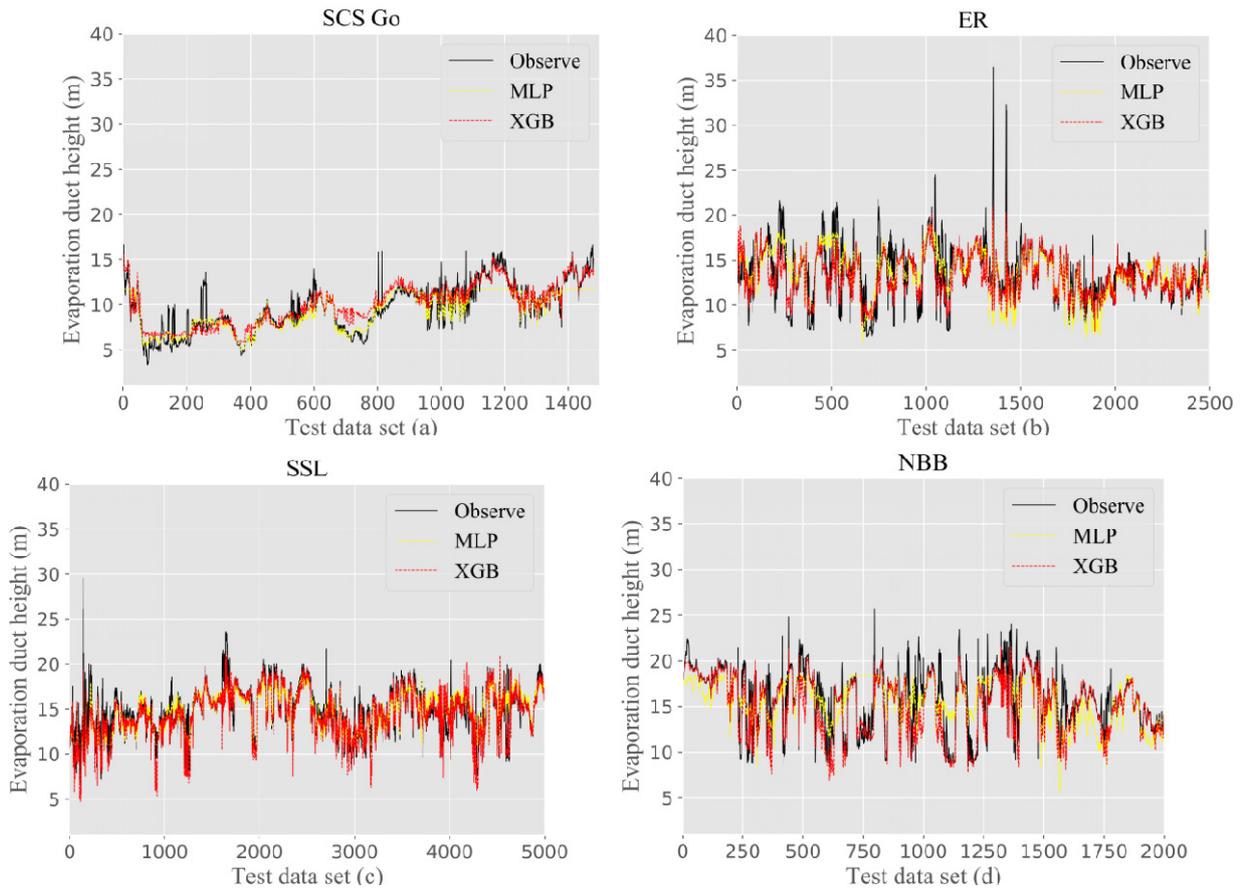


Fig. 10. Prediction of the other areas based on SCS Back data. Prediction of (a) SCS Go, (b) ER, (c) SSL and (d) NBB.

Training area		Test area											
		Go		ER		SSL		NBB		Back		Average value	
		XGB	MLP	XGB	MLP	XGB	MLP	XGB	MLP	XGB	MLP	XGB	MLP
RMSE	Go	-	-	2.12	2.23	1.5	1.65	2.52	2.83	1.40	1.46	1.88	2.04
	ER	3.77	2.55	-	-	0.94	1.33	0.73	1.21	1.36	1.70	1.70	1.70
	SSL	1.59	1.65	1.73	1.79	-	-	1.69	2.63	1.15	1.34	1.54	1.85
	NBB	2.80	2.84	1.28	1.47	1.57	2.55	-	-	1.41	2.28	1.76	2.28
	Back	1.27	1.33	1.93	2.15	1.46	1.48	2.06	2.79	-	-	1.68	1.94
SCC	Go	-	-	0.68	0.64	0.74	0.70	0.63	0.56	0.84	0.82	0.72	0.68
	ER	0.47	0.53	-	-	0.88	0.78	0.99	0.87	0.84	0.77	0.79	0.74
	SSL	0.74	0.72	0.75	0.73	-	-	0.78	0.59	0.88	0.84	0.72	0.72
	NBB	0.56	0.47	0.85	0.80	0.72	0.52	-	-	0.82	0.65	0.74	0.61
	Back	0.78	0.80	0.71	0.65	0.75	0.74	0.72	0.57	-	-	0.74	0.69

Tab. 3. Comparison of RMSE and SCC in cross learning and prediction.

### 5. Conclusions

Due to the limitation of the traditional evaporation duct model, which relies on the similarity theory of the air-sea boundary layer with the large prediction error, we propose the XGBoost evaporation duct prediction model based on the hydrometeorological observation data by using XGBoost algorithm. Through the comprehensive performance evaluation experiments of XGB model, MLP model and traditional PJ model and cross-learning contrast experiments, we draw the following conclusions:

(1) Compared with the traditional PJ model, the comprehensive performance of XGB model and MLP model based on machine learning has been significantly improved, and the prediction results of both XGB model and MLP model are excellent. The prediction accuracy and goodness of fit of XGB model are optimal in all regions with stable comprehensive performance. Compared to MLP model, XGB model reduces RMSE by 57.1% and promotes SCC by 11.5%. It shows the learning and prediction ability of XGBoost algorithm are better than feed-forward deep neural network for EDH prediction. Com-

pared with the traditional PJ model, the performance of XGB model is improved more sharply, RMSE is reduced by 92% and SCC is improved by 110%.

(2) In cross-learning experiment, XGB model has better prediction accuracy and goodness of fit than MLP model except for the special case of SCS Go predicted by ER. XGB model's comprehensive ability of cross-learning prediction is better than MLP's. From the mean values of RMSE and SCC in other sub-regions predicted by one region, it shows the RMSE of XGB model is less than that of MLP model and the SCC value is larger than that of MLP model. It fully shows that XGB model has stronger generalization ability, better stability and universality than MLP model, which is suitable for the prediction of EDH.

Evaporation duct possesses strong characteristics of relating space with physics, fully verified by XGB model and MLP model from the experimental and technical level, which shows the consistency of the research method and theory in this paper. SCS Go and SCS Back, ER and NBB are respectively the best learning prediction regions for each other. Because the geographic location is relatively close and the meteorological and marine environment are comparatively similar, which leads to the spatial correlation of evaporation duct in the two regions and the high accuracy of mutual prediction. The worst prediction region for SCS Go and SCS back is NBB. Similarly, due to the different geographical locations of SCS and the Indian Ocean, the meteorological and marine environment differs greatly, and the spatial correlation of evaporation duct is relatively low. The best prediction area for learning SSL is SCS Back, followed by SCS Go. The secondly best prediction area for both learning SCS Go and SCS Back is SSL. The geographical location of the two areas is far from each other, but because of the similarity of meteorological and marine environment, similar spatial correlation of evaporation duct is also produced.

## Acknowledgments

This work was supported by the National Key Research and Development Program of China (2016YFC1401800), the National Natural Science Foundation of China (Nos.41775027) of Xiaofeng Zhao.

## References

- [1] KANG, S. F., ZHANG, Y. S., WANG, H. G. *Tropospheric Atmospheric Duct*. Zhang Z. 1st ed. Beijing (China): Science Press, 2014. ISBN: 9787030422576 (in Chinese)
- [2] JIANG, B., WANG, H. Q., LI, X., et al. A novel method of target detection based on the sea clutter. *Acta Physica Sinica*, 2006, vol. 55, no. 8, p. 3985–3991. DOI: 10.7498/aps.55.3985 (In Chinese)
- [3] LI, Y. B., ZHANG, Y. G., TANG, H. C., et al. Application of air-sea flux algorithms in diagnosis of evaporation duct at sea. *Journal of Applied Meteorology*, 2009, vol. 20, no. 5, p. 628–633. DOI: 10.11898/1001-7313.20090515 (in Chinese)
- [4] YAO, Z. Y., ZHAO, B. L., LI, W. B., et al. The analysis on characteristics of atmospheric duct and its effect on the propagation of electromagnetic wave. *Journal of Meteorology*, 2000, vol. 58, no. 5, p. 605–616. DOI: 10.11676/qxxb2000.062 (in Chinese)
- [5] GUO, X. M., KANG, S. F., ZHANG, Y. S., et al. Study on characteristics and applicability of evaporation duct model. *Marine Forecast*, 2013, vol. 30, no. 5, p. 75–83. DOI: 10.11737/j.issn.1003-0239.2013.05.012 (in Chinese)
- [6] YANG, K. D., MA, Y. L., SHI, Y. Spatial-temporal distributions of evaporation duct for the West Pacific Ocean. *Acta Physica Sinica*, 2009, vol. 58, no. 10, p. 7339–7350. DOI: 10.7498/aps.58.7339 (in Chinese)
- [7] FAN, J. Y., GUO, S. H., KANG, S. F., et al. Evaporation duct detection based on meteorological grads tower. *Radio Engineering*, 2012, vol. 42, no. 11, p. 32–33+47. DOI: 10.3969/j.issn.1003-3106.2012.11.011 (in Chinese)
- [8] CRAIN, C. M. Survey of airborne microwave refractometer measurements. *Proceedings of the IRE*, 2007, vol. 43, no. 10, p. 1405–1411. DOI: 10.1109/JRPROC.1955.277956
- [9] BIN, T., CHA, H., ZHANG, Y. S., et al. Study on the applicability of evaporation duct Model A in Chinese sea areas. *Chinese Journal of Radio Science*, 2009, vol. 37, no. 5, p. 1100–1103. (in Chinese)
- [10] ZHANG, J. P., ZHANG, Y. S., WU, Z. S., et al. Inversion of regional range-dependent evaporation duct from radar sea clutter. *Acta Physica Sinica*, 2015, vol. 64, no. 12, p. 136–146. DOI: 10.7498/aps.64.124101 (in Chinese)
- [11] SHENG, Z., CHEN, J. Q., XU, R. H. Tracking refractivity from radar clutter using particle filter. *Acta Physica Sinica*, 2012, vol. 61, no. 6, p. 523–528. DOI: 10.7498/aps.61.069301 (in Chinese)
- [12] LOWRY, A. R., ROCKEN, C., SOKOLOVSKIY, S. V., et al. Vertical profiling of atmospheric refractivity from ground-based GPS. *Radio Science*, 2002, vol. 37, no. 3, p. 13-1–13-19. DOI: 10.1029/2000RS002565
- [13] LIU, L. J., XIA, J. M., BAI, W. H., et al. Effect of evaporation duct on effective scattering zone of GNSS sea surface reflection signal. *Chinese Journal of Geophysics*, 2019, vol. 62, no. 2, p. 59 to 67. DOI: 10.6038/cjg2019L0689 (in Chinese)
- [14] LI, S. M., CHEN, Z., QIAO, R., et al. Progress and problems of evaporation duct modes at sea. *Marine Forecast*, 2005, vol. 22, no. z1, p. 128–139. DOI: 10.3969/j.issn.1003-0239.2005.z1.019
- [15] WESELY, M. L. Comments on “Bulk parameterization of air-sea exchanges of heat and water vapor including the molecular constraints at the interface”. *Journal of the Atmospheric Sciences*, 1980, vol. 37, no. 12, p. 2798–2800. DOI: 10.1175/1520-0469(1980)037<2798:COPOAS>2.0.CO;2
- [16] PAULUS, R. A. Practical application of an evaporation duct model. *Radio Science*, 1985, vol. 20, no. 4, p. 887–896. DOI: 10.1029/RS020i004p00887
- [17] COOK, J. A sensitivity study of weather data inaccuracies on evaporation duct height algorithms. *Radio Science*, 1991, vol. 26, no. 3, p. 731–746. DOI: 10.1029/91rs00835
- [18] COOK, J., BURK, S. Potential refractivity as a similarity variable. *Boundary-Layer Meteorology*, 1992, vol. 58, no. 1–2, p. 151 to 159. DOI: 10.1007/BF00120756
- [19] MUSSON-GENON, L., GAUTHIER, S., BRUTH, E. A simple method to determine evaporation duct height in the sea surface boundary layer. *Radio Science*, 1992, vol. 27, no. 5, p. 635–644. DOI: 10.1029/92rs00926

- [20] BABIN, S. M., YOUNG, G. S., CARTON, J. A. A new model of the oceanic evaporation duct. *Journal of Applied Meteorology*, 1997, vol. 36, no. 3, p. 193–204. DOI: 10.1175/1520-0450(1997)036<0193:ANMOTO>2.0.CO;2
- [21] BABIN, S. M., DOCKERY, G. D. LKB-based evaporation duct model comparison with buoy data. *Journal of Applied Meteorology*, 2002, vol. 41, no. 4, p. 434–446. DOI: 10.1175/1520-0450(2002)041<0434:LBEDMC>2.0.CO;2
- [22] FAIRALL, C. W., BRADLEY, E. F., HARE, J. E., et al. Bulk parameterization of air-sea fluxes: updates and verification for the COARE algorithm. *Journal of Climate*, 2003, vol. 16, no. 4, p. 571–591. DOI: 10.1175/1520-0442(2003)016<0571:bpoasf>2.0.co;2
- [23] LIU, C. G., HUANG, J. Y., JIANG, C. Y., et al. Modeling evaporation duct over sea with pseudo-refractivity and similarity theory. *Journal of Electronic Science*, 2001, vol. 29, no. 7, p. 970 to 972. DOI: 10.3321/j.issn:0372-2112.2001.07.030 (in Chinese)
- [24] DING, J. L., FEI, J. F., HUANG, X. G. Development and validation of an evaporation duct model. Part I: Model establishment and sensitivity experiments. *Journal of Meteorological Research*, 2015, vol. 29, no. 3, p. 467–481. DOI: 10.1007/s13351-015-3238-3
- [25] LIU, L. H., LI, Y. B., GAO, Z. Q., et al. Research on evaporation duct prediction model based on non-iterative air-sea flux algorithms. *Journal of Applied Oceanography*, 2017, vol. 4, p. 23 to 35. DOI: 10.3969/J.ISSN.2095-4972.2017.04.003 (in Chinese)
- [26] HITNEY, H. V., RICHTER, J. H., SCHEFER, M. H. Integrated refractive effects prediction system. *US. Naval Engineers Journal*, 1976, vol. 88, no. 2, p. 257–262. DOI: 10.1111/j.1559-3584.1976.tb03831.x
- [27] BAUMGARTNER, G.B., HITNEY, H. V., PAPPERT, R. A. Duct propagation modelling for the integrated-refractive-effects prediction system (Ireps). *Communications Radar*, 1983, vol. 130, no. 7, p. 630–642. DOI: 10.1049/ip-f-1.1983.0096
- [28] LI, J., WANG, H., ZHAO, Z. Statistical method of evaporation duct propagation based on marine meteorological data. *Chinese Journal of Radio Science*, 2013, vol. 28, no.5, p. 891–896. DOI: 10.3969/j.issn.1005-0388.2013.05.015 (in Chinese)
- [29] TIAN, B., YU, S. J., LI, J., et al. Study on the adaptability of PJ model of evaporation duct in subtropical sea area. *Ship Science and Technology*, 2009, vol. 9, p. 96–99. DOI: 10.3404/j.issn.1672—7649.2009.09.017 (in Chinese)
- [30] YAO, J. S., YANG, S. Y. Application of PJ evaporation duct model in coastal sea area. *Fire and Command Control*, 2010, vol. 35, no. 6, p. 121–124. DOI: 10.3969/j.issn.1002-0640.2010.06.034 (in Chinese)
- [31] BREIMAN, L. Stacked regressions. *Machine Learning*, 1996, vol. 24, no. 1, p. 49–64. DOI: 10.1023/a:1018046112532
- [32] PROVOST, F., FAWCETT, T. Robust classification for imprecise environment. *Machine Learning*, 2001, vol. 42, no. 3, p. 203–231. DOI: 10.1023/a:1007601015854
- [33] DOUVENOT, R., FABBRO, V., GERSTOFT, P., et al. A duct mapping method using least squares support vector machines. *Radio Science*, 2008, vol. 43, no. 6, p. 1–12. DOI: 10.1029/2008rs003842
- [34] DOUVENOT, R., FABBRO, V., BOURLIER, C., et al. Retrieve the evaporation duct height by least-squares support vector machine algorithm. *Journal of Applied Remote Sensing*, 2009, vol. 3, no. 1, p. 1–15. DOI: 10.1117/1.3081546
- [35] YANG, C. A comparison of the machine learning algorithm for evaporation duct estimation. *Radioengineering*, 2013, vol. 22, no. 2, p. 657–661. ISSN: 1210-2512
- [36] GEORGANOS, S., GRIPPA, T., VANHUYSSE, S., et al. Very high resolution object-based land use – land cover urban classification using extreme gradient boosting. *IEEE Geoscience and Remote Sensing Letters*, 2018, vol. 15, no. 4, p. 607–611. DOI: 10.1109/LGRS.2018.2803259
- [37] ZHU, X. Y., LI, J. C., ZHU, M., et al. An evaporation duct height prediction method based on deep learning. *IEEE Geoscience and Remote Sensing Letters*, 2018, vol. 15, no. 9, p. 1307–1311. DOI: 10.1109/LGRS.2018.2842235
- [38] ZHU, X. Y., ZHU, M., LI, J. C., et al. An optimization research of evaporation duct prediction models based on a deep learning method. In *IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC 2018)*. Xi'an (China), 2018, p. 1818–1822. DOI: 10.1109/IMCEC.2018.8469552
- [39] CHEN, W., FU, K., ZUO, J., et al. Radar emitter classification for large data set based on weighted-xgboost. *IET Radar, Sonar & Navigation*, 2017, vol. 11, no. 8, p. 1203–1207. DOI: 10.1049/iet-rsn.2016.0632
- [40] NATEKIN, A., KNOLL, A. Gradient boosting machines, a tutorial. *Frontiers in Neurobotics*, 2013, vol. 7, no. 7, p. 1–21. DOI: 10.3389/fnbot.2013.00021
- [41] JESKE, H. State and Limits of Prediction Methods of Radar Wave Propagation Conditions Over Sea. *Modern Topics in Microwave Propagation and Air-Sea Interaction*, 1973, vol. 5, p. 130–148. DOI: 10.1007/978-94-010-2681-9\_13

## About the Authors ...

**Wenpeng ZHAO** is currently a master student of the National University of Defense Technology, China. His research interests include atmospheric duct and numerical weather forecast.

**Jincai LI** (corresponding author) is currently an Associate Research Fellow of the National University of Defense Technology, China. His research interests include atmospheric duct and large scale scientific computing.