

RESEARCH ARTICLE

Heave compensation prediction based on echo state network with correntropy induced loss function

Xiaogang Huang^{1,2*}, Dongge Lei², Lulu Cai², Tianhao Tang¹, Zhibin Wang³

1 Department of Electrical Automation, College of Logistics Engineering, Shanghai Maritime University, Shanghai, China, **2** College of Electrical and Information Engineering, Quzhou University, Quzhou, China, **3** Institute of Electrical Engineering, Yanshan University, Qinhuangdao, China

* huangxiaogang@stu.shmtu.edu.cn



OPEN ACCESS

Citation: Huang X, Lei D, Cai L, Tang T, Wang Z (2019) Heave compensation prediction based on echo state network with correntropy induced loss function. PLoS ONE 14(6): e0217361. <https://doi.org/10.1371/journal.pone.0217361>

Editor: Lixiang Li, Beijing University of Posts and Telecommunications, CHINA

Received: February 13, 2019

Accepted: May 9, 2019

Published: June 13, 2019

Copyright: © 2019 Huang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research was supported by Zhejiang Province Public Welfare Project (LGG19F050003 to XH) and Project of Shanghai Maritime University Graduate Student Innovation (2015YCX073 to XH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

In this paper, a new prediction approach is proposed for ocean vessel heave compensation based on echo state network (ESN). To improve the prediction accuracy and enhance the robustness against noise and outliers, a generalized similarity measure called correntropy is introduced into ESN training, which is referred as corr-ESN. An iterative method based on half-quadratic minimization is derived to train corr-ESN. The proposed corr-ESN is used for the heave motion prediction. The experimental results verify its effectiveness.

Introduction

When operating on sea, a vessel is inevitably affected by waves, wind and ocean currents, thereby moving away from the desired position horizontally and vertically [1]. The vertical motion of the vessel, also called heave motion, which is undesirable for offshore installations, offshore drilling and other tasks on sea because it reduces the work efficiency, causes damage to safety manufacturing system, facility and operation. To reduce this passive effects, heave compensation technologies were proposed to remove vessel's heave motion from the load, which results in the decoupling of load motion from ship motion [2]. Now, heave compensation is popular in underwater conveying systems for oil and gas fields, payload transfer between vessels. Heave compensation technology can be classified into two classes, namely passive heave compensation (PHC) and active heave compensation (AHC). Compared to PHC, AHC can provide higher decoupling efficiency. AHC system is a close-loop system, in which the ship's heave motion is measured and fed back to a controller. Then, the controller drives an actuator to move in an opposite direction of the heave motion. Some research result show that a controller with heave motion prediction is helpful in creating an AHC system, which results in 100% effectiveness in heave motion decoupling [2]. Furthermore, heave motion prediction can be used to partially correct a large phase lag within the controller structure [3]. Hence, heave motion prediction is an important issue to AHC.

Though it is important, the research works on heave motion prediction are not so much. In the past years, many works focused on the ship's roll motion prediction [4, 5]. In the literature,

researchers utilize autoregressive (AR), autoregressive moving average (ARMA) and moving average (MA) models to construct prediction model from time series for heave motion prediction. In [6], a heave motion model was constructed from time series based on autoregressive (AR) model, the model's parameters were estimated using a robust estimation, i.e., iteratively reweighted least squares techniques. In [7], a method was proposed to predict vessel's vertical motion for the purpose of forming an active compensation system. The proposed method firstly formulated a linear model of the wave-induced motion based on the dominant modes, which was obtained from fast Fourier transformation and peak detection algorithm. Then, the amplitude and phase of each mode was estimated using Kalman filter. In [1], support vector regression (SVR) combined particle swarm optimization (PSO) was adopted for heave motion prediction. In the proposed method, PSO was used to optimize the super parameter of SVR. However, the above models are only effective for linear system prediction. Since ship's heave motion is a complex nonlinear system, it is necessary to develop a nonlinear prediction model for heave motion.

In this paper, a nonlinear model is developed for heave motion prediction based on each state network (ESN). ESN is a class of recurrent neural network (RNN), whose hidden layer (also called "dynamic reservoir") contains many randomly and sparsely connected neural units. In ESN, only the output weights need to be trained and other weights are randomly set, and thus the training complexity of ESN is reduced. Furthermore, ESN can be found many applications in system identification and control [8, 9], wind speed and direction forecasting [10], emotion recognition [11], etc. Apart from successful applications, many research works had been done to improve the performance of ESN, which mainly focus on constructing more efficient dynamic reservoir. The leaky integrator echo state network [12], reservoirs with biological properties [13], hierarchical reservoirs [14], and simple cycle reservoir [14] are just a few examples. When training ESN, not only the basic ESN but also the improved variants, the mean square error (MSE) criteria is used. The advantage of adopting MSE is that it can obtain the least square solution and lead to training simplicity. However, MSE is sensitive to noises and outliers, and inefficient for non-Gaussian error distribution. An alternative solution to this problem is to utilize the criterion based on correntropy. Correntropy, proposed by Santamaria [15], is recognized as more flexible and robust to noise or outlier than MSE. Owing to this, a correntropy based ESN is proposed to predict the heave motion.

The rest of this paper is organized as follows. In Section 2, the basic of ESN is introduced. The correntropy based ESN is given in Section 3. The experiments for heave motion using correntropy based ESN are presented in Section 4. Finally, the concluding remarks are presented in Section 5.

The basic of ESN

ESN is a recurrent neural network, whose structure is shown in Fig 1. An ESN consists of an input layer, a hidden layer and an output layer. The hidden layer is also called dynamic reservoir. The neural units in reservoir are sparsely connected each other. Different from other RNNs, the input weights W^{in} , the weights between reservoir units W^{x} and the feedback weight W^{fb} are predetermined randomly without being trained, only the output weights W^{out} should be trained. This characteristic greatly reduces the computation complexity. The training of ESN is divided into two stages. Firstly, the training data is fed into ESN and the state of reservoir $X(t)$ is calculated and updated as

$$x(t+1) = f(W^{\text{x}}X(t) + W^{\text{in}}u(t+1)), \quad (1)$$

where $X(t) \in R^N$ is the state of reservoir at time instant t , $u(t) \in R^L$ is the external input at time

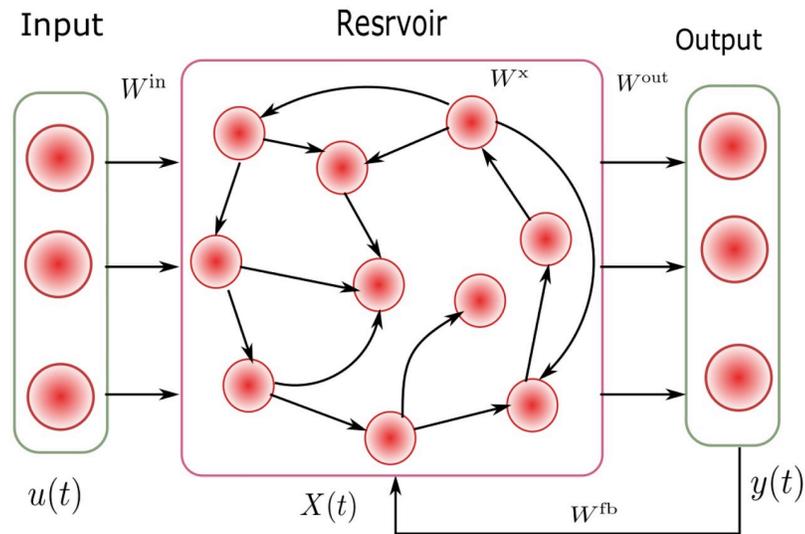


Fig 1. Structure of ESN.

<https://doi.org/10.1371/journal.pone.0217361.g001>

instant t , $W^x \in R^{N \times N}$ is the reservoir weight matrix and $W^{in} \in R^{N \times L}$ is the input weight matrix. $f(\cdot)$ is the activation function, usually the **tanh** function is adopted. For leaky integrator ESN, the state is updated as

$$X(t + 1) = \alpha X(t) + (1 - \alpha) * f(W^x X(t) + W^{in} u(t + 1)), \tag{2}$$

where α is called leaking rate. The output of ESN is computed as

$$y(t) = g(z(t) W^{out}), \tag{3}$$

where $y(t) \in R^{1 \times M}$ is the output of ESN at time instant t , $z(t) = [X^T(t) u^T(t + 1)] \in R^{1 \times (N+L)}$ is the concatenation of reservoir states and input vectors and $W^{out} \in R^{(N+L) \times M}$ the output weight matrix, $g(\cdot)$ is a nonlinear mapping function. In practice, the nonlinear function $g(\cdot)$ is selected as linear function. Therefore, the output can be written as

$$y(t) = z(t) W^{out}. \tag{4}$$

Let

$$Y = \begin{bmatrix} y(t) \\ y(t + 1) \\ \vdots \\ y(t + N - 1) \end{bmatrix}, Z = \begin{bmatrix} z(t) \\ z(t + 1) \\ \vdots \\ z(t + N - 1) \end{bmatrix}, \tag{5}$$

one can get

$$Y = Z W^{out}, \tag{6}$$

then, the optimal output weight matrix is obtained as

$$W^{out} = (Z^T Z)^{-1} Z^T Y. \tag{7}$$

Correntropy based ESN

Correntropy

In information theoretic learning (ITL), correntropy is a local similarity measure between two random vectors. It is regarded as a generalization of correlation function. Given two arbitrary random variables A and B, their correntropy is defined as

$$V_{\sigma}(A, B) = E[k_{\sigma}(A, B)], \tag{8}$$

where $k_{\sigma}(\cdot)$ is the kernel function that satisfies Mercer’s theorem with kernel size σ and $E[\cdot]$ denotes the mathematical expectation. In practice, only a finite number of samples are available and the real joint probability density of A and B is unknown. Therefore, the mathematical expectation is approximated by arithmetical average as

$$\hat{V}_{\sigma}(A, B) = \frac{1}{M} \sum_{i=1}^M k_{\sigma}(a_i, b_i). \tag{9}$$

In this paper, the Gaussian kernel function is selected, its expression is as follows

$$k_{\sigma}(a_i, b_i) \triangleq G(a_i, b_i) = \exp\left(-\frac{(a_i - b_i)^2}{2\sigma^2}\right). \tag{10}$$

Therefore, Eq (9) can be rewritten as

$$\hat{V}_{\sigma}(A, B) = \frac{1}{M} \sum_{i=1}^M \exp\left(-\frac{(a_i - b_i)^2}{2\sigma^2}\right). \tag{11}$$

The closer between A and B is, the larger the correntropy is. Compared with MSE, correntropy is not sensitive to noises or outliers and lead to more robust estimation.

Training ESN based on regularized correntropy criterion

In the training phase of ESN, the MSE is replaced by correntropy in order to improve the performance of ESN. Furthermore, to enhance the generalization of ESN, a L_2 norm regularization term is added. The new criteria for training ESN is given as

$$J(W^{\text{out}}) = \max_{W^{\text{out}}} \left[\sum_{i=1}^M \exp\left(-\frac{\|y(i) - z(i)W^{\text{out}}\|^2}{2\sigma^2}\right) - \gamma \|W^{\text{out}}\|^2 \right]. \tag{12}$$

The new criteria (12) is not quadratic any more. In this paper, the half-quadratic optimization is used to solve the optimization problem (12).

Proposition 1 For $G(z) = \exp\left(-\frac{\|z\|^2}{2\sigma^2}\right)$, there exists a convex conjugated function φ , such that

$$G(z) = \sup_{\alpha \in R^-} \left(\alpha \frac{\|z\|^2}{2\sigma^2} - \varphi(\alpha) \right). \tag{13}$$

Moreover, for a fixed z , the supremum is reached at $\alpha = -G(z)$.

Hence, introducing (13) into the objective function (12), the following augmented objective function can be obtained,

$$\hat{J}(W^{\text{out}}, \boldsymbol{\alpha}) = \max_{W^{\text{out}}, \boldsymbol{\alpha}} \left[\sum_{i=1}^M \left(\alpha_i \frac{\|y(i) - z(i)W^{\text{out}}\|_2^2}{2\sigma^2} - \varphi(\alpha_i) \right) - \lambda \|W^{\text{out}}\|^2 \right], \tag{14}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_M)$ stores the auxiliary variables appeared in the half-quadratic optimization. Moreover, for a fixed W^{out} , the following equation holds

$$J(W^{\text{out}}) = \hat{J}(W^{\text{out}}, \boldsymbol{\alpha}). \tag{15}$$

The optimal problem (14) can be solved via the following iterative manner,

$$\alpha_i^{\tau+1} = -\exp \left\{ -\frac{\|y(i) - z(i)W^{\text{out}}\|_2^2}{2\sigma^2} \right\}, \tag{16}$$

and

$$W_{\tau+1}^{\text{out}} = \arg \max_{W^{\text{out}}} (ZW^{\text{out}} - Y)^T \Lambda (ZW^{\text{out}} - Y) - \gamma W^{\text{out}} W^{\text{out}T} \tag{17}$$

where τ denote the τ th iteration and Λ is a diagonal matrix with its primary elements $\Lambda_{ii} = -\alpha_i^{\tau+1}$. The optimal problem (17) is easy to be solved, one can set the partial derivative of $\hat{J}(W^{\text{out}}, \boldsymbol{\alpha})$ with respect to W to zero, and yields

$$\frac{\partial \hat{J}(W^{\text{out}}, \boldsymbol{\alpha}^{\tau+1})}{\partial W^{\text{out}}} = 2(Z^T \Lambda Z - \gamma I) W^{\text{out}} - 2Z^T \Lambda Y = 0. \tag{18}$$

Therefore,

$$W_{\tau+1}^{\text{out}} = (Z^T \Lambda Z - \gamma I)^{-1} Z^T \Lambda Y. \tag{19}$$

After some iterations, the objective function (14) converges. The strict proof can be referred to [16].

Experimental results

In this section, the simulation studies are conducted to verify the effectiveness of the proposed method. All the algorithms are implemented in Matlab 2016b programming language and run in a ThinkPad T440 notebook computer with Intel Core™ i5-4200U processor, 8G random access memory (RAM). The heave motion data is taken from [17]. The data is measured from a simulation platform of wave movement with an accelerometer. The sampling frequency is 100Hz. The measured data is normalized into [0, 1], which is shown in Fig 2 (The data can be referred to S1 Fig).

Data preprocessing

The measured data contains many components including sudden vibration, high frequency component generated by fluctuating oil pressure and direct component [18]. Therefore, the original data is preprocessed by a simple filtering operation. The frequency domain filtering method based on Fast Fourier Transform (FFT) is applied to the original measured signal. The filtering is performed in frequency domain. A low-pass filter between 0 and 15 Hz is adopted

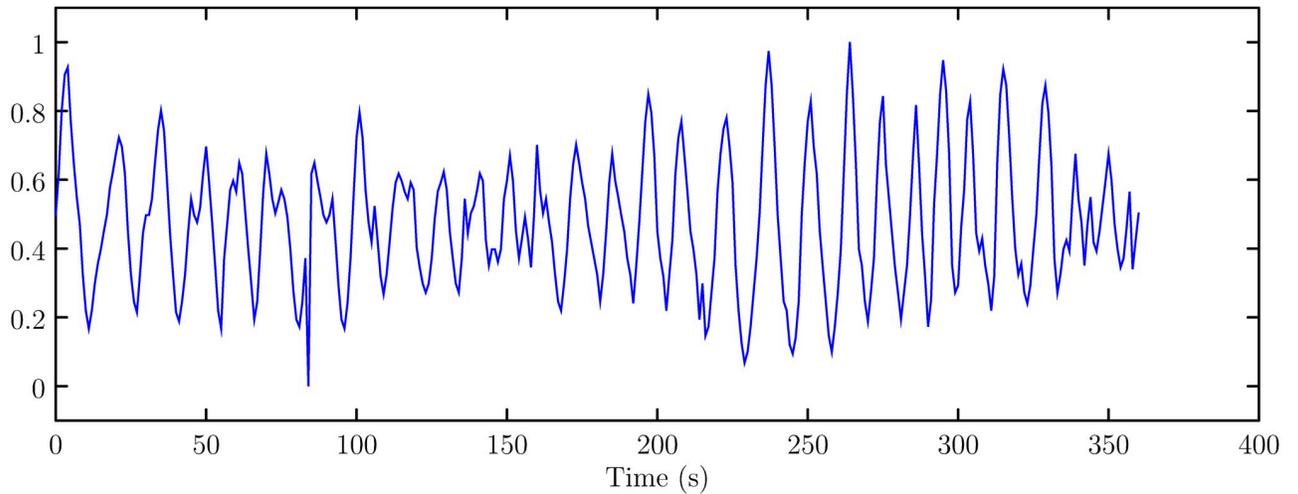


Fig 2. Heave motion data.

<https://doi.org/10.1371/journal.pone.0217361.g002>

and its frequency response is as follows,

$$H(f) = \begin{cases} 1, & 0 < f < 15 \\ 0, & \text{others} \end{cases} \quad (20)$$

The filtered data is shown in Fig 3 (The data can be referred to S2 Fig).

Prediction results

The filtered data is used for prediction experiments. Let $y(k)$, $k = 1, 2, \dots, n$ denotes the heave motion time series. In this paper, a three-order model is used for prediction, i.e., $y(k - 1)$, $y(k - 2)$, $y(k - 3)$ is used to predict $y(k)$. For the purpose of comparison, an autoregression (AR) model and original ESN are also implemented and used for prediction. The RMSE (root mean

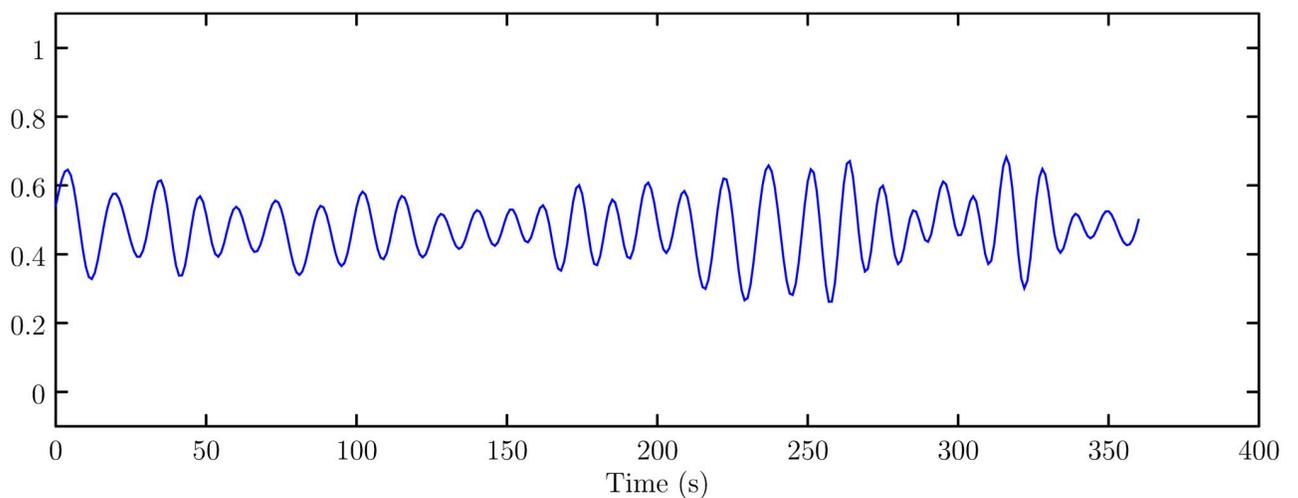


Fig 3. Heave motion data after filtering.

<https://doi.org/10.1371/journal.pone.0217361.g003>

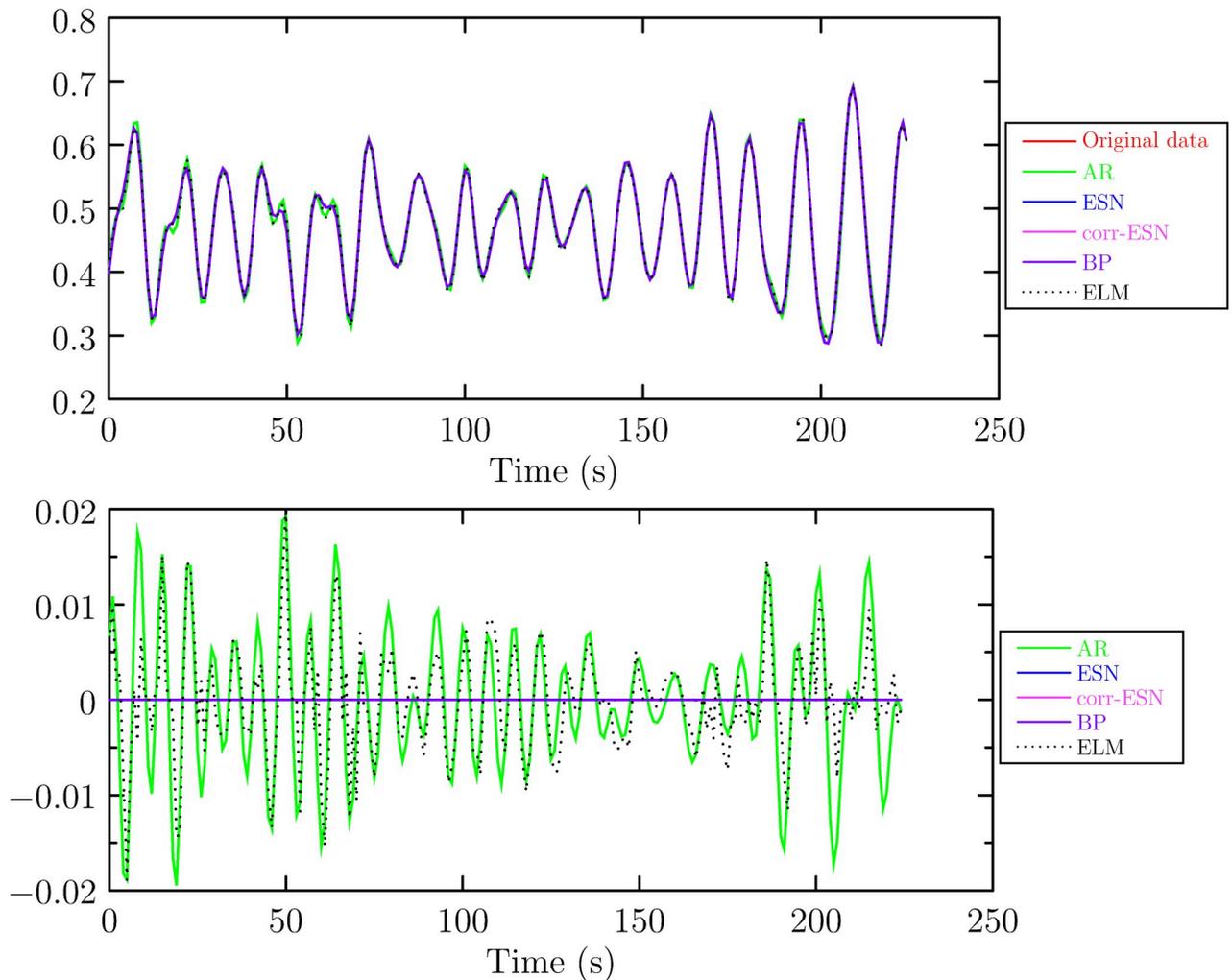


Fig 4. The training results and error. Top: training result, Bottom: the training error.

<https://doi.org/10.1371/journal.pone.0217361.g004>

square error) is used to evaluate the prediction performance, which is defined as

$$RMSE = \sqrt{\frac{\sum_{k=1}^N [y(k) - \hat{y}(k)]^2}{N}}, \tag{21}$$

where $y(k)$ is the real measured value at time k and $\hat{y}(k)$ is the predicted value, N is the number of data points used for training or test.

In experiments, the whole data is partitioned into two parts, the first 70% is used for training and the rest 30% is used for testing. The parameters of ESN are set as follows. The size of reservoir is 1000, λ in (17) is set to 10^{-6} and σ in (10) to 0.1. The maximum number of iteration in correntropy based ESN is 200. If the error of objective function value between successive iteration is less than a given tolerance, then the iterations break out and the algorithm terminates. The tolerance is set to 10^{-5} . The training results and error are shown in Fig 4 (The data can be referred to S3 Fig). Fig 5 (The data can be referred to S4 Fig) shows the one-step prediction results and prediction error. It can be seen that the training and prediction error of the proposed method are the smallest.

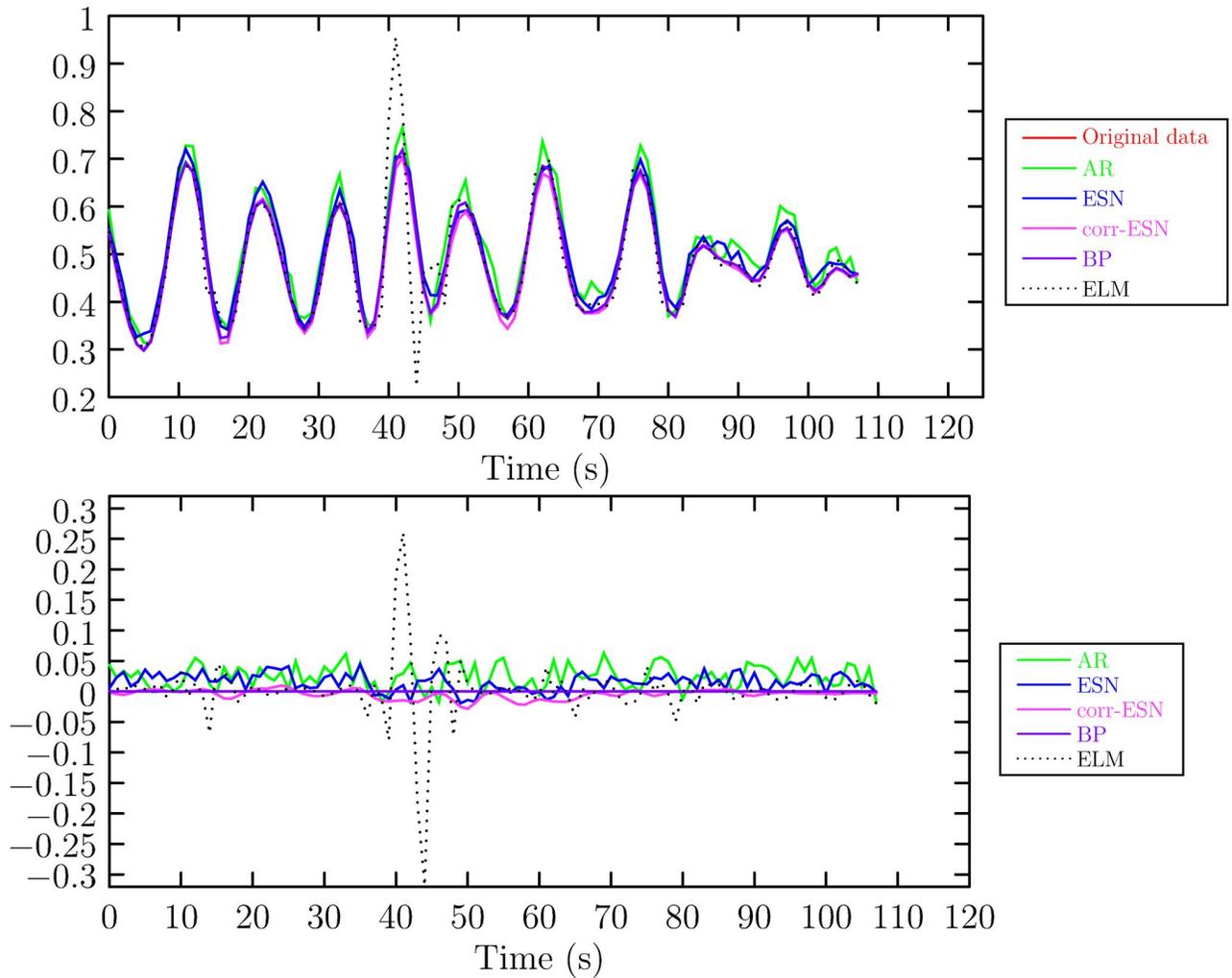


Fig 5. The one-step prediction results and error. Top: prediction result, Bottom: the prediction error.

<https://doi.org/10.1371/journal.pone.0217361.g005>

Table 1. RMSEs in training and prediction phase of different methods.

	AR	ESN	corr-ESN	BP	ELM
Training	0.0073	5.7923E-6	3.8741E-13	0.0067	0.0056
Prediction	0.0310	0.0206	0.0088	0.0224	0.0531

<https://doi.org/10.1371/journal.pone.0217361.t001>

To objectively evaluate the performance of the referenced methods, Table 1 lists the RMSE in training and prediction phase of the referenced methods.

Conclusions

A correntropy based ESN is proposed to predict heave motion for the purpose of heave compensation. The proposed approach adopts correntropy instead of MSE as the error criterion for ESN training, which is called corr-ESN. An iterative training algorithm is derived using half quadratic optimization theory. Since the correntropy is insensitive to noise and outliers,

the corr-ESN is more accurate than ESN for heave motion prediction. Simulation results validate the effectiveness of the proposed method.

Supporting information

S1 Fig. The data used to plot the Fig 2.

(CSV)

S2 Fig. The data used to plot the Fig 3.

(CSV)

S3 Fig. The data used to plot the Fig 4.

(CSV)

S4 Fig. The data used to plot the Fig 5.

(CSV)

Author Contributions

Conceptualization: Xiaogang Huang, Tianhao Tang.

Data curation: Lulu Cai.

Formal analysis: Dongge Lei, Lulu Cai, Zhibin Wang.

Writing – original draft: Xiaogang Huang.

Writing – review & editing: Zhibin Wang.

References

1. Shi BH, Xian L, Wu QP, Zhang YL. Active heave compensation prediction research for deep sea home-work crane based on KPSO—SVR. In: Control Conference; 2014. p. 7637–7642.
2. Neupert J, Mahl T, Haessig B, Sawodny O, Schneider K. A heave compensation approach for offshore cranes. In: 2008 American Control Conference; 2008. p. 538–543.
3. Woodacre JK, Bauer RJ, Irani RA. A review of vertical motion heave compensation systems. *Ocean Engineering*. 2015; 104:140–154. <https://doi.org/10.1016/j.oceaneng.2015.05.004>
4. Huang BG, Zou ZJ, Ding WW. Online prediction of ship roll motion based on a coarse and fine tuning fixed grid wavelet network. *Ocean Engineering*. 2018; 160:425–437. <https://doi.org/10.1016/j.oceaneng.2018.04.065>
5. Yin JC, Perakis AN, Wang N. A real-time ship roll motion prediction using wavelet transform and variable RBF network. *Ocean Engineering*. 2018; 160:10–19. <https://doi.org/10.1016/j.oceaneng.2018.04.058>
6. El-Hawary F, Mbamalu Gan. Dynamic heave compensation using robust estimation techniques. *Computers & Electrical Engineering*. 1996; 22(4):257–273. [https://doi.org/10.1016/0045-7906\(96\)00003-1](https://doi.org/10.1016/0045-7906(96)00003-1)
7. Kuchler S, Mahl T, Neupert J, Schneider K, Sawodny O. Active Control for an Offshore Crane Using Prediction of the Vessel's Motion. *IEEE/ASME Transactions on Mechatronics*. 2011; 16(2):297–309. <https://doi.org/10.1109/TMECH.2010.2041933>
8. Park J, Lee B, Awan AU, Kim HJ. Application of Echo-State Networks to the Position Control of Shape-Memory Alloys. *IFAC Proceedings Volumes*. 2013; 46(11):712–717. <https://doi.org/10.3182/20130703-3-FR-4038.00054>
9. Jaeger H. Adaptive Nonlinear System Identification with Echo State Networks. *Nips*. 2002; p. 609–616.
10. Chitsazan MA, Fadali MS, Trzynadlowski AM. Wind speed and wind direction forecasting using echo state network with nonlinear functions. *Renewable Energy*. 2019; 131:879–889. <https://doi.org/10.1016/j.renene.2018.07.060>
11. Trentin E, Scherer S, Schwenker F. Emotion recognition from speech signals via a probabilistic echo-state network. *Pattern Recognition Letters*. 2015; 66:4–12. <https://doi.org/10.1016/j.patrec.2014.10.015>

12. Jaeger H, Lukoševičius M, Dan P, Siewert U. Optimization and applications of echo state networks with leaky- integrator neurons. *Neural Netw.* 2007; 20(3):335–352. <https://doi.org/10.1016/j.neunet.2007.04.016> PMID: 17517495
13. Ozturk MC, Xu D, Principe JC. Analysis and design of echo state networks. *Neural Computation.* 2007; 19(1):111–138. <https://doi.org/10.1162/neco.2007.19.1.111> PMID: 17134319
14. Zhidong D, Yi Z. Collective behavior of a small-world recurrent neural system with scale-free distribution. *IEEE Transactions on Neural Networks.* 2007; 18(5):1364. <https://doi.org/10.1109/TNN.2007.894082>
15. Santamaria I, Pokharel PP, Principe JC. Generalized correlation function: definition, properties, and application to blind equalization. *IEEE Transactions on Signal Processing.* 2006; 54(6):2187–2197. <https://doi.org/10.1109/TSP.2006.872524>
16. Liu YJ, Xing HJ. Training RBFNN with reglarized correntropy criterion and its application to foreign exchange rate forecasting. In: 26th Chinese Control And Decision Conference (CCDC); 2014. p. 178–183.
17. Shi P. Study on the key technique of new wave-movement compensative. South China University of Technology; 2013.
18. Chen A. Research on wave motion compensation of two side-by-side positioned ships. South China University of Technology; 2011.