Ocean Science
Discussions

1   **Hybrid improved EMD-BPNN model for the prediction of sea surface temperature**

2   Zhiyuan Wu [a,b,c], Changbo Jiang [a,c,*], Mack Conde [d], Bin Deng [a,c], Jie Chen [a,c]

3   a. School of Hydraulic Engineering, Changsha University of Science & Technology, Changsha, 410004, China;

4   b. School for Marine Science and Technology, University of Massachusetts Dartmouth, New Bedford, MA 02744, USA;

5   c. Key Laboratory of Water-Sediment Sciences and Water Disaster Prevention of Hunan Province, Changsha, 410004, China;

6   d. Department of Mathematics, University of Massachusetts Dartmouth, North Dartmouth, MA 02747, USA.

7

8   **Highlights**

9   •   A novel SST predicting method based on the hybrid improved EMD algorithms and BP neural network

10   method are proposed in this paper.

11   •   SST prediction results based on the hybrid EEMD-BPNN and CEEMD-BPNN models are compared and

12   discussed.

13   •   Cases study of SST in the North Pacific shows that the proposed hybrid CEEMD-BPNN model can

14   effectively predict the time-series SST.

15

16   **Abstract:** Sea surface temperature (SST) is the major factor that affects the ocean-atmosphere interaction,

17   and in turn the accurate prediction of SST is the key to ocean dynamic prediction. In this paper, an SST

18   predicting method based on improved empirical mode decomposition (EMD) algorithms and back-

19   propagation neural network (BPNN) is proposed. Two different EMD algorithms have been applied

20   extensively for analyzing time-series SST data and some nonlinear stochastic signals. Ensemble empirical

21   mode decomposition (EEMD) algorithm and Complementary Ensemble Empirical Mode Decomposition

22   (CEEMD) algorithm are two improved algorithms of EMD, which can effectively handle the mode-mixing

23   problem and decompose the original data into more stationary signals with different frequencies. Each

24   Intrinsic Mode Function (IMF) has been taken as an input data to the back-propagation neural network model.

25   The final predicted SST data is obtained by aggregating the predicted data of individual IMF. A case study,

26   of the monthly mean sea surface temperature anomaly (SSTA) in the northeastern region of the North Pacific,

27   shows that the proposed hybrid CEEMD-BPNN model is much more accurate than the hybrid EEMD-BPNN

28   model, and the prediction accuracy based on BP neural network is improved by the CEEMD method.

29   Statistical analysis of the case study demonstrates that applying the proposed hybrid CEEMD-BPNN model

30   is effective for the SST prediction.

31

32   **Keywords.**

33   Sea Surface Temperature; Back-Propagation Neural Network; Empirical Mode Decomposition; Prediction;

34   Machine Learning Algorithms.

35

36   **1 Introduction**

37   The Sea Surface Temperature (SST) is a main factor in the interaction between the ocean and the

38   atmosphere (Wiedermann et al., 2017; He et al., 2017), and it characterizes the combined results of ocean

39   heat (Buckley et al., 2014; Griffies et al., 2015), dynamic processes (Takakura et al., 2018). It is a very

40   important parameter for climate change and ocean dynamics process, reflects sea-air heat and water vapor

41   exchange. Small changes in sea temperature can have a huge impact on the global climate. The well-known

42   El Niño and La Niña phenomena are caused by abnormal changes in SST (Chen et al., 2016a; Zheng et al.,

43   2016).

44   Therefore, scholars have begun to observe the SST in recent years, the observation of the SST is

45   important (Kumar et al., 2017; Sukresno et al., 2018). Accurate observation and effective prediction of the

46   SST are very important (Hudson et al., 2010). Predicting the SST in advance can enable people to take

47   appropriate measures to reduce the impact on daily life and reduce unnecessary losses. However, due to the

48   high randomness of the monthly mean sea surface temperature anomaly (SSTA), the nonlinear and non-

49   stationary characteristics are obvious. At present, there is no clear and feasible method with high accuracy to

50   effectively predict the SST (Zhu et al., 2015; Chen et al., 2016b; Khan et al., 2017).

51   In mathematics and science, a nonlinear system is a system in which the change of the output is not

52   proportional to the change of the input. Nonlinear dynamical systems, describing changes in variables over

53   time, may appear chaotic, unpredictable, or counterintuitive, contrasting with much simpler linear systems.

54   A stationary process is a stochastic process whose unconditional joint probability distribution does not change

55   when shifted in time. Consequently, parameters such as mean and variance also do not change over time. The

56   variation of SST is a deterministic non-linear dynamic system and a non-stationary time series data. Empirical

57   Mode Decomposition (EMD) is a state-of-the-art signal processing method proposed by Huang et al. (1998).

58   This method can decompose the signal data of different frequencies step by step according to the

59   characteristics of the data and obtain several periodic and trending signals orthogonal to each other, the

Ocean Science
Discussions

Open Access

60 method can decompose the stronger nonlinear and non-stationary signals into weaker nonlinear and non-

61 stationary signals (Wang et al., 2015; Amezquita-Sanchez and Adeli,2015; Wang et al., 2016; Kim and Cho,

62 2016). However, there were some problems of the EMD method, such as mode mixing (Huang and Wu, 2008;

63 Wu and Huang, 2009).

64   To solve this problem, Wu and Huang (2009) proposed the Ensemble Empirical Mode Decomposition

65 (EEMD) method by adding different white noise in each ensemble member to suppress mode mixing. Yeh et

66 al. (2010) added two opposite-signal white noises to the time-series data sequence, and proposed an improved

67 algorithm for EEMD, Complete Ensemble Empirical Mode Decomposition (CEEMD). The decomposition

68 effect is equivalent to EEMD, and the reconstruction error caused by adding white noise is reduced (Tang et

69 al., 2015). At present, the EMD model and its improved algorithms had been widely used in many fields on

70 ocean science, such as storm surge and sea level rise (Wu et al., 2011; Lee, 2013; Ezer and Atkinson, 2014),

71 tidal amplitude (Cheng et al., 2017; Pan et al., 2018) and wave height (Duan et al., 2016; Sadeghifar et al.,

72 2017; López et al., 2017). These studies and applications reflected that the EMD model and its improved

73 algorithms can effectively reduce the non-stationarity of the time-series data, which helps further analysis

74 and processing.

75   For nonlinear prediction, the more commonly used methods are curve fitting (Motulsky and Ransnas,

76 1987), gray-box model (Pearson and Pottmann, 2000), homogenization function model (Monteiro et al.,

77 2008), neural network (Deo et al., 2001; Wang et al 2015; Kim et al., 2016) and so on. Among them, Back-

78 Propagation Neural Network (BPNN) (Lee, 2004; Jain and Deo, 2006; Savitha and Al, 2017; Wang et al.,

79 2018) has certain advantages in dealing with nonlinear problems, it is a basic machine learning algorithm

80 and its principle is simple and operability is strong, so in ocean science and engineering it has been widely

81 used.

82   In view of non-stationary and nonlinear monthly mean SST, the EEMD, CEEMD and BP neural network

83 will be used here to study how to improve the accuracy of SST prediction. The improved hybrid EMD-BPNN

84 models will be established for the prediction of SSTA in the northeastern region of the Pacific Ocean.

85

86 **2 Data collection**

87   The SST time-series data in this study is from NOAA Optimum Interpolation Sea Surface Temperature

88 (OISST) official website (Reynolds et al., 2007; Banzon et al., 2016; https://www.ncdc.noaa.gov/oisst/data-

89 access). The NOAA 1/4°daily OISST is an analysis constructed by combining observations from different

90   platforms (satellites, ships, buoys) on a regular global grid. There are two kinds of OISST, named after the

91   relevant satellite SST sensors. These are the Advanced Very High Resolution Radiometer (AVHRR) and

92   Advanced Microwave Scanning Radiometer on the Earth Observing System (AMSR-E); the AVHRR dataset

93   is used in this study. The average annual sea surface temperature in North Pacific (0°N-60°N, 100°E-100°W)

94   during January 1982 to December 2016 is shown in Fig.1.
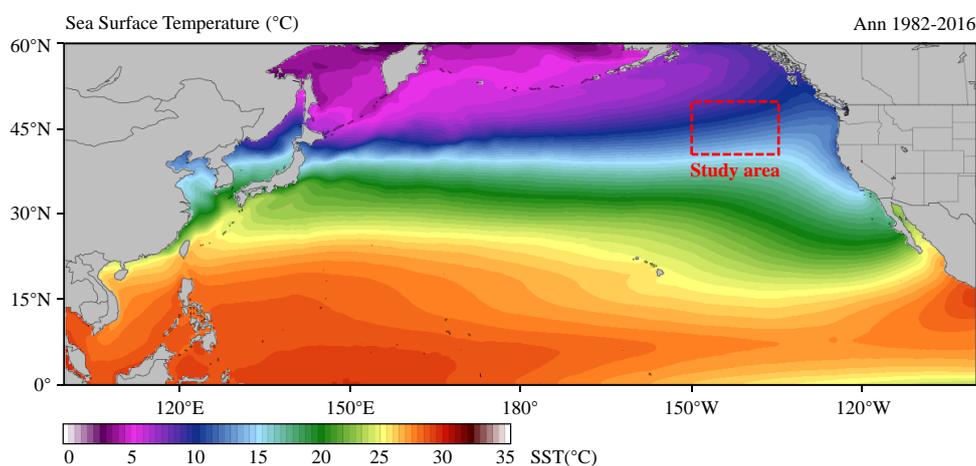
96   **Fig.1** Average annual sea surface temperature in North Pacific during Jan 1982 to Dec 2016 (35-years).

97

98   It has been shown that the sea surface temperature anomaly in the northeastern Pacific in the ten years

99   2006-2016 was 2.0°C warmer than in previous ten years 1996-2006. Previous studies (Bond et al., 2015)

100   showed that in the spring and summer of 2014, the high SST area of the northeastern Pacific had expanded

101   to coastal ocean waters, which affected the weather in coastal areas and the lives of fishermen, and even

102   affected the temperature in Washington, USA, causing interference to daily life.

103   In this study, we select the northeastern region of the North Pacific Ocean (in Fig.1, 40°N-50°N, 150°W-

104   135°W) to measure sea surface temperature. The time-series data of SST for the study area from January

105   1982 to December 2016 with a data length of 420 months was obtained from OISST-V2 (Fig. 2). The monthly

106   mean sea surface temperature anomaly (SSTA) was used in the analysis and calculation. As shown in Fig.

107   2(a), it can be found the overall time-series data is very messy, nonlinear and random from the perspective

108   of the image.

109

Ocean Science
Discussions

Open Access
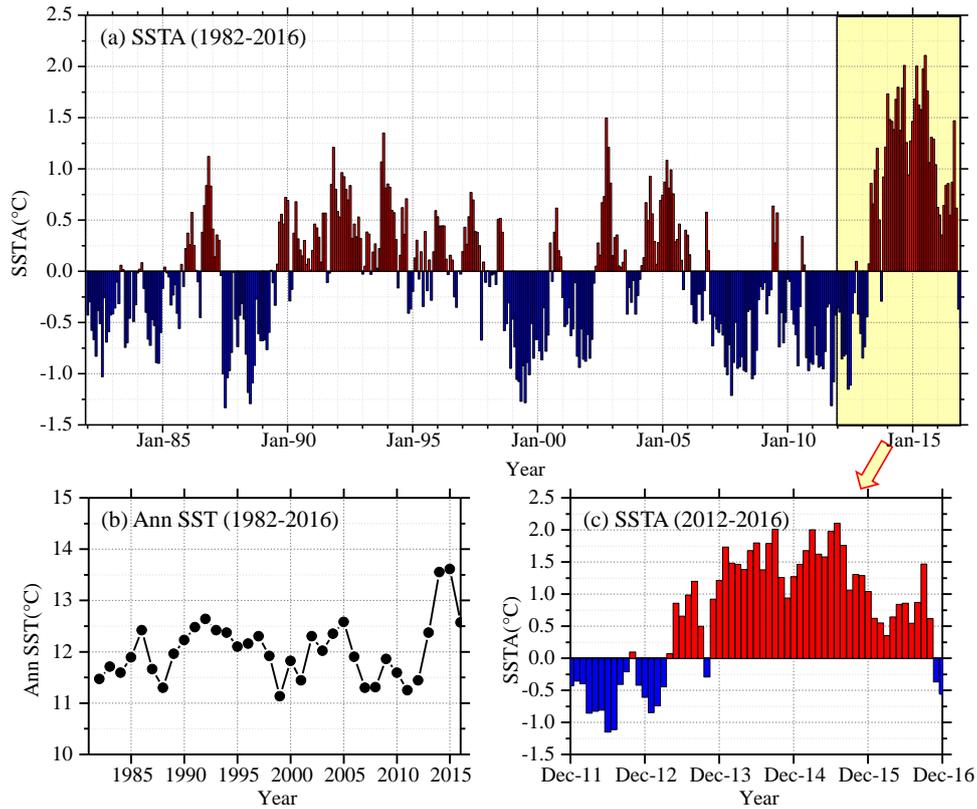
Fig.2 The time-series of sea surface temperature in the study area. (a) SST anomaly (1982-2016, 35 years);

(b) Annual SST (1982-2016, 35 years); (c) SST anomaly (2012-2016, 5 years).

3 Decomposition of SSTA

The purpose of this study is to combine the EEMD algorithm and the CEEMD decomposition algorithm

respectively with the BP neural network algorithm to establish a new prediction model, an improved hybrid

EMD-BPNN model. The EEMD and CEEMD algorithms are performed on the monthly mean SSTA data to

obtain a series of intrinsic mode functions (IMFi). Each IMFi is predicted by a BP neural network and then

each IMFi is reconstructed to obtain the predicted value of SSTA.

3.1 Decomposition by the EEMD algorithm

The SSTA in Fig. 2(a) has been decomposed based on the ensemble empirical mode decomposition

(EEMD algorithm), and seven IMF components and a residual component RES (Residue) are obtained as

shown in Fig. 3. It can be seen from Fig. 3 that the first three intrinsic mode function components IMF1,

124   IMF2, and IMF3 still exhibit strong nonlinearity and non-stationarity. The IMF4 to IMF7 and the final trend

125   term RES have some periodicity and relatively regular fluctuation, and the non-stationary and nonlinear

126   properties are less than the first three components. The trend term RES reflects that the overall trend of SSTA

127   has gradually increased since 1982. As the non-stationarity of each IMFi is gradually reduced, the EEMD

128   algorithm will reduce the influence of non-stationarity on prediction. The absolute error (ERR) of the

129   decomposition can been calculated by the following Formula (1).

130
$$a(t) = \left| S(t) - \left[\sum_{i=1}^{7} I_i(t) + R(t) \right] \right| \tag{1}$$

131   where, $a(t)$ is the absolute error (ERR), $S(t)$ the original SSTA observation data, $I_i(t)$ the i-th component

132   of the IMF (IMFi), and $R(t)$ the trend term (RES).
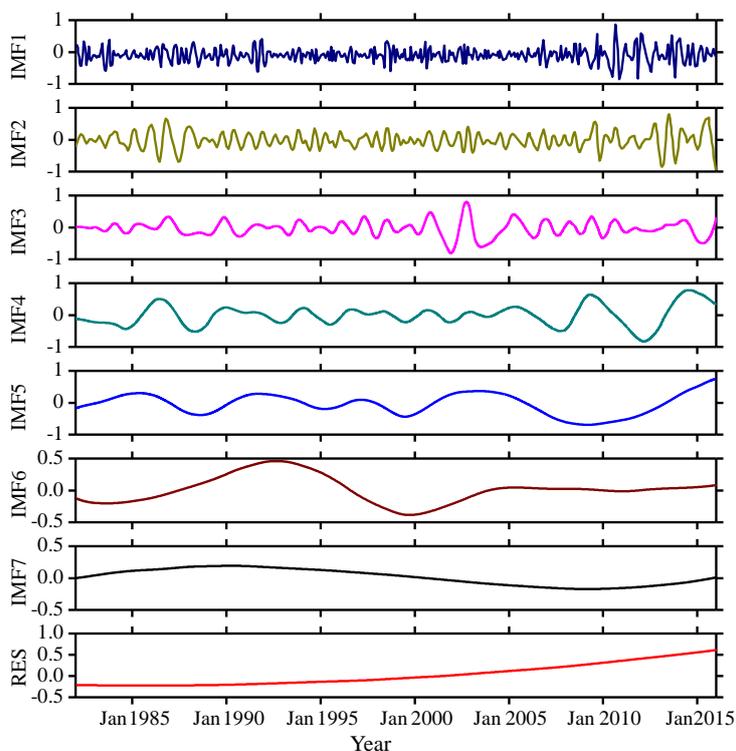
133

134   **Fig.3** IMF components and the trend item RES of monthly mean SSTA over the study area based on the

135   EEMD algorithm during 1982-2016.

136

137   The absolute error (ERR) based on EEMD algorithm is shown in Fig. 4. It can be seen from the figure

Ocean Science
Discussions

138   that the ERR of 420 months after decomposition is basically below 0.01 °C, and the ERR exceeds 0.01 °C in

139   five months: June 1989, September 1993, July 1998, May 1999 and March 2010.

140   In addition to June 1989, the other four monthly data with a large ERR occurred during the El Niño

141   period. The maximum error is in March 2010, the actual value is -0.1204 °C, the result based on EEMD

142   algorithm is -0.1325 °C, the ERR of decomposition is 0.0121 °C; the minimum error, in April 1987, is

143   1.73×10^{-5} °C. The overall mean ERR based on EEMD algorithm is 0.0035 °C and the order of magnitude is

144   10^{-3}.
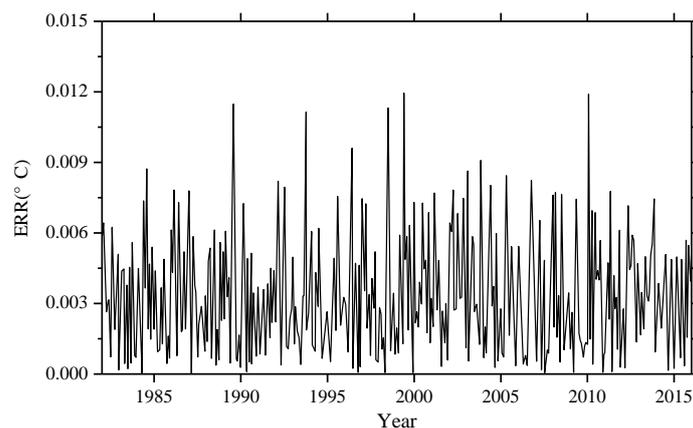
145

146   **Fig. 4** The ERR of monthly mean SSTA over the study area based on the EEMD algorithm during 1982-2016.

147

148   **3.2 Decomposition by the CEEMD algorithm**

149   The SSTA has been decomposed based on the complementary ensemble empirical mode decomposition

150   (CEEMD algorithm) and seven IMF components and a residual component RES (Residue) are obtained as

151   shown in Fig. 5. It can be seen when comparing the decomposition results based on EEMD and CEEMD

152   algorithms that although the mode components decomposed by CEEMD algorithm are different from the

153   corresponding results decomposed by EEMD, the nonlinearities and non-stationarities of the eight modes

154   decomposed by the two decomposition algorithms are gradually decreasing, and the final trend term RES is

155   an upward trend. Both decomposition algorithms confirm the characteristic of gradual increase for the overall

156   trend of the data series.

157   The absolute error (ERR) obtained based on the CEEMD algorithm is shown in Fig. 6. It can be seen

158   from the figure that the ERR of 420 months data after decomposition is less than 5×10^{-16} °C, and the accuracy

Ocean Science
Discussions

159   is very better. The maximum error is 4.48×10^{-16} °C in March 2016; the minimum error is zero. The overall

160   mean ERR based on CEEMD algorithm is 6.10×10^{-17} °C and the order of magnitude is 10^{-17}. By comparing

161   the results and errors of the above two decomposition algorithms, it can be seen that the error based on the

162   improved algorithm (CEEMD) is much smaller than the error based on EEMD algorithm. Because more

163   white noise with opposite sign had been added in CEEMD algorithm, the reconstruction error caused by

164   the white noise has been reduced over it in EEMD algorithm.
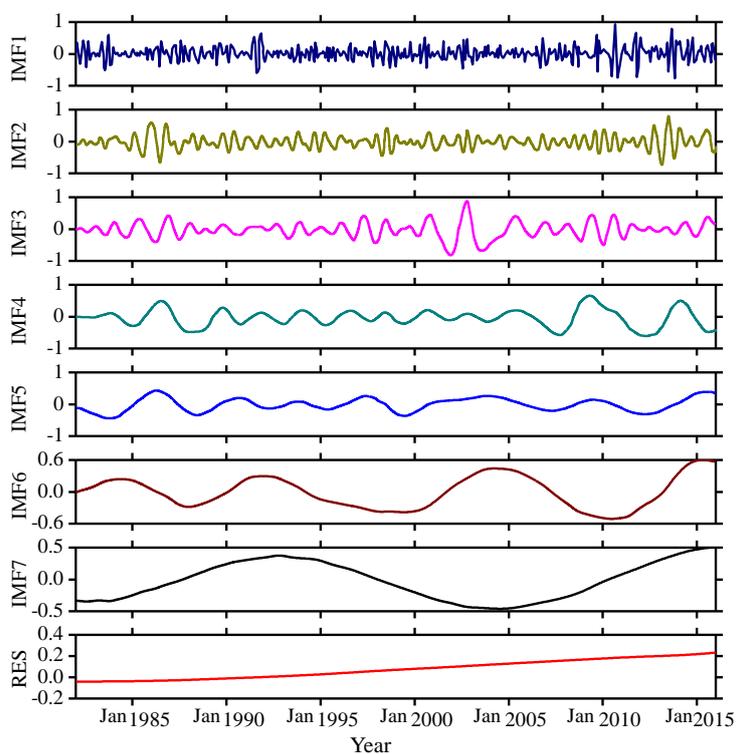
165

166   **Fig.5** IMF components and the trend item RES of monthly mean SSTA over the study area based on the

167   CEEMD algorithm during 1982-2016.

168

169
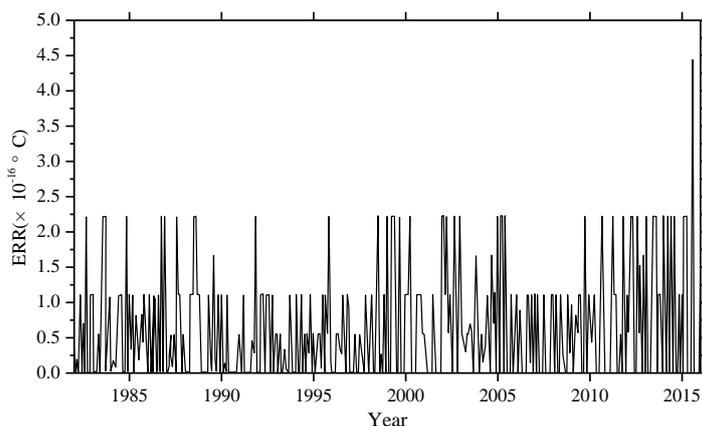
Fig. 6 The ERR of monthly mean SSTA over the study area based on the CEEMD algorithm during 1982-

2016.

4 SSTA prediction model

4.1 The BP neural network

Artificial Neural Network (ANN) is an information processing approach based on the biological neural

network (López et al., 2015; Kim et al., 2016). In theory, ANN can simulate any complex nonlinear

relationship through nonlinear units (neurons) and has been widely used in the prediction area, such as wave

height and storm surge. The most basic structure of ANN consists of input layers, hidden layers and output

layers. One of the most widely used ANN models is the back propagation neural network (BPNN, Wang et

al., 2018) algorithm based on the BP algorithm.

The BPNN algorithm is a multi-layer feedforward network trained according to the error back

propagation algorithm and is one of the most widely used deep learning algorithms. The BP network can be

used to learn and store a large number of mappings of input and output models without the need to publicly

describe the mathematical equations of these mapping relationships. The learning rule is to use the steepest

descent method. When applied to SST predicting, the input data are monthly mean SST in previous months

and the output data are predicted SST time-series data. The desired data for comparison is the observed actual

SST.

4.2 SSTA prediction model based on hybrid improved EMD-BPNN algorithm

The proposed monthly mean sea surface temperature anomaly (SSTA) predicting model includes three

Ocean Science
Discussions

190   steps as follows. First, original SST datasets are decomposed into certain more stationary signals with

191   different frequencies by EEMD. Second, BP neural network is used to predict each IMF and the residue RES.

192   A rolling forecasting process is studied. The prediction is made using the previous data for one step ahead.

193   Finally, the prediction results of each IMF and the residue RES are aggregated to obtain the final SST

194   prediction results. The flowchart of SST prediction model based on hybrid improved empirical mode

195   decomposition algorithm (improved EMD algorithm) and back-propagation neural network (BPNN)is shown

196   in Fig. 7. The SST prediction model has been abbreviated as hybrid improved EMD-BPNN model in the

197   following article.
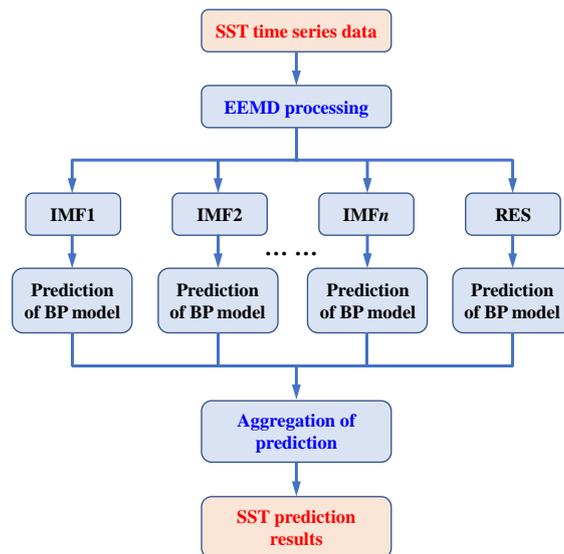
198
199   **Fig.7** The flowchart of SST prediction model based on hybrid improved empirical mode decomposition

200   algorithm (improved EMD algorithm) and back-propagation neural network (BPNN).

201

202   **5 Case study: SSTA prediction based on the hybrid improved EMD-BPNN models**

203   In order to study the effects of the two improved EMD algorithms (EEMD and CEEMD) on the

204   prediction results, and to analyze the prediction ability of BP neural network, the following experiments were

205   carried out. Predict SSTA results in 2017 and analyze the prediction abilities of different mode decomposition

206   data based on EEMD and CEEMD algorithms. The experiment content is as follows: the BP neural network

207   is trained with the decomposition data of each mode from 1982 to 2016, and the SSTA in 2017 is predicted

208   by the trained neural network, and the observation results of 12 months in 2017 is used to compare and
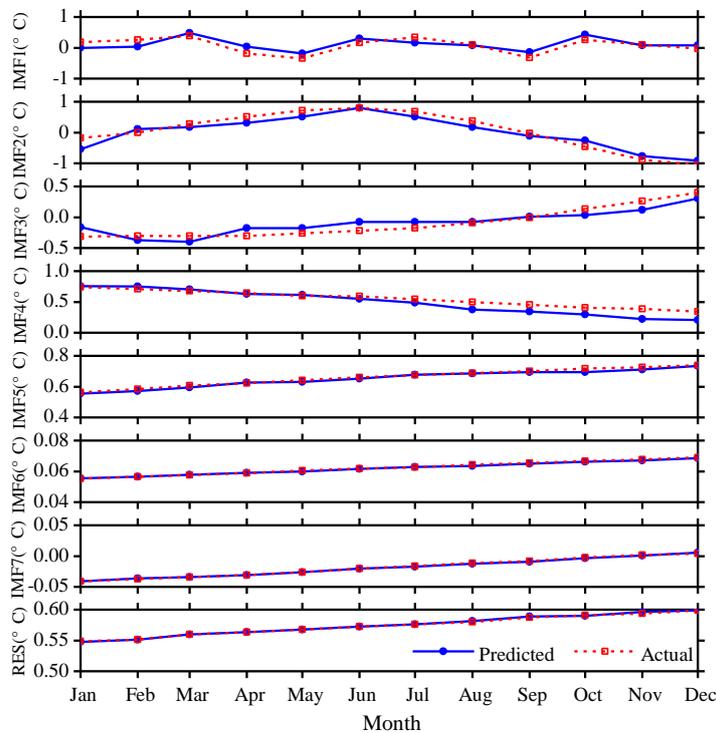
209   analyze with the prediction results.

210   Since the nonlinearity of the IMF1 to IMF3 is still relatively strong, a three-layer BP neural network

211   structure has been chosen and independently analyze and predict each month. For the IMF4 and subsequent

212   modes, since the nonlinearity and non-stationarity have been degraded relative to the first three modes, a BP

213   neural network with 12 nodes at input layer and output layer has been used to train and predict SSTA.

214   The prediction results of each mode decomposition component based on the EEMD algorithm are shown

215   in Fig. 8. The absolute errors of the predicted value and the actual value are shown in Table 1. Root mean

216   square error (RMSE) is used as metrics to access the performance of the two different models.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(x_n - y_n\right)^2} \qquad (2)$$

218   where, x_n and y_n are the observed and the predicted values respectively, N is the number of data used for

219   the performance evaluation. Results are shown in Table 1.

Fig. 8 SSTA prediction results based on the hybrid EEMD-BPNN model of each individual component in 2017.

224 **Table 1.** The absolute errors ERRs of the SSTA prediction results of each individual component based on the

225 hybrid EEMD-BPNN model (unit: °C).

|	Max ERR	Min ERR	Mean ERR	RMSE
IMF1	0.2197	0.0014	0.1424	0.1486
IMF2	0.2166	0.0323	0.1297	0.1673
IMF3	0.1872	0.0051	0.1070	0.1245
IMF4	0.1602	$1.6869×10^{-4}$	0.0663	0.0857
IMF5	0.0158	0.0010	0.0089	0.0104
IMF6	$3.8766×10^{-4}$	$1.9752×10^{-4}$	$2.7221×10^{-4}$	0.0003
IMF7	$5.2662×10^{-4}$	$1.6387×10^{-4}$	$1.7907×10^{-4}$	0.0002
RES	$5.4859×10^{-4}$	$2.2308×10^{-4}$	$2.4766×10^{-4}$	0.0002

226

227   It can be seen from Fig. 8 and Table 1 that the maximum absolute error (Max ERR) of the first

228 decomposition component IMF1 based on the hybrid EEMD-BPNN model is 0.2197 °C in January. The

229 minimum absolute error (Min ERR) is 0.0014 °C, which is in August. The prediction ability of the second

230 mode decomposition component IMF2 is roughly equivalent to the IMF1, and the mean absolute error (Mean

231 ERR) of the first three intrinsic mode function components IMF1, IMF2, and IMF3 are between 0.10 °C and

232 0.15 °C. The mean absolute errors of the IMF4 and IMF5 are 0.0663 °C and 0.0089 °C, respectively, and the

233 prediction accuracy based on the hybrid EEMD-BPNN model is roughly equivalent to the decomposition

234 accuracy of the EEMD algorithm. The prediction errors of the last two intrinsic mode function components

235 and the residue RES are on the order of 10^{-4}. It can be seen that as the nonlinearity and non-stationarity of

236 the series data decreases, the error of the prediction results becomes smaller and smaller.

237   According to the same method, the eight mode components decomposed by CEEMD algorithm have

238 been analyzed and predicted. The prediction results and error analysis have been shown in Fig. 9 and Table

239 2. It can be seen from Fig. 9 and Table 2 that the maximum error of the first decomposition component IMF1

240 based on the hybrid CEEMD-BPNN model is 0.1779 °C in May. The minimum error is 0.0068 °C, which is

241 in June.

242   The prediction ability of the second mode decomposition component IMF2 is roughly equivalent to the

243 IMF1. Except for the four months of May, September, October, and November, the accuracies of prediction

244 results of other months are satisfactory. The prediction results of the first three intrinsic mode function

Ocean Science
Discussions

Open Access

245   components IMF1, IMF2, and IMF3 are basically the same as the actual data. In the prediction results of the

246   fourth mode component IMF4, except for slight error in December, the prediction ability is better. The

247   predicted results of the last three intrinsic mode function components IMF5, IMF6, IMF7 and the residue

248   RES are basically consistent with the observation results.
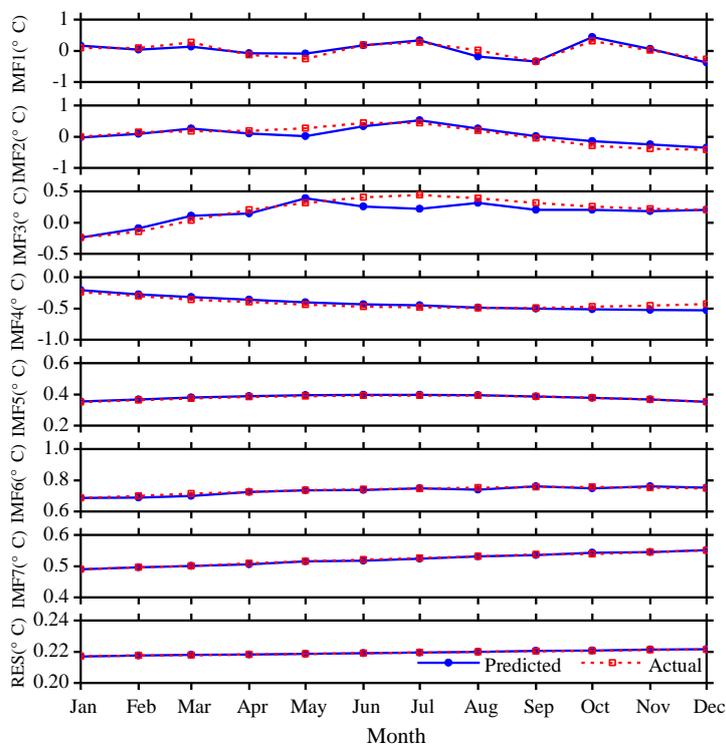
249

250   **Fig. 9** SSTA prediction results based on the hybrid CEEMD-BPNN model of each individual component in

251   2017.

252

253   **Table 2.** The absolute errors ERRs of the SSTA prediction results of each individual component based on the

254   hybrid CEEMD-BPNN model (unit: °C).

	Max ERR	Min ERR	Mean ERR	RMSE
IMF1	0.1779	0.0068	0.0827	0.0987
IMF2	0.1643	0.0413	0.0811	0.1124
IMF3	0.1521	0.0160	0.0713	0.1006
IMF4	0.0851	0.0211	0.0324	0.0427
IMF5	0.0052	8.7694×10^{-5}	0.0021	0.0029
IMF6	0.0103	5.7748×10^{-5}	0.0043	0.0056
IMF7	0.0017	3.6026×10^{-5}	9.1374×10^{-4}	0.0010
RES	3.0342×10^{-5}	2.0163×10^{-6}	1.1572×10^{-5}	1.5017×10^{-5}

255

256   The prediction results of the monthly mean SSTA in 2017 are obtained by reconstructing the mode

257   decomposition components (Fig. 10) and the absolute error (ERR) of prediction results has been shown in

258   Table 3. It can be seen from the figure and table that the prediction results based on the EEMD-BPNN model

259   have larger ERRs in January and August, exceeding 0.3 °C, and the accuracies of prediction results in other

260   months are satisfactory (the ERR is less than 0.3). The prediction accuracy based on the CEEMD-BPNN

261   model is satisfactory, except for the ERR exceeding 0.1 °C in October, and the prediction ability based on

262   the CEEMD-BPNN model is generally better than that of the EEMD-BPNN model.
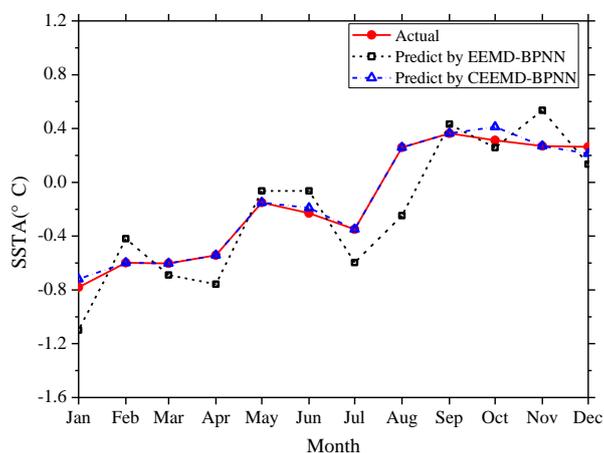
263

264   **Fig. 10** Monthly SSTA prediction results based on the hybrid improved EMD-BPNN models in 2017.

Ocean Science
Discussions

265

Table 3. The absolute errors ERRs of the SSTA prediction results based on the two different hybrid improved

EMD-BPNN models (unit: °C).

	EEMD-BPNN model	CEEMD-BPNN model		EEMD-BPNN model	CEEMD-BPNN model
Jan	0.3188	0.0623	Sep	0.0687	0.0132
Feb	0.1780	0.0103	Oct	0.0545	0.1607
Mar	0.0867	0.0063	Nov	0.2651	0.0101
Apr	0.2153	0.0137	Dec	0.1290	0.0183
May	0.0854	0.0102	**Min ERR**	**0.0545**	**0.0063**
Jun	0.1662	0.0224	**Max ERR**	**0.5068**	**0.1607**
Jul	0.2474	0.0077	**Mean ERR**	**0.1935**	**0.0289**
Aug	0.5068	0.0112	**RMSE**	**0.2299**	**0.0512**

268

269   Correlation coefficient between the prediction values based on the CEEMD-BPNN model and

270 observations is shown that the value of the correlation coefficient that indicates a significance level of 0.001

271 and the correlation coefficient reached 0.97. The result indicates that SSTA in 2017 had been predicted

272 accurately by the CEEMD-BPNN model. As can be seen from the above discussions, the ERR of

273 decomposition components based on the EEMD and CEEMD algorithms will affect the accuracy of the final

274 prediction results. Table 3 shows that predicting results of the hybrid CEEMD and BPNN model are

275 ameliorated a lot as compared to the EEMD-BPNN direct predicting model. This is because after CEEMD,

276 the original unsteady and nonlinear data are changed into certain components that have fixed frequency and

277 periodicity. The CEEMD algorithm with less decomposition error has less error in the final prediction results,

278 which proves that the CEEMD method has more advantages in data decomposition than the EEMD method.

279 At the same time, we can find that the final prediction error of the two prediction models mainly comes from

280 the first three mode decomposition components, and the error of the last five components has little effect on

281 the accuracy of the final prediction results.

282

283 **6 Conclusions**

284   This paper presents a novel SST predicting method based on the hybrid improved EMD algorithms and

285 BP neural network method to process the SST data with strong nonlinearity and non-stationarity. Through

286   EEMD and CEEMD algorithms, SSTA time-series data are decomposed into different IMFs and a residue

287   RES. BP neural network is applied to predict individual IMFs and the residue RES. Final results can be

288   obtained by adding the predicting results of individual IMFs and RES.

289   In order to illustrate the effectiveness of the proposed approach, a case study was carried out. SSTA

290   prediction results based on the hybrid EEMD-BPNN model and hybrid CEEMD-BPNN model are discussed

291   respectively. In comparison, the proposed hybrid CEEMD-BPNN model is much better and its prediction

292   results are more accurate.

293   From the absolute error of the prediction results of each component IMF and the absolute error of the

294   predicted SSTA, the prediction error of SSTA mainly comes from the prediction of the first three mode

295   decomposition component (IMF1, IMF2 and IMF3), because the first three mode components still have

296   strong nonlinearity and non-stationarity. As the nonlinearity gradually decreases, the absolute error of the

297   prediction results gradually decreases.

298   SST prediction has been only preliminary carried out based on the two improved EMD algorithms and

299   BP neural network in this paper. The results show that the hybrid CEEMD-BPNN model is more accurate in

300   predicting SST. This work can provide a reference for predicting SST and El Niño in the future. In the follow-

301   up study, how to improve the forecast duration is the focus of this work.

302

306

307   **References:**

308   Amezquita-Sanchez, J. P. and Adeli, H.: A new music-empirical wavelet transform methodology for time–

309   frequency analysis of noisy nonlinear and non-stationary signals, Digit. Signal Process., 45, 55-68,

310   https://doi.org/10.1016/j.dsp.2015.06.013, 2015.

311   Banzon, V., Smith, T. M., Chin, T. M., Liu, C., and Hankins, W.: A long-term record of blended satellite and

312   in situ sea-surface temperature for climate monitoring, modeling and environmental studies, Earth Syst.

313   Sci. Data, 8, 165-176, https://doi.org/10.5194/essd-8-165-2016, 2016.

314   Bond, N. A., Cronin, M. F., Freeland, H., and Mantua N.: Causes and impacts of the 2014 warm anomaly in

315   the NE Pacific. Geophys. Res. Lett., 42, 3414-3420, https://doi.org/10.1002/2015GL063306, 2015.

Ocean Science
Discussions

316   Buckley, M. W., Ponte, R. M., Forget, G., and Heimbach, P.: Low-frequency SST and upper-ocean heat

317   content variability in the North Atlantic, J. Climate, 27, 4996-5018, https://doi.org/10.1175/JCLI-D-13-

318   00316.1, 2014.

319   Chen, C., Cane, M. A., Henderson, N., Lee, D. E., Chapman, D., Kondrashov D., and Chekroun, M. D.:

320   Diversity, nonlinearity, seasonality, and memory effect in ENSO simulation and prediction using

321   empirical model reduction, J. Climate, 29: 1809-1830, https://doi.org/10.1175/JCLI-D-15-0372.1,

322   2016b.

323   Chen, Z., Wen, Z., Wu, R., Lin X., and Wang J.: Relative importance of tropical SST anomalies in maintaining

324   the Western North Pacific anomalous anticyclone during El Niño to La Niña transition years, Clim.

325   Dynam., 46, 1027-1041, https://doi.org/10.1007/s00382-015-2630-1, 2016a.

326   Cheng, Y., Ezer, T., Atkinson, L. P., and Xu, Q.: Analysis of tidal amplitude changes using the EMD method,

327   Cont. Shelf Res., 148: 44-52, https://doi.org/10.1016/j.csr.2017.09.009, 2017.

328   Deo, M. C., Jha, A., Chaphekar, A. S., and Ravikant, K.: Neural networks for wave forecasting, Ocean Eng.,

329   28: 889-898, https://doi.org/10.1016/S0029-8018(00)00027-5, 2001.

330   Duan, W. Y., Han, Y., Huang, L. M., Zhao, B. B., and Wang, M. H.: A hybrid EMD-SVR model for the short-

331   term   prediction   of   significant   wave   height,   Ocean   Eng.,   124,   54-73,

332   https://doi.org/10.1016/j.oceaneng.2016.05.049, 2016.

333   Ezer, T. and Atkinson, L. P.: Accelerated flooding along the US East Coast: on the impact of sea‑level rise,

334   tides, storms, the Gulf Stream, and the North Atlantic oscillations, Earths Future, 2, 362-382,

335   https://doi.org/10.1002/2014EF000252, 2014.

336   Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., Dunne, J. P.,

337   Goddard, P., Morrison, A. K., Rosati, A., Wittenberg, A. T., Yin, J., and Zhang, R.: Impacts on ocean

338   heat from transient mesoscale eddies in a hierarchy of climate models. J. Climate, 28, 952-977,

339   https://doi.org/10.1175/JCLI-D-14-00353.1, 2015.

340   He, J., Deser, C., and Soden, B. J.: Atmospheric and oceanic origins of tropical precipitation variability. J.

341   Climate, 30, 3197-3217, https://doi.org/10.1175/JCLI-D-16-0714.1, 2017.

342   Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung, C. C., and Liu, H. H.:

343   The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time

344   series analysis, P. Roy. Soc. A-Math. Phy., 454, 903-995. https://doi.org/10.1098/rspa.1998.0193, 1998.

345   Huang, N. E. and Wu, Z.: A review on Hilbert‑Huang transform: Method and its applications to geophysical

346   studies, Rev. Geophys., 46, RG2006, https://doi.org/10.1029/2007RG000228, 2008.

347   Hudson, D., Alves, O., Hendon, H. H., Wang, G.: The impact of atmospheric initialisation on seasonal

348   prediction of tropical Pacific SST, Clim. Dynam., 36, 1155-1171, https://doi.org/10.1007/s00382-010-

349   0763-9, 2011.

350   Jain, P. and Deo, M. C.: Neural networks in ocean engineering, Ships Offshore Struc., 1, 25-35,

351   https://doi.org/10.1533/saos.2004.0005, 2006.

352   Khan, M. Z. K., Sharma, A., and Mehrotra, R.: Global seasonal precipitation forecasts using improved sea

353   surface   temperature   predictions,   J   Geophys.   Res.   -Atmos.,   122,   4773-4785,

354   https://doi.org/10.1002/2016JD025953, 2017,

355   Kim, Y., Kim, H., and Ahn, I. G.: A study on the fatigue damage model for Gaussian wideband process of

356   two   peaks   by   an   artificial   neural   network,   Ocean   Eng.,   111,   310-322,

357   https://doi.org/10.1016/j.oceaneng.2015.11.008, 2016.

358   Kumar, M., Parmar, C., Chaudhary, V., Kumar, A., and SST-1 team.: Observation of plasma shift in SST-1

359   using optical imaging diagnostics, J Phys. Conf. Ser., 823, 012056, https://doi.org/10.1088/1742-

360   6596/823/1/012056, 2017.

361   Lee, H. S.: Estimation of extreme sea levels along the Bangladesh coast due to storm surge and sea level rise

362   using EEMD and EVA, J Geophys. Res.-Oceans, 118, 4273-4285, https://doi.org/10.1002/jgrc.20310,

363   2013,

364   Lee, T. L.: Back-propagation neural network for long-term tidal predictions, Ocean Eng., 31, 225-238,

365   https://doi.org/10.1016/S0029-8018(03)00115-X, 2004.

366   López, I., Aragonés, L., Villacampa, Y., and Serra, J. C.: Neural network for determining the characteristic

367   points of the bars, Ocean Eng., 136: 141-151, https://doi.org/10.1016/j.oceaneng.2017.03.033, 2017.

368   Monteiro, E., Yvonnet, J., He, Q. C.: Computational homogenization for nonlinear conduction in

369   heterogeneous   materials   using   model   reduction.   Comp.   Mater.   Sci.,   42,   704-712,

370   https://doi.org/10.1016/j.commatsci.2007.11.001, 2008.

371   Motulsky, H. J. and Ransnas, L. A.: Fitting curves to data using nonlinear regression: a practical and

372   nonmathematical review, Faseb J., 1, 365-374. https://doi.org/10.1096/fasebj.1.5.3315805, 1987.

373   Pan, H., Guo, Z., Wang, Y., and Lv, X.: Application of the EMD method to river tides, J. Atmos. Ocean. Tech.,

374   35, 809-819, https://doi.org/10.1175/JTECH-D-17-0185.1, 2018.

375   Pearson, R. K. and Pottmann, M.: Gray-box identification of block-oriented nonlinear models, J. Process

376   Contr., 10, 301-315, https://doi.org/10.1016/S0959-1524(99)00055-4, 2000.

377   Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax., M. G.: Daily high-

378   resolution-blended analyses for sea surface temperature, J. Climate, 20, 5473-5496,

379   https://doi.org/10.1175/2007JCLI1824.1, 2007.

380   Sadeghifar, T., Motlagh, M. N., Azad, M. T., and Mahdizadeh, M. M.: Coastal wave height prediction using

381   Recurrent Neural Networks (RNNs) in the south Caspian Sea, Mar. Geod., 40, 454-465,

382   https://doi.org/10.1080/01490419.2017.1359220, 2017.

383   Savitha, R. and Mamun, A. A,: Regional ocean wave height prediction using sequential learning neural

384   networks, Ocean Eng., 129: 605-612, https://doi.org/10.1016/j.oceaneng.2016.10.033, 2017.

385   Sukresno, B., Hanintyo, R., Kusuma, D. W., Jatisworo, D., and Murdimanto., A.: Three-way error analysis

386   of sea surface temperature (SST) between HIMAWARI-8, buoy, and mur SST in SAVU Sea, Int. J.

387   Remote Sens. Earth Sci., 15, 25-36, https://doi.org/10.30536/j.ijreses.2018.v15.a2855, 2018,

388   Takakura, T., Kawamura, R., Kawano, T., Ichiyanagi, K., Tanoue, M., and Yoshimura, K.: An estimation of

389   water origins in the vicinity of a tropical cyclone's center and associated dynamic processes, Clim.

390   Dynam., 50, 555-569, https://doi.org/10.1007/s00382-017-3626-9, 2018.

391   Tang, L., Dai, W., Yu, L., and Wang, S.: A novel CEEMD-based EELM ensemble learning paradigm for crude

392   oil price forecasting, Int. J. Inf. Tech. Decis., 14, 141-169, https://doi.org/10.1142/S0219622015400015,

393   2015.

394   Wang, S., Zhang, N., Wu, L., and Wang, Y.: Wind speed forecasting based on the hybrid ensemble empirical

395   mode decomposition and GA-BP neural network method, Renew. Energ., 94, 629-636,

396   https://doi.org/10.1016/j.renene.2016.03.103, 2016.

397   Wang, W., Chau, K., Xu, D., and Chen, X.: Improving forecasting accuracy of annual runoff time series using

398   ARIMA based on EEMD decomposition, Water Resour. Manag., 29, 2655-2675,

399   https://doi.org/10.1007/s11269-015-0962-6, 2015.

400   Wang, W., Tang, R., Li, C., Liu, P., and Luo, L.: A BP neural network model optimized by Mind Evolutionary

401   Algorithm for predicting the ocean wave heights, Ocean Eng., 162, 98-107,

402   https://doi.org/10.1016/j.oceaneng.2018.04.039, 2018.

403   Wang, Y., Wilson, P. A., Zhang, M., and Liu, X.: Adaptive neural network-based backstepping fault tolerant

404   control for underwater vehicles with thruster fault, Ocean Eng., 110, 15-24,

405   https://doi.org/10.1016/j.oceaneng.2015.09.035, 2015.

Ocean Science
Discussions

406   Wiedermann, M., Donges, J. F., Handorf, D., Kurths, J., and Donner, R. V.: Hierarchical structures in

407   Northern Hemispheric extratropical winter ocean–atmosphere interactions, Int. J. Climatol., 37, 3821-

408   3836, https://doi.org/10.1002/joc.4956, 2017.

409   Wu, L. C., Kao, C. C., Hsu, T. W., Jao K. C. and Wang, Y. F.: Ensemble empirical mode decomposition on

410   storm   surge   separation   from   sea   level   data,   Coast.   Eng.   J.,   53,   223-243,

411   https://doi.org/10.1142/S0578563411002343, 2011.

412   Wu, Z. and Huang, N. E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method[J].

413   Adv. Adap. Data Anal., 1, 1-41, https://doi.org/10.1142/S1793536909000047, 2009.

414   Yeh, J. R., Shieh, J. S., and Huang, N. E.: Complementary ensemble empirical mode decomposition: A novel

415   noise   enhanced   data   analysis   method,   Adv.   Adap.   Data   Anal.,   2,   135-156,

416   https://doi.org/10.1142/S1793536910000422, 2010.

417   Zheng, X. T., Xie, S. P., Lv, L. H., and Zhou, Z. Q.: Intermodel uncertainty in ENSO amplitude change tied

418   to Pacific Ocean warming pattern, J. Climate, 29, 7265-7279, https://doi.org/10.1175/JCLI-D-16-0039.1,

419   2016.

420   Zhu, J., Huang, B., Kumar, A., and Kinter, J. L.: Seasonality in prediction skill and predictable pattern of

421   tropical Indian Ocean SST, J. Climate, 28, 7962-7984, https://doi.org/10.1175/JCLI-D-15-0067.1, 2015.