

# A SIMULATION MODEL OF SEAWATER VERTICAL TEMPERATURE BY USING BACK-PROPAGATION NEURAL NETWORK

Ning Zhao<sup>1,2)</sup>, M. S.

Zhen Han<sup>1,3)</sup>, Ph. D

<sup>1)</sup> College of Marine Sciences, Shanghai Ocean University, Shanghai 201306  
China

<sup>2)</sup> Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga 8168580  
Japan

<sup>3)</sup> Collaborative Innovation Center for Distant-water Fisheries, Shanghai Ocean University, Shanghai  
201306, China

## ABSTRACT

*This study proposed a neural-network-based model to estimate the ocean vertical water temperature from the surface temperature in the northwest Pacific Ocean. The performance of the model and the sources of errors were assessed using the Gridded Argo dataset including 576 stations with 26 vertical levels from surface (0 m)–2,000 m over the period of 2007–2009. The parameter selection, model building, stability of the neural network were also investigated. According to the results, the averaged root mean square error (RMSE) of estimated temperature was 0.7378 °C and the correlation coefficient R was 0.9967. More than 67% of the estimates from the four selected months (January, April, July and October) lay within  $\pm 0.5$  °C. When counting with errors lower than  $\pm 1$  °C, the lowest percentage was 83%.*

**Keywords:** neural network; Argo data; vertical structure; surface temperature

## INTRODUCTION

As development of remote sensing progressed, large amounts of sea surface information could be obtained daily. However, few data were available for the subsurface and deeper ocean. For many reasons, obtaining in-situ data has always posed a difficult problem. Obtaining such data costs exorbitant amounts of time and money. Even over recent years, the problem of lack of in-situ observations of ocean subsurface data is still not solved [1].

Swain et al. (2006) introduced an artificial neural network to estimate the mixed layer depth from surface parameters [2]. Ballabrera-Poy et al. (2009) compared linear and non-linear models of the vertical salinity structure based on temperature observations and proposed that the neural network method performs better than the linear models when the models introduced the surface observations [3]. Comparing with the traditional methods, as Ballabrera-Poy et al. and Swain et al. mentioned [2, 3], the neural network method seems to have a great potential ability in estimating ocean structure.

In accordance to previous studies, this study introduced a back-propagation neural network model to simulate the sea vertical temperature structure. After training with the

historical temperature data, this model could use ocean surface measurements (SST) only as input parameters and estimate the unknown current subsurface temperature structure (the available range of depth depends on the initial field for model building). Following this subsection, data and methods are presented in the section 2; results and discussion are described in section 3; the section 4 is devoted to summary and conclusion, respectively.

## DATA

The main dataset of this paper was obtained from the China Argo Real-time Data Centre (<http://www.argo.org.cn>) covering the region of 20–35°N and 145–180°E in the period 2007-2009. This dataset is a gridded monthly average temperature product with a spatial resolution of 1°×1°. The dataset contains data for 576 stations with 26 vertical levels from surface (0 m)–2,000 m. Each level in every station and every month was picked as one set of data. As a result, 538,632 data sets were generated (the stations at 32°N, 173°E and 35°N, 172°N only had 19 levels from 0–1000 m).

For the model building, two other datasets were used. The original Argo profiles data were obtained from the database of the Global Ocean Data Assimilation Experiment (<http://www.usgodae.org>). The reference series of the weekly maps of absolute dynamic topography (ADT) products were obtained from the website of AVISO and this product contains gridded sea surface heights above geoid [4]. The weekly data were obtained and averaged to match the time resolution of the main dataset in this study.

For the comparison and discussion of results, the reanalysis data obtained from the database of the MyOcean2 project were used [5]. Because of the differences of resolution between the results and the reanalysis data, the weekly data had been averaged by month and processed to match the spatial resolution of estimated data in this paper using the Ordinary Kriging method [6]. Following the former processing, the reanalysis data were also divided in the same way as the Gridded Argo data and a total of 13,238 sets were obtained within each month.

## METHODS

In order to simulate the vertical temperature structure, an initial field of temperature was required. In this research, the data from 2007 were selected as the initial field for building the model, while the rest (data in 2008 and 2009) were used for simulation. During the model building phase, 70% of data in 2007 were used as training data and 30% for testing. According to the complexity of the vertical temperature structure, the model was based on a simplified mapping relationship and built by a back-propagation neural network (BP-NN). In this network, one single hidden layer between the input layer and the output layer was used [7]. As the main aim of this study was to reconstruct the unknown vertical temperature structure from surface parameters which could be directly or easily obtained from remote sensing, some parameters, such as the subsurface heat advection, radiation and surface heat fluxes, should be excluded from the list. Subsequently, eight parameters were selected as candidates including: geographical location (longitude and latitude), sea surface temperature (SST), Depth (D), sea surface height anomalies (SSHA), absolute dynamic topography (ADT) and noisy data (N1 and N2). The noisy data N1 and N2 were random numbers to ensure the result of the method is reliable.

The Mean Impact Value (MIV) method was used in this experiment which could determine the impacts of the each input parameter on output [8, 9]. The degree of impact could be described by the absolute value of the MIV. The experimental model had eight input neurons for the eight parameters including the geo-geographical location, ten hidden neurons and one output neuron. It should be noted that all parameters were normalized to avoid unwanted influences and outliers by using the Standardized Moment. For example, for a set of numbers  $X(x_1, x_2, x_3, \dots, x_n)$ , the basic equation of the standardized moment method could be described as follows:

(1)

where:

$a_r$  is the  $r$ th standardized moment ( $r=1$  in this study)

$\mu_r$  is the  $r$ th moment about the mean

$\sigma$  is the standard deviation

$z_i$  is calculated by:

$$z_i = (x_i - \bar{x}) / \sigma$$

$\bar{x}$  is the mean value of  $X$  and  $a_r$  should be zero.

To reduce the uncertainty of this simple neural network, this model was trained ten times and the absolute values of the result were averaged to show the impact degrees of each parameter. As shown in Tab. 1, the randomly created N1 and N2 played insignificant roles in the model. In addition, it was not surprising that SST and depth showed great impact on the model output. Interestingly, two other parameters (SSHA and ADT) did not have considerable distinct impacts relative to SST and depth. Their MIVs were quite small.

Tab. 1. The averaged MIVs of six parameters after ten times trained.

Parameters	SST	SSHA	Depth	ADT	N1	N2
Average MIV	0.0169	0.0004	0.0885	0.0011	0.0012	0.0013

Concerning the results of the selection experiment, the mapping relation between water temperature and other parameters in this study could be simply shown as follows:

$$F(\text{SST}, \text{D}, \text{Lon}, \text{Lat}) \rightarrow T \quad (2)$$

In this network, the input layer is composed of: 1) four neurons for the four input parameters (SST, depth, longitude and latitude); 2) the hidden layer (layer 1) neurons are set to have a hyperbolic tangent activation function (the numbers of the neurons in hidden layer will be discussed in the later section); 3) the output layer (layer 2) has a single neuron whose activation function is the identity function and its value is equal to water temperature ( $T$ ). Each function has its own set of coefficients (weights  $w$  and biases  $b$ ). The values of each neuron  $i$  in the hidden layer (layer 1) are calculated by:

$$H_i = f_i^1 \left( \sum_{j=1}^4 w_{ij}^1 x_j + b_i^1 \right) \quad (3)$$

$$i=1, 2, 3, \dots, n; j = 1, 2, 3, 4$$

where:

$H_i$  is the value of neuron  $i$  in the hidden layer

$f_i^1$  is the transfer function (hyperbolic tangent sigmoid function)

$x_j$  is the value of input parameter

$j$  is the number of neurons in the input layer

$i$  is the number of neurons in the hidden layer

The value of the neuron in the output layer is calculated by:

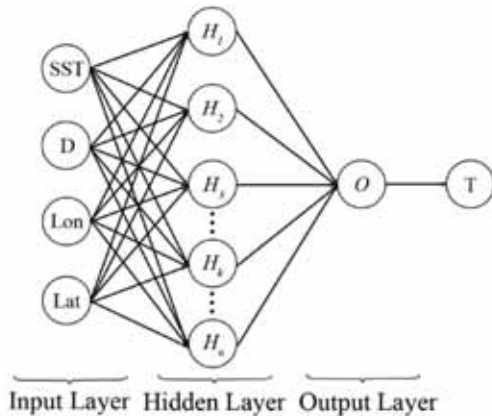
$$O = f\left(\sum_{i=1}^n v_i H_i + b\right) \quad (4)$$

where:

O is the value of the neurons in the output layer (layer 2)  
f is the transfer function (linear function).

All the parameters and transfer functions of this BP-NN are shown in Table 2. In this network, the Levenberg-Marquardt back-propagation algorithm was used for training networks [10].

Fig. 1. Structure of the back-propagation neural network.



The dataset was divided by months and years. The data for 2007 was set for model building and others for estimation (the data for 2008 and 2009). Two test indicators were used to assess the models: the root mean square error (RMSE) and the Pearson's correlation coefficient (R):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (T_{1i} - T_{2i})^2} \quad (5)$$

$$R = \frac{\sum_{i=1}^n (T_{1i} - \bar{T}_{1i})(T_{2i} - \bar{T}_{2i})}{\sqrt{\sum_{i=1}^n (T_{1i} - \bar{T}_{1i})^2 \sum_{i=1}^n (T_{2i} - \bar{T}_{2i})^2}} \quad (6)$$

where:

n is the number of datasets

$T_{1i}$  is the estimated temperature

$T_{2i}$  is the observed temperature

Before the models were built, the number of neurons in the hidden layer required a solution first by an experimental model. Fig. 2 shows the RMSE and the Pearson's product-moment coefficient (R) of the experimental model within different number of neurons in the hidden layer. It is shown that the model of this experiment had over fitted the training data when using more than fourteen neurons in the hidden layer. In Fig.2b, the same result could be found. For this reason, the neural network of this study was built by using fourteen neurons in the hidden layer.

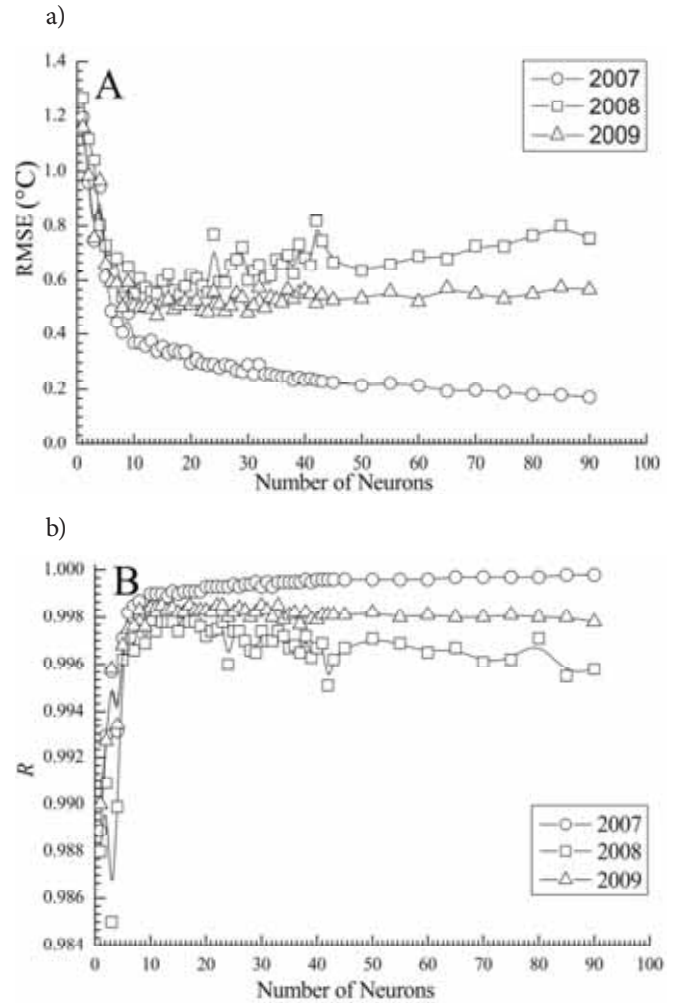


Fig. 2. Comparison of networks with different number of neurons in the hidden layer: (a) RMSE. (b) coefficient (R).

## RESULTS

After the model was built, the data were calculated month by month and year by year. The results are presented in Table 2. Two indicators show that the BP-NN model was fluctuating narrowly on a monthly time scale (RMSE < 1 °C and R > 0.99).

The quantities of errors over the four months were also counted and the results are shown in Tab. 3 (14,962 sets data in each month). During all four months, over 67% of data errors were lower than 0.5 °C. When counting with errors lower than 1°C, the lowest percentage was 83%.

## DISCUSSION AND CONCLUSION

In this study, a simulation model of ocean vertical water temperature was performed by using the back-propagation neural network. After training the model with the historical temperature data, it could use ocean surface measurements (SST) only as input parameters and estimate the unknown current subsurface temperature structure. In this model, there were 538,632 sets of data calculated month by month and year by year. The total RMSE is 0.7378 °C and the correlation

Tab.2. Results of BP-NN.

Month	2007		2008		2009	
	RMSE (°C)	R	RMSE (°C)	R	RMSE (°C)	R
1	0.3693	0.9989	0.5815	0.9976	0.5127	0.9984
2	0.3657	0.9989	0.6167	0.9972	0.5711	0.9975
3	0.3739	0.9989	0.6582	0.9961	0.5612	0.9975
4	0.3752	0.9990	0.8937	0.9952	0.5255	0.9980
5	0.4245	0.9988	0.7847	0.9959	0.7957	0.9963
6	0.5319	0.9983	0.6984	0.9972	0.9404	0.9943
7	0.5580	0.9982	0.8205	0.9962	0.8922	0.9959
8	0.4640	0.9988	0.9324	0.9962	0.8728	0.9960
9	0.4811	0.9987	0.7933	0.9966	0.8380	0.9967
10	0.4695	0.9987	0.6396	0.9977	0.8355	0.9967
11	0.5076	0.9983	0.6045	0.9977	0.9295	0.9950
12	0.5081	0.9982	0.5448	0.9979	0.8633	0.9951
Average	0.4524	0.9986	0.7140	0.9968	0.7615	0.9965

coefficient R is 0.9967. All the results shown this BP-NN model has a good performance.

**SELECTION EXPERIMENT**

During the first part of model building, a selection experiment was made to decide which parameter should be chosen as the input parameter. An interesting aspect of this experiment was that the SSHA and ADT did not show any considerable impact on the output data and the impact was even lower than that of the noisy data N1 and N2. Indeed, as mentioned in many studies concerning the upper ocean, the sea surface height is an important parameter. This parameter could be used to calculate the geo-strophic current or estimate the mixed layer depth [2]. And it could also provide a way to estimate the upper ocean heat content which could greatly influence the temperature profiles [11]. However, in this study, it did not play an important role, contrary to previous studies. One possible explanation is: the neural network was focused on the relationship between the input and the output and reconstructed the whole system via numerical experiments which used the different weights and biases in the equations to

try to rebuild that relation. Obviously, this approach simplified the complex inner processes. And it might also because of that the historical vertical structures had already been input when the model was building.

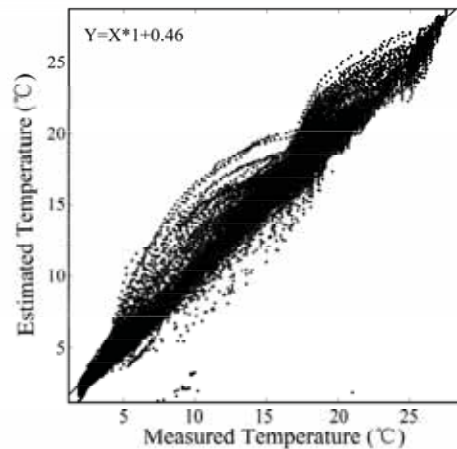


Fig. 3. The model performance when the original Argo profiles data were input.

Tab. 3. Errors over four months during 2008 and 2009.

Error	2008		2009	
	< 0.5 °C	0.5 ~ 1 °C	< 0.5 °C	0.5 ~ 1 °C
Jan	12338(82.46%)	1770(11.83%)	12175(81.37%)	1974(13.19%)
Apr	10074(67.33%)	2449 (16.37%)	12549(83.87%)	1916(12.81%)
Jul	11137(74.44%)	2612(17.45%)	10916(72.95%)	2551(17.05%)
Oct	12422(83.02%)	1901(12.71%)	10628(71.03%)	2365(15.81%)

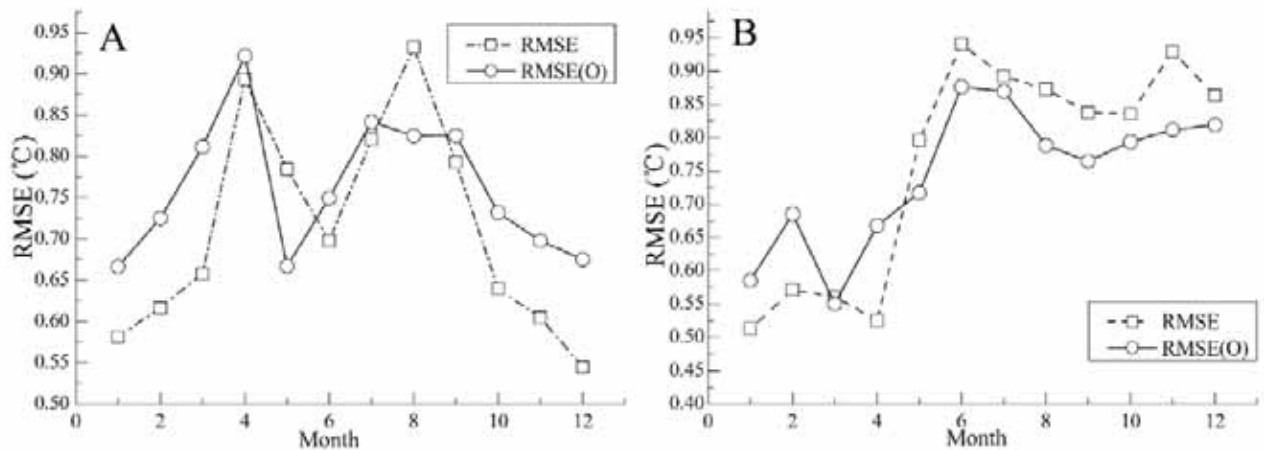


Fig. 4. Model performances and the differences between initial fields and test fields: (a) 2008. (b) 2009

### GRIDDED ARGO PRODUCT AND ORIGINAL PROFILES

It is obvious that the regular dataset had a stabilizing effect on the neuron network. However, the model would have to be assessed for suitability for practical applications. In addition, this is a good opportunity to assess the potentiality of the artificial neural network. Thus, a further experiment was performed to assess the model when using the original Argo profile data. During this experiment, the initial temperature field was still based on the gridded Argo data during 2007. The original Argo profile data during January 2008 was picked for the simulation including 320 profiles within 26,512 sets of data.

The RMSE was 1.2605 °C and the coefficient R was 0.9880. Fig. 3 gives the linear regression of the results in this experiment. The slope was still 1 and the intercept was slightly higher than before (intercept=0.46). These indicated the BP-NN model was still effective.

### INTER-ANNUAL VARIATION AND INITIAL FIELD

In the simulation model, a reliable initial field is obligatory.

During this study, the initial temperature field was based on the gridded dataset (monthly averaged and vertically delimited). Both initial and simulated data were set in the same month but over different years to ensure the external conditions are similar.

But it should be pointed out that the inter-annual variation and its influences on ocean vertical structure were not considered in this model. This could give a great impact on the model performance. To confirm that, the differences between the data in 2007 and data in 2008-2009 are given in Fig. 4 (as the RMSE(O)).

As expected, a high correlation was found between those two lines:  $R=0.7517$  during 2008 and  $R=0.9129$  during 2009. Some studies indicated that it could further ascribe these errors to the deficiency of the initial field [12]. A reliable reason might be seen from the comparison between estimated temperature and reanalysis data. Fig. 5 gives this comparison. This reveals that a major part of the error in this model is due to the differences between the data for model building (initial fields) and model simulation. The solution is to expand the samples in the initial field and adding the temporal parameters into the model.

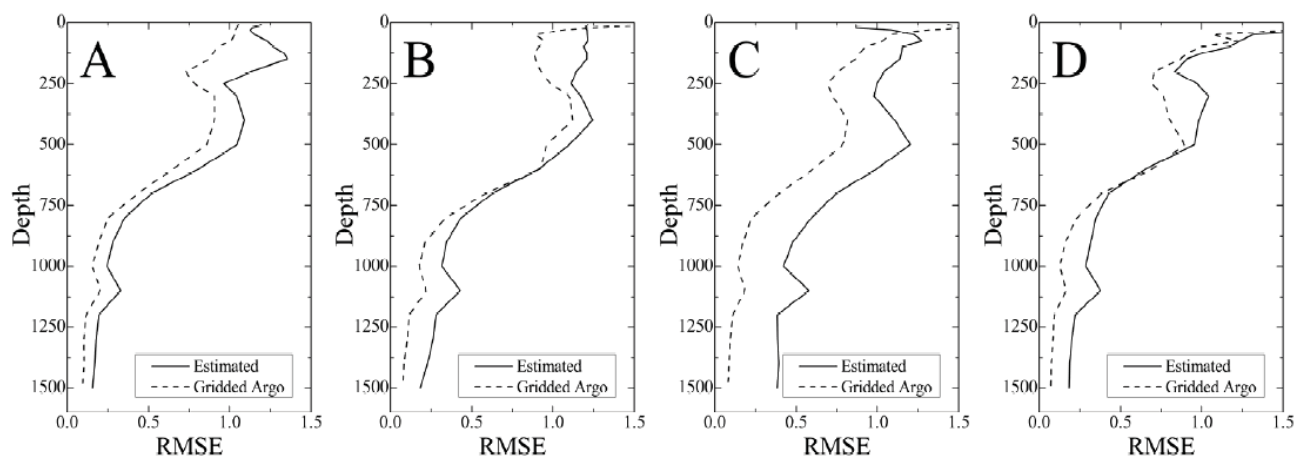


Fig. 5. The relation between RMSE and differences of two kinds of data (Gridded Argo and reanalysis) as a function of depth: (a) January. (b) April. (c) July. (d) October.

## ERRORS WITH DEPTH

The RMSE values over four months as a function of depth is shown in Fig. 6. The lowest error appeared at the surface and the depth over 800 m while some depths between 0–800 m showed higher errors.

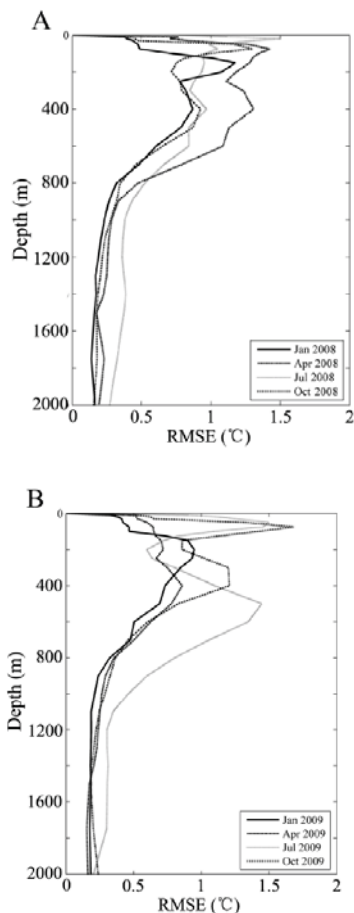


Fig. 6. RMSE of four months as a function of depth: (a) 2008. (b) 2009.

At some depth between 0–800 m, the errors became higher than 1 °C. The first peak values appeared at the surface part and the second appeared at the depth of about 400 m. To reduce these errors, the segment-based model from the study of Chu et al. (2000) might be helpful [13]. For further research, the possible procedure might be described as follows: divide the historical temperature data into several layers based on the vertical parameters (e.g. historical mixed layer depth and thermocline depth) and train the model layer by layer. This could not only reduce the quantity of the samples allowing the model to run faster, but also make the training data more representative.

## ACKNOWLEDGMENTS

We thanked the Argo project, AVISO group and MyOcean2 project for access to their climatological and observation database. This paper was financially supported by the

Foundation of Key Laboratory of Integrated Monitoring and Applied Technology for Marine Harmful Algal Blooms (MATHAB) 201401.

## BIBLIOGRAPHY

1. Nardelli B.B. and Santoleri R.: Methods for the reconstruction of vertical profiles from surface data: Multivariate analyses residual GEM and variable temporal signals in the North Pacific Ocean, *J. Atmos. Ocean. Technol.*, 22 (11), 1762–1781, 2005.
2. Swain D., Ali M.M., Weller R.A.: Estimation of mixed layer depth from surface parameters, *J. Mar. Res.*, 64, 745–758, 2006.
3. Ballabrera-Poy J., Mourre B., Garcia-Ladona E., et al.: Linear and non-linear T-S models for the eastern North Atlantic from Argo data: Role of surface salinity observations, *Deep-Sea Res. I*, 56, 1605–1614, 2009.
4. AVISO: Ssalto/Duacs User handbook: (M)SLA and (M) ADT Near-Real Time and Delayed Time Product. Ref: CLS-DOS-NT-O6-034, 39–41, 2012.
5. Cressie N.: The Origins of Kriging, *Math. Geol.*, 22(3), 239–252, 1990.
6. MyOcean: MyOcean catalogue of products v2.1: the ocean in one click, 2012.
7. Elsken T.: Even on finite test sets smaller nets may perform better, *Neural Netw.*, 10(2), 369–385, 1997.
8. Dombi G.W.; Nandi P.; Saxe J.M., et al.: Prediction of rib fracture injury outcome by an artificial neural network. *J. Trauma*, 39(5), 915–921, 1995.
9. Li Z.D. and Sun W.: A new method of calculate weights of attributes in spectral clustering Algorithms, *IEEE: Int. Conf. Inf. Technol. Comput. Eng. Manag. Sci.*, 8–60, 2011.
10. Hagan M.T., Demuth H.B., Beale M.N.: *Neural Network Design*, PWS Publ. Co. Int. Thomson Publ. Inc., Boston, 1996.
11. Kelly K.A., Thompson L., Vivier F.: Heat content changes in the Pacific Ocean, *Sci.*, 284, 1735, 1999.
12. Hashemi M.R., Ghadampour Z., Neill S.P.: Using and artificial neural network to model seasonal changes in beach profiles, *Ocean Eng.*, 37, 1345–1356, 2010.
13. Chu P.C., Fan C., Liu W.T.: Determination of vertical thermal structure from sea surface temperature, *J. Atmos. Ocean. Technol.*, 17 (7), 971–979, 2000.

## CONTACT WITH THE AUTHORS

Ning Zhao, M. S.  
Zhen Han, Ph.D.

College of Marine  
Sciences, Shanghai  
Ocean University,  
Shanghai 201306  
**CHINA**

e-mail: ningz1120@riam.kyushu-u.ac.jp  
e-mail: zhhan@shou.edu.cn