Prediction of Ship Heave Motion Using Regularized BP Neural Network with Cross Entropy Error Function

Hailun Wang¹ · Fei Wu¹ · Dongge Lei¹

Received: 22 February 2021 / Accepted: 2 November 2021
© The Author(s) 2021

Abstract

Accurate prediction of ship's heave motion can greatly enhance the safety of offshore operation. Due to its complexity and nonlinearity, however, ship's heave motion prediction is a difficult task to be solved. In this paper, a new method for predicting ship's heave motion is proposed based on an improved back propagation neural network (IBPNN). To overcome the gradient saturation phenomenon of traditional BPNN, the mean square error (MSE) loss function is replaced with a cross entropy (CE) loss function in IBPNN. Meanwhile, the weights of IBPNN is regularized by L_2 norm to enhance the generalization ability of traditional BPNN. Finally, conjugate gradient method is adopted to train IBPNN. The IBPNN is used to predict ship's heave motion and the prediction results prove its effectiveness.

Keywords Ship heave motion · Prediction · Neural network · Conjugate gradient

1 Introduction

Recently, the offshore operations, for example, underwater conveying systems for oil and gas field, windmills installation offshore crane and so on, are growing rapidly [1]. The offshore operations require the ships are stationary in the sea. Unfortunately, due to the effect of wind, wave and ocean currents, a vessel in the sea usually occurs heave motion, making it deviates the ideal position in vertical direction. The heave motion of a vessel in the sea destroys the safety and reduces the efficiency of offshore lifting operation. To reduce the effect of heave motion on offshore operations, researchers make an effort to decouple the vertical motion of load from that of vessel.

In the literature, heave motion compensation techniques had been proposed to achieve the purpose of heave motion decoupling. Passive heave compensation (PHC) and active heave compensation (AHC) are two primary techniques. Compared to PHC, AHC has higher decoupling efficiency [2] because AHC is equipped with a prediction controller, which receives the measurement of ship heave motion and

generates an instruction signal to the actuator to produce an opposite motion to achieve the purpose of compensation. In AHC, a key component is the heave motion prediction. Precise prediction of heave motion can not only reduce the control error but also correct the phase lag of controller, resulting in an improved response speed [3]. Therefore, developing an accurate model or method for heave motion prediction is an important issue for prediction controller design in AHC system.

Though it is very important, heave motion prediction had received little attention of researchers. A limited works had been done on heave motion prediction. In Ref. [1], a support vector regression (SVR) model was constructed to predict heave motion combined with particle swarm optimization (PSO), which is used to select the optimal values of super parameters of SVR. In Ref. [4], an autoregressive (AR) model was used as a predictor for heave motion, and the parameters of the constructed AR model was estimated using iteratively re-weighted least squares method, which is more robust than ordinary least squares. In Ref. [5], the heave motion is regarded as periodic signals, which consists of a set of sine waves with different amplitudes, frequencies and phases. The amplitudes, frequencies and phases of different sine components are identified using fast Fourier transformation (FFT) and peak detection algorithm. Then, the identified sine model was used to predict the heave motion.

✉ Hailun Wang
wanghl@qztc.edu.cn

¹ College of Electrical and Information Engineering, Quzhou University, No. 78, Jiuhua North Road, Quzhou 324000, Zhejiang, China

In the literature, many methods had been developed for the purpose of prediction. For example, neural network models [6] and fuzzy models [7, 8]. Neural network (NN), a universal function approximator, had been proven to be a powerful tool for modeling complex dynamical systems. Among different NN structures, multi-layer perceptron (MLP) is a popular NN model. MLP trained by error back propagation algorithm is called back propagation neural network (BPNN). BPNN had gained much attention and been applied to solve various engineering problems. In Ref. [9], BPNN was used to simulate and predict the low-temperature oxidation process of coal, as a result, the odds of coal spontaneous combustion is lowered. In Ref. [10], a new fault detection and diagnostic method was proposed based on BPNN, in which the fault features were extracted using wavelet and ensemble empirical mode decomposition. In Ref. [11], a mathematical model relating the concentration of amphoteric surfactant as an independent variable and surface finish was developed using BPNN and regression analysis.

Though BPNN had gained great success in many applications, however, BPNN still suffers several drawbacks. First, BPNN is trained using gradient-based algorithm, as a result, it is sensitive initial value and easily traps into local optimal. To overcome these drawbacks, some researchers proposed to used computational intelligent algorithms, such as particle swarm optimization (PSO) [12], genetic algorithm (GA) [13], to train BPNN. Second, gradient saturation occurs when sigmoid function is used in BPNN. Furthermore, the MSE loss function is not robust to noise and can not lead to better generalization ability.

Motivated by the aforementioned statements, we developed a new method for accurately and efficiently predicting heave motion. The proposed method is based on an improved BPNN, called regularized BPNN with cross entropy loss. A regularization term is imposed on the weights of BPNN to enhance the generalization ability, and meanwhile, the cross entropy loss function is adopted to replace the ordinary mean square error loss function to overcome the gradient saturation and sensitivity to noise. Finally, a conjugate gradient is adopted to train the BPNN. The proposed BPNN is used as a prediction model of ship heave motion. The main contribution of this paper are twofold. First, an improved BPNN with fast convergence and global optimum weights is developed. Second, the proposed BPNN can generate more accurate heave motion prediction since regularized weights put better trade-off between train and test accuracy.

The organization of this paper is arranged as follows. In Sect. 2, the basics of BPNN are reviewed. The proposed regularized BPNN with cross entropy error function is illustrated in Sect. 3. In Sect. 4, the training algorithm using conjugate gradient is explained. The prediction experiments for ship heave motion are conducted in Sect. 5. In the last, the conclusive remarks are given in Sect. 6.

2 BPNN

Usually, BPNN is a three layers neural network which consists of input layer, hidden layer and output layer. Figure 1 shows the structure of BPNN, which has n neurons in the input layer, l neurons in the hidden layer and m neurons in the output layer. The input variable is denoted as $u = [u_1, u_2, \dots, u_n]$ while the output variable $y = [y_1, y_2, \dots, y_m]$. The weights connecting each node in the input layer to that of hidden layer are w_{ij} , the weights connecting each node in hidden layer to that of output layer are w_{jk} . The output of the j th neuron in the hidden layer is

$$o_j = f\left(\sum_{i=1}^n w_{ij}u_i + d_j\right), j = 1, 2, \dots, l \quad (1)$$

where d_j is the bias of the j th neuron in the hidden layer and $f(\cdot)$ is the activation function of each neuron in hidden layer. Usually, the activation function is selected as sigmoid function, which is defined as

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

As the same way, the output of the k th neuron in the output layer is calculated as

$$y_k = f\left(\sum_{j=1}^l o_j w_{jk} + d_k\right), k = 1, 2, \dots, m, \quad (3)$$

where d_k is the bias of the k th neuron in the output layer, and $f(\cdot)$ is defined as in (2).

3 Regularized BPNN with Cross Entropy Error Function

3.1 Cross Entropy

Cross entropy is a mathematical tool that measures the difference between two probability distributions P and Q ,

Fig. 1 The structure of BPNN

where $P = \{p_1, p_2, \dots, p_n\}$ and $Q = \{q_1, q_2, \dots, q_n\}$. Cross entropy between P and Q is defined as

$$H(P, Q) = \sum_{i=1}^n -p_i \log(q_i). \quad (4)$$

3.2 BPNN with Cross Entropy Error Function

Let the training samples are $\{(u(i), t(i)), r = 1, 2, \dots, N\}$, and $u(i)$ be the i th input and $t(i)$ be the output. In the training phase of BPNN, the goal is to adjust the weights w_{ij} and w_{jk} such that the output of BPNN is as close to the desired output as possible. Usually, the mean square error (MSE) loss is adopted as the objective function of BPNN. The MSE is defines as

$$E = \frac{1}{2} \sum_{r=1}^N \sum_{k=1}^m (t_k(r) - y_k(r))^2, \quad (5)$$

where $t_k(r)$ represents desired output the k th dimension of the r th training sample. The objective function (5) is a function of weights w_{ij} and w_{jk} . To minimize the objective function (5), the steepest gradient descent method is used. That is to say, the weights are updated according to the following formula,

$$w_{jk}(\tau + 1) = w_{jk}(\tau) - \alpha \frac{\partial E}{\partial w_{jk}}, \quad (6)$$

and

$$w_{ij}(\tau + 1) = w_{ij}(\tau) - \alpha \frac{\partial E}{\partial w_{ij}}, \quad (7)$$

where α is a constant, called learning rate, τ represents the time instance in the iterative updating process, and

$$\frac{\partial E}{\partial w_{jk}} = \sum_{r=1}^N \delta_k(r) o_j(r), \quad (8)$$

where $\delta_k(r) = (t_k(r) - y_k(r))f'(\sum_{i=1}^n w_{ij}u_i + d_j)$ and

$$\frac{\partial E}{\partial w_{ij}} = - \sum_{r=1}^N \delta_j(r) u_i. \quad (9)$$

with $\delta_j(r) = \sum_{k=1}^m (t_k(r) - y_k(r))f'(\sum_{i=1}^n w_{ij}u_i + d_j)w_{jk}f'(\sum_{i=1}^n w_{ij}u_i + d_j)$. From Eqs. (8) and (9), one can see that when the estimated error is relatively large, weights sum to any output node is near the incorrectly saturated extreme value, and thus, the reduction of error function is non-significant in the last phase of training process [14]. To overcome this drawback, many researchers proposed to use other error functions. In

Ref. [15], the cross entropy (CE) is adopted as the error function, which is defined as

$$E = \frac{1}{2} \sum_{r=1}^N \sum_{k=1}^m [t_k(r) \ln y_k(r) + (1 - t_k(r)) \ln(1 - y_k(r))]. \quad (10)$$

In this paper, the CE error loss function (10) is used for the objective function of BPNN. In addition, the weights are regularized by L_2 norm to improve the generalization ability of BPNN, thus, the final error loss function is

$$E = \frac{1}{2} \sum_{r=1}^N \sum_{k=1}^m [t_k(r) \ln y_k(r) + (1 - t_k(r)) \ln(1 - y_k(r))] + \left[\sum_{i=1}^n \sum_{j=1}^l (w_{ij})^2 + \sum_{j=1}^l \sum_{k=1}^m (w_{jk})^2 \right]. \quad (11)$$

In Eq. (11), the last term is the weights regularization term, which is used to enhance the generalization ability of BPNN. For convenience, regularized BPNN with cross entropy error function is abbreviated as RBPNN-CE.

4 Training RBPNN-CE Using Conjugate Gradient Method

Usually, the BPNN is trained using gradient decent method. However, gradient decent is slow and easily trapped into local optimal. In this paper, the conjugate gradient (CG) method is used to train RBPNN-CE. CG method is a class of optimization method with less computation burden and less storage memory. It has been widely used to train neural networks [16, 17].

In CG method, the primary work is to determine a search direction to minimize an objective function, i.e., the error function (11) in this paper. Different from gradient descent method, the search direction of CG method is the linear combination of the negative gradient vector at the current iteration with the previous search direction, i.e.,

$$d_k = \begin{cases} -g_k, & \text{if } k=0; \\ -g_k + \beta_k d_{k-1}, & \text{otherwise.} \end{cases} \quad (12)$$

where β_k is called scalar parameter, which gives rise to distinct CG method. In this paper, the Polak-Ribière (P-R) method is adopted. The P-R method selects the scalar parameter as

$$\beta_k^{PR} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2}. \quad (13)$$

In Eqs. (12) and (13), g_k is the gradient of objective function. In this paper, g_k equals to the partial derivative of E in (11) with respect to weights w_{ij} or w_{jk} . The more details using

CG method to train neural network is omitted here, one can refer to Refs. [16, 18].

5 Simulation Results

To demonstrate the effectiveness of the proposed RBPNN-CE, a prediction experiment of ship heave motion is performed. The experimental data are taken from Ref. [19], which are acquired in a simulation platform of wave movement with sampling frequency 100Hz. In the experiments, 360 data points are used, and these data are firstly normalized into [0, 1]. Figure 2 shows the normalized data. The data set is denoted as a time series $\{y(k), k = 1, 2, \dots, N\}$.

In this paper, $y(k-1), y(k-2), y(k-3)$ are used to predict $y(k)$. The original data are divided into two groups, the first 60% data are used to train the model and the last 40% to test the built model. For the purpose of comparison, BPNN with gradient descent training algorithm, AR model, extreme learning machine (ELM) and adaptive neural-fuzzy inference system (ANFIS) and the proposed RBPNN-CE with CG training algorithm are all implemented. The number of hidden nodes of BPNN-CE, BPNN, ELM and ANFIS are all set to 5. The regularization parameter of BPNN-CE is set to 0.01, i.e., $\lambda = 0.01$. The activation function of ELM is selected as sigmoid function. In ANFIS, the 'gridPartition' method is used to generate the fuzzy inference system and the membership function is selected as Gaussian function.

To objectively evaluate the prediction performance of the referenced algorithms, four evaluation indexes, including coefficient of determination, coefficient of efficiency, root mean square error (RMSE) and mean absolute error (MAE), are adopted. They are defined as follows.

Coefficient of determination:

$$R^2 = \left[\frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}} \right]^2, \quad (14)$$

The larger the R^2 is, the more variability is explained by the model.

Fig. 2 The experimental data

Coefficient of efficiency:

$$E_{sn} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (15)$$

The larger the coefficient of efficiency is, the prediction value is more close to the real value.

Mean absolute error:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (16)$$

and root mean square error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N [y_i - \hat{y}_i]^2}{N}}, \quad (17)$$

where y_i and $\hat{y}_i$ are the real and the predicted value, respectively, N is the total number of data points used in the experiments. The smaller the MAE and RMSE are, the more accurate the model is.

First, the case that the heave motion signal is free of noise is considered. Figures 3 and 4 show the training and test output of RBPNN-CE, BPNN, AR, ELM and ANFIS. Table 1 shows the four evaluation indexes value in training phase and Table 2 shows those in test phase. It can be seen from Table 1 that, in training phase, the prediction performance of BPNN-CE indicated by the four indexes, i.e., R^2 , E_{sn} , MAE and RMSE, is not good as BPNN, ELM and ANFIS, but is better than AR. The ANFIS gives the

Fig. 3 The train output of referenced algorithms without noise case

Fig. 4 The test output of referenced algorithms without noise case

Table 1 Evaluation indices in train phase-without noise case

	BPNN-CE	BPNN	AR	ELM	ANFIS
R^2	0.7800	0.8248	0.7097	0.8401	0.9540
E_{sn}	0.7799	0.8169	0.6542	0.8401	0.9539
MAE	0.0561	0.0523	0.0694	0.0469	0.0236
RMSE	0.0775	0.0707	0.0971	0.0660	0.0354

Table 2 Evaluation indices in test phase-without noise case

	BPNN-CE	BPNN	AR	ELM	ANFIS
R^2	0.8416	0.7835	0.7971	0.7004	0.0094
E_{sn}	0.8351	0.7596	0.7646	0.6760	-18.4165
MAE	0.0541	0.0851	0.0916	0.0964	0.4283
RMSE	0.0611	0.1101	0.1089	0.1278	0.9892

best performance indexes in training phase. However, in test phase, from Table 2, the performance of BPNN-CE is better than other four prediction methods. The reason behind this phenomenon is that a L_2 norm regularized term is added to the weights of BPNN. The regularized term emphasizes trade-off between training variance and test variance. That is to say, if the training variance is smaller, then the test variance is larger, and vice versa. Since BPNN, ELM, AR and ANFIS adds no regularized term to model parameters, therefore, these models will obtain smaller variance in training phase, but larger variance in test phase. It implies that BPNN, ELM, AR and ANFIS will probably over fit the training data and produce poor prediction results. This can be seen from Fig. 4 and Table 2. In Fig. 4, the test output of ANFIS seriously deviates the true output. However, it fits the true data well in training phase, see in Fig. 3. This phenomenon also occurs in BPNN, ELM, AR. To further validate the superior of the BPNN-CE, regression plots of different algorithms are shown in Fig. 5. Obviously, the R value of regression of BPNN-CE is the highest, which means the BPNN-CE fit the test data better, and gives the highest prediction accuracy.

Second, to further check the performance of the proposed BPNN-CE, the case that the heave motion signal is corrupted by noise is considered. In this case, a 30 dB Gaussian noise is added to the heave motion data. Figures 6 and 7 show the outputs of the referenced algorithms in training and testing phase. Tables 3 and 4 show the performance indexes of all the referenced algorithms in training and test phase. From the experimental results, similar conclusion can be drawn as the case that the heave motion is free of noise. From Fig. 7, one can see that the predicted outputs of BPNN, AR, ELM and ANFIS are poorer than that of BPNN-CE. Especially, the predicted output of ANFIS seriously deviates from the actual output. This phenomenon demonstrates

Fig. 5 The regression plots of the referenced algorithms in test phase without noise**Table 3** Evaluation indices in train phase-with noise case

	BPNN-CE	BPNN	AR	ELM	ANFIS
R^2	0.7078	0.7512	0.6211	0.7792	0.9805
E_{sn}	0.7076	0.7506	0.5496	0.7792	0.9805
MAE	0.0704	0.0637	0.0864	0.0624	0.0144
RMSE	0.0920	0.0850	0.1142	0.0800	0.0238

Table 4 Evaluation indices in test phase-with noise case

	BPNN-CE	BPNN	AR	ELM	ANFIS
R^2	0.7937	0.7182	0.7279	0.4109	0.1715
E_{sn}	0.7841	0.7129	0.6873	0.1041	-5.8201
MAE	0.0658	0.0964	0.1055	0.1307	0.3153
RMSE	0.0763	0.1225	0.1279	0.2165	0.5973

that over-fitting occurs in BPNN, AR, ELM and ANFIS and their generalizations are poor. However, our proposed BPNN-CE produces a better prediction result, and it shows

Fig. 6 The train output of referenced algorithms with noise case

Fig. 7 The test output of referenced algorithms with noise case

that the generalization of BPNN-CE is better than PNN, AR, ELM and ANFIS. To further validate the superior of the proposed BPNN-CE, regression plots of different algorithms are shown in Fig. 8. It can be seen from Fig. 8 that the R value of BPNN-CE is the best, which indicates that the test output of BPNN-CE is more close to the actual output.

In a word, from the above two cases, our proposed BPNN-CE has obtained the best prediction performance. This attributes to the two improvements on ordinary BPNN, which greatly improves the generalization ability of BPNN. However, as ordinary BPNN, the structure selection of BPNN-CE is still open. In practice, the structure of BPNN-CE is determined by manual test.

6 Conclusions

Accurate prediction of heave motion is very important for AHC system of vessel. In this paper, we developed a new and accurate prediction model for vessel's heave motion based on BPNN. To avoid gradient saturation in ordinary BPNN, a cross entropy loss function is adopted to replace MSE loss function. Furthermore, a L_2 norm regularized term is added to the weights of BPNN to enhance the generalization ability of BPNN. The improved BPNN is called BPNN-CE. The proposed BPNN-CE is compared with ordinary BPNN, AR, ELM and ANFIS models and comparative experiments are performed. Experimental results show that the proposed model can give better prediction performance than BPNN, AR, ELM and ANFIS.

Fig. 8 The regression plots of the referenced algorithms in test phase with noise

Though the proposed BPNN-CE had gained satisfactory prediction performance, however, it is still a difficult problem to select an appropriate structure for BPNN-CE. Relevant researches show that 1-norm regularization of weights can select model structure. However, 1-norm regularization of weights can only be used in single layer structure. How to use the 1-norm regularization technique in BPNN is our future research work.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (61403229, 61503213) and Natural Science Foundation of Zhejiang Province (LGN20C050002).

Author Contributions HW conceives the idea and writes the draft of the paper. FW implements the algorithm and DL performs analysis.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Shi, B.H., Xian, L., Wu, Q.P., Zhang, Y.L.: Active heave compensation prediction research for deep sea homework crane based on KPSO-SVR. In: Control Conference, pp. 7637–7642 (2014)
- Neupert, J., Mahl, T., Haessig, B., Sawodny, O., Schneider, K.: A heave compensation approach for offshore cranes, in: American Control Conference, pp. 538–543 (2008)
- Woodacre, J.K., Bauer, R.J., Irani, R.A.: A review of vertical motion heave compensation systems. *Ocean Eng.* **104**, 140–154 (2015)
- El-Hawary, F., Mbamalu, G.A.N.: Dynamic heave compensation using robust estimation techniques. *Comput. Electr. Eng.* **22**(4), 257–273 (1996)
- Küchler, S., Mahl, T., Neupert, J., Schneider, K., Sawodny, O.: Active control for an offshore crane using prediction of the vessel's motion. *IEEE/ASME Trans. Mechatron.* **16**(2), 297–309 (2011)
- Ghosh, I., Chaudhuri, T.D.: FEB-Stacking and FEB-DNN models for stock trend prediction: a performance analysis for pre and post Covid-19 periods. *Decis. Mak. Appl. Manag. Eng.* **4**(1), 51–84 (2021)
- Stojčić, M., Stjepanović, A., Stjepanović, D.: ANFIS model for the prediction of generated electricity of photovoltaic modules. *Decis. Mak. Appl. Manag. Eng.* **2**(1), 35–48 (2019)
- Precup, R.E., Preitl, S., Petriu, E.M., et al.: Model-based fuzzy control results for networked control systems. *Rep. Mech. Eng.* **1**(1), 10–25 (2020)
- Zhao, J.Q., Yang, D.G., Meng, X.L., et al.: Prediction of temperature and CO concentration fields based on BPNN in low temperature coal oxidation. *Thermochim. Acta* **695**, 178820–178833 (2021)
- Cai, B.P., Sun, X.T., Wang, J.X., et al.: Fault detection and diagnostic method of diesel engine by combining rule-based algorithm and BNs/BPNNs. *J. Manuf. Syst.* **57**, 148–157 (2020)
- Vijayanand, M., Varahamoorthi, R., Kumaradhas, P., Kulkarni, M.V.: regression-BPNN modeling of surfactant concentration effects in electroless Ni=B coating and optimization using genetic algorithm. *Surf. Coat. Technol.* **409**, 126878 (2021)
- An, J., He, G.Q., Qin, F., Li, R., Huang, Z.W.: A new framework of global sensitivity analysis for the chemical kinetic model using PSO-BPNN. *Comput. Chem. Eng.* **112**, 154–164 (2018)
- Wang, L., Zhou, H., Yang, J., et al.: A decision support system for tobacco cultivation measures based on BPNN and GA. *Comput. Electron. Agric.* **181**, 105928 (2021)
- Jiang, M., Deng, B., Wang, B., Zhang, B.: A fast learning algorithm of neural networks by changing error functions. In: Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, vol.1, pp. 249–252 (2003)
- Ooyen, A.V., Nienhuis, B.: Improving the convergence of the back-propagation algorithm. *Neural Netw.* **5**, 465–471 (1992)
- Zhang, Y., Le, J., Liao, X., Zheng, F., Li, Y.: A novel combination forecasting model for wind power integrating least-squares support vector machine, deep belief network, singular spectrum analysis and locality-sensitive hashing. *Energy* **168**, 558–572 (2019)
- Wang, J., Zhang, B., Sun, Z., Hao, W., Sun, Q.: A novel conjugate gradient method with generalized ARMIJO search for efficient training of feedforward neural networks. *Neurocomputing* **275**, 308–316 (2018)
- Xing, X., Zhang, Q., Yang, P., Chao, Li., Chen, Z.: Research on bp algorithm based on conjugate gradient. In: The 2nd International Conference on Information Science and Engineering, pp. 5620–5623 (2010)
- Shi, P.: Study on the key technique of new wave-movement compensation. Ph.D. thesis, South China University of Technology (2013)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.