A Comparison of the Machine Learning Algorithm for Evaporation Duct Estimation

Chao YANG

School of Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

yang_chaomail@163.com

Abstract. *In this research, a comparison of the relevance vector machine (RVM), least square support vector machine (LSSVM) and the radial basis function neural network (RBFNN) for evaporation duct estimation are presented. The parabolic equation model is adopted as the forward propagation model, and which is used to establish the training database between the radar sea clutter power and the evaporation duct height. The comparison of the RVM, LSSVM and RBFNN for evaporation duct estimation is investigated via the experimental and the simulation study, and the statistical analysis method is employed to analyze the performance of the three machine learning algorithms in the simulation study. The analysis demonstrate that the M profile of RBFNN estimation has a relatively good match to the measured profile for the experimental study; for the simulation study, the LSSVM is the most precise one among the three machine learning algorithms, besides, the performance of RVM is basically identical to the RBFNN.*

Keywords

Machine learning algorithm, evaporation duct, radar sea clutter, parameter estimation.

1. Introduction

Evaporation duct in a sea environment is nearly everlasting phenomenon in low-altitude troposphere, and caused by the abrupt changes in the vertical atmospheric temperature and humidity profiles above the sea surface [1]. It may result in an abnormal atmospheric structure compared with the standard atmosphere, and the modified refractivity profile (M profile) rapidly decreases with the increasing of altitude. Under the atmosphere duct environment, the non-standard electromagnetic propagation can be observed, and the fundamental parameters and performance of radar system can be greatly affected, such as the maximum operation range, creation of radar holes, and strengthened radar sea clutter, etc [2]. Therefore, it is of significance to detect the evaporation duct under the sea environment. The evaporation duct is usually described by modified refractivity profile, and it can be directly measured by the traditional ways, such as radiosondes, rocketsondes, microwave refractometers, lidar, etc [2]. However, these methods have the shortcomings of high cost and difficult in practice. In recent years, refractivity from clutter (RFC) [1-11] technique has been an active research area to estimate the refractivity profile in sea environment instead of using the traditional ways. Estimation of atmosphere duct using RFC technique is an inverse problem, and the relationship between the forward propagation model and atmosphere parameters is a complex nonlinear model. Recent years, many researchers dedicated to estimate the atmosphere duct with different optimization methods and analyze the performance of the optimization algorithm. Among them, [4] gives the detailed steps involved in RFC. [11] presents an overview of research progress in the field of RFC. Owing to the evaporation duct occur with extremely high probability, there is great interest in studies of estimating the evaporation duct [3], [12], [13] with the novel and promising RFC technique. The machine learning algorithm is a type of optimization algorithm, which is based on the training database. Compared with the global intelligent algorithm, such as genetic algorithm, particle swarm algorithm, etc, the machine learning algorithm has the advantage of real time [8].

In this study, a comparison of the relevance vector machine (RVM), least square support vector machine (LSSVM) and the radial basis function neural network (RBFNN) for evaporation duct estimation are presented. For the comparison, the experimental data and simulation data are used to evaluate the performance of the machine learning algorithm. In addition, the statistical analyses for the simulation study are also provided with different machine learning algorithms.

2. Refractivity and Radar Sea Clutter Model

The atmosphere duct is characterized by the modified refractivity profile. In this paper, the following log-linear evaporation duct profile model is considered [14]

$$M(z) = M_0 + 0.125z - 0.125d \ln \frac{z + z_0}{z_0} \qquad (1)$$

where z is the height above the mean height of sea surface, d is the evaporation duct height, z_0 is roughness factor whose typical value is 0.00015, and the constant M_0 is usually taken as 330.0M units. It is clearly seen that the evaporation duct height is the only parameter in (1).

To study the inversion problem, the electromagnetic signal propagation problem under the evaporation duct must be taken into account. In this work, the most widely used parabolic equation model [15-19] is adopted due to its accuracy and stability.

According to the fundamental theory of radar, considering the influence of atmosphere condition, the radar equation can be expressed by [4]

$$P_c = \frac{P_t G_t^2 \lambda^2 F^4 \sigma}{(4\pi)^3 r^4} \tag{2}$$

where P_t is the transmitted power, G_t is the gain of transmitting antenna, λ is the wavelength, r is the distance between radar and illumination area, F is the propagation factor, and σ is the sea surface radar cross section, which can be expressed in term of normalized sea surface radar cross section σ_0.

Finally, the received radar sea clutter power can be obtained by representing in dB

$$P_{c\,,\,\mathrm{dB}} = -2L + \sigma^{\circ} + 10\lg(r) + C \tag{3}$$

where L is the propagation loss obtained from the parabolic equation mentioned above, C accounts for the constant terms in (2). In general, the normalized sea surface radar cross section σ_0 is of difficulty to calculate with analytical or numerical methods at low grazing angle, so it can be acquired by the GIT model [20].

3. Introduction of the Estimation Algorithm

In this section, the RVM and RBFNN regression model [21-23] are briefly introduced in the following. For the LSSVM regression model, one can refer to [8].

3.1 RVM Algorithm

Given a training data set with N input vectors $\{\boldsymbol{x}_k\}_{k=1}^N$ and their corresponding scalar-valued output $\{t_k\}_{k=1}^N$. Considering the output t_k is sampled from the model with additive noise ε_k, we have

$$t_k = y(\boldsymbol{x}_k;\boldsymbol{\omega}) + \varepsilon_k \tag{4}$$

where ε_k is assumed to be independent zero-mean Gaussian distribution with variance σ^2. Hence, the output is a Gaussian distribution over t_k with mean $y(\boldsymbol{x}_k)$ and variance σ^2, it can be obtained by

$$p(t_k|\boldsymbol{x}_k) = N(t_k|y(\boldsymbol{x}_k),\sigma^2) \tag{5}$$

where $y(\boldsymbol{x}_k, \boldsymbol{\omega})$ is the output prediction of the true value t_k and $\boldsymbol{\omega} = [\omega_0, \omega_1, \ldots, \omega_N]$ is the weight vector for the RVM regression model.

The regression model of RVM can be given by

$$y(\boldsymbol{x};\boldsymbol{\omega}) = \omega_0 + \sum_{k=1}^{N} \omega_k K(\boldsymbol{x},\boldsymbol{x}_k) = \boldsymbol{\omega}\phi(\boldsymbol{x}) \tag{6}$$

where $K(\boldsymbol{x}, \boldsymbol{x}_k)$ is a Kernel function and $\phi(\boldsymbol{x}) = [1, K(\boldsymbol{x},\boldsymbol{x}_1),\ldots,K(\boldsymbol{x},\boldsymbol{x}_N)]^T$.

3.2 RBFNN Algorithm

Given a training set $X = [X_1, X_2,\ldots,X_k,\ldots,X_N]^{\mathrm{T}}$, where $X_k = [x_{k1}, x_{k2},\ldots,x_{km},\ldots,x_{kM}]$ denote any one of the training sample, and it's corresponding actual output $Y_k = [y_{k1}, y_{k2},\ldots,y_{kj},\ldots,y_{kJ}]^{\mathrm{T}}$.

When the input sample is X_k, the actual output of the jth output neuron can be expressed by

$$y_{kj}(X_k) = \sum_{i=1}^{N} w_{ij}\varphi(X_k, X_i), \quad j = 1,2,\cdots,J\,. \tag{7}$$

In this paper, the Gaussian Kernel Function is adopted as the basis function

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j) = \exp\left(-\frac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2}{\tau^2}\right) \tag{8}$$

where norm $\|\cdot\|$ denotes the Euclidean distance, and τ is the width of Kernel function.

4. Estimation Results and Discussions

In the following, a comparison of the RVM, LSSVM and RBFNN for evaporation duct estimation are investigated via the experimental and the simulation study.

In order to train the machine learning algorithm, the training database is constructed by the parabolic equation model. As we all know, the evaporation duct height commonly range from 0 m to 40 m. Without loss of generality, the Latin hypercube sampling [8] is used to randomly extract 14 evaporation duct height samples from the interval range. Hence, the corresponding 14 discrete radar sea clutter power from 10 km to 50 km with an interval of 200 m can be obtained by parabolic equation at a receiving altitude of 0.3 m [20]. And those discrete radar sea clutter power and evaporation duct heights are treated as the input vector and output vector in the machine learning algorithm, respectively.

For the purpose of comparison of the RVM, LSSVM and RBFNN, those three machine learning algorithms are firstly compared by a set of experimental data measured in East China Sea [24]. During the experiment, the radar system works at a frequency of 10 GHz, antenna height of 10 m, beam width of 0.7°, and the polarization mode is horizontal polarization.

The dependence of measured radar sea clutter power on the propagation distance is given in Fig. 1. Here, taking the radar sea clutter power described in Fig. 1 as the input vector to estimate the evaporation duct with the RVM, LSSVM and RBFNN algorithm, and their corresponding modified refractivity profile can be obtained by (1).

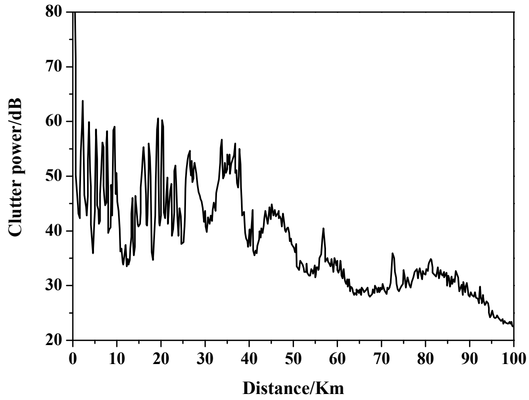

Fig.1. The dependence of measured radar sea clutter power on the propagation distance.

The comparison of the machine learning algorithm estimation results with the measured profile are shown in Fig. 2, it is observed that the M profile of RBFNN estimation has a relatively good match to the measured one as a whole.

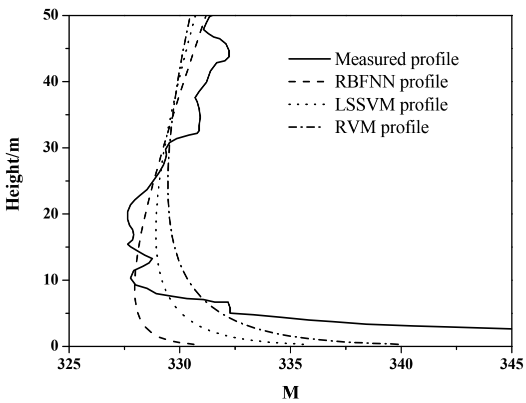

Fig. 2. The comparison of the estimation results of three machine learning algorithms with the measured profile.

Then, simulation studies are also implemented with synthesized radar sea clutter power. In order to take the influence of range dependence sea clutter radar cross section into consideration. The synthesized radar sea clutter power can be expressed as [5]

$$\boldsymbol{P}_s = \boldsymbol{P}_c + N \tag{9}$$

where $\boldsymbol{P}_s$ is synthesized radar sea clutter power, $N = [n_1, n_2, n_3\ldots]$ is the noise. And the noise can be produced by

$$n_1 = 0\ , \tag{10}$$

$$n_{i+1} = n_i + \gamma_i \tag{11}$$

with γ_i is an independent draw from a zero-mean Gaussian distribution with variance $\sigma^2{}_\gamma$, and the standard deviation σ_γ is used to denote the noise level.

To train the RVM, LSSVM and RBFNN algorithm in the simulation study, the 14 sets of synthesized radar sea clutter power are generated with the sampling evaporation duct height parameters, at a frequency of 7 GHz, power of 91.4 dBm, antenna gain of 52.8 dB, antenna height of 10 m, beam width of 0.7°, 600 m range bins, and the polarization mode is horizontal polarization. Here, we take the

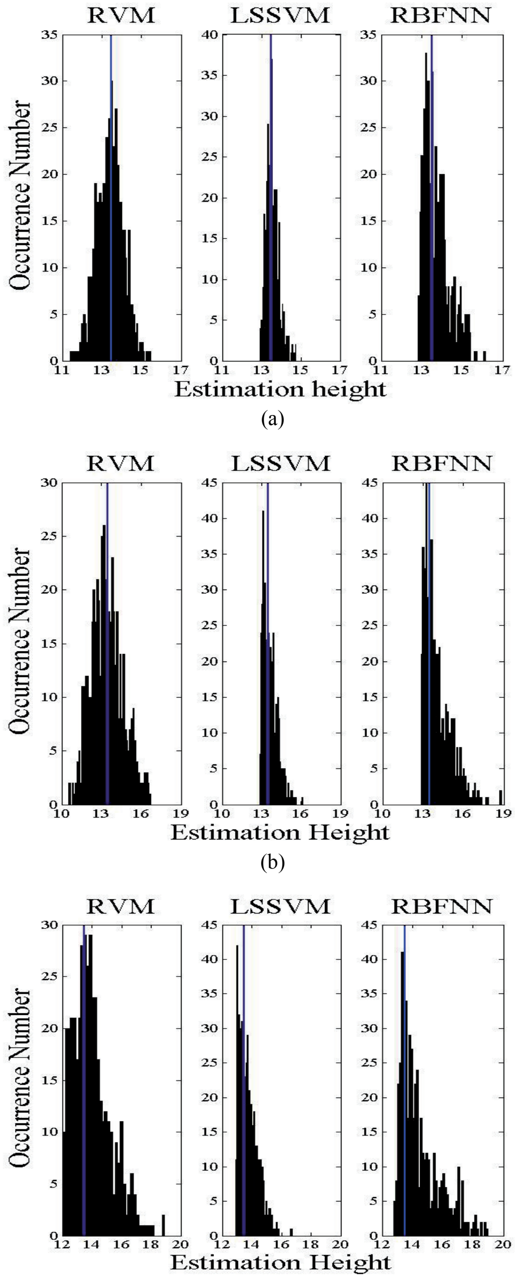

Fig. 3. The histograms of different optimization algorithms with different noise level (a) 1% (b) 3% (c) 5% .

synthesized radar sea clutter power of evaporation duct height of 13.5 m as an example to analyze the RVM, LSSVM and RBFNN algorithm in detail. Considering the influence of random noise, 500 runs are simulated to evaluate the performance of the three machine learning algorithms. In addition, the effects of noise level on the estimation results are also investigated. The histograms of estimation evaporation duct height with different machine learning algorithms and noise level are illustrated in Fig. 3, and the blue vertical lines indicate the actual evaporation duct height of synthesized radar sea clutter power. It can be seen that the results of LSSVM is more accurate than RVM and RBFNN for different noise level and the statistical results of RVM is similar to a Gaussian distribution. Also, the statistical results of RBFNN method on the right side of the vertical lines deviate greatly from the actual value.

To quantitatively analyze the performance of different machine learning algorithms, the following evaluation indexes are introduced [25]

$$\text{Mean Squared Error (MSE)} = \frac{1}{N}\sum_{i=1}^{N}(d_i^a - d_i^e)^2 \qquad (12)$$

$$\text{Mean Absolute Deviation (MAD)} = \frac{1}{N}\sum_{i=1}^{N}\left|d_i^a - d_i^e\right| \qquad (13)$$

$$\text{Mean Relative Error (MRE)} = \frac{100}{N}\sum_{i=1}^{N}\frac{\left|d_i^a - d_i^e\right|}{d_i^e} \qquad (14)$$

where N is number of simulations, d_i^a and d_i^e are the actual and estimated evaporation duct height, respectively.

Noise level	RVM			LSSVM			RBFNN		
	MSE	MAD	MRE	MSE	MAD	MRE	MSE	MAD	MRE
1%	0.457	0.537	3.98%	0.105	0.248	1.83%	0.438	0.489	3.63%
3%	1.453	0.974	7.21%	0.319	0.427	3.16%	1.384	0.797	5.90%
5%	2.354	1.190	8.81%	0.467	0.508	3.76%	2.739	1.131	8.38%

Tab. 1. Comparison of the statistical analysis results of the RVM, LSSVM and BRFNN.

In order to further compare the performance of RVM, LSSVM and BRFNN, the statistical analysis results are provided in terms of MSE, MAD and MRE defined by (12)- (14), are shown in Tab. 1. From Tab. 1, we can clearly see that the MSE, MAD and MRE of LSSVM are smaller than the others, that is to say, the evaporation duct height estimation of LSSVM is the most precise one, which is consistent with the conclusion obtained from Fig. 3. Furthermore, the statistical results show that the performance of RVM is basically identical to the RBFNN in the simulation study from the view of the above evaluation index.

5. Conclusion

In summary, a comparison of the RVM, LSSVM and RBFNN for evaporation duct estimation in sea environment are presented in this paper. Firstly, the parabolic equation model is utilized to construct the training database between the radar sea clutter power and the evaporation duct height. Then, the comparison of the RVM, LSSVM and RBFNN estimation results obtained by the experimental data with the measured profile gathered in East China Sea are provided. In addition, simulation studies are also implemented with synthesized radar sea clutter power generated by the parabolic equation model. The analysis demonstrate that the M profile of RBFNN estimation has a relatively good match to the measured profile for the experimental study; for the simulation study, the LSSVM is the most precise one among the three machine learning algorithms, besides, the performance of RVM is basically identical to the RBFNN. It should be noted that the model parameter of the machine learning algorithm is determined by experience in this paper, the good way of the selection of model parameter will be investigated in the future work.

Acknowledgements

This work was supported by the Youth Foundation (Nos.1050410 and 0001292) of Xi'an University of Posts and Telecommunications, and the Scientific Research Program Funded by Shaanxi Provincial Education Department (No. 12JK0964).

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, vol. 54, no. 4, p. 1318-1327.

[2] YARDIM, C., GERSTOFT, P., HODGKISS, W. S. Tracking refractivity from clutter using Kalman and particle filters. *IEEE Trans. Antennas Propag*, 2008, vol. 56, no. 4, p. 1058-1070.

[3] ROGER, L. T., HATTAN, C. P. Estimating evaporation duct heights from radar sea echo. *Radio Sci.*, 2000, vol. 35, no. 4, p. 955-966.

[4] GERSTOFT, P., ROGER, L. T., KROLIK, J. L., HODGKISS, W. S. Inversion for refractivity parameters from radar sea clutter. *Radio Sci.*, 2003, vol. 38, no. 3, p. 8053.

[5] ROGER, L. T., JABLECKI, M. Posterior distributions of a statistic of propagation loss inferred from radar sea clutter. *Radio Sci.*, 2005, vol. 40, p. RS6005.

[6] VASUDEVAN, S., ANDERSON, R. H., KRAUT, S., GERSTOFT, P., ROGERS, L. T., KROLIK, J. L. Recursive Bayesian electromagnetic refractivity estimation from radar sea clutter. *Radio Sci.*, 2007, vol. 42, p. RS2104.

[7] YARDIM, C., GERSTOFT, P., HODGKISS, W. S. Statistical maritime radar duct estimation using a hybrid genetic algorithms–Markov chain Monte Carlo method. *Radio Sci.*, 2007, vol. 42, p. RS3014.

[8] DOUVENOT, R., FABBRO, V., GERSTOFT, P., BOURLIER, C., SAILLARD, J. A duct mapping method using least squares support vector machines. *Radio Sci.*, 2008, vol. 43, p. RS6005.

[9] YARDIM, C., GERSTOFT, P., HODGKISS, W. S. Sensitivity analysis and performance estimation of refractivity from clutter techniques. *Radio Sci.*, 2009, vol. 44, p. RS1008.

[10] ZHAO, X. F., HUANG, S. X., DU, H. D. Theoretical analysis and numerical experiments of variational adjoint approach for refractivity estimation. *Radio Sci.*, 2011, vol. 46, p. RS1006.

[11] KARIMIAN, A., YARDIM, C., GERSTOFT, P., HODGKISS, W. S., BARRIOS, A. E. Refractivity estimation from sea clutter: An invited review. *Radio Sci.*, 2011, vol. 46, p. RS6013.

[12] WANG, B., WU, Z. S., ZHAO, Z. W.,WANG, H. G. Retrieving evaporation duct heights from radar sea clutter using particle swarm optimization (PSO) algorithm. *Prog. Electromagn. Res.*, 2009, vol. 9, p.79-91.

[13] 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, vol. 114, p.353-368.

[14] PAULUS, R. A. Evaporation duct effects on sea clutter. *IEEE Trans. Antennas Propag.*, 1990, vol. 38, no. 11, p. 1765-1771.

[15] KUTTLER, J. R., JANASWAMY, R. Improved Fourier transform methods for solving the parabolic wave equation. *Radio Sci.*, 2002, vol. 37, no. 2, p. RS6013.

[16] VALTR, P., PECHAC, P. Tropospheric refraction modeling using ray-tracing and parabolic equation. *Radioengineering*, 2005, vol. 14, no. 4, p. 98-104.

[17] VALTR, P., PECHAC, P., KVICERA, V., GRABNER, M. A terrestrial multiple-receiver radio link experiment at 10.7 GHz – comparisons of results with parabolic equation calculations. *Radioengineering*, 2010, vol. 19, no. 1, p. 117-121.

[18] ZHAO, X., HUANG, S. Influence of sea surface roughness on the electromagnetic wave propagation in the duct environment. *Radioengineering*, 2010, vol. 19, no. 4, p. 601 – 605.

[19] SIRKOVA, I. Brief review on PE method application to propagation channel modeling in sea environment. *Central European Journal of Engineering*, 2012, vol. 2, no. 1, p. 19-38.

[20] REILLY, J. P., DOCKERY, G. D. Influence of evaporation ducts on radar sea return. *IEE Proc. Radar and Signal Processing*, 1990, p. 80-88.

[21] TIPPING, M. E. Sparse Bayesian learning and the relevance vector machine. *The Journal of Machine Learning Research*, 2001, vol. 1, p. 211-244.

[22] WONG, P. K., XU, Q. S., VONG, C. M., WONG, H. C. Rate-dependent hysteresis modeling and control of a postage using online support vector machine and relevance vector machine. *IEEE Trans. on Industrial Electronics,* 2012, vol. 59, no. 4, p. 1988 to 2001.

[23] GAO, J. *The Principles of Neural Network and Its Simulation Example*. Beijing: China Machine Press, 2003.

[24] HAN, J. Study on inversion the low-altitude atmospheric refractivity profile from radar sea clutter. *Master Dissertation,* China Academy of Electronics and Information Technology, Beijing, 2008.

[25] NOTARNICOLA, C., ANGIULLI, M., POSA, F. Soil moisture retrieval from remotely sensed data: neural network approach versus Bayesian method. *IEEE Trans. Geosci. Remote Sens.,* 2008, vol. 46, no.2, p.547-557.

About Authors ...

Chao YANG is currently a lecturer at Xi'an University of Posts and Telecommunications, China. His research interests mainly include: electromagnetic wave propagation and scattering in random media.