



## Article

# A Random Forest Algorithm Combined with Bayesian Optimization for Atmospheric Duct Estimation

Chao Yang <sup>1,\*</sup>, Yulu Wang <sup>1</sup>, Aoxiang Zhang <sup>1</sup>, Hualei Fan <sup>1</sup> and Lixin Guo <sup>2</sup>

<sup>1</sup> School of Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China; ylwang@stu.xupt.edu.cn (Y.W.); 15518307806@stu.xupt.edu.cn (A.Z.); hlfan@stu.xupt.edu.cn (H.F.)

<sup>2</sup> School of Physics, Xidian University, Xi'an 710071, China; lxguo@xidian.edu.cn

\* Correspondence: yangchao@xupt.edu.cn

**Abstract:** Inversion of atmospheric ducts is of great importance in the field of performance evaluation for radar and communication systems. Since the model parameters in machine learning play a crucial role in prediction performance, this paper develops a random forest (RF) model integrated with Bayesian optimization (BO) called BO-RF for atmospheric duct prediction, and the BO is adopted to determine appropriate model parameters during the training process. In addition, the K-fold cross-validation (CV) method is also incorporated into the model to obtain the best model partition and overcome the overfitting problem. To test the performance of the proposed model, the results obtained by the BO-RF are compared with other commonly used methods, such as classical RF, extreme gradient boosting (XGBoost) with/without BO, and K-nearest neighbor (KNN) with/without BO. Comparisons demonstrate that BO-RF has the best accuracy and anti-noise ability for the estimation of duct parameters.

**Keywords:** atmospheric duct estimation; random forest; radar sea clutter; parabolic equation; orthogonal experiment design



**Citation:** Yang, C.; Wang, Y.; Zhang, A.; Fan, H.; Guo, L. A Random Forest Algorithm Combined with Bayesian Optimization for Atmospheric Duct Estimation. *Remote Sens.* **2023**, *15*, 4296. <https://doi.org/10.3390/rs15174296>

Academic Editor: Weimin Huang

Received: 15 June 2023

Revised: 15 July 2023

Accepted: 26 July 2023

Published: 31 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

An atmospheric duct is an anomalous electromagnetic propagation phenomenon that commonly appears in the maritime environment and can have a significant impact on radar and communication systems, such as communication blind areas, increased radar sea surface clutter, etc. Therefore, accurate analysis of radar echoes and efficient estimation of atmospheric duct parameters have become active research fields. Atmospheric ducts generally consist of three types, including evaporation ducts, surface-based ducts (SBDs), and elevated ducts. Evaporation ducts and SBDs are two common types in maritime environments due to the rapid changes in vertical atmospheric and humidity profiles with altitude above the sea surface [1]. It should be noted that SBDs exhibit a lower probability of occurrence than evaporation ducts, but their influence on radar and communication systems is much greater than that of evaporation ducts.

Estimation of atmospheric duct refractivity parameters with the refractivity from clutter (RFC) technique has attracted considerable interest in the field of optimization theory. Numerous methods have been proposed to retrieve atmospheric refractivity parameters. Gerstoft et al. [2,3] published simulation results for an eleven-parameter refractivity profile with five parameters in the vertical direction and six parameters in the horizontal direction, solved using a genetic algorithm, and presented the key steps and the latest research progress regarding the RFC technique. Douvenot et al. [4] first employed the least squares support vector, which belongs to a class of machine learning based on the pregenerated radar sea clutter power (RSCP) database produced by the parabolic wave equation (PWE), to estimate SBD. Wang et al. [5] inverted the evaporation duct using particle swarm optimization with RFC. Zhang et al. [6] proposed a four-parameter evaporation duct model

and retrieved refractivity parameters from it. Zhao et al. [7] applied classical simulated annealing to the remote sensing of an atmospheric duct. Yang et al. [8–11] performed the inversion of an atmospheric duct with standard and improved swarm intelligence optimization and presented a comparison of the machine learning algorithm for evaporation duct estimation. Tepecik et al. [12] introduced a novel hybrid model based on an artificial neural network and a genetic algorithm for the inversion problem of refractivity estimation. Lentini et al. [13] calculated the global sensitivity of radar wave propagation to environmental parameters at the maritime atmospheric boundary layer. Penton et al. [14] studied the influence of rough ocean surfaces on evaporation duct atmospheric refractivity inversion with a genetic algorithm. Pozderac et al. [15] adopted a novel transmit–receive array system named the X-band beacon-receiver array to determine the maritime evaporation duct. Yan et al. [16] proposed an estimation method for evaporation ducts using a neural network. Zhu et al. [17] predicted evaporation duct height based on a multilayer perceptron network. Sit et al. [18,19] showed that deep learning and artificial neural networks can be efficiently used to classify and predict atmospheric ducts. It is of great significance to obtain instantaneous environmental information for assessing the performance of radar and communication systems. It can be distinctly seen from the above study that more and more attention has been paid to the field of monitoring the atmospheric environment in the maritime boundary layer with machine learning due to its superiority in real-time compared with swarm intelligence optimization. Random forest (RF) [20,21] is a powerful machine learning algorithm inspired by the ensemble learning method that works by incorporating many decision trees into one algorithm model to enhance its performance. However, the hyper-parameter in RF that directly relates to the performance of machine learning is difficult to determine. In this paper, a combination of RF with Bayesian optimization (BO) is presented to predict atmospheric duct parameters, where BO is adopted to search for the most appropriate hyper-parameter during the training process. To test the performance of the proposed BO-RF model, the RSCP versus propagation distance is numerically simulated by the PWE method with the combinations of refractivity parameters generated with an orthogonal experiment design (OED) method; the results obtained by the BO-RF are compared with the commonly used methods, such as classical RF, extreme gradient boosting (XGBoost) [22] with/without BO, and K-nearest neighbor (KNN) [23] with/without BO. Comparisons demonstrate that BO-RF has the best accuracy and noise immunity for the estimation of atmospheric duct parameters.

## 2. Forward Model

The forward model commonly includes two parts: the refractivity model and the radar propagation model. In the subsequent section, we will introduce the basic concept in the field of atmospheric ducts.

### 2.1. Refractivity Model

As the radio wave travels in the electromagnetic duct, the propagation features of the radio wave are controlled by the index of refraction:

$$n = c/v, \quad (1)$$

where  $c$  is the speed of light in a vacuum and  $v$  is the speed of light in the medium. The value of  $n$  is close to unity, typically 1.00035; however, even a tiny variation in the refractive index may cause a huge change in the propagation features. Hence, a useful dimensionless physical variable  $N$ , which is the part-per-million change in the index of refraction, named refractivity, is introduced:

$$N = (n - 1) \times 10^6. \quad (2)$$

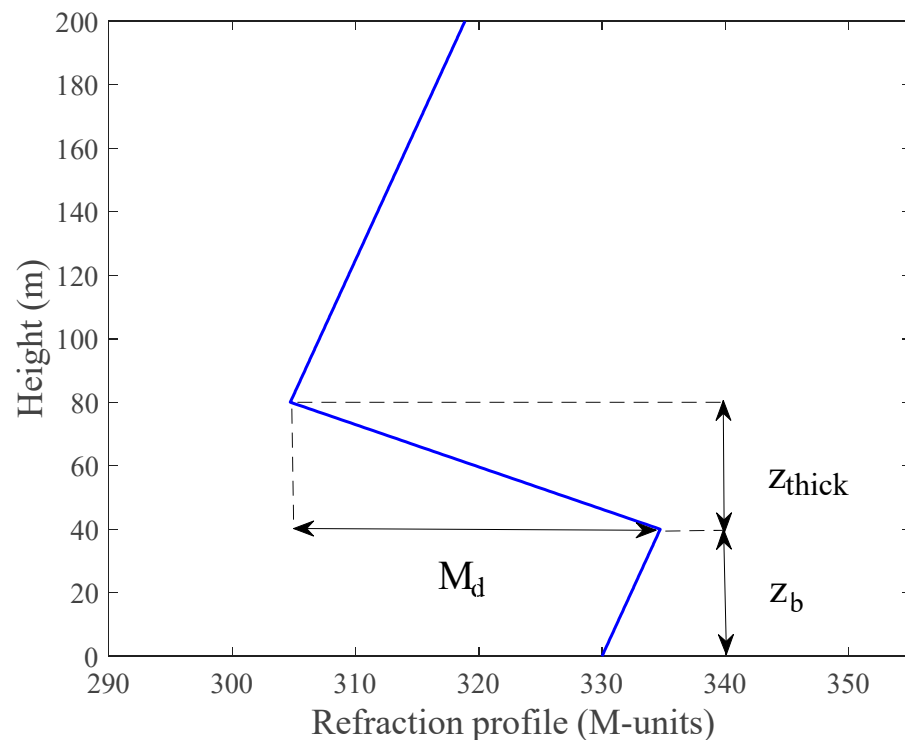
To further account for the curvature of the earth, another variable  $M$  called modified refractivity is provided by

$$M(z) = N + \frac{z}{R_e} \times 10^6 = N + 0.157z, \quad (3)$$

where  $R_e$  is the radius of the earth and  $z$  is the height above the earth's surface. As  $dM/dz < 0$ , an electromagnetic duct that can confine radio wave propagation and act like a waveguide is formed [24,25]. In this paper, our emphasis is placed on the three parameters of SBD shown in Figure 1, according to [26].

$$M(z) = M(0) + \begin{cases} c_1 z & z \leq z_b \\ c_1 z_b - M_d \frac{z - z_b}{z_{thick}} & z_b < z < z_{thick} \\ c_1 z_b - M_d + c_2 (z - z_b - z_{thick}) & z \geq z_{thick} \end{cases} \quad (4)$$

where  $M(0)$  is the modified refractivity at the mean sea level, typically taken as 330 M-units,  $z$  describes the height from the mean sea level,  $z_b$  and  $z_{thick}$  represent the thicknesses of the base layer and trapping layer,  $c_1$  and  $c_2$  denote the slopes of the base layer and top layer, and  $M_d$  is the duct strength. Obviously, there are three parameters in the SBD refractivity profile model, and we will employ the parametrization refractivity model of SBD via  $\mathbf{M} = (z_b, z_{thick}, M_d)$ . The tri-linear refractivity profile for SBD is shown in Figure 1.



**Figure 1.** Tri-linear refractivity profile for SBD.

## 2.2. PWE Method

PWE can be divided into a narrow-angle and a wide-angle PWE. The propagation elevation of the narrow-angle PWE has good accuracy when it does not exceed  $15^\circ$ . An atmospheric duct usually occurs when the trapping angle does not exceed  $1^\circ$ . Therefore, the narrow-angle PWE is selected for studying the radio wave for long-range propagation in atmospheric ducts [27]. In order to analyze the two-dimensional tropospheric propagation

problem, the scalar wave equation taking the index of refraction  $n$  into account is provided by [28]:

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial z^2} + k_0^2 n^2 \varphi = 0, \quad (5)$$

where  $\varphi(x, z)$  represents the electric or magnetic field component,  $k_0$  is the wavenumber in a vacuum, and  $x$  and  $z$  are the range and height, respectively. It is a good approximation if the index of refraction varies slowly with the propagation direction. With the help of the reduced function  $\varphi(x, z) = u(x, z)e^{-ik_0 x}$  and the paraxial approximation, the PWE is expressed by

$$\frac{\partial^2 u}{\partial z^2} + 2ik_0 \frac{\partial u}{\partial x} + k_0^2 (n^2 - 1)u = 0. \quad (6)$$

Then, the split-step Fourier-transform solution to the PWE can be written as [29]

$$u(x_0 + \Delta x, z) = e^{ik_0 M(z)10^{-6}\Delta x} \mathcal{F}^{-1} \left\{ e^{-ip^2 \Delta x / 2k_0} \mathcal{F} \{ u(x_0, z) \} \right\}, \quad (7)$$

where  $p$  is the transform variable, and  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  denote the forward Fourier transform and inverse Fourier transform, respectively. If the initial field  $u(x_0, z)$  is provided, the field distribution along the vertical direction at each range interval can be easily marched in a horizontal range.

Once the field distribution of the entire computational region is obtained with Equation (7), the RSCP can be easily expressed in dB at range  $x$  by radar theory [2],

$$P_c(x, \mathbf{M}) = -2L(x, \mathbf{M}) + 10 \log(x) + \sigma^\circ + C_0, \quad (8)$$

where  $\mathbf{M}$  is the parameterized vector of the atmospheric environment,  $\sigma^\circ$  is the normalized scattering coefficient from the sea surface,  $C_0$  is a constant that contains the parameters of the radar system, and  $L(x, \mathbf{M})$  is the propagation loss in the atmospheric duct environment that can be conveniently computed from the field distribution.

### 3. Optimization Algorithm

#### 3.1. Principle of RF Regression

Decision tree is the primary component of RF and belongs to a class of machine learning that generally includes three main steps: feature selection, decision tree generation, and pruning of the decision tree. RF commonly adopts the decision tree that has not completed the pruning optimization operation. Each decision tree makes a prediction for the sample, and the final prediction can be achieved by taking the average of the combined predictions of all decision trees [30]. It has strong robustness and flexibility in dealing with regression problems. However, it has the shortcomings of falling into the local optimum, leading to high complexity and overfitting.

RF is an integrated learning algorithm using decision tree modeling, consisting of multiple independent decision trees. Different decision trees are constructed from random samples, but there is no relationship between them, and the output of the model is determined by each decision tree. In the process of training decision trees, randomly selected features further improve the model's accuracy and avoid insufficient generalization ability. The detailed steps are given as follows:

1. A bootstrap method is adopted to randomly selected  $M$  sample points from the original sample dataset  $S$  to construct a decision tree and its training subset.
2. The training subset is used to train each decision tree. Assume that the number of input features to the dataset is  $N$ . The split rule for each node is to randomly select  $k'$  features from the  $N$  features as alternative branch features, and then select the best split node from the  $k'$  features to divide the left and right subspaces. Each split node performs a binary test on each subset, and the test result is sent to the left or right sub-node. The test randomly chooses a subset of features and finds a value with the



lowest mean square error  $\kappa$  to group and determine its optimal branch, and it can be expressed as

$$\kappa = \frac{1}{D_1} \sum_{i \in D_1} (E_i - C_1)^2 + \frac{1}{D_2} \sum_{i \in D_2} (E_i - C_2)^2, \quad (9)$$

where  $E_i$  is the true value of the  $i$ th sample, and  $C_1$  and  $C_2$  are the sample predictions for the left and right subspaces  $D_1$  and  $D_2$ , respectively.

3. Multiple decision trees can be generated via the first two steps above. The final prediction of each decision tree depends on the average of the leaf nodes where the sample points are located. The accuracy of RF is further improved by optimizing parameters such as the number of decision trees, the maximum depth of the decision tree, the minimum number of samples required for node division, the minimum sample number of the leaf nodes, and the number of features in the tree.
4. Finally, a RF model is produced by taking the average of the multiple decision trees as follows:

$$f(x) = \frac{1}{k'} \sum_{i=1}^{k'} E_i(x), \quad (10)$$

where  $f(x)$  denotes the combined regression model.

### 3.2. BO Algorithm

BO is a novel and efficient hyper-parameter tuning algorithm [31,32]. In BO, the objective function distribution is first evaluated by Bayes' theorem, and the new hyper-parameter for the algorithm are then selected according to the optimal sampling point distribution of the objective function. In addition, the update of the hyper-parameter for the next time depends on the objective function value of the optimal sample point. BO is mainly composed of the probabilistic proxy model and the acquisition function. In the following, we will briefly introduce the two parts.

#### 3.2.1. Probabilistic Proxy Model

The updated formula in the probabilistic proxy model is given by

$$p(f|D_{1:t}) = \frac{p(D_{1:t}|f)p(f)}{p(D_{1:t})}, \quad (11)$$

where  $f$  represents the objective function,  $D_{1:t} = \{(x_1, f_1), (x_2, f_2), L(x_t, f_t)\}$  denotes the collection of sampling point,  $p(D_{1:t}|f)$  is the likelihood distribution of observation set,  $p(f)$  is the prior probability distribution of  $f$ , and  $p(f|D_{1:t})$  represents the posterior probability distribution of  $f$  and is used to modify and test the unknown target function.

The probabilistic proxy model can be divided into parametric and non-parametric models. The Gaussian process is a commonly adopted non-parametric model that has a strong ability to fit functions and is widely used in the field of regression.

A Gaussian process is entirely defined by the mean function  $m(x)$  and the covariance function  $K(x, x')$  as

$$f(x) \sim GP(m(x), K(x, x')), \quad (12)$$

where the mean function is a vector and the covariance function is a matrix.

If we assume that there is a collection of sample points  $D = \{(x_{1:t}, y_{1:t})\}$ , its covariance matrix is given by

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_t) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_t) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_t, x_1) & k(x_t, x_2) & \cdots & k(x_t, x_t) \end{bmatrix}. \quad (13)$$

When a set of sample  $(x_{t+1}, f_{t+1})$  is added to the collection of historical evaluation points, supposing that  $k = [k(x_{t+1}, x_1) \quad k(x_{t+1}, x_2) \quad \dots \quad k(x_{t+1}, x_t)]$ , the corresponding covariance matrix can be updated by

$$K = \begin{bmatrix} K & k^T \\ k & k(x_{t+1}, x_{t+1}) \end{bmatrix}. \quad (14)$$

Using the updated covariance matrix, the posterior probability distribution of  $f_{t+1}$  is computed from the first  $t$  samples

$$\begin{cases} p(f_{t+1}|D_{1:t}, x_{t+1}) \sim N(\mu, \delta^2) \\ \mu = k^T K^{-1} f_{1:t} \\ \delta^2 = k(x_{t+1}, x_{t+1}) - k^T K^{-1} k \end{cases}. \quad (15)$$

### 3.2.2. Acquisition Function

In the Gaussian process, the acquisition function is adopted to balance the ability of exploration and exploitation and search for the optimal objective function value for the next iteration to obtain the optimal hyper-parameter. Currently, the probability of improvement (PI), expected improvement (EI), and upper confidence bound (UCB) are widely used for acquisition functions. In this paper, PI is selected as the acquisition function, and it can be expressed as follows:

$$\partial_t(x; D_{1:t}) = p(f(x) \leq v^* - \xi) = \phi\left(\frac{v^* - \xi - \mu_t(x)}{\delta_t(x)}\right), \quad (16)$$

where  $v^*$  represents the current optimal function value,  $\phi(\cdot)$  is the cumulative density function of the standard normal distribution, and  $\xi$  is the balance parameter. The problem of local optimum can be overcome to a certain extent and the optimal value can be globally searched by adjusting the parameter  $\xi$ .

## 4. A Hybrid BO-RF Model

### 4.1. Orthogonal Experimental Design Method

In order to train the RF model, an OED method is used to generate the combinations of refractivity parameters. It is well known that OED is a robust way to effectively handle multi-dimension and multi-factor problems because it can select a few representative combinations in the experimental database to obtain the optimal result. The key in OED is the orthogonal array (OA) that can reflect the characteristics of each dimension. Thus, we first need to construct an appropriate OA. Here, we use  $L_M(Q^N)$  to represent an OA with  $N$  factors and  $Q$  levels per factor, where  $M$  is the number of combinations of levels. Then, we take a three-factor problem as an example by setting  $Q = 3$  and  $N = 3$  to obtain an orthogonal table  $L_9(3^3)$  [33–35]:

$$L_9(3^3) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 3 \\ 2 & 1 & 2 \\ 2 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 3 \\ 3 & 2 & 1 \\ 3 & 3 & 2 \end{bmatrix}. \quad (17)$$

In  $L_9(3^3)$ , there are three columns, showing that it is suitable for the problem with three parameters at most, and nine rows, demonstrating that one can do nine experiments for a combination with different levels. Each row in (17) represents an experiment with

different levels. Normally, we should carry out  $3^3 = 27$  combinations of experiments for a three-factor and three-levels-per-factor problem. However, this task can be reduced to  $3^2 = 9$  combinations of experiments with the help of OA.

An orthogonal expansion is made based on OED. Assuming the lower and upper boundaries of a  $D$ -dimension parameter are  $u = (u_1, u_2, \dots, u_D)$  and  $t = (t_1, t_2, \dots, t_D)$ , the range of the corresponding cross-solutions are determined by

$$low = (\min(u_1, t_1), \min(u_2, t_2), \dots, \min(u_D, t_D)), \quad (18)$$

$$up = (\max(u_1, t_1), \max(u_2, t_2), \dots, \max(u_D, t_D)). \quad (19)$$

In the cross-solution, the range of the  $i$ th dimension variable  $x_i$  is  $[low_i, up_i] = [\min(u_i, t_i), \max(u_i, t_i)]$ . The  $i$ th dimension variable  $x_i$  can be quantized into  $Q$  levels according to the following formula:

$$l_{i,j} = \begin{cases} low_i & j = 1 \\ low_i + \frac{j-1}{Q-1}(up_i - low_i) & 2 \leq j \leq Q-1 \\ up_i & j = Q \end{cases} \quad (20)$$

Then, we use  $L_M(Q^N)$  to construct  $M$  feasible combinations. More detailed information can be found in [33–35].

The inversion of SBD is a three-parameter problem, and the lower and upper boundaries are given by  $t = (0, 20, 20)$  and  $u = (100, 100, 100)$ , respectively [4]. Let us first construct an appropriate OA  $L_M(Q^N)$  with  $Q = 100$ ,  $N = 3$ , and  $M = 10,000$  for the purpose of generating 10,000 training datasets [4].

#### 4.2. Range of the Hyper-Parameter in RF

The hyper-parameter in RF mainly includes the number of decision trees, maximum depth, minimum samples of split, minimum samples of leaf, and maximum features in the tree. The range of hyper-parameter is shown in Table 1.

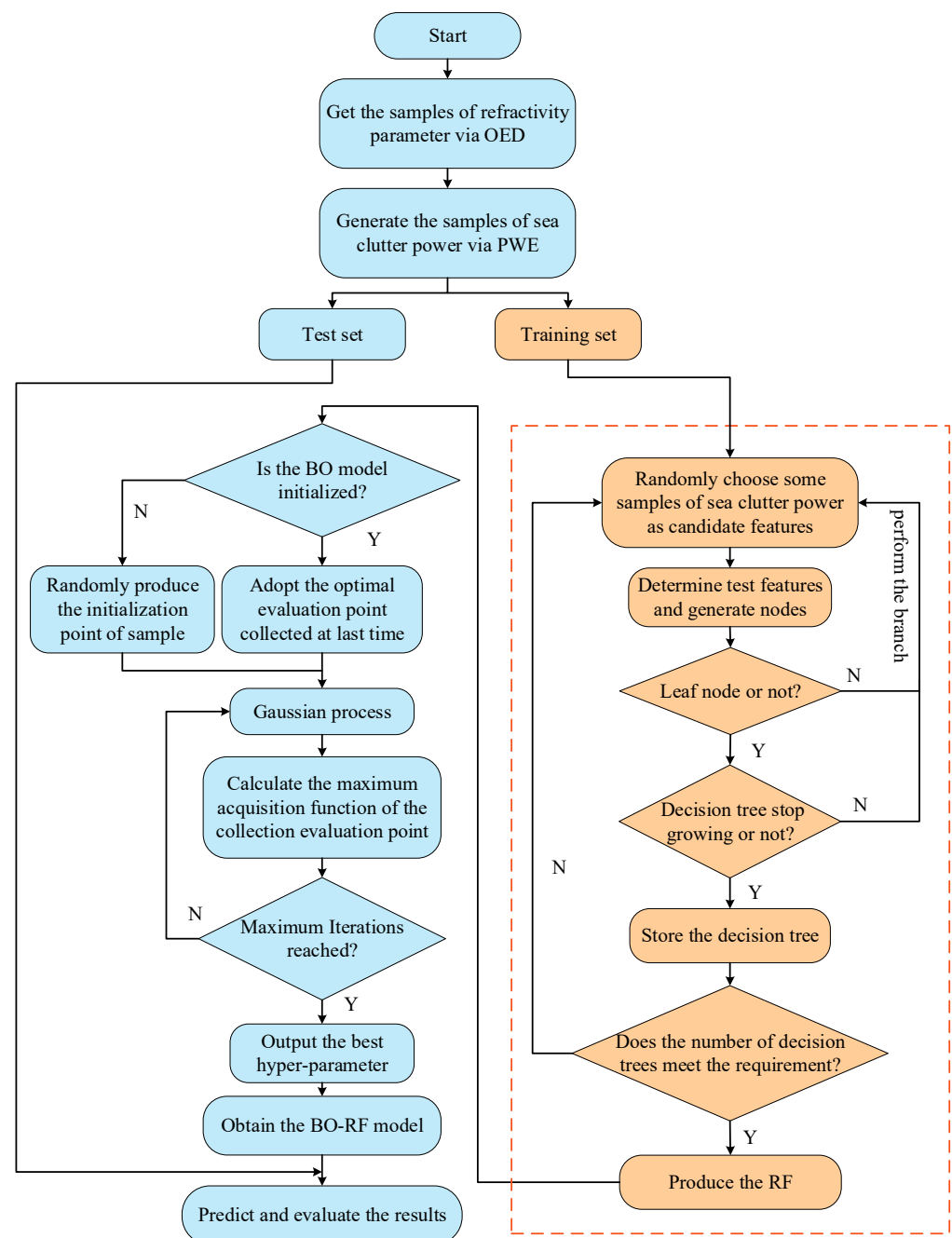
**Table 1.** Range of hyper-parameter in RF.

Description	Hyper-Parameter	Search Range
number of decision trees	n_estimator	(10, 500)
maximum depth of trees	max_depth	(5, 50)
minimum samples of split	min_sample_split	(1, 50)
minimum samples of leaf	min_sample_leaf	(1, 50)
maximum number of features	max_features	(4, 100)

#### 4.3. BO-RF Model

The core of the RF model lies in the selection of the hyper-parameter. In this paper, a BO-RF is adopted to predict the refractivity parameter of an atmospheric duct, and the BO is incorporated into the RF to find the best hyper-parameter. In addition, we apply K-fold cross-validation (CV) technology [36] to the BO-RF, which is commonly used to evaluate the training effect and avoid overfitting phenomenon. The CV makes full use of limited data, and the evaluation results may be as close as possible to the performance of the model on the test set, which can be used as an indicator for model optimization.

In BO-RF, the input and output datasets are first used to train the RF model and optimize the RF parameters. Then, the probabilistic proxy model for the estimation of atmospheric duct is constructed and optimized to obtain the best hyper-parameter in RF. Up to this point, the BO-RF model has been successfully established. The flowchart of the BO-RF is shown in Figure 2 for atmospheric duct estimation. The corresponding optimization process is given as follows:



**Figure 2.** Flowchart of the BO-RF for atmospheric duct estimation.

1. An OED method is used to generate the combinations of refractivity parameters.
2. The split-step Fourier-transform solution to the PWE is adopted to generate the RSCP samples according to the combinations of refractivity parameters.
3. A bootstrap method is used to produce decision trees from the training set and then form a random forest.
4. A group of hyper-parameter is randomly selected within the scope of the parameter to produce the corresponding initialization point of the sample, then substituted into the RF model to obtain the corresponding objective function  $f$  and the initial sample set  $D$ .
5. The probabilistic proxy model in BO is applied to fit  $(x, f)$  and find the most possible evaluation point  $x_t$  in the sample, which can be used to obtain the optimal acquisition function. Furthermore, the hyper-parameter is then assigned to RF model to find

the objective function value  $f_t$  of the optimal sample point and use it as the basis for selecting the hyper-parameter next time.

6. A new set of sampling points  $(x_t, f_t)$  is added to the historical sample set  $D_{t-1}$ , and the Gaussian process is adjusted to optimize the objective function in it.
7. The BO stops updating as soon as it reaches the maximum number of iterations. That is to say, the best hyper-parameter, the optimal value of the objective function, and the corresponding BO-RF model are determined.
8. The test set is then used to examine the BO-RF model.
9. The results are evaluated and analyzed.

## 5. Results and Discussion

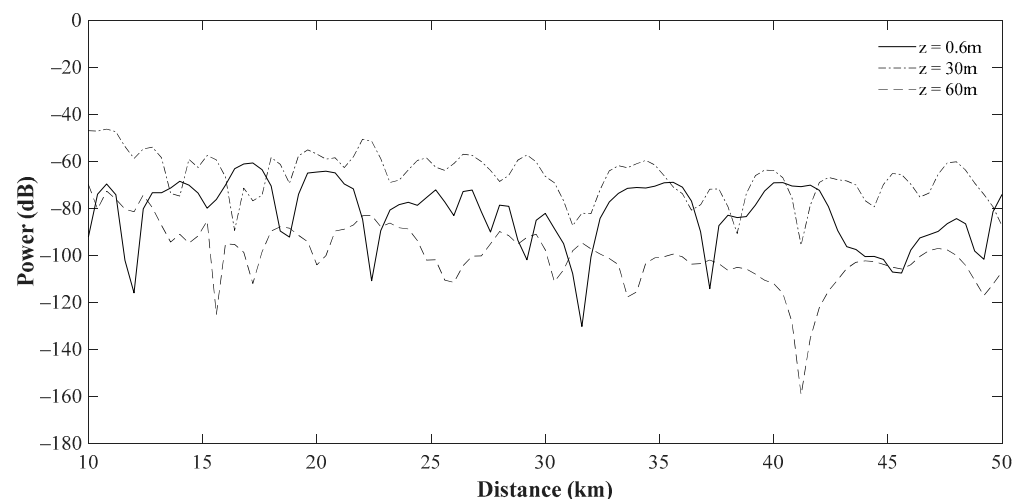
### 5.1. Simulations of the RSCP in SBD

In this section, the RSCP versus propagation distance are numerically simulated by Equation (8) via the PWE method. In order to estimate the refractivity profile with the RFC technique, the RSCP should be sensitive to the refractivity parameters of SBD. In the simulations, the radar system parameters are given in Table 2, and the parameterized vector of refractivity profile for SBD is  $M = (30, 42, 71)$ .

**Table 2.** Radar system parameters.

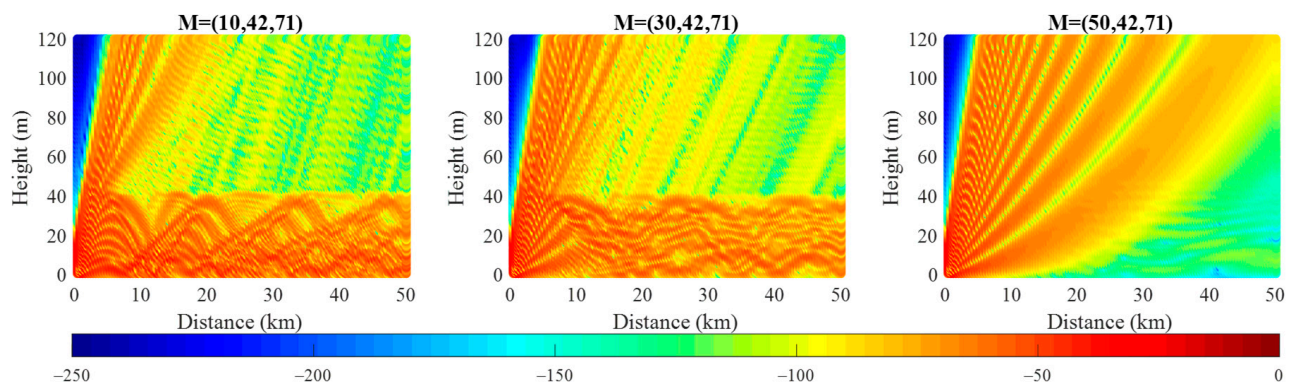
Parameters	Value
Frequency	2.84 GHz
Transmitting power	91.4 dBm
Beamwidth	0.4°
Antenna height	30.78 m
Transmitting antenna gain	52.8 dB
Polarization	VV

Firstly, the curve of RSCP versus the propagation distance at different receive heights are shown in Figure 3. It can be clearly seen that the RSCP in SBD is distinctly sensitive to the receive height.

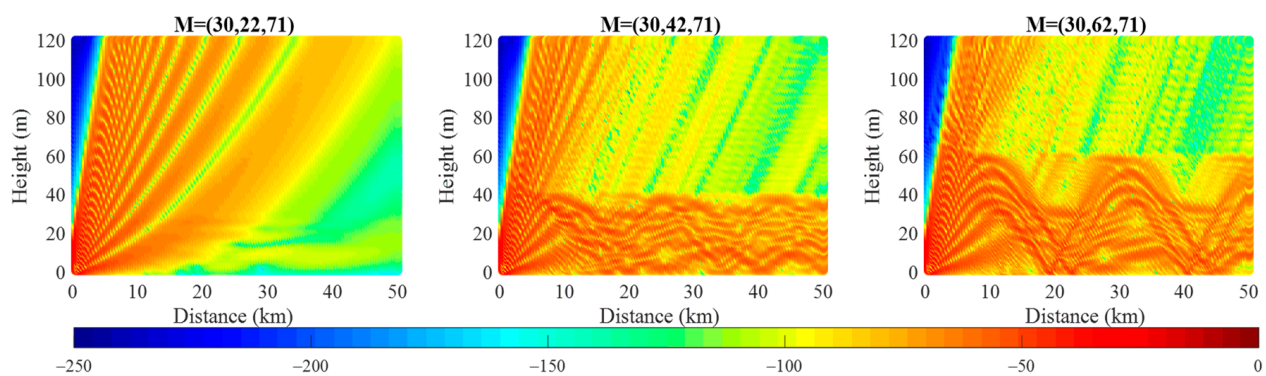


**Figure 3.** Curve of the RSCP versus the propagation distance at different receiving heights.

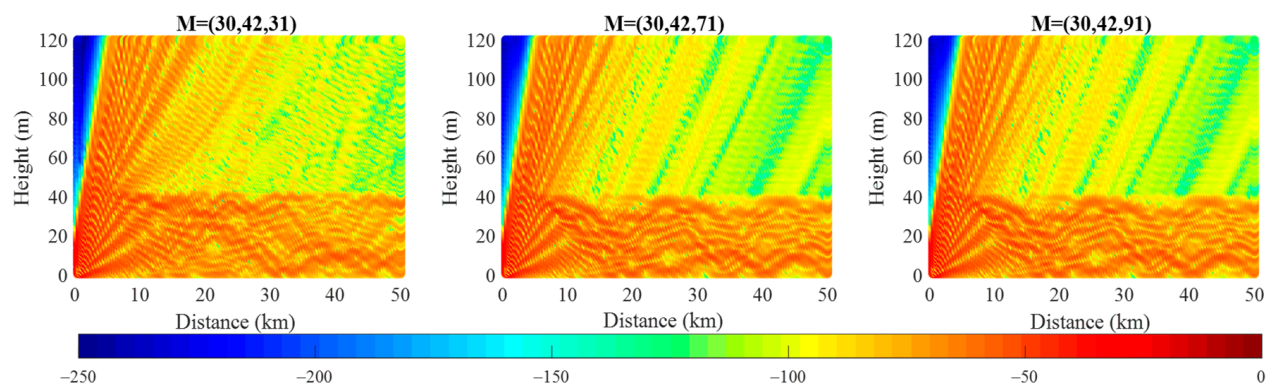
Furthermore, the coverage diagrams are presented in Figures 4–6 to examine the change in RSCP in SBD with every dimension of the parameterized vector. It is obvious that the spatial distributions of RSCP change as each parameter in refractivity profile changes. In other words, the spatial distributions of RSCP vary obviously with the parameters in SBD.



**Figure 4.** Comparison of the coverage diagrams for different thicknesses of the base layer.



**Figure 5.** Comparison of the coverage diagrams for different thicknesses of the trapping layer.



**Figure 6.** Comparison of the coverage diagrams for different duct strength.

### 5.2. Estimation of SBD

To test the performance of the BO-RF, the estimation study of three parameters SBD with RFC is carried out, and the results obtained by the BO-RF are compared with the results obtained from commonly used methods, such as classical RF, XGBoost with/without BO, and KNN with/without BO. The radar system parameters are identical to those given in Table 2. In order to fully consider the propagation characteristics of the SBD and reduce the number of training samples, the OED method is adopted to randomly select 10,000 sets of samples of the refractivity parameter  $(z_b, z_{thick}, M_d)$ , namely a matrix with ten thousand rows and three columns, and 10,000 sets of RSCP from 10.2 km to 50.0 km at an interval of 200.0 m, namely a matrix with ten thousand rows and two hundred columns, generated by the PWE method. It should be noted that we randomly selected 80% of the samples used for training the model, and the other 20% were used to test the regression model. Furthermore, 5% Gaussian noise is also added to 20% of the samples to examine the anti-noise ability.



Finally, the chosen RSCP matrix with eight thousand rows and two hundred columns and the refractivity parameter matrix with eight thousand rows and three columns are treated as the input and output matrix to train the regression model, respectively. In this paper, the default and BO-optimized hyper-parameter are given in Table 3.

**Table 3.** Default and BO-optimized hyper-parameter.

Hyper-Parameter	Default Value	Optimized Value
n_estimator	30	382
max_depth	6	42
min_sample_split	4	2
min_sample_leaf	4	1
max_features	auto	50

In addition, the following three evaluation indicators, namely the coefficient of determination  $R^2$ , mean absolute error (MAE), and mean square error (MSE), are introduced to test the performance of the BO-RF model for atmospheric duct estimation. Smaller MAE, MSE, and closer  $R^2$  value to one mean that the algorithm has better performance. The evaluation indicators are defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{m'} (X_i - Y_i)^2}{\sum_{i=1}^{m'} (\bar{Y} - Y_i)^2}, \quad (21)$$

$$MAE = \frac{1}{m'} \sum_{i=1}^{m'} |X_i - Y_i|, \quad (22)$$

$$MSE = \frac{1}{m'} \sum_{i=1}^{m'} (X_i - Y_i)^2, \quad (23)$$

where  $X_i$  is the predicted  $i$ th value,  $Y$  is the actual  $i$ th value,  $\bar{Y}$  is the mean of the true values, and  $m'$  is the number of datasets.

Firstly, a more informative and truthful evaluation indicator, the coefficient of determination  $R^2$  is adopted to assess the generalization ability of the regression model [37] based on the test sets with different regression models with/without noise, and the values of the coefficient of determination  $R^2$  are shown in Table 4, in which the best results are highlighted in bold. From Table 4, we can see that the BO-RF model obtains the best results in the with/without noise case. That is to say, the BO-RF model exhibits good generalization performance and anti-noise ability compared with the other regression models.

**Table 4.** Coefficient of determination  $R^2$  for the regression model.

Algorithms	$R^2$ without Noise	$R^2$ with Noise
RF	98.92%	98.46%
BO-RF	<b>99.93%</b>	<b>99.82%</b>
KNN	90.67%	87.51%
BO-KNN	91.76%	88.72%
XGBoost	94.68%	93.73%
BO-XGBoost	95.96%	94.83%

Then, the corresponding statistical analysis with respect to MAE and MSE is used to further test the accuracy of the regression model. It can be observed from Table 5 that the BO-RF outperforms the other models, such as RF, KNN, BO-KNN, XGBoost, and BO-XGBoost, for the estimation of SBD, and the rank order of the prediction accuracy is followed by BO-XGBoost and BO-KNN from greatest to least. Overall, the regression models optimized

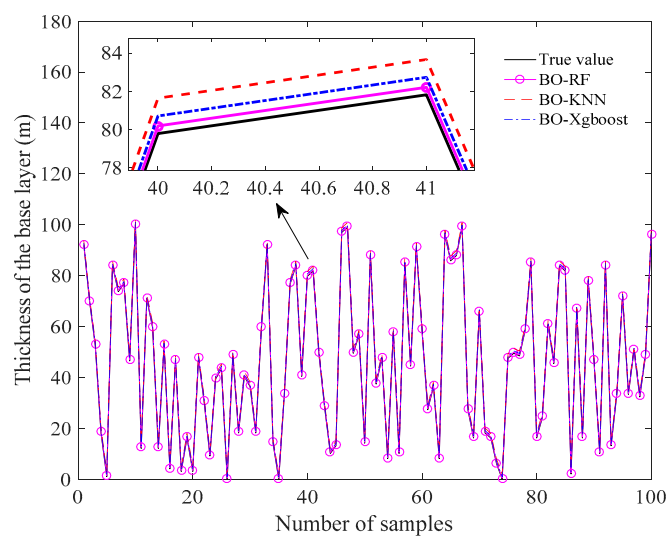
with BO are obviously better than those without BO regardless of whether the 5% Gaussian noise is taken into consideration or not. These results suggest that BO is an effective way to find the best hyper-parameter in the regression model.

**Table 5.** Comparison of the MAE and MSE for the test results without noise and with 5% Gaussian noise.

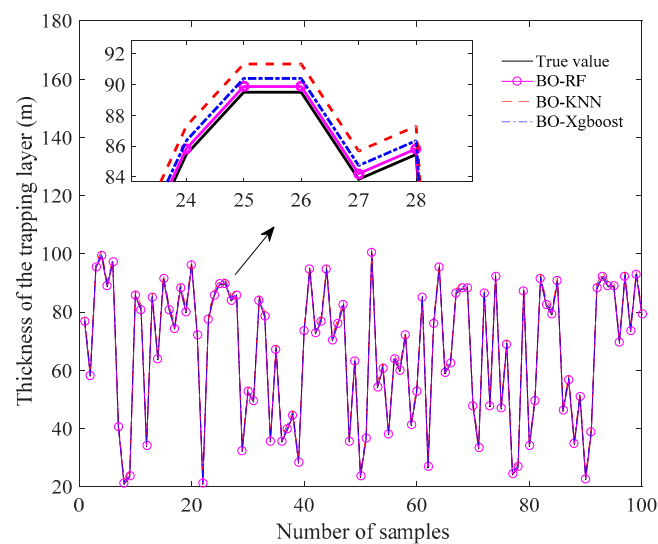
Algorithms	Parameters	MAE		MSE	
		Without Noise	With 5% Gaussian Noise	Without Noise	With 5% Gaussian Noise
RF	$z_b/m$	0.48	0.49	0.44	0.82
	$z_{thick}/m$	0.35	0.67	0.71	1.18
	$M_d/M\text{-units}$	0.48	0.50	0.40	0.95
BO-RF	$z_b/m$	<b>0.14</b>	<b>0.39</b>	<b>0.05</b>	<b>0.62</b>
	$z_{thick}/m$	<b>0.09</b>	<b>0.36</b>	<b>0.06</b>	<b>0.51</b>
	$M_d/M\text{-units}$	<b>0.16</b>	<b>0.40</b>	<b>0.09</b>	<b>0.84</b>
KNN	$z_b/m$	1.57	1.91	3.63	7.24
	$z_{thick}/m$	1.96	2.09	8.44	8.37
	$M_d/M\text{-units}$	2.61	2.85	9.57	9.91
BO-KNN	$z_b/m$	1.47	1.85	3.45	5.30
	$z_{thick}/m$	1.76	1.84	5.64	6.03
	$M_d/M\text{-units}$	2.46	2.70	6.30	8.80
XGBoost	$z_b/m$	0.93	1.06	2.78	4.00
	$z_{thick}/m$	1.42	1.69	4.07	4.15
	$M_d/M\text{-units}$	1.12	1.67	3.55	5.06
BO-XGBoost	$z_b/m$	0.64	0.91	1.16	3.14
	$z_{thick}/m$	0.75	0.90	2.11	2.79
	$M_d/M\text{-units}$	0.98	0.97	2.08	2.97

Furthermore, we randomly select 100 sets of predicted vectors of the refractivity parameter with 5% Gaussian noise from 2000 sets of predicted results to show the distributions of each parameter in SBD. From Figures 7–9, it can be roughly observed that the predicted parameters, such as the thickness of the base layer, the thickness of the trapping layer, and the duct strength obtained with BO-RF, BO-KNN, and BO-XGBoost, are in good agreement with the true values. In order to carefully examine the performance of the three methods, the partial enlarged drawing in each figure is also provided. It is clearly shown that the BO-RF obtains the best performance compared with BO-KNN and BO-XGBoost.

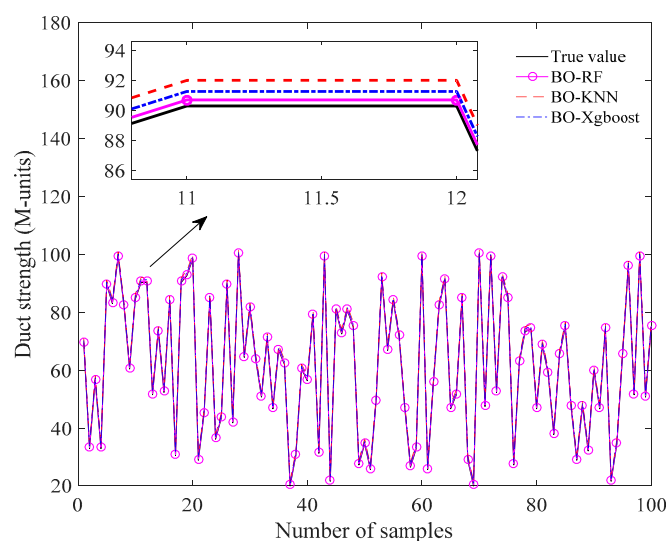
Another analysis focused on the accuracy and stability based on 50 independent runs for the inversion of SBD with/without 5% Gaussian noise is provided. Figures 10 and 11 demonstrate the comparison of the histograms of the estimation results achieved by BO-RF, BO-KNN, and BO-XGBoost. It should be pointed out that the red lines denote the position of the true parameter in SBD. It can be seen that the estimation parameters of BO-RF distribute closely to the true parameter compared with the other two models; these mean that the performance of BO-RF performs better than those of BO-KNN and BO-XGBoost in the case of with/without noise. Additionally, the difference between the true and predicted parameters, named inversion error, for the thickness of the base layer, the thickness of the trapping layer, and the duct strength based on the above 50 independent runs is given in Figures 12–14. It is clearly shown that the inversion error of BO-RF is also much lower than those of the two models; these also suggest that the accuracy of BO-RF is superior to those of BO-KNN and BO-XGBoost.



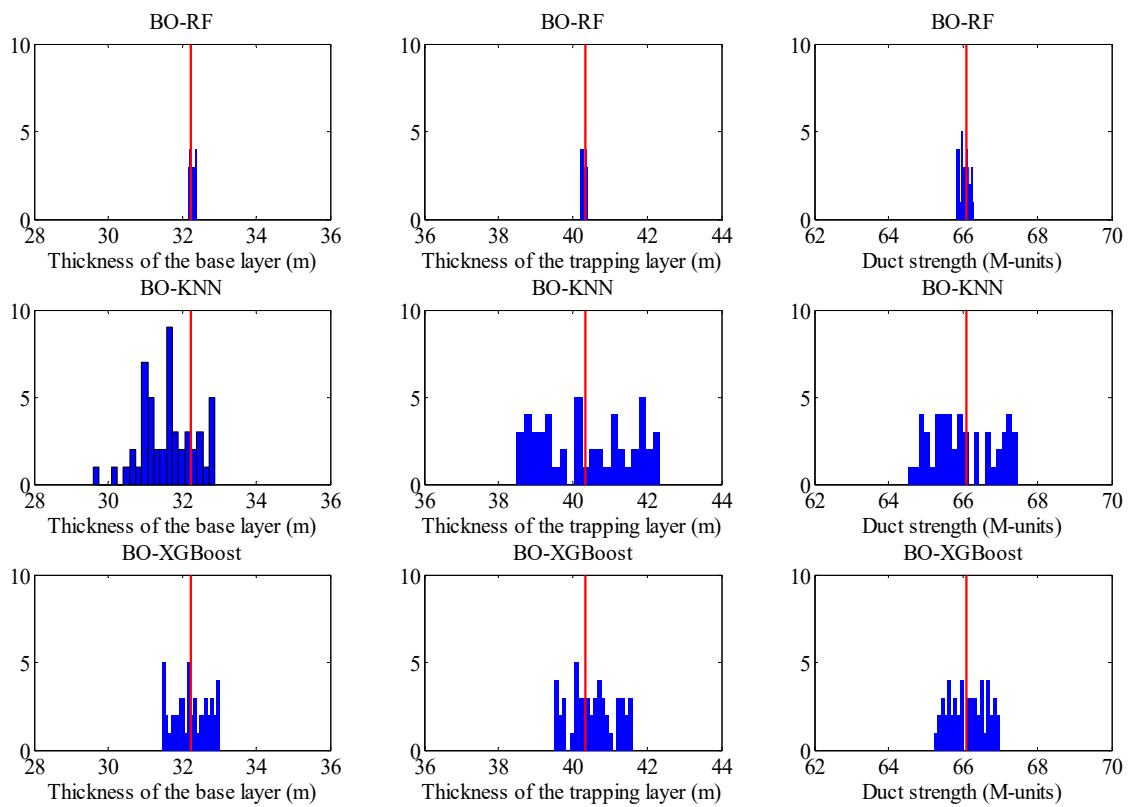
**Figure 7.** Comparison of the predicted thicknesses of the base layer for the BO-optimized models.



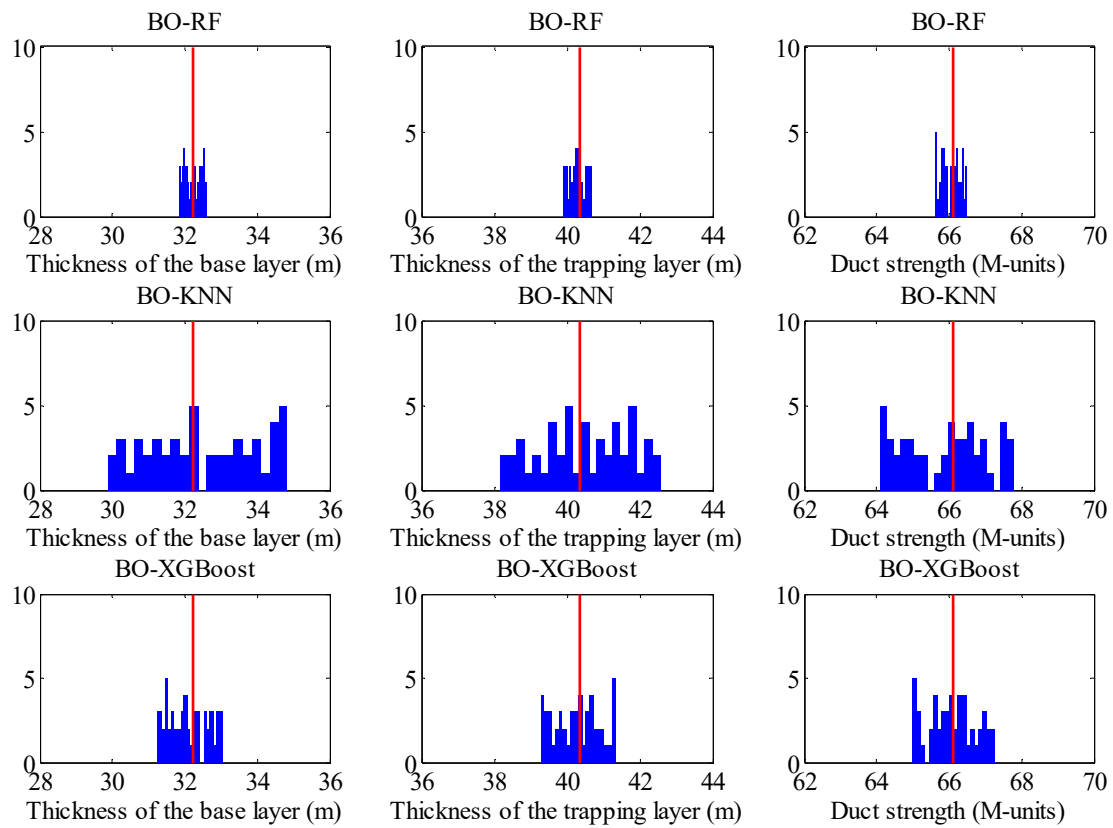
**Figure 8.** Comparison of the predicted thicknesses of the trapping layer for the BO-optimized models.



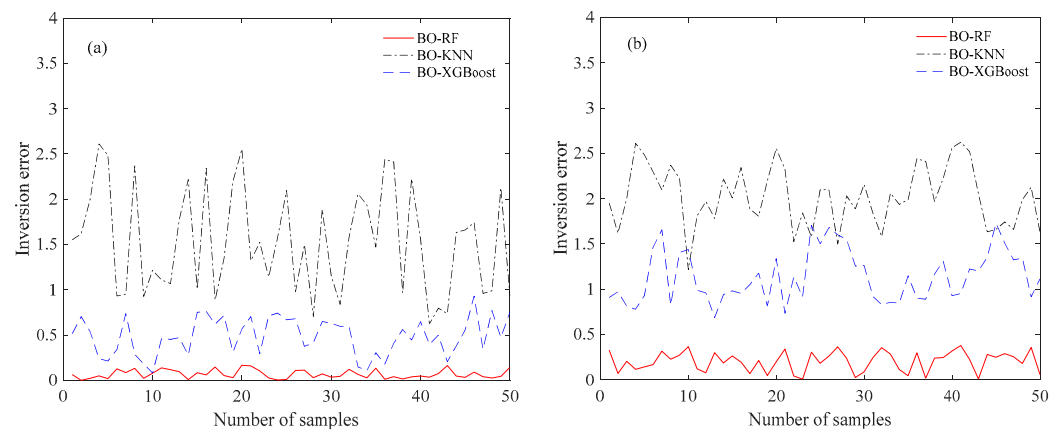
**Figure 9.** Comparison of the predicted duct strength for the BO-optimized models.



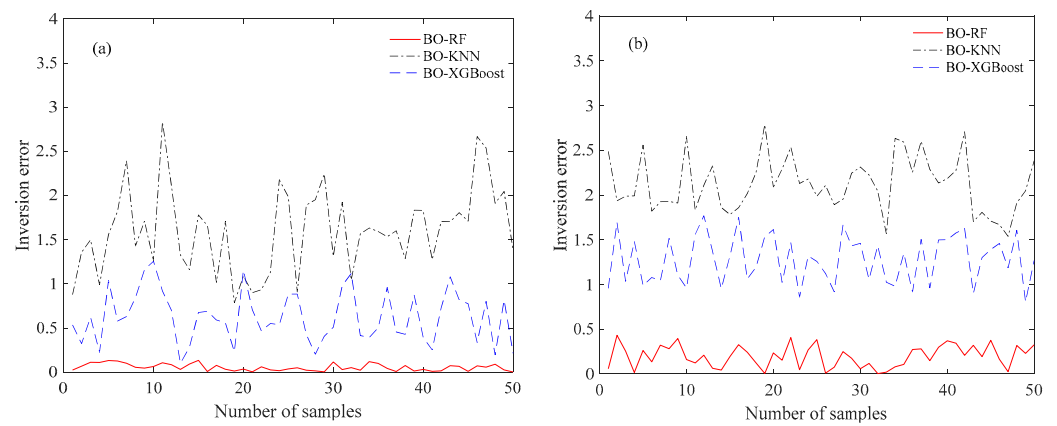
**Figure 10.** Comparison of the histograms of BO-RF, BO-KNN, and BO-XGBoost without noise.



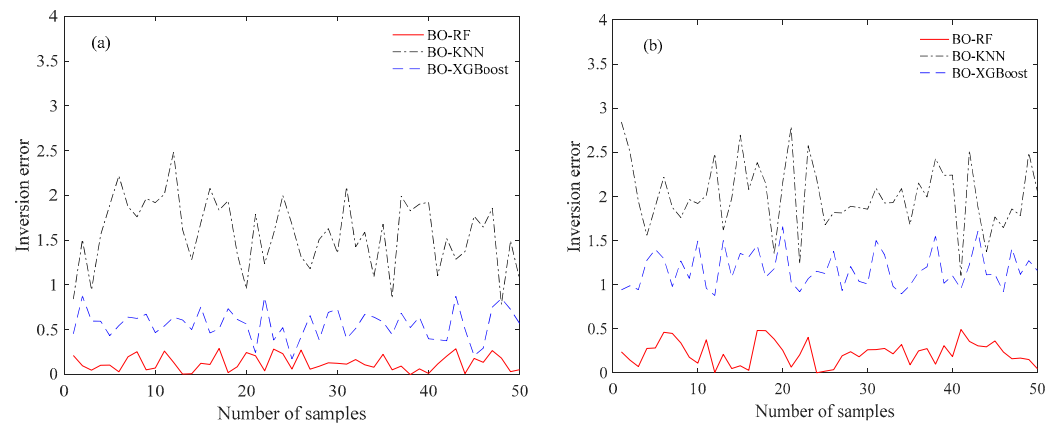
**Figure 11.** Comparison of the histograms of BO-RF, BO-KNN, and BO-XGBoost with 5% Gaussian noise.



**Figure 12.** Comparison of the inversion error with the thickness of the base layer: (a) without noise and (b) with 5% Gaussian noise.



**Figure 13.** Comparison of the inversion error with the thickness of the trapping layer: (a) without noise and (b) with 5% Gaussian noise.



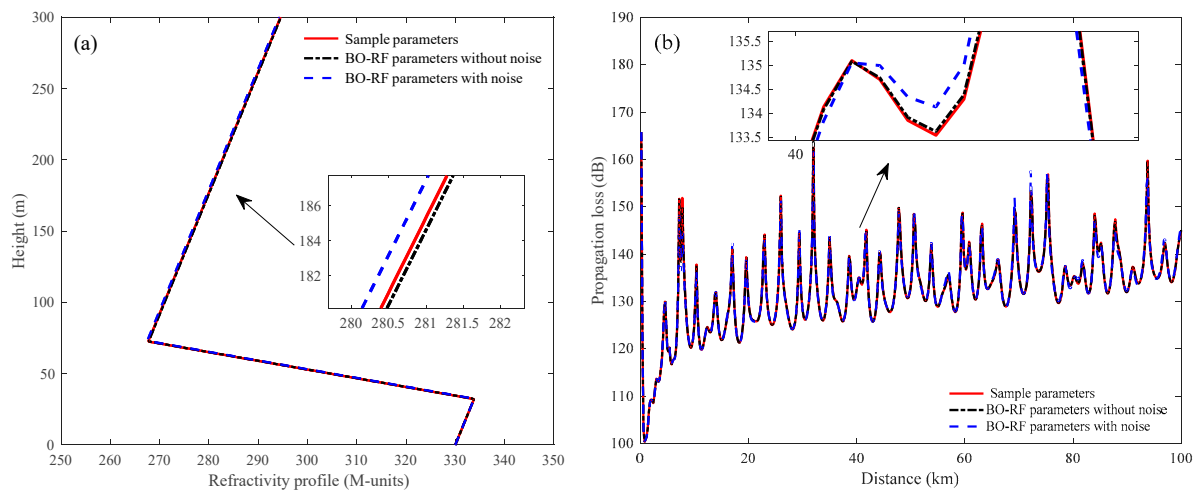
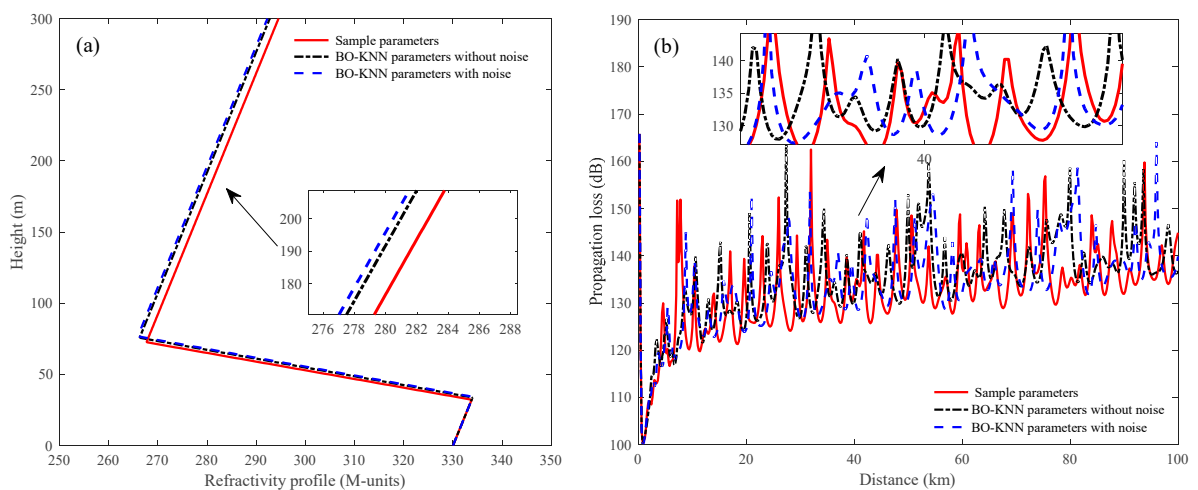
**Figure 14.** Comparison of the inversion error with the duct strength: (a) without noise and (b) with 5% Gaussian noise.

In addition, a comparative study focused on the refractivity profile and propagation loss between the true and predicted parameters in SBD is given in the subsequent section. The average inversion parameters of SBD with BO-RF, BO-KNN, and BO-XGBoost are given in Table 6.

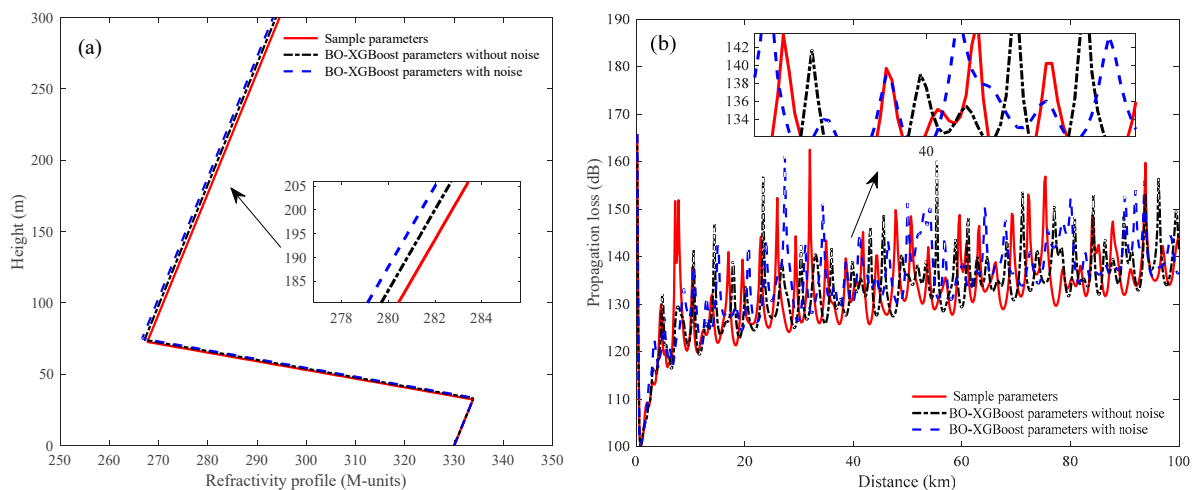
**Table 6.** Average inversion results of SBD.

Parameters	BO-RF		BO-KNN		BO-XGBoost	
	Without Noise	Noise	Without Noise	Noise	Without Noise	Noise
$z_b$ /m	32.28	32.43	33.77	34.26	32.92	33.37
$z_{thick}$ /m	40.30	40.54	41.95	42.44	41.10	41.62
$M_d$ /M-units	66.03	66.33	67.69	68.11	66.76	67.28

Figures 15–17 give comparisons of the refractivity profile and propagation loss between the true and predicted parameters. Furthermore, a partial enlarged drawing of each figure is presented to exhibit the details. Compared with BO-KNN and BO-XGBoost, we can conclude that the predicted refractivity profiles and their propagation loss obtained with BO-RF exhibit good agreement with the results simulated with true parameters with/without noise. In addition, we can see from Figures 16 and 17 that there are somewhat large deviations between the propagation loss simulated by the true and predicted parameters. This is due to the fact that a small variation in the index of refraction can cause a relatively large change in propagation loss.

**Figure 15.** Comparison of the (a) refractivity profile and (b) propagation loss between the true and predicted parameters obtained by BO-RF.**Figure 16.** Comparison of the (a) refractivity profile and (b) propagation loss between the true and predicted parameters obtained by BO-KNN.





**Figure 17.** Comparison of the (a) refractivity profile and (b) propagation loss between the true and predicted parameters obtained by BO-XGBoost.

## 6. Conclusions

The accurate inversion of atmospheric ducts is of great significance to the performance evaluation and design of radar and communication systems. In this paper, a BO-RF model is presented to predict the atmospheric duct parameters. Since the hyper-parameter plays a crucial role in the prediction performance of machine learning, BO is adopted to find the appropriate hyper-parameter in RF, and the results suggest that BO is an effective way to find the best hyper-parameter in RF. In addition, OED is used to sample the combinations of refractivity parameters from a range of hyper-parameter since it is a robust way to effectively handle multi-dimension sampling problems. To test the performance of the proposed BO-RF model, the results obtained by the BO-RF are compared with the results obtained by commonly used methods, such as classical RF, XGBoost with/without BO, and KNN with/without BO, with respect to the coefficient of determination  $R^2$ , MAE, and MSE, as well as inversion error. The statistical and comparative analysis indicates that the BO-RF obtains the best results in the case with/without noise and outperforms the other algorithms adopted in this paper for the estimation of atmospheric duct parameters. Future work will mainly focus on the estimation study based on real data with machine learning and the improvement of the existing refractivity model.

**Author Contributions:** Conceptualization, C.Y.; methodology, A.Z.; software, A.Z. and Y.W.; validation, C.Y., Y.W. and A.Z.; formal analysis, Y.W.; investigation, Y.W.; resources, C.Y. and L.G.; data curation, Y.W. and A.Z.; writing—original draft preparation, A.Z.; writing—review and editing, C.Y. and Y.W.; visualization, Y.W., A.Z. and H.F.; supervision, C.Y.; project administration, C.Y. and L.G.; funding acquisition, C.Y. and L.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant Nos. 61871457, 61302050) and the Natural Science Foundation of Shaanxi Province (Grant No. 2019JQ-200).

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yardim, C.; Gerstoft, P.; Hodgkiss, W.S. Estimation of radio refractivity from radar clutter using Bayesian Monte Carlo analysis. *IEEE Trans. Antennas Propag.* **2006**, *54*, 1318–1327. [\[CrossRef\]](#)
2. Gerstoft, P.; Rogers, L.T.; Krolik, J.L.; Hodgkiss, W.S. Inversion for refractivity parameters from radar sea clutter. *Radio Sci.* **2003**, *38*, 1801–1822. [\[CrossRef\]](#)
3. Karimian, A.; Yardim, C.; Gerstoft, P.; Hodgkiss, W.S.; Barrios, A.E. Refractivity estimation from sea clutter: An invited review. *Radio Sci.* **2011**, *46*, RS6013.

4. Douvenot, R.; Fabbro, V.; Gerstoft, P.; Bourlier, C.; Saillard, J. A duct mapping method using least squares support vector machines. *Radio Sci.* **2008**, *43*, RS6005. [\[CrossRef\]](#)
5. Wang, B.; Wu, Z.S.; Zhao, Z.; Wang, H.G. Retrieving evaporation duct heights from radar sea clutter using particle swarm optimization (PSO) algorithm. *Prog. Electromagn. Res. M* **2009**, *9*, 79–91. [\[CrossRef\]](#)
6. Zhang, J.P.; Wu, Z.S.; Zhu, Q.L.; Wang, B. A four-parameter M-profile model for the evaporation duct estimation from radar clutter. *Prog. Electromagn. Res.* **2011**, *114*, 353–368. [\[CrossRef\]](#)
7. Zhao, X. Evaporation duct height estimation and source localization from field measurements at an array of radio receivers. *IEEE Trans. Antennas Propag.* **2011**, *60*, 1020–1025. [\[CrossRef\]](#)
8. Yang, C. Estimation of the atmospheric duct from radar sea clutter using artificial bee colony optimization algorithm. *Prog. Electromagn. Res.* **2013**, *135*, 183–199. [\[CrossRef\]](#)
9. Yang, C.; Guo, L. Inferring the atmospheric duct from radar sea clutter using the improved artificial bee colony algorithm. *Int. J. Microw. Wireless Technol.* **2018**, *10*, 437–445.
10. Yang, C.; Wang, Y. Inversion of the surface duct from radar sea clutter using the improved whale optimization algorithm. *Electromagnetics* **2019**, *39*, 611–627. [\[CrossRef\]](#)
11. Yang, C. A comparison of the machine learning algorithm for evaporation duct estimation. *Radioengineering* **2013**, *22*, 657–661.
12. Tepecik, C.; Navruz, I. A novel hybrid model for inversion problem of atmospheric refractivity estimation. *Int. J. Electron. Commun.* **2018**, *84*, 258–264. [\[CrossRef\]](#)
13. Lentini, N.E.; Hackett, E.E. Global sensitivity of parabolic equation radar wave propagation simulation to sea state and atmospheric refractivity structure. *Radio Sci.* **2015**, *50*, 1027–1049. [\[CrossRef\]](#)
14. Penton, S.E.; Hackett, E.E. Rough ocean surface effects on evaporative duct atmospheric refractivity inversions using genetic algorithms. *Radio Sci.* **2018**, *53*, 804–819. [\[CrossRef\]](#)
15. Pozderac, J.; Johnson, J.; Yardim, C.; Merrill, C.; Paolo, T.; Terrill, E.; Ryan, F.; Frederickson, P. X-band Beacon-receiver array evaporation duct height estimation. *IEEE Trans. Antennas Propag.* **2018**, *66*, 2545–2556. [\[CrossRef\]](#)
16. Yan, X.; Yang, K.; Ma, Y. Calculation method for evaporation duct profiles based on artificial neural network. *IEEE Antennas Wireless Propag. Lett.* **2018**, *17*, 2274–2278. [\[CrossRef\]](#)
17. Zhu, X.; Li, J.; Zhu, M.; Jiang, Z.; Li, Y. An evaporation duct height prediction method based on deep learning. *IEEE Trans. Geosci. Remote Sens.* **2018**, *15*, 1307–1311. [\[CrossRef\]](#)
18. Sit, H.; Earls, C.J. Characterizing evaporation ducts within the marine atmospheric boundary layer using artificial neural networks. *Radio Sci.* **2019**, *54*, 1181–1191. [\[CrossRef\]](#)
19. Sit, H.; Earls, C.J. Deep Learning for Classifying and Characterizing Atmospheric Ducting within the Maritime Setting. *Comput. Geosci.* **2021**, *157*, 104919. [\[CrossRef\]](#)
20. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [\[CrossRef\]](#)
21. Onan, A.; Korukoğlu, S.; Bulut, H. Ensemble of keyword extraction methods and classifiers in text classification. *Expert Syst. Appl.* **2016**, *57*, 232–247. [\[CrossRef\]](#)
22. Kumari, P.; Toshniwal, D. Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. *J. Clean. Prod.* **2021**, *279*, 123285. [\[CrossRef\]](#)
23. Song, Y.; Liang, J.; Lu, J.; Zhao, X. An efficient instance selection algorithm for k nearest neighbor regression. *Neurocomputing* **2017**, *251*, 26–34. [\[CrossRef\]](#)
24. Alberoni, P.P.; Andersson, T.; Mezzasalma, P.; Michelson, D.B.; Nanni, S. Use of the vertical reflectivity profile for identification of anomalous propagation. *Meteorol. Appl.* **2001**, *8*, 257–266. [\[CrossRef\]](#)
25. Bech, J.; Sairouni, A.; Codina, B.; Lorente, J.; Bebbington, D. Weather radar anaprop conditions at a Mediterranean coastal site. *Phys. Chem. Earth B* **2000**, *25*, 829–832. [\[CrossRef\]](#)
26. Karimian, A.; Yardim, C.; Hodgkiss, W.S.; Gerstoft, P.; Barrios, A.E. Estimation of radio refractivity using a multiple angle clutter model. *Radio Sci.* **2012**, *47*, 1–9. [\[CrossRef\]](#)
27. Sirkova, I. Brief review on PE method application to propagation channel modeling in sea environment. *Open Eng.* **2012**, *2*, 19–38. [\[CrossRef\]](#)
28. Levy, M.F. *Parabolic Equation Methods for Electromagnetic Wave Propagation*; Institution of Engineering and Technology: London, UK, 2000.
29. Dockery, G.D. Modeling electromagnetic wave propagation in the troposphere using the parabolic equation. *IEEE Trans. Antennas Propag.* **1988**, *36*, 1464–1470. [\[CrossRef\]](#)
30. Adusumilli, S.; Bhatt, D.; Wang, H.; Bhattacharya, P.; Devabhaktuni, V. A low-cost INS/GPS integration methodology based on random forest regression. *Expert Syst. Appl.* **2013**, *40*, 4653–4659. [\[CrossRef\]](#)
31. Rafe, V.; Mohammady, S.; Cuevas, E. Using Bayesian optimization algorithm for model-based integration testing. *Soft Comput.* **2021**, *26*, 3503–3525. [\[CrossRef\]](#)
32. Ziatdinov, M.A.; Ghosh, A.; Kalinin, S.V. Physics makes the difference: Bayesian optimization and active learning via augmented Gaussian process. *Mach. Learn. Sci. Technol.* **2022**, *3*, 015022. [\[CrossRef\]](#)
33. Gao, W.; Liu, S.; Huang, L. A novel artificial bee colony algorithm based on modified search equation and orthogonal learning. *IEEE Trans. Cybern.* **2013**, *43*, 1011–1024.

34. Wang, Y.; Cai, Z.X.; Zhang, Q.F. Enhancing the search ability of differential evolution through orthogonal crossover. *Inf. Sci.* **2012**, *185*, 153–177. [[CrossRef](#)]
35. Zhan, Z.H.; Zhang, J.; Li, Y.; Shi, Y.H. Orthogonal learning particle swarm optimization. *IEEE Trans. Evol. Comput.* **2011**, *15*, 832–847. [[CrossRef](#)]
36. Moreta, C.E.G.; Acosta, M.R.C.; Koo, I. Prediction of digital terrestrial television coverage using machine learning regression. *IEEE Trans. Broadcast.* **2019**, *65*, 702–712. [[CrossRef](#)]
37. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.