Journal of Computational Information Systems 10: 1 (2014) 1–8
Available at http://www.Jofcis.com

An Extreme Learning Machine Application in Geophysical Modeling for Scatterometer

Boheng Duan [1,*], Weimin Zhang [1] Chengzhang Zhu [1]

[1] *College of Computer, National University of Defense Technology, Changsha, China, 410073*

Abstract

Scatterometer observations can provide accurate and spatially wind information near the sea surface which have proven significant for the forecasting of dynamical weather, such as tropical cyclones. One important step for the wind retrieval based on scatterometer observations is using the geophysical model function (GMF) which depicts the relationship between the normalized radar cross-section measurements and the sea surface wind speed, wind direction, radar parameters and environmental parameters. Since the establishment of strict theoretical geophysical model function is very difficult, at present most of the models are empirical models which are built by statistical methods. In this paper, we use a method called Extreme Learning Machine (ELM) to develop a unified GMF respectively using the simulated training data-set generated by the empirical GMF-CMOD5.N and the wind data retrieved from the advanced scatterometer (ASCAT). And we compare our method with the neural network approach. The results indicate that the ELM method shows a good inversion result with higher accuracy compared with the neural network technique. This new method could provide a novel feasible way for sea surface wind inversion.

Keywords: ELM; scatterometer; neural network; geophysical model function

1 Introduction

A wind scatterometer is a kind of spaceborne radar remote sensing instrument also known as active strabismal microwave sounding unit, which measures the radar backscatter of the oceans surface. Scatterometer observations cover about 70% of the world's sea surface, and it can penetrate clouds, provides accurate information on wind speed and direction makes it possible for monitoring the wind field of the sea at any time and any weather. The high resolution, timeliness and coverage of scatterometer observations can effectively compensate for the deficiencies of conventional observations of the sea, which makes it the primary means of ocean surface wind observation.

The wind information can be obtained using multiple measurements (usually at least three) of radar cross sections σ^0. The basis of the wind inversion is the geophysical model function

*Corresponding author.
Email address: bhduan@foxmail.com (Boheng Duan).

January 2014

(GMF) which gives the σ^0 as a function of the wind vector, incidence angle, azimuth angle, radar frequency and polarization. Nowadays the wind retrieval is mainly through some empirical models due to a lack of sufficiently accurate theory to describe the relationship between short sea waves and wind vectors [1]. Different kinds of scatterometer instruments usually use particular empirical GMF models for wind retrieval, e.g. SASS-1/2, NSCAT-1/2 and QSCAT-1 models for Ku-band sensors, and CMOD series for C-band scatterometers [2]. These empirical models are mostly established using statistical methods which require a lot of scatterometer observations. However, a complete data set needs a long time monitoring and is expensive to get. Therefore, a GMF making full use of insufficient spaceborne scatterometer observations to meet the accuracy requirements for wind retrieval is of great importance.

Based on the fact that the artificial neural network has a good ability of incomplete nonlinear mappings and noise reduction, Hongju Zhou [2] established a unified C-band and Ku-band geophysical model using the neural network, and it showed a good result compared to the empirical models. However, it is known that neural networks face some challenging issues: (1) slow learning speed, (2) trivial human intervention, (3) poor computational scalability [3]. A new learning algorithm called extreme learning machine (ELM) for single-hidden layer feed forward neural networks (SLFNs) is proposed by Guangbin Huang et al. 2006, which performs much better than conventional popular learning algorithm for feedforward neural networks [4]. In this paper, we use ELM to build the unified geophysical model with both simulated data of empirical model CMOD5.N and ASCAT wind products and compare with the neural network method. The results show that ELM method can effectively reflect the true state of the backscatter and ocean surface wind and gets a better accuracy than the neural network method.

This paper is organized as follows: First, a brief background information about the geophysical model function and the theory of ELM are provided. Next, we describe the ELM modeling process for GMF. Then we show the experiments and analysis. Finally, conclusions are provided in the last section.

2 Related work

2.1 The geophysical model function

The geophysical model function, providing a normalized radar backscatter coefficient, is based on the measurement geometry (relative azimuth and incidence angle), sensing parameters (frequency, polarization) and the wind vector. Since a rigorous theoretical model needs a thorough understanding of the theory of interaction between the sea surface wind and sea surface roughness as well as the electromagnetic scattering mechanism which is very difficult at present. Therefore, an alternative approach is to use the statistical methods to establish an empirical model.

The general form of empirical model gives the function between the backscatter σ^0 and parameters such as polarization $p(VV, HH)$, incidence θ, wind speed U and azimuth ϕ_R. This function can be expressed as follows:

$$\sigma^0 \quad = F(P, U, \theta, \phi_R) \tag{1}$$

Currently, the general expression of the CMOD series models can be given as:

$$\sigma^0 \quad = b_0(1 + b_1 cos\phi + b_2 cos\phi)^{1.6} \tag{2}$$

Where ϕ is the wind direction (which is a function of azimuth ϕ_R), and b_0, b_1, b_2 are complex functions of wind speed, wind direction and incidence [5, 6], which we do not discuss in detail here.

The CMOD5.N is a geophysical model function used to retrieve ASCAT neutral winds [6]. Under the condition of fixed incidence (47°) and polarization (VV), the σ^0 derived from the CMOD5.N has a relation with the azimuth as showd in Fig. 1. The σ^0 reaches the maximum at the point of upwind and downwind, while the minimum at the point of cross-wind.

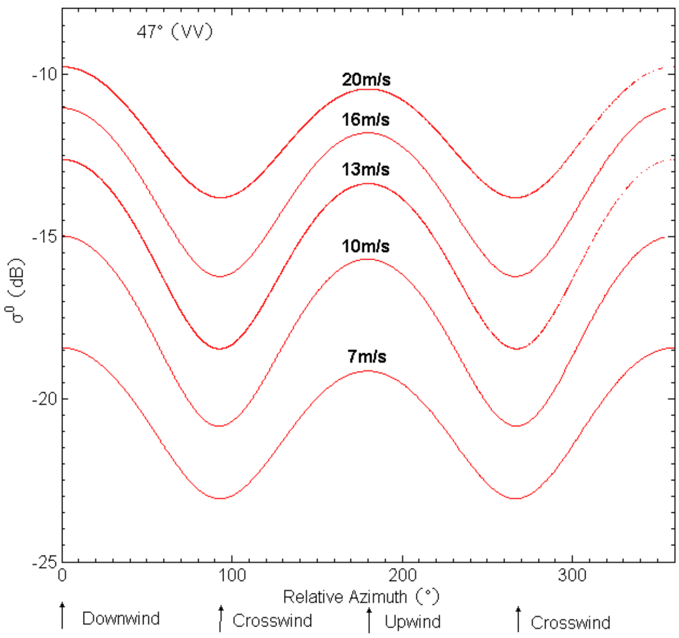

Fig. 1: C-band normalized radar cross sections at vertical polarizations as a function of angle between wind and radar look directions for various wind speeds at an incidence angle of 47°

2.2 The extreme learning machine

Extreme learning machine (ELM) is a learning algorithm for single-hidden layer feedforward neural networks (SLFNs) which randomly chooses hidden nodes and analytically determines the output weights of SLFNs [4]. Traditional neural network learning algorithm (such as BP algorithm) requires a large amount of parameters which are set manually and it is prone to local optima.While the ELM algorithm only needs to set the number of node of hidden layer, and does not need to adjust the input weights and the network hidden element during the execution of the algorithm, and it is proved to generate a unique optimal solution.

The ELM consists of three layers, namely the input layer, hidden layer and output layer. Generally, we train this neural network by adjusting the connection weights and the deviation between layers.

Denoted input layer, hidden layer and output layer nodes as n_1,n_2,n_3, then the former single hidden layer feedforward neural networks can be expressed as:

$$\overrightarrow{t_r} = f_r(\overrightarrow{x_j}) = \sum_{i=1}^{n_2} \beta_{ir} G_i(\overrightarrow{a_i}, b_i, \overrightarrow{x_j}) \qquad (j = 1, 2, \ldots, n_1; r = 1, 2, \ldots, n_3) \tag{3}$$

Where $\overrightarrow{t_r} = [t_{r1}, t_{r2}, \ldots, t_{rn}]^T$ is the output vector and $\overrightarrow{x_j} = [x_{j1}, x_{j2}, \ldots, x_{jn_1}]^T$ is the input vector; $\overrightarrow{a_i} = [a_{1i}, a_{2i}, \ldots, a_{n_1 i}]$ denotes the connection weight vector of the input layer and the i-th hidden layer while $\overrightarrow{\beta_i} = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{in_2}]^T$ the connection weight vector of the output layer; b_i is the offset of the i-th hidden layer; $G_i = (\overrightarrow{a_i}, b_i, \overrightarrow{x_j}) = g(\overrightarrow{a_i} \cdot \overrightarrow{x_j} + b_i)$ is the value computed through the hidden layer, where $g(\cdot)$ represents the hidden layer activation function, such as sigmoid function, sine function, hard limit function, triangular basis functions and radial basis functions and so on.

3 ELM modeling

In this section, we detail the procedure of the ELM technique we use for the GMF modeling. It consists of the following steps as is shown in Fig.2. Before we use the training samples (input data), we need to do the normalization, and we also need to do the inverse normalization before the accuracy evaluation. The structure of the training samples is described in the following subsection. The input vectors for this ELM training are wind speed V, azimuth ϕ and incidence angle θ, while the target vector is radar cross sections σ^0. After that, we obtain a training model which could be used for the latter prediction. The specific process of the prediction is that we use a number of test data as the input, use the trained model to get an output and compare with the target to evaluate the accuracy.

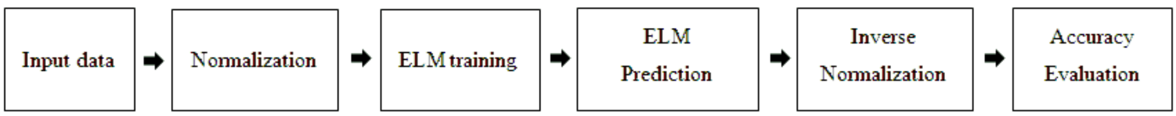

Fig. 2: The ELM modeling process

3.1 Training sample set

Same as the empirical models, the GMF established by ELM is also a kind of statistical model, the accuracy of which greatly depends on the training samples. Since it is difficult to obtain a complete scatterometer wind product in a specific period, we first use the CMOD5.N simulated data to demonstrate the effectiveness of this method, then we select a certain amount of ASCAT scatterometer wind data for training and analysis. This paper aims to explore a new method for fast and efficient modeling of geophysical model, rather than get a more practicable GMF.

3.1.1 Simulated data

The way to get the simulation sample is to give the value of wind speed, azimuth and incidence, we can compute the correspondent σ^0 using the CMOD5.N. In fact, the observation of scatterometer is often contaminated with noise due to various factors. In order to make the simulated data closer to the real state, we add about 0.2dB random perturbation to the σ^0 of the simulated data [7]. In this experiment, the polarization is fixed at VV polarization, wind speed sampling interval of 1m/s, in the range of $5 \sim 25m/s$; the range of incidence is $20° \sim 60°$ with a sampling interval $2°$; azimuth in the range of $0° \sim 360°$, the sampling interval is $10°$. Thus, a total of 16,317 samples can be generated.

3.1.2 Scatterometer observation

The scatterometer observation (contains information such as wind speed, wind direction, azimuth, incidence and normalized radar cross sections, etc.) used in this experiment is obtained from the ASCAT L2B product which provided by KNMI (Koninklijk Nederlands Meteorologisch Instituut). The model used in wind retrieval for the ASCAT product is CMOD5.N. The σ^0 values are compared to the CMOD5.N by means of a Maximum Likelihood Estimator (MLE) [8]. The wind vectors that give the best description of the σ^0 values (the solutions) are retained. Usually we get two or more solutions (which are called ambiguities) for each wind vector cell (WVC). To identify the most probable solution, we use a two-dimensional variational scheme (2DVAR) [9], which introduced Numerical Weather Prediction (NWP) wind background information [10]. Thus, the solution obtained does not strictly represent the optimal solution derived from the GMF. Therefore, we need to do some screening before using the ASCAT L2B product. First, we map a wind vector (specified in term of wind speed and wind direction) along with the measurement geometry (relative azimuth and incidence angle) and beam parameters (frequency and polarization) to a σ^0 value, and then compare it to the value of the correspond σ^0 observation. If the deviation of the two values falls into certain threshold, then the observation is remained. Assume the incidence, wind speed and azimuth of the sample data are θ,V,ϕ, then the σ^0 of the model can be expressed as:

$$\sigma^0_{ascat} = f(\theta, V, \phi) \tag{4}$$

Where f denotes the mapping between the parameters and the modeling σ^0. Supposed the maximum tolerated wind speed deviation difference set between the model and observed values is caused by ΔV, then the corresponding minimum and maximum values of the modeling σ^0 are [?]:

$$\sigma^0_{min} = f(\theta, V - \Delta V, \phi) \tag{5}$$

$$\sigma^0_{max} = f(\theta, V + \Delta V, \phi) \tag{6}$$

In the screening process, the observation is retained if the σ^0 of observation falls in the interval $[\sigma^0_{min}, \sigma^0_{max}]$, otherwise abandon it. In this experiment, the value of ΔV is set as 1 .09m/s based on the statistical precision of the ASCAT wind product [7]. After the screening, we get 29000 sample data of the ASCAT observations.

3.2 Data normalization

Before the training process using ELM method, we need to normalize the data to eliminate the effects caused by different dimensions. In the training set, there are three input vectors (θ, V, ϕ) and a target vector (σ^0). Suppose one of these vectors is $X = (x_1, x_2, x_n)$, all the data in X need to be normalized to a range of $[0, 1]$,

$$x_i^{normalize} = \frac{x_i - min(x)}{max(x) - min(x)} \tag{7}$$

In addition, we need to translate the output vector Y to its original state, which we call this procedure as inverse normalization:

$$y_i^{denormalize} = y_i(max(T) - min(T)) + min(T) \tag{8}$$

Where T represents the target vector.

3.3 The ELM modeling process

We need a certain amount of samples to train the new GMF using the ELM method. Assume the target vector of the sample is $T_{train} = [\sigma_1^0, \sigma_2^0, ...\sigma_n^0]^T$, the feature vector is $F_{train} = [\theta, V, \phi]$, where V is the wind vector, ϕ is the azimuth vector, θ is the incidence vector (all after normalization).

The weight matrix W_{train} is generated randomly according to the number of the hidden layer nodes which depends on the scale of the samples. The weight matrix W_{train} can be expressed as $W_{train} = [W_\theta, W_V, W_\phi]$, where W_θ,W_V and W_ϕ denote the weight vector of θ,V and ϕ respectively. Then we can obtain the I_{train} through:

$$I_{train} = F_{train} \circ W_{train} \tag{9}$$

Where $\circ$ is the Hadamard product.

Use I_{train} as the input matrix, T_{train} as the target vector, after training in ELM, we get a mapping matrix between I_{train} and T_{train}. Then we can compute the model target vector T_{model}, which after inverse normalization we can get the model σ^0.

4 Experiment

We compare the ELM method with the neural network (NN for short) method. We use two kinds of data set, one is the simulated data, and the other is the ASCAT data. Fig. 3 shows a comparison of σ^0 between the two model output value and the samples. The training accuracy of the two methods is rather high, while the ELM approach is better for both data sets. Although the number of ASCAT samples is bigger than simulated data, the training accuracy is lower than the latter for both methods. The reason may be that the samples of ASCAT observation are mainly distributed at the low wind while the simulated data is artificially well-distributed.

Compare the experimental results of two data sets, the model using the simulated data is in line with the CMOD5.N with a higher degree than the ASCAT observations. In Fig. 3(a), under the condition of wind speed 10m/s, 13m/s, 16m/s, 20m/s, the ELM prediction value is above the curve near the upwind direction. In 7m/s wind condition, ELM model has obvious systematic bias to the CMOD5.N, which the ELM model is totally beneath the CMOD5.N especially in the crosswind area. This is mainly because σ^0 is very sensitive to the noise at low wind, which also verified the theory mentioned in section 2 that the measurement accuracy of the σ^0 increases as the wind speed. However, the pattern occurs in figure 4(c) may due to an effect of overfitting of

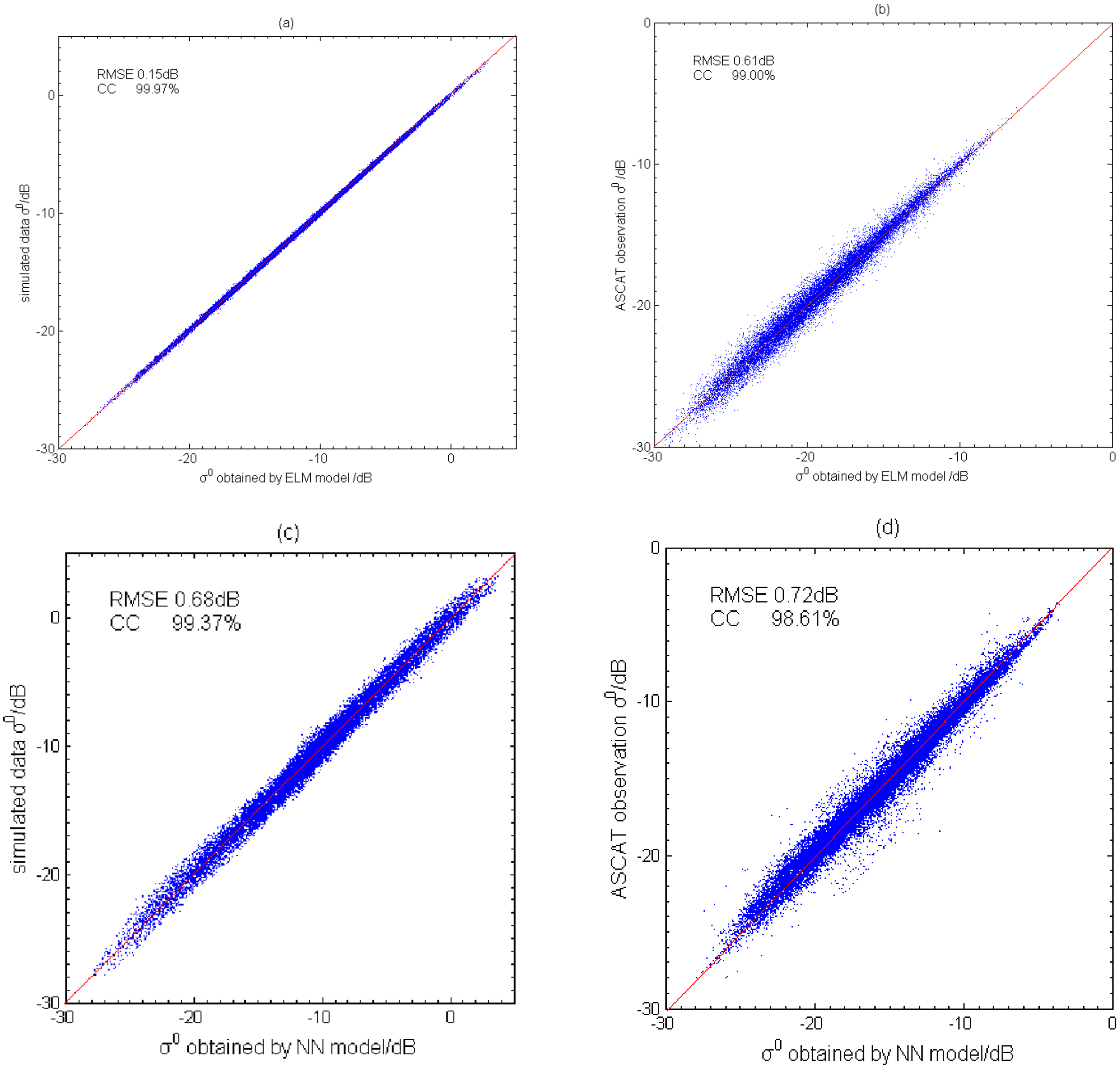

Fig. 3: Comparison of σ^0 between the model output and the samples: (a) the ELM simulated data experiment; (b) the ELM ASCAT observation experiment; (c)the nerual network simulated data experiment; (d)the nerual network ASCAT observation experiment

neural network training; The result of the ASCAT observation is a little divergent which is due to the uneven distribution of samples. Since the ratio of low wind less than 10m/s accounted for about 60% of the ASCAT data set, the predicted values fit well with the CMOD5.N curve at 4m/s, 7m/s wind speed in Fig. 4(b) and Fig. 4(d), while it is divergent at high speed wind due to lack of training data. But in general, the predicted value is substantially located in the vicinity of the curve.

5 Conclusion

In this paper, we use a new method called extreme learning machine for the modeling of the geophysical model function which is important for the wind retrieval of the scatterometer observations. We compare our method with the neural network approach and the result shows that the ELM works better than the latter in accuracy. The result for the simulated data indicates that using ELM for the modeling of GMF is feasible. However, due to the sparse distribution of the ASCAT data, the modeling result demonstrated in Fig 4(b) may not effectively reflect the real state of GMF. Besides, the number of the training data is quite small compared with the real ASCAT observations which usually reach a magnitude of millions. These problems may need further research in our next step.

References

[1] Naderi, FM, Freilich, MH. Space borne radar measurement of wind velocity over the ocean: An overview of the NSCAT scatterometer system[J]. Proceedings of the IEEE, 1991, 79(6): 850 - 866.

[2] Zou Juhong, LIN Mingsen, Pan Delu. A unified geophysical model of C -band and Ku-band based on neural networks[J]. Oceanography, 2008, 30(5): 23-28.

[3] Guang-Bin Huang, Dian Hui Wang, Yuan Lan. Extreme learning machines: a survey[J]. Int. J. Mach. Learn. Cyber. 2011, 2:107122.

[4] Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew. Extreme learning machine: Theory and applications[J]. Neurocomputing, 2006, 70:489501.

[5] Hans Hersbach. CMOD5-An improved geophysical model function for ERS C-band scatterometry[R]. European Centre for Medium Range Weather Forecasts.2003:47

[6] Anton Verhoef, Marcos Portabella and Ad Stoffelen. CMOD5.n - the CMOD5 GMF for neutral winds [T]. Ocean and Sea Ice SAF.2008:13

[7] Jeroen Verspeek, Anton Verhoef, Ad Stoffelen. ASCAT-B NWP Ocean Calibration and Validation[R]. Ocean and Sea Ice SAF. 2013: 5-6.

[8] Chi, Li F K. A comparative study of several wind estimation algorithms for space borne scatterometers. IEEE Transactions on Geoscience and Remote Sensing, 1988, 26(2): 115-121.

[9] Jur Vogelzang. Two dimensional variational ambiguity removal (2DVAR)[T]. Satellite Application Facility for Numerical Weather Prediction.2007

[10] Anton Verhoef, Jur Vogelzang, Jeroen Verspeek, Ad Stoffelen. AWDP User Manual and Reference Guide[R]. Satellite Application Facility for Numerical Weather Prediction. (2013)6:7-9

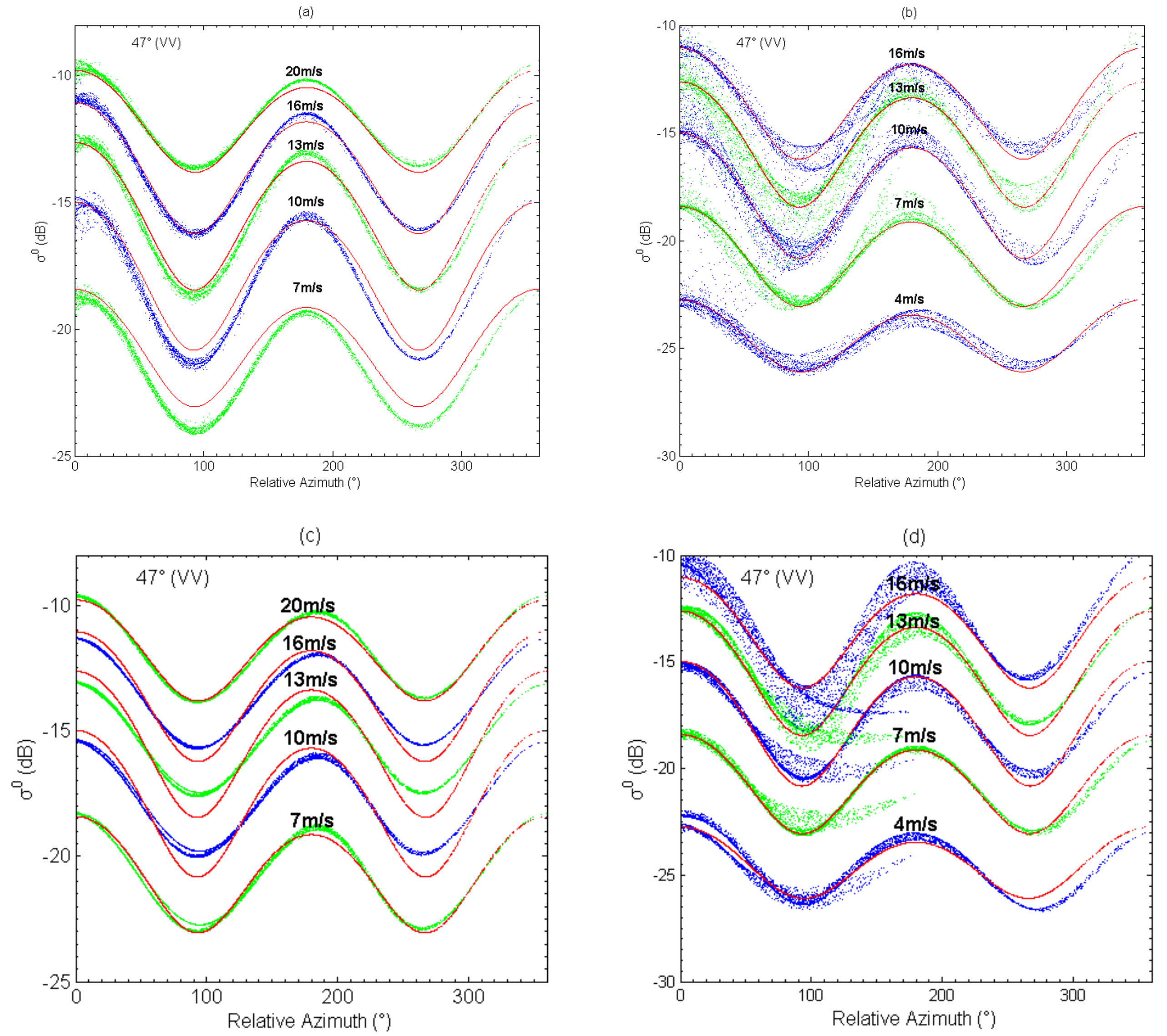

Fig. 4: Comparison of σ^0 between model output and CMOD5.N for various wind speeds at an incidence angle of 47°: (a) the ELM simulated data experiment; (b) the ELM ASCAT observation experiment; (c) the nerual network simulated data experiment; (d) the nerual network ASCAT observation experiment. Red curve denotes σ^0 of CMOD5.N. Blue and green scattering dots represent σ^0 of model at different wind speeds.