(Received 25 September 2023, Accepted 15 January 2024, Scheduled 13 February 2024)

# Weather Radar High-Resolution Spectral Moment Estimation Using Bidirectional Extreme Learning Machine

Zhongyuan Wang[1,*], Ling Qiao[1], Yu Jiang[2], Mingwei Shen[1], and Guodong Han[3]

[1] College of Information Science and Engineering, Hohai University, Nanjing, China
[2] Unit 95438 of PLA Air Force, Meishan, China
[3] The 54th Research Institute of CETC, Shijiazhuang, China

**ABSTRACT:** Since the performance of the spectral moment estimation algorithm commonly used in engineering degrades under the conditions of low SNR, this paper introduces the Extreme Learning Machine (ELM) to the spectral moment estimation of weather signals based on the correlation of the signals of adjacent range cells. To solve the problem that the hidden layer nodes of ELM algorithm are difficult to be determined, the Bidirectional Extreme Learning Machine (B-ELM) algorithm is applied to achieve the high resolution of spectral moments. Firstly, to improve the SNR of the training samples, time-domain pulse signals are converted into weather power spectrum by Welch method. Then, the parameters of the B-ELM hidden layer nodes are directly calculated by backpropagation of network residuals. The model parameters are optimized according to the least-squares solution, where the optimal number of hidden layer nodes is determined adaptively. Finally, the optimized B-ELM model is employed for the spectral moment estimation of weather signals. The algorithm is validated to be fast and accurate for spectral moment estimation using the measured IDRA weather radar data and is easy to implement in engineering.

Keywords: Weather radar, Spectral moment estimation, B-ELM

## 1. INTRODUCTION

With the further development of weather radar, radar has become an important tool for atmospheric remote sensing detection and weather detection. It plays an important role in severe weather warnings, short-time weather monitoring and forecasting, public service, disaster prevention and mitigation [1–3]. Weather target information in weather echoes is critical in weather radar systems. The spectral moment information of weather signals is an essential parameter of weather target information, which is an indispensable basis for judging the type of weather targets and the need to issue warnings for weather states. Spectral moments mainly include average power, average Doppler velocity, and Doppler spectral width [4]. Average power corresponds to the reflectivity of the weather target, which reflects the content of liquid water and rainfall rate in the region. Average Doppler velocity is the sum of the average velocity of each scattered particle. Doppler spectral width represents the dispersion of the scattered particle velocity and indicates the intensity of the weather target motion.

Traditional algorithms such as Pulse Pair Processing method [5] and fast Fourier transform method [6] under the conditions of low signal-to-noise ratio (SNR) are unable to meet the high accuracy requirements of weather radar systems for spectral moments because of their seriously degraded performance. In recent years, spectral moment estimation methods based on parametric model [7, 8] and sparse reconstruction [9] have been applied and studied. These two algorithms have greatly improved the accuracy of spectral moment estimation. However, the computational complexity of the two mentioned algorithms is high, and the timeliness is not strong.

At present, machine learning has a great application prospect in the field of meteorology [10, 11]. This paper introduces an extreme learning machine in the field of machine learning into the spectral moment estimation of weather signals and studies a spectral moment estimation algorithm based on ELM. The algorithm firstly constructs training samples, then selects and optimizes the model parameters, and finally uses the optimized ELM algorithm to build a prediction model. For the issue that the number of hidden layer nodes is difficult to determine, the ELM algorithm is further improved, and the Bidirectional Extreme Learning Machine (B-ELM) algorithm is investigated. The network structure is modified by gradually adding hidden layer nodes. The parameters of some hidden layer neurons are calculated directly using the network residuals backpropagation. A model for estimating the spectral moments of weather signals is established to improve the accuracy of spectral moment estimation. The contributions of this paper are as follows:

(1) ELM is applied to the estimation of spectral moments in weather signals, and a spectral moment estimation model is established.

(2) Time-domain pulses are transformed into Welch power spectrum to improve the SNR of subsequent B-ELM input weather signals.

* Corresponding author: Zhongyuan Wang (wzy_200101@163.com).

(3) For the difficulty of determining the hidden layer nodes of ELM algorithm, the B-ELM algorithm is further investigated to achieve the high accuracy estimation of spectral moments.

This paper is organized as follows. In Section 2, we preprocess the signal to the B-ELM. The principles and theoretical derivation of ELM and B-ELM algorithms are introduced in Section 3. Some simulation experiments are presented in Section 4. Finally, the conclusion of this paper is given in Section 5.

## 2. SIGNAL PREPROCESSING FOR B-ELM

### 2.1. Estimation of the Power Spectrum

To estimate the spectral moment information of the weather signals with high accuracy, the time-domain pulse signals of each range cell are transformed into the weather power spectrum by Welch method, which significantly improves the input SNR of the weather signals.

Suppose that the weather pulse data $u(n)$ is divided into $L$ segments, and each segment length is $K$. The offset of the segment signal is $D$. The $i$th segment signal is

$$u_i(k) = u(k + iD), \quad (0 \le k \le K-1, 0 \le i \le L-1) \quad (1)$$

where the case of overlapping data splitting occurs when $D < K$.

An appropriate window function is introduced before calculating the power spectrum. The energy of the window function is

$$U = \frac{1}{K}\sum_{k=0}^{K-1} w^2(k) \quad (2)$$

The power spectrum corresponding to each segment of data is calculated by

$$P_{per}^{i}(e^{j\omega}) = \frac{1}{KU}\left|\sum_{k=0}^{K-1} u_i(k)w(k)e^{-j\omega k}\right|^2 \quad (3)$$

The power spectrum estimate of $u(n)$ is obtained by adding and averaging the power spectra of the $L$ segments

$$P(e^{j\omega}) = \frac{1}{L}\sum_{i=0}^{L-1} P_{per}^{i}(e^{j\omega}) \quad (4)$$

After transforming the weather signals into a power spectrum by Welch method, the input SNR is improved, which enhances the accuracy of the spectral moment estimation. Both the traditional spectral moment estimation algorithm and its improved algorithm utilize the impulse information within each range cell for numerical solution, ignoring the correlation between different range cells. In order to further improve the accuracy of spectral moment estimation, it is necessary to discuss how to use correlation to complete the estimation of spectral moment information.

### 2.2. Construction of Training Samples for B-ELM

This paper utilizes the correlation of Welch power spectrum of adjacent range cells to estimate the spectral moments of the current cell. Therefore, in this section, Pearson correlation coefficient is used to verify the correlation of Welch power spectrum. Pearson correlation coefficient is one of the most commonly used expressions for correlation coefficient, which reflects the degree of direct correlation between two different variables, with the following formula:

$$cov(P_i, P_j) = \frac{\sum_{n=1}^{N}(P_i(n) - \bar{P}_i)(P_j(n) - \bar{P}_j)}{N-1} \quad (5)$$

$$r = \frac{cov(P_i, P_j)}{\zeta_{P_i}\zeta_{P_j}} \quad (6)$$

where $r$ is the Pearson correlation coefficient; $P_i = [P_i(1), P_i(2), \ldots, P_i(N)]$ is the Welch power spectrum of the $i$th range cell; $P_j = [P_j(1), P_j(2), \ldots, P_j(N)]$ is the Welch power spectrum of the $j$th range cell; $\bar{P}_i$ and $\bar{P}_j$ are the mean values of $P_i$ and $P_j$, respectively. $\zeta_{P_i}$ and $\zeta_{P_j}$ are the standard deviation of $P_i$ and $P_j$, respectively. The equation represents the correlation between the weather signal of the $i$th range cell and the weather signal of the $j$th range cell. The coefficient ranges from $-1$ to 1, and the absolute value of the coefficient is proportional to its correlation. In addition, when $r \ge 0.6$, it shows a strong correlation.

Taking the 30th range cell as an example, the correlation results of weather signals and Welch power spectrum are shown in Fig. 1(a) and Fig. 1(b), respectively. Each value on the ordinate of the figure represents the degree of correlation between the 30th range cell and the 1st to 60th range cells. It can be clearly seen from Fig. 1 that Welch power spectrum of different range cells has strong correlation characteristics compared with the initial weather signals. Moreover, the correlation between the current range cell and the nearby range cell is stronger.

The prerequisite for the application of B-ELM is a high correlation between the training samples. According to the above correlation analysis, it can be seen that the current range cell has a stronger correlation with the adjacent range cell, and more information can be utilized to the traditional spectral moment estimation algorithms. Therefore, in order to obtain high resolution spectral moment estimation, the power spectra of the adjacent range cells are taken as training samples, and the estimated spectral moments of the current range cell are taken as the output.

Assuming that the $a$th range cell requires spectral moment estimation, the training samples can be taken as $a-N_L, \ldots, a-1, a+1, \ldots, a+N_L$th range cell. For $2N_L$ different training samples $(t_i, y_i)$ $(i = 1, 2, \ldots, 2N_L)$, their construction form is

$$t_i = [t_{i1}, t_{i2}, \ldots, t_{is}]^T \quad (7)$$

$$y_i = [y_{i1}, y_{i2}, \ldots, y_{im}]^T \quad (8)$$

where $t_i$ is the input, $y_i$ the output, $s$ the number of pulses, i.e., the number of nodes in the input layer, and $m$ the number of nodes in the output layer. In this algorithm, $m = 1$, so the

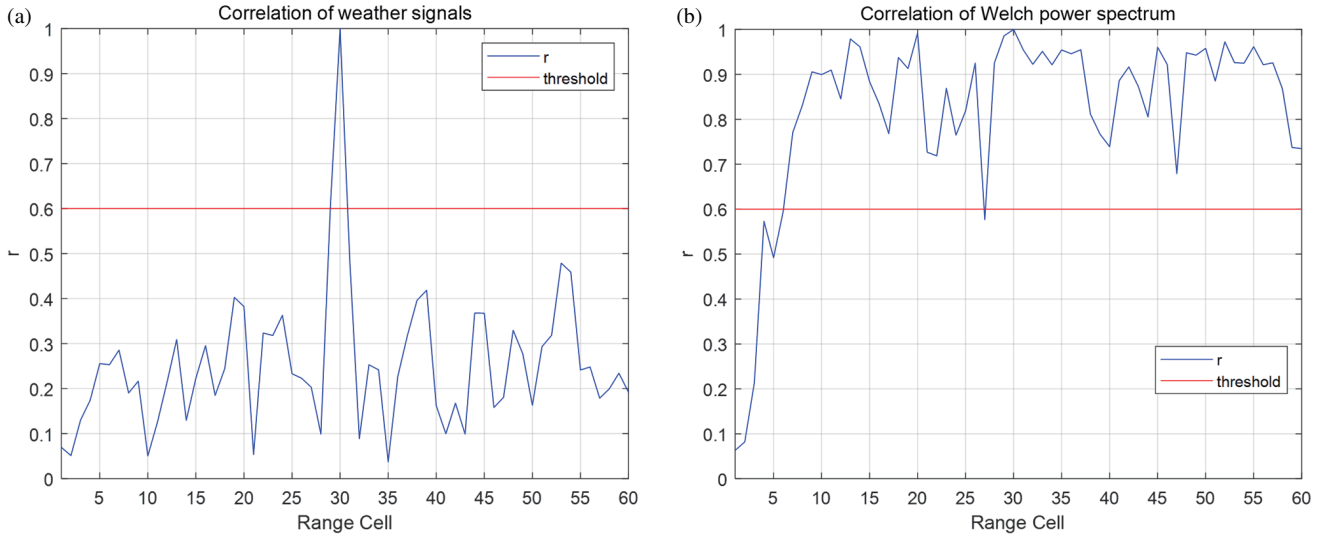


**FIGURE 1**. Correlation of each range cell with the 30th range cell: (a) Correlation of weather signals. (b) Correlation of Welch power spectrum.

input becomes

$$t_i = [t_{i1}, t_{i2}, \ldots, t_{is}]^T = \begin{cases} [P_1^{a-N_L+(i-1)}, \ldots, P_s^{a-N_L+(i-1)}]^T & 1 \le i \le N_L \\ [P_1^{a-N_L+i}, \ldots, P_s^{a-N_L+i}]^T & N_L < i \le 2N_L \end{cases} \quad (9)$$

The spectral moment as an output can be expressed as

$$y_i = \begin{cases} \kappa_{a-N_L+(i-1)} & 1 \le i \le N_L \\ \kappa_{a-N_L+i} & N_L < i \le 2N_L \end{cases} \quad (10)$$

where $P_s^{a-N_L+(i-1)}$ is the spectrum of the $s$th pulse of the $a - N_L + (i-1)$th range cell, and $\kappa_{a-N_L+(i-1)}$ is the spectral moment of the $a - N_L + (i-1)$th range cell.

In this paper, the Welch power spectra of different range cells under low SNR conditions are used as the training samples for the algorithm. Note that the Welch power spectrum contains information on radial velocity and spectral width.

## 3. B-ELM-BASED SPECTRAL MOMENT ESTIMATION ALGORITHM

### 3.1. ELM Model

ELM is a new type of single-hidden layer feedforward neural network (SLFN) [12], which is used in many fields [13, 14]. The ELM network structure is shown in Fig. 2.

For $2N_L$ training samples $(t_i, y_i)$ $(i = 1, 2, \ldots, 2N_L)$, the mathematical model of ELM with $J$ hidden layer nodes and the activation function of $g(x)$ has the following form

$$\sum_{j=1}^{J} \beta_j g(w_j \cdot t_i + b_j) = o_i \quad i = 1, 2, \ldots, 2N_L \quad (11)$$

where $w_j = [w_{j1}, w_{j2}, \ldots, w_{js}]^T$ is the weight vector of the $j$th hidden layer node and the input layer; $\beta_j = [\beta_{j1}, \beta_{j2}, \ldots, \beta_{jm}]^T$ is the weight vector of the $j$th hidden layer node and the output layer; $b_j$ is the bias of the $j$th hidden layer node; $o_i$ is the network output value of the $i$th sample; and $w_j \cdot t_i$ denotes the inner product of $w_j$ and $t_i$.

For SLFN with activation function $g(x)$ containing $J$ hidden layer nodes, these $2N_L$ training samples can be approximated as zero error as follows

$$\sum_{j=1}^{J} ||o_j - y_j|| = 0 \quad (12)$$

Then there exist suitable $w_j$, $b_j$, $\beta_j$, such that

$$\sum_{j=1}^{J} \beta_j g(w_j t_i + b_j) = y_i \quad i = 1, 2, \ldots, 2N_L \quad (13)$$

Eq. (11) can be abbreviated as

$$H\beta = Y \quad (14)$$

where

$$H = \begin{bmatrix} g(w_1 t_1 + b_1) & \cdots & g(w_J t_1 + b_J) \\ \vdots & \ddots & \vdots \\ g(w_1 t_{2N_L} + b_1) & \cdots & g(w_J t_{2N_L} + b_J) \end{bmatrix} \quad (15)$$

$H$ is called the hidden layer output matrix. Huang [15] has proved that when the activation function is infinitely differentiable, it is not necessary to completely adjust the network parameters. At the beginning of training, the weight vector connecting the hidden node and the input node and the threshold of the hidden node can be randomly selected. The output connection weight can be obtained by solving the least-squares solution of the linear equation. By solving the linear equation with the least-squares solution, the following can be obtained

$$\hat{\beta} = H^+ Y \quad (16)$$

where $H^+$ is the Moore-Penrose generalized inverse of $H$.

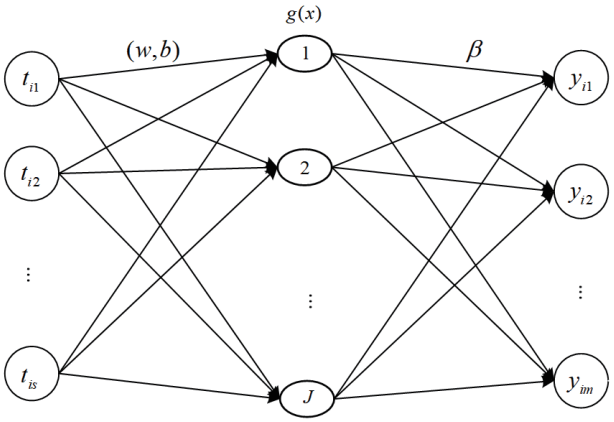


**FIGURE 2**. ELM network structure.

### 3.2. B-ELM Model

ELM algorithm has the advantages of fast learning speed and good generalization performance, but its biggest drawback is that it must manually set the hidden layer nodes. The number of hidden layer nodes increases the performance of the algorithm, but not the more the number of hidden nodes the better the performance of the algorithm, which often requires a lot of experiments to determine the number of hidden nodes. Therefore, it is necessary to improve the ELM algorithm. On the basis of the ELM algorithm, the spectral moment estimation algorithm based on the Bidirectional Extreme Learning Machine (B-ELM) is investigated.

Unlike the traditional ELM and I-ELM [16] algorithms, B-ELM finds the parameters of the hidden layer nodes by error feedback. Both B-ELM and I-ELM algorithms modify the network structure by adding new hidden layer nodes. B-ELM algorithm is consistent with I-ELM algorithm when the number of hidden layer nodes is odd. When the number of hidden layer nodes is even, the hidden layer node parameters are obtained by calculating the error feedback [17, 18].

The network residuals of B-ELM are calculated as

$$E_J = E_{J-1} - \beta_J H_J \tag{17}$$

where $E_{J-1}$ is the network error before adding the $J$th hidden layer node. When the hidden layer nodes are odd

$$\beta_J = \frac{E_{J-1} \cdot H^T}{H \cdot H^T} = \frac{\sum_{i=1}^{2N_L} e(i) \cdot g(w_J \cdot t_i + b_J)}{\sum_{i=1}^{2N_L} g^2(w_J \cdot t_i + b_J)} \tag{18}$$

$$H = \begin{bmatrix} g(w_J \cdot t_1 + b_J) \\ \vdots \\ g(w_J \cdot t_{2N_L} + b_J) \end{bmatrix} \tag{19}$$

where $e(i)$ is the error of the $i$th training sample before the new hidden layer node is added, and $H$ is the output matrix of the new hidden layer node.

When the hidden layer nodes are even, the error feedback matrix from the last iteration is calculated

$$H^e = E_{J-1} \cdot (\beta_{J-1})^{-1} \tag{20}$$

where $\beta_{J-1}$ is the weight between the hidden layer nodes and the output layer after the last iteration. The error feedback matrix is then used to calculate the hidden layer node parameters $w_J, b_J$

$$w_J = g^{-1}\left(u\left(H^e\right)\right) \cdot t^{-1} \tag{21}$$

$$b_J = \sqrt{mse\left(g^{-1}\left(u\left(H^e\right)\right) - w_J \cdot t\right)} \tag{22}$$

where $mse(\cdot)$ denotes the variance, and $u : R \to [0, 1]$ is the normalization function. The output matrix is updated by the above calculation

$$H = u^{-1}\left(g\left(w_J \cdot t + b_J\right)\right) \tag{23}$$

Then the weight $\beta_J$ between the hidden layer nodes and output layer is calculated as

$$\beta_J = \frac{E_{J-1} \cdot H^T}{H \cdot H^T} \tag{24}$$

Figure 3 shows the flow of the B-ELM algorithm. In the B-ELM algorithm, when the number of hidden layer nodes is odd, the relevant parameters are obtained randomly, which is the same as the I-ELM algorithm. When the number of hidden layer nodes is even, the parameters of some hidden layer nodes are calculated directly by using the back propagation of network residuals. This avoids the appearance of "useless hidden layer nodes", minimizes the computational effort and time caused by searching for the optimal hidden layer node parameters, and greatly improves the computational speed of the model.

### 3.3. Optimization of B-ELM Parameters

Different from ELM networks, B-ELM does not require manual setting of the number of hidden layer nodes. B-ELM algorithm can achieve optimization of the network by self-increasing hidden layer nodes during the operation of the algorithm. The main

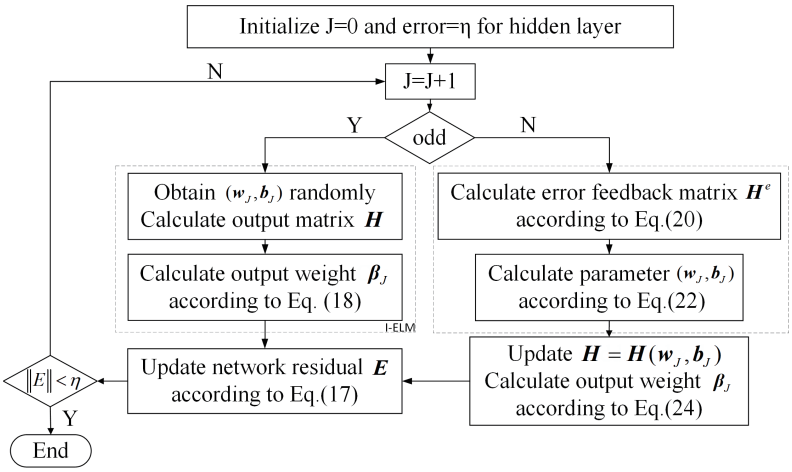


**FIGURE 3**. Flow chart of B-ELM algorithm.

**TABLE 1**. IDRA radar parameters table.

| Parameter name | Symbol | Parameter value |
|---|---|---|
| Scanning time | $T_p$ | 409.6 μs |
| Carrier frequency | $f_c$ | 9.475 GHz |
| Wavelength | $\lambda$ | 0.03 m |
| Number of pulses | $K$ | 512 |
| Bandwidth | $B$ | 5 MHz |
| Number of range cells | $L$ | 100 |

parameter that affects the performance of the B-ELM algorithm is the activation function, which is studied in this section.

The activation function is a function that runs on the neurons of an artificial neural network and is responsible for mapping the input to the output of the neuron. The commonly used activation functions are Sigmoid [19], Hardlim [20], ReLu [21], etc. This section will analyze the effect of different activation functions on the performance of the B-ELM model.

In this paper, absolute error (error) and root mean square error (RMSE) are used to analyze the accuracy. The formula for error and RMSE can be expressed as:

$$error = |y_i - y_{obj}| \tag{25}$$

$$RMSE = \sqrt{\frac{1}{N_m}\sum_{i=1}^{N_m}(y_i - y_{obj})^2} \tag{26}$$

where $N_m$ is the number of range cells, and $y_i$ and $y_{obj}$ are the estimated and true values of the $i$th range cell, respectively.

Figure 4(a) depicts the effects of different activation functions and the number of training samples on the estimation performance of B-ELM. It can be seen from the figure that when the Sigmoid function is used as the activation function, the estimation error of this network model is the smallest. It can also be seen that when the number of training samples $N_L$ is 1–3, the estimation performance of the B-ELM network under the same activation function is basically the same.

Figure 4(b) is similar to Fig. 4(a), which shows the spectral width of the weather signal. As seen in the figure, the difference in the performance of the B-ELM algorithm for spectral width estimation is large for different activation functions. Among them, the Sigmoid function has the best B-ELM estimation performance, and the ReLu function is the worst. Combined with the above analysis, the Sigmoid function is the most suitable activation function for B-ELM for spectral moment estimation. Therefore, the Sigmoid function is chosen as the activation function of both B-ELM and ELM algorithms in this paper to estimate spectral moments. In addition, the spectral width estimation is optimal when training samples $N_L$ is set to 3, so the total number of training samples $2N_L$ is set to 6.

## 4. PERFORMANCE ANALYSIS

To verify the effectiveness of the algorithm, IRCTR drizzle radar (IDRA) [22] measured data are used in this paper to compare the accuracy of spectral moment estimation of different algorithms. IDRA is a weather radar system designed and developed by the International Research Centre for Telecommunications and Radar (IRCTR) of Delft University of Technology. The data selected in this paper are from September 10, 2011, 21:00–22:00 UTC, and in standard range mode, using measured data from sector 25. Table 1 shows the parameters of the IDRA radar in standard mode.

Traditional spectral moment estimation algorithms have high estimation accuracy under high SNR conditions. To test the effectiveness of the algorithm, the spectral moments obtained by using the FFT method for 512 weather radar pulses are used as the real values. 64 pulses are taken from 512 pulses as raw data

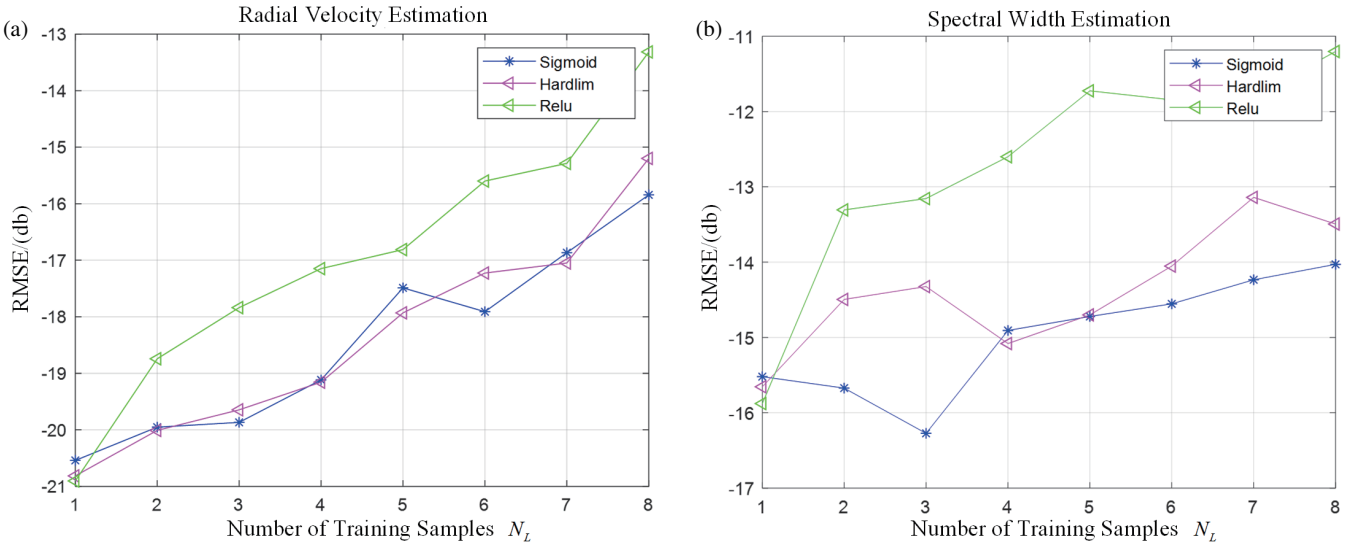


**FIGURE 4**. Performance of activation functions: (a) radial velocity, (b) spectral width.

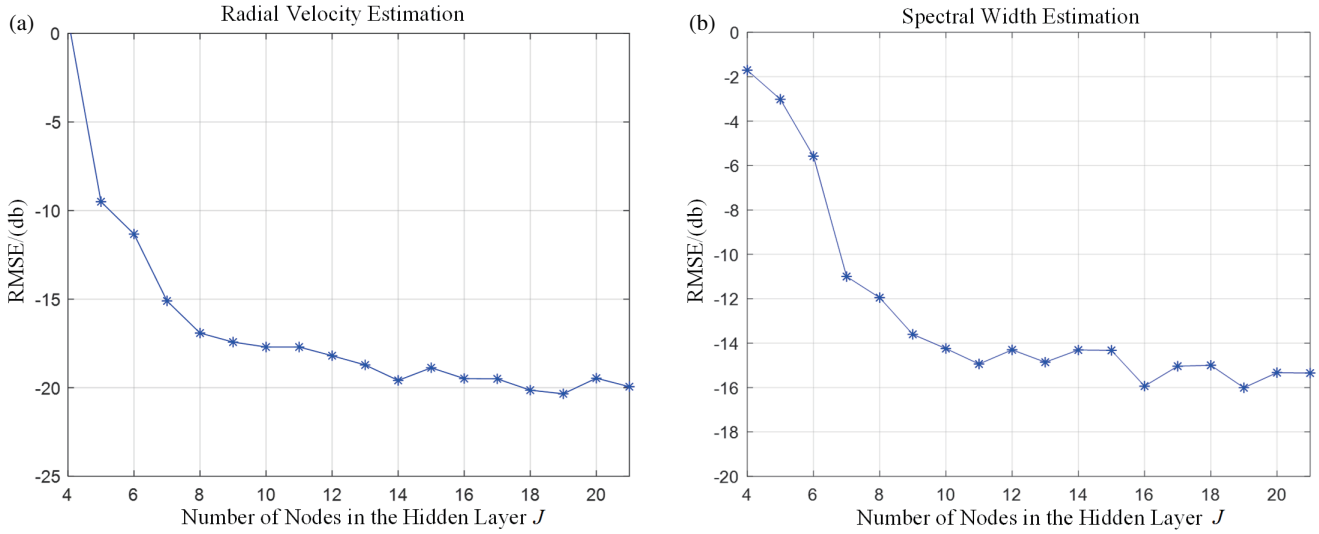


**FIGURE 5**. Effect of the number of hidden layer nodes: (a) radial velocity estimation error, (b) spectral width estimation error.

to verify the algorithm's performance under low SNR conditions. The SNR is reduced by a factor of 8 by taking 64 pulses from 512 pulses. The following experimental results are based on 100 measured weather radar data.

### 4.1. Impact of Hidden Layer Nodes on ELM Performance

Two main parameters affecting the performance of the ELM algorithm: the activation function and the number of hidden layer nodes, of which the activation function has been studied in Section 3. In the following, the effect of the number of hidden layer nodes on the performance of ELM will be briefly studied.

Figure 5(a) demonstrates the radial velocity estimation after applying the ELM algorithm. As can be seen, when the value of $J$ is less than 14, the RMAE error of radial velocity estimation decreases significantly with the increase of the number of hidden layer nodes. When the value of $J$ is more than or equal to 18, the RMAE error of radial velocity estimation is basically controlled to be around $-20$ dB. Therefore, when the output of ELM is radial velocity, the number of nodes in the hidden layer $J$ is set to 18.

Figure 5(b) illustrates the spectral width estimation after applying the ELM algorithm. When $J$ is less than 11, the RMAE error of the spectral width estimation decreases significantly with the increase in the number of nodes in the hidden layer. When $J$ is more than or equal to 16, the RMAE error is basically controlled at about $-15$ dB. So, when the output of ELM algorithm is spectral width, $J$ is set to 16.

### 4.2. Performance Analysis of B-ELM Algorithm

Figures 6(a) and 6(b) show the comparison between the estimators and the true values of spectral moments for the ELM and B-ELM algorithms, respectively. The true values of radial

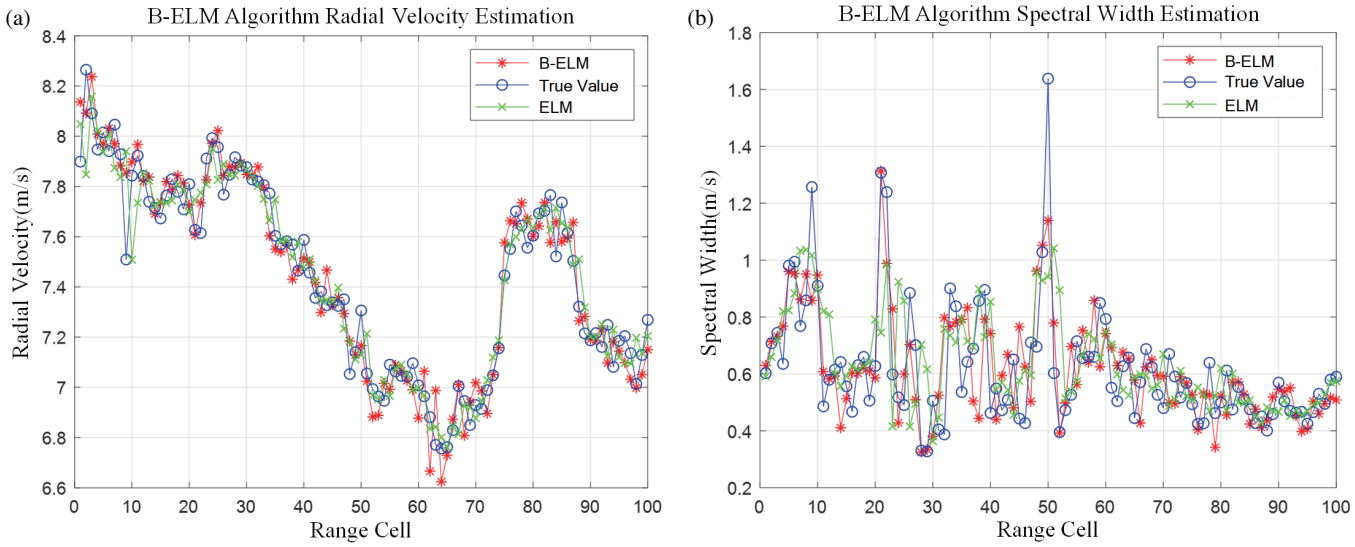


**FIGURE 6**. Performance of B-ELM algorithm: (a) radial velocity, (b) spectral width.

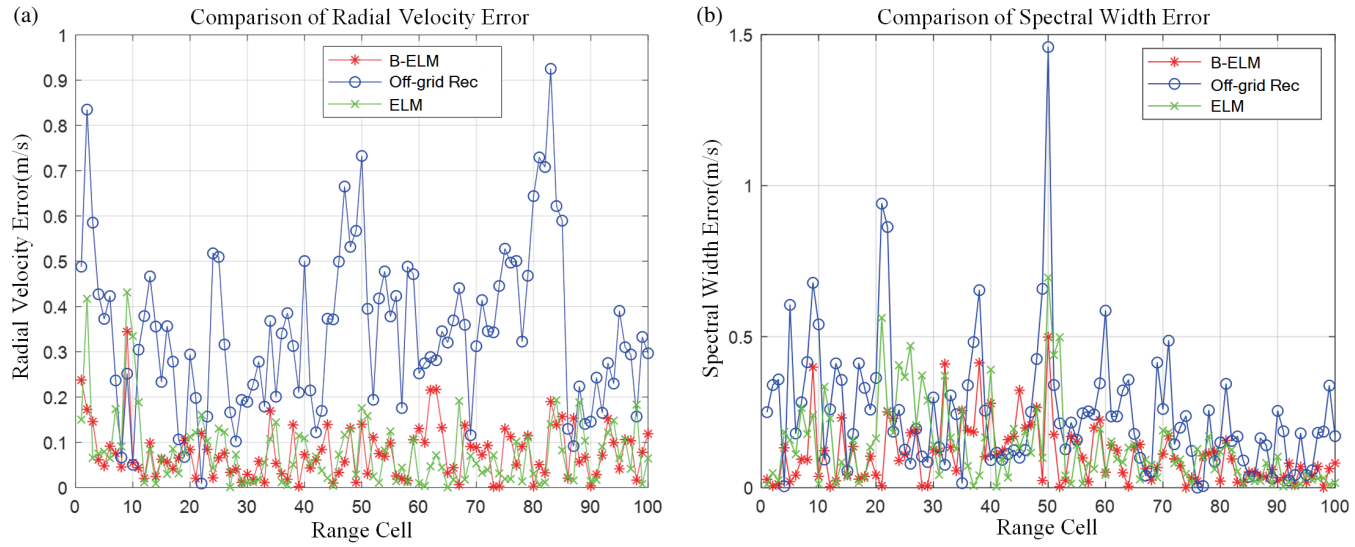


**FIGURE 7**. Performance comparison of three algorithms: (a) radial velocity error, (b) spectral width error.

velocity vary slowly between 6.6 and 8.3 m $s^{-1}$. It is obvious from Fig. 6(a) that the radial velocity estimators of these two algorithms are very close to the true values and vary slow in that range. The radial velocity slowly decreases at 30–60 range cells and increases rapidly at 70–80 range cells. These two algorithms can perform accurate spectral moment estimation when the radial velocity changes rapidly. From Fig. 6(b), it can be seen that the true values of spectral width change slowly between 0.3 and 1.7 m $s^{-1}$. Most of the spectral width values are distributed between 0.4 and 0.8 m $s^{-1}$. The deviation of ELM and B-ELM algorithms from the true values is very small.

Figures 7(a) and 7(b) show the absolute errors between the estimators and true values of the three algorithms. The off-grid reconstruction algorithm [23] is introduced to compare with the algorithm of this paper. The blue, red, and green lines indicate the absolute values of the errors of the off-grid reconstruction, B-ELM and ELM algorithms, respectively.

Figure 7(a) shows the comparison of the absolute value of the error of the radial velocity. It can be seen that the estimation error of the off-grid reconstruction algorithm is much larger than that of the ELM and B-ELM algorithms in most of the range cells. And the estimation error value of the B-ELM algorithm is basically less than 0.2 m/s with the smallest error value and the most accurate estimation. Fig. 7(b) compares the absolute value of the error of the spectral width. It can be seen that the estimation error of the algorithm in the off-grid reconstruction algorithm is much larger than that of the ELM and B-ELM algorithms in most of the range cells. And the estimation error value of the B-ELM algorithm is basically less than 0.5 m $s^{-1}$, which is close to the estimation performance of the ELM algorithm.

**TABLE 2**. Computational complexity analysis.

| Different algorithm | Radial velocity average error ($m\ s^{-1}$) | Spectral width average error ($m\ s^{-1}$) |
|---|---|---|
| Off-Grid reconstruction | 0.3238 | 0.2501 |
| ELM | 0.0825 | 0.1198 |
| B-ELM | 0.0766 | 0.1063 |

Table 2 compares the average error of spectral moment estimation of different algorithms. In this paper, the spectral moments obtained by the fast Fourier transform (FFT) method with 512 pulses are taken as the true values. Table 2 shows that the estimation performance of the B-ELM algorithm is slightly better than that of the ELM algorithm, and the estimation performance of the off-grid reconstruction algorithm is the worst. The average error of radial velocity estimation of the B-ELM algorithm is 0.0766 m $s^{-1}$, and the error of spectral width estimation is 0.1063 m $s^{-1}$, improving performance by 7% and 11%, respectively, compared with the ELM algorithm. The radial velocity estimation of B-ELM algorithm improves 76.3%, and spectral width estimation improves by 57.4% compared with the off-grid reconstruction algorithm.

## 5. CONCLUSION

In this paper, an extreme learning machine is introduced into the estimation of spectral moments of weather signals. For the problem that the hidden layer nodes of the ELM algorithm are difficult to determine, Bidirectional Extreme Learning Machine (B-ELM) algorithm is investigated. B-ELM algorithm increases the hidden layer nodes one by one. When the number of nodes is odd, the current-hidden-layer node parameters are set randomly. When the number of nodes is even, the parameters of some hidden layer nodes are calculated directly by using the network residual backpropagation. The model parameters are optimized according to the least-squares solution. The experimental results show that B-ELM algorithm can accurately estimate spectral moment. Compared with other algorithms, the algorithm in this paper has the smallest error and low computational complexity, which is easy for engineering implementation.

## ACKNOWLEDGEMENT

This work was partly supported by the National Natural Science Foundation of China [No. 62271190], Natural Science and Foundation of Jiangsu Province of China [BK20221499] and Science and Technology on Electronic Information Control Laboratory.

## REFERENCES

[1] Klimenko, D. E., "Studying the areal rainfall reduction in the Urals based on radar data," *Russian Meteorology and Hydrology*, Vol. 44, 484–493, 2019.

[2] Da Silva, F. P., O. C. R. Filho, M. G. A. J. d. Silva, R. J. Sampaio, G. D. Pires, and A. A. M. d. Araújo, "Real-time river level estimation based on variations of radar reflectivity — A case study of the Quitandinha River watershed, Petrópolis, Rio de Janeiro (Brazil)," *Bulletin of Atmospheric Science and Technology*, Vol. 2, No. 1, 2021.

[3] Li, B., "Development and challenge of weather radar technology in China," *Progress in Meteorological Science and Technology*, Vol. 12, No. 5, 37–46, 2022.

[4] Zhao, J., X. G. Lu, H. Li, and Z. Zhang, "A fast spectral moments estimation approach of radar echoes with low signal-to-noise ratio," *Signal Processing*, Vol. 36, No. 5, 703–709, 2020.

[5] Miller, K. and M. Rochwarger, "A covariance approach to spectral moment estimation," *IEEE Transactions on Information Theory*, Vol. 18, No. 5, 588–596, 1972.

[6] Dias, J. M. B. and J. M. N. Leitao, "Nonparametric estimation of mean doppler and spectral width," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No. 1, 271–282, 2000.

[7] Lu, X., Z. Zhang, and P. Han, "Low complexity spectral moments estimator under low SNR condition," in *2016 Integrated Communications Navigation and Surveillance (ICNS)*, Herndon, VA, 2016.

[8] Boyer, E., M. Petitdidier, and P. Larzabal, "Stochastic Maximum Likelihood (SML) parametric estimation of overlapped Doppler echoes," in *Annales Geophysicae*, Vol. 22, No. 11, 3983–3993, 2004.

[9] Li, F., S. Hong, Y. Gu, and L. Wang, "An optimization-oriented algorithm for sparse signal reconstruction," *IEEE Signal Processing Letters*, Vol. 26, No. 3, 515–519, 2019.

[10] Zhang, C., H. Wang, J. Zeng, L. Ma, and L. Guan, "Short-term dynamic radar quantitative precipitation estimation based on wavelet transform and support vector machine," *Journal of Meteorological Research*, Vol. 34, 413–426, 2020.

[11] Zhang, Y., Z. Ji, B. Xue, and P. Wang, "A novel fusion forecast model for hail weather in plateau areas based on machine learning," *Journal of Meteorological Research*, Vol. 35, No. 5, 896–910, 2021.

[12] Nilesh, R. and W. Sunil, "Improving extreme learning machine through optimization a review," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 906–912, Coimbatore, India, 2021.

[13] Yaseen, Z. M., S. O. Sulaiman, R. C. Deo, and K.-W. Chau, "An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction," *Journal of Hydrology*, Vol. 569, 387–408, 2019.

[14] Ahmad, I., M. Basheri, M. J. Iqbal, and A. Rahim, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE Access*, Vol. 6, 33 789–33 795, 2018.

[15] Huang, G.-B., "Learning capability and storage capacity of two-hidden-layer feedforward networks," *IEEE Transactions on Neural Networks*, Vol. 14, No. 2, 274–281, 2003.

[16] Abramova, E. S., A. A. Orlov, and K. V. Makarov, "Research of the extreme learning machine as incremental learning," in *2022 International Conference on Industrial Engineering, Applica-*

*tions and Manufacturing (ICIEAM)*, 1068–1072, Sochi, Russian Federation, 2022.

[17] Yang, Y., Y. Wang, and X. Yuan, "Bidirectional extreme learning machine for regression problem and its learning effectiveness," *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 23, No. 9, 1498–1505, 2012.

[18] Cao, W., Z. Ming, X. Wang, and S. Cai, "Improved bidirectional extreme learning machine based on enhanced random search," *Memetic Computing*, Vol. 11, 19–26, 2019.

[19] Kaloev, M. and G. Krastev, "Comparative analysis of activation functions used in the hidden layers of deep neural networks," in *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1–5, Ankara, Turkey, 2021.

[20] Zhang, L. and P. N. Suganthan, "A comprehensive evaluation of random vector functional link networks," *Information Sciences*, Vol. 367-368, 1094–1105, 2016.

[21] Nair, V. and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning*, 807–814, 2010.

[22] Ventura, J. F. I. and H. W. J. Russchenberg, "IDRA: IRCTR drizzle radar," in *2006 European Radar Conference*, 174–177, Manchester, UK, 2006.

[23] Xu, X., M. Shen, X. Wu, D. Wu, and D. Zhu, "Direction of arrival estimation based on modified fast off-grid L1-SVD," *Electronics Letters*, Vol. 58, No. 1, 32–34, 2022.