Final response in the interactive discussion

Dear Referees, dear Editor,

We would like to thank you very much for your positive comments and constructive suggestions to our manuscript "*Retrieving monthly and interannual pH_T on the East China Sea shelf using an artificial neural network: ANN-pH_T-v1*".

In this document, we would like to provide our responses to the comments of each of the referees in one single document and to outline the corresponding changes to the manuscript. We will represent the referee comment in **bold** font, and our response in normal font. Quotations from the original manuscript will be in *italics*, changes as part of the manuscript revision will be highlighted as underlined. For the sake of clarity and brevity, we have omitted the introductory parts of the referee report (this omittance is marked as [...]).

We hope that our response together with the revision of the manuscript sufficiently addresses the referee' concerns.

Sincerely,

Xiaoshuang Li (on behalf of the author team)

Referee comment #1 (by Richard Mills)

[…] I came away from my reading of the paper with the following major questions/concerns which, if addressed, will greatly improve the quality of the paper:

1. **First, since the paper has been submitted to a model development journal, I would like to see more information on how and why the authors arrived at the particular form of the machine-learning model they used, and how this model performed against some other possible model architectures. The authors have used a feed-forward multilayer perceptron network with two hidden layers (with 40 neurons in the first layer and 16 in the second) and full connectivity between the layers. Why did the authors decide on two layers, and how did they choose the number of neurons in each layer? (They do state that they tried varying the number of neurons in each layer, but don't give further details.) How did they choose the activation function? And why did they choose a neural network, instead of another approach such as k-nearest neighbors, random forest regression, or support vector regression? When I first started working in machine learning, around two decades ago, it would not have been expected for authors to try a variety of different types of models, as this would likely involve substantial code development effort, as well as possibly significant computational expense for training models. Today, however, it is easy to try many different models, as code provided in many easily obtained packages such as Scikit-learn or those provided by Matlab (the environment that the authors use for this study), and it is becoming the norm for papers presenting the development of machine-learning models to compare several types to determine the one that performs best for the chosen task. I would like to see some comparison against other models (some of the ones easily constructed using Matlab) to demonstrate that the ANN is the most appropriate choice.**
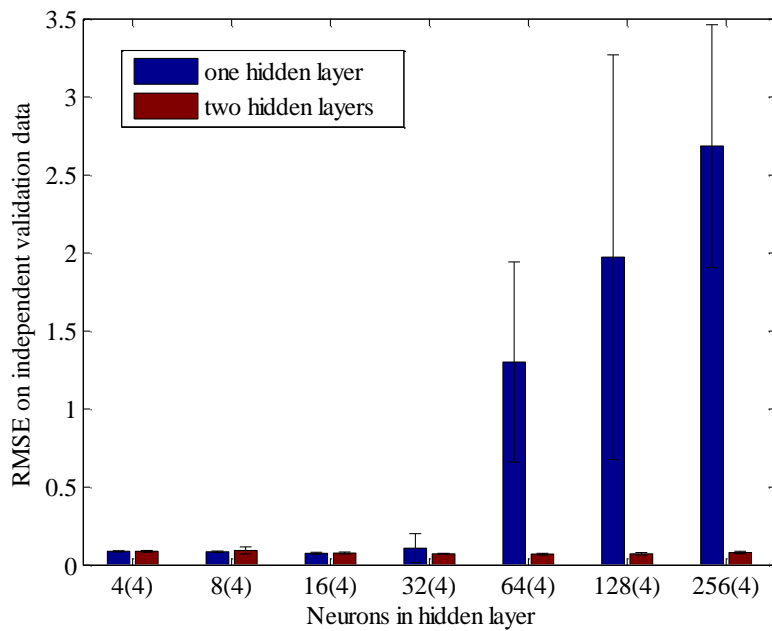
We thank the referee for the suggestion: the required details should, in fact, be provided to the reader. We will add, in the revised manuscript, the corresponding information (Why did the authors decide on two layers, and how did they choose the number of neurons in each layer? How did they choose the activation function? And why did they choose a neural network, instead of another approach such as k-nearest neighbors, random forest regression, or support vector regression?)—as follows ll. 111-127 and 160-163 of the revised manuscript:
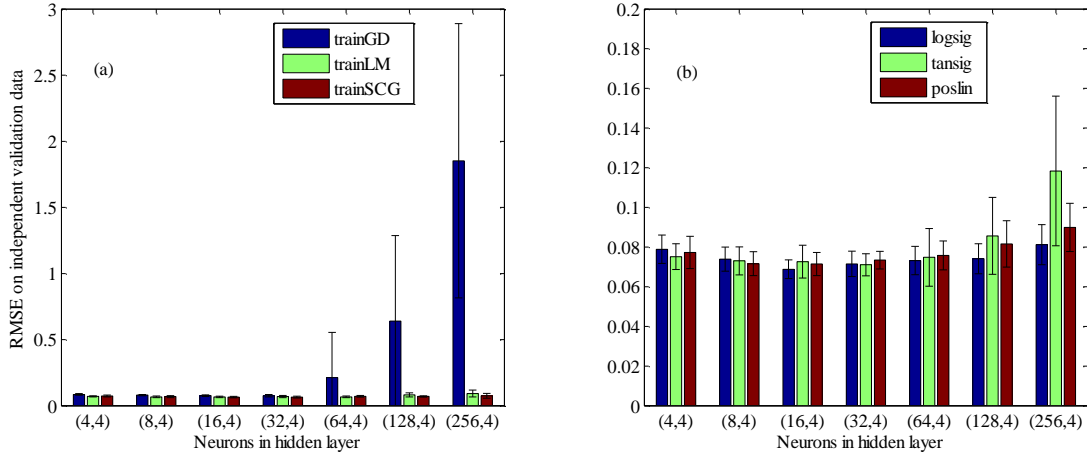
ll. 111-127 of the revised manuscript:

In our study, calculations were done in the MathWorks Matlab environment, using the Deep Learning Toolbox.
First, we compared the performance of one hidden layer vs. two hidden layers in predicting independent validation data. The number of neurons varied from 2^2 to 2^8 for the first hidden layer and was fixed at four in the second hidden layer for the two
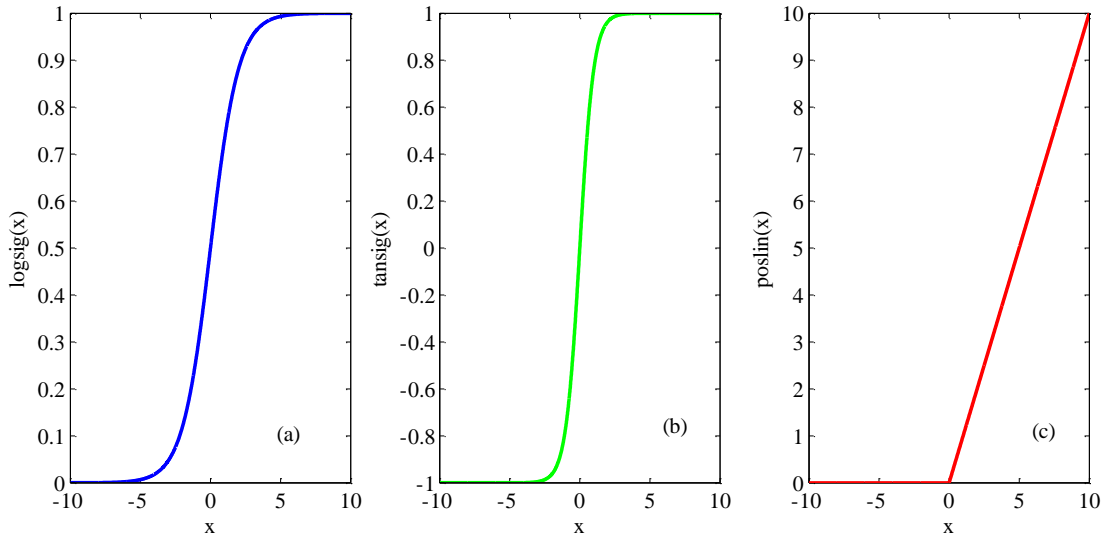
hidden layers model; the number of neurons in the first layer was the same in the one hidden layer vs. two hidden layers model (Fig. 4). The ten-fold cross-validation showed that the model with two hidden layers performed better as the number of neurons increased. Second, in order to choose suitable training techniques and activation functions of the ANN model with two hidden layers, we tested three training functions (Gradient descent backpropagation (trainGD), Levenberg-Marquardt backpropagation (trainLM), and Scaled conjugate gradient backpropagation (trainSCG)), which differed in how the weights are modified, and three transfer functions (Log-sigmoid transfer function (logsig), Hyperbolic tangent sigmoid transfer function (tansig), and Positive linear transfer function (poslin)) (Fig. 5). The output values of logsig, tansig and poslin were compressed onto [0, 1], [-1, 1], and [0, +∞], respectively (Fig. S1). As the number of neurons increased, the performances of trainGD and tansig became poor. Although there was no obvious difference between trainLM and trainSCG, the training technique trainSCG was selected and the transfer function logsig was applied to two hidden layers considering the overall performance (Fig. 5). Third, in the training phase of the ANN model, the number of neurons was tested, varying from 4 to 128 for two hidden layers (Table S1). Best performance for both training data and independent validation data was obtained with 40 neurons in the first hidden layer and 16 neurons in the second layer. Finally, different combinations of input variables were tested to choose the optimal architecture of the ANN model (Table 2); best performance was obtained using longitude, latitude, month, T, S, DO, N, P and Si as input variables.

Figure 4 (revised): Comparison of the performance of one hidden layer vs. two hidden layers in predicting independent validation data. The number of neurons in the first hidden layer was the same in the one hidden layer vs. two hidden layers model, numbers in parentheses show the number of neurons in the second hidden layer (for the two hidden layers model). Bars show the mean and standard deviation of the Root-Mean-Square-Error over a ten-fold cross-validation, for different numbers of neurons in the first hidden layer.

Figure 5 (revised): Comparison of the performance of different training functions and transfer functions on independent validation data. (a)-three training functions: Gradient descent backpropagation (trainGD), Levenberg-Marquardt backpropagation (trainLM), and Scaled conjugate gradient backpropagation (trainSCG); (b) three transfer functions: Log-sigmoid transfer function (logsig), Hyperbolic tangent sigmoid transfer function (tansig), and Positive linear transfer function (poslin). Bars show the mean and standard deviation of the Root-Mean-Square-Error over a ten-fold cross-validation, for different numbers of neurons in the first hidden layer.

Figure S1 (revised): Comparison of three transfer functions. (a)-Log-sigmoid transfer function (logsig); (b) Hyperbolic tangent sigmoid transfer function (tansig); (c)-Positive linear transfer function (poslin).

Table S1 (revised): The performance of different number of neurons for two hidden layers in the training step. Three statistics are the coefficient of determination (R^2), the root mean squared error (RMSE), and the mean absolute error (MAE).

Model	Number of neurons		Training data			Independent validation data		
	first hidden	second hidden	R^2	RMSE	MAE	R^2	RMSE	MAE
1	4	4	0.68	0.071	0.054	0.67	0.072	0.057
2	8	4	0.70	0.070	0.050	0.67	0.069	0.050
3	16	4	0.76	0.062	0.045	0.76	0.062	0.045
4	32	4	0.74	0.063	0.046	0.79	0.062	0.048
5	40	4	0.76	0.062	0.044	0.76	0.061	0.045
6	64	4	0.79	0.058	0.041	0.78	0.056	0.043
7	128	4	0.76	0.062	0.045	0.74	0.062	0.044
8	8	8	0.73	0.065	0.047	0.73	0.065	0.048
9	16	8	0.78	0.059	0.042	0.78	0.058	0.044

10	32	8	0.78	0.059	0.042	0.83	0.053	0.039
11	40	8	0.79	0.059	0.042	0.77	0.055	0.040
12	64	8	0.77	0.061	0.044	0.76	0.059	0.042
13	128	8	0.77	0.060	0.042	0.79	0.059	0.043
14	16	16	0.79	0.057	0.041	0.85	0.054	0.041
15	32	16	0.80	0.057	0.040	0.69	0.059	0.043
16	40	16	0.82	0.054	0.039	0.81	0.053	0.039
17	64	16	0.79	0.059	0.041	0.76	0.057	0.040
18	128	16	0.79	0.058	0.040	0.78	0.059	0.043
19	32	32	0.78	0.059	0.042	0.75	0.058	0.039
20	40	32	0.79	0.058	0.041	0.79	0.055	0.040
21	64	32	0.78	0.059	0.042	0.83	0.052	0.040
22	128	32	0.79	0.058	0.041	0.79	0.056	0.041
23	40	40	0.77	0.060	0.043	0.77	0.060	0.044
24	64	40	0.79	0.058	0.042	0.75	0.060	0.043
25	128	40	0.80	0.057	0.040	0.78	0.057	0.042
26	64	64	0.78	0.060	0.042	0.78	0.057	0.040
27	128	64	0.72	0.068	0.050	0.65	0.067	0.048
28	128	128	0.72	0.067	0.049	0.65	0.072	0.051

II. 160-163 of the revised manuscript:

(Dai and Trenberth, 2002). As a reference, the performance of some other empirical approaches, including MLR, multi-variate nonlinear regression (MNR), decision tree, random forest, and Support Vector Machine (SVM) regression, is shown in Table 3. The selected ANN model (Table 2, Model#10) showed better performance than the other tested approaches using the same input variables (Table 3).

Table 3 (revised): Model comparison between traditional empirical methods (MLR and MNR) and mechine-learning based empirical methods (Decision tree, Random Forest, and SVM). The statistics was derived from confimatory dataset (training data independent validation data) using input variables: T, S, DO, N, P, and Si. Note R^2 statistics in our study was based on the calculation of coefficient of determination, therefore negative R^2 could be derived if there were strong bias.

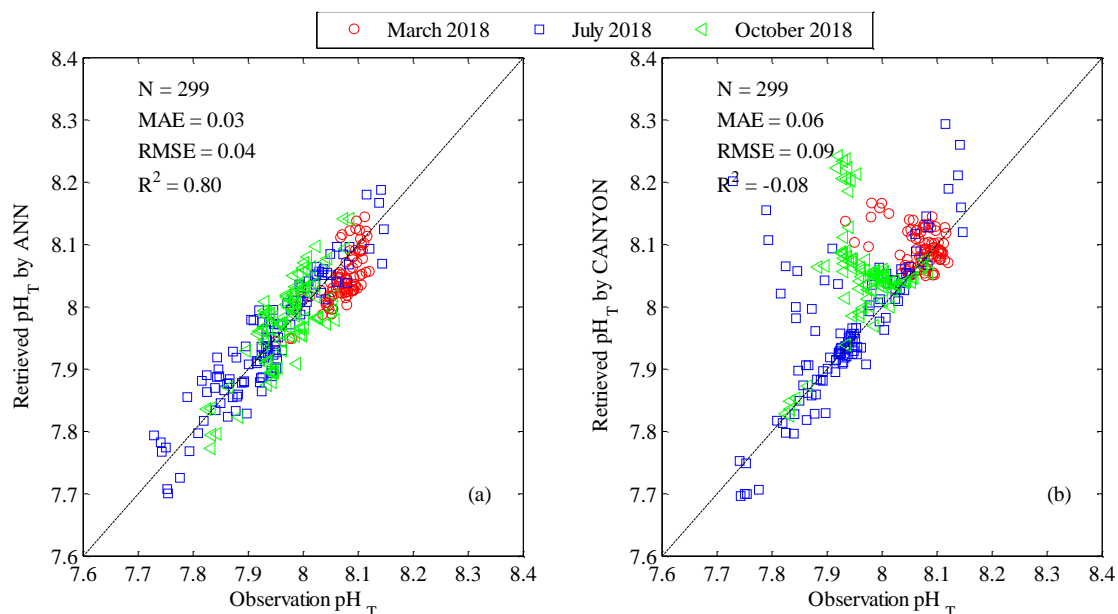| Model	Kernel Function	Input variables	RMSE	R^2	MAE
MLR	-	T, S, DO, N, P, Si	0.078	0.56	0.062
MNR	-	T, S, DO, N, P, Si	0.060	0.74	0.047
Decision Tree	Simple Tree	T, S, DO, N, P, Si	0.064	0.71	0.047
	Medium Tree	T, S, DO, N, P, Si	0.060	0.74	0.044
	Complex Tree	T, S, DO, N, P, Si	0.061	0.73	0.043
Random Forest	Boosted Trees	T, S, DO, N, P, Si	0.340	-7.51	0.339
	Bagged Trees	T, S, DO, N, P, Si	0.056	0.77	0.04
SVM	Linear	T, S, DO, N, P, Si	0.079	0.55	0.061
	Quadratic	T, S, DO, N, P, Si	0.061	0.73	0.046
	Cubic	T, S, DO, N, P, Si	0.060	0.74	0.043
	Fine Gaussian	T, S, DO, N, P, Si	0.064	0.70	0.042
	Medium Gaussian	T, S, DO, N, P, Si	0.054	0.79	0.041
	Coarse Gaussian	T, S, DO, N, P, Si	0.069	0.65	0.054

Referee comment #1 (by Richard Mills)

2. **Second, the authors do a good job of citing other papers in which authors have used similar ANN approaches for similar biogeochemical prediction tasks in marine waters, and compare the RMSE of their model with published values from other models. I think that the paper would be greatly improved if the authors could do a direct comparison. For instance, the authors cite the CANYON neural network model of Sauzede et al., 2017, which has**

been developed for the global ocean, but note that "coastal seas tend to show greater temporal and spatial variability than open oceans", which I believe is an argument for why they developed the model presented in their paper. I can easily imagine that the model presented here will outperform the CANYON model for prediction on the East China Sea shelf, but I think it would be interesting for the authors to demonstrate this: The CANYON model appears to be freely available online, and it would be interesting to see how much better a model trained speficially for the East China Sea shelf will outperform one developed for the global ocean.

We would like to thank the referee very much for his suggestion. We applied the CANYON model developed by Sauzède et al. (2017) to exploratory dataset, result showed that the ANN model presented here outperformed the CANYON model developed for the global ocean for predicting pH_T on the ECS shelf. We will add the corresponding information in the revised manuscript—as follows II. 187-194 of the revised manuscript:

Sauzède et al. (2017) developed a neural network method to estimate pH with RMSE of 0.02 in the global ocean. As a further comparison we applied the CANYON model developed by Sauzède et al. (2017) to our coastal exploratory dataset (Fig. 8b), and obtained an RMSE of 0.09 and MAE of 0.06. It is not surprising that the ANN model (developed here for the ECS shelf) outperforms the CANYON model (developed for the global ocean) for predicting pH_T on the ECS shelf. The carbon chemistry parameters in this region are not only under the direct impact of Taiwan Warm Current and remote control of the Kuroshio water intrusion into the shelf, but are also significantly controlled by seasonal variations of the Changjiang discharge (e.g., Isobe and Matsuno, 2008; Chen et al., 2008; Chou et al., 2009). Taking into account the highly complex hydrographic, biological and chemical conditions, the accuracy of pH_T presented is promising.

Figure 8 (revised): Comparison of retrieved pH_T with corresponding observations for exploratory dataset. (a)-pH_T retrieved by the ANN model vs observations; (b)-pH_T retrieved by CANYON (Sauzède et al., 2017) vs observations. The red circles represent March 2018, the blue squares represent July 2018, the green triangles represent October 2018. The 1:1 line is shown in the plot as visual reference. Three statistics approaches used are the mean absolute error (MAE), the coefficient of determination (R^2), and the root mean squared error (RMSE). N represents the number of data points.

Referee comment #1 (by Richard Mills)

3. **Finally, the authors perform an intersting study in which they use prognostic variables from the Changjian Biology Finite-Volume Coastal Ocean Model (FVCOM) as input to their ANN model in order to recover the pH_T. I am not a marine biogeochemistry modeler, so perhaps I am missing something obvious, but I am guessing that mechanistic models like FVCOM can provide prognostic pH_T. Is this available from the FVCOM runs that were used, or could it be obtained using FVCOM, or ROMS, or another, similar model? If so, how would the prognostic pH_T from FVCOM (or similar) compare to the pH_T from the authors' own ANN model? And what is the motivation for using the ANN? Is it because it can potentially provide a more accurate pH_T, or because it can provide pH_T for situations**

in which it is not desirable to run a forward simulation or reanalysis to get the pH$_T$, or some other reason? This may be obvious to an marine biogeochemist, but I and many of the readers of GMD don't have this expertise. The motivation needs to be explained for the general GMD audience.

We fully agree with the referee. We will compare the prognostic pH$_T$ from FVCOM with retrieved pH$_T$ from our ANN model and add the following sentence to the paragraph in ll. 222-225 of the revised manuscript:

Comparisons of monthly average pH$_T$ from the Changjiang Biology FVCOM model with pH$_T$ retrieved by the ANN model suggested that the ANN model can potentially provide a more accurate pH$_T$ (Fig. S3). The possible reason was that the carbonate system from the Changjiang Biology FVCOM was not optimized due to challenges obtaining sufficient boundary information.
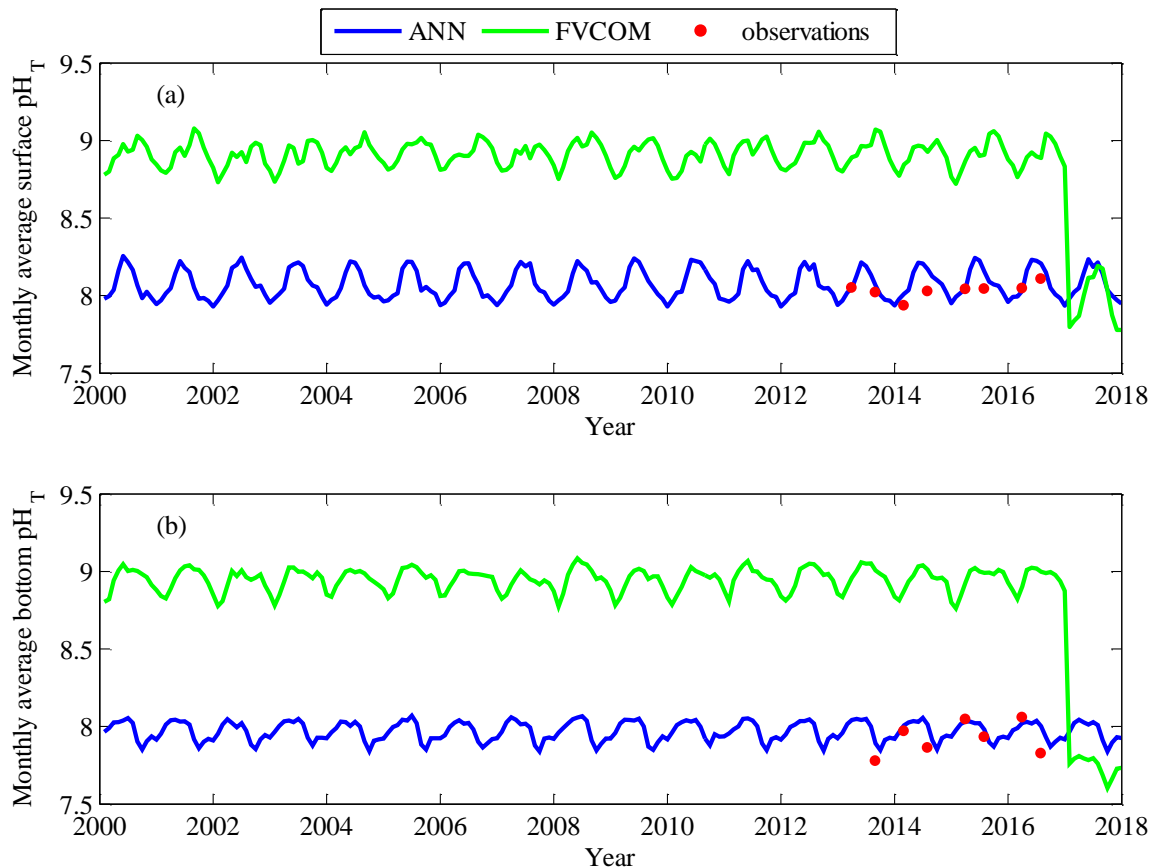
Figure S3 (revised): Comparison of monthly average pH$_T$ on the East China Sea shelf. Blue solid line represents retrieved pH$_T$ by the ANN model using Changjiang Biology FVCOM output; green solid line represents simulated pH$_T$ by the Changjiang Biology FVCOM; red points show monthly average pH$_T$ observations from 2013-2016. (a)-surface; (b)-bottom.

Referee comment #1 (by Richard Mills)

4. Detailed comments:
Lines 34-35: The authors state, while comparing ANNs to multiple linear regression, that ANNs have the advantage of not requiring 'an a priori model but rather "learn" the model from existing data'. I think it would be more precise to say that they are nonparametric models and do not require assuming any underlying statistical distribution.

We agree with the referee. See ll. 37-38 of the revised manuscript:

~~they do not require an a priori model but rather "learn" the model from existing data~~ may be a greater flexibility and versatility in modelling complex nonlinear relationships.

Lines 75 and 78: The authors say that samples were "poisoned" by addition of HgCl2. I think it may be more idiomatic to say "sterilized".

We agree with the referee. See Lines 79 and 86 of the revised manuscript:

~~poisoned~~ *sterilized*

Line 81: "The final number of data used by the ANN model was 1854". I would say the final number of "observations" or "records", to be precise.

We agree with the referee. See Lines 93-94 of the revised manuscript:

The final number of observations in the confirmatory dataset was 1854 (see Table 1 for more detailed information on the field survey).

Line 94: The authors talk about a model being "over-matched". I believe that "overfitted" is the term they mean.

We agree with the referee. See Line 107 of the revised manuscript:

the testing set was used to monitor whether the model was over-fitted~~matched~~

Referee comment #2

1. **Choice of inputs: A total of nine variables were used as inputs to the ANN, six of which were direct measurements (T, S, DO, N, P, and Si). Lines 105-107 notes "We found geographical information to be a powerful addition in improving the skill of the method (see Table 2), allowing the network to learn spatio-temporal patterns that could not be explained by other input variables (Sasse et al., 2013)." Adding geographical information does appear to improve the performance for the initial model training (Table 2). However, the cruise tracks are only sampling certain latitudes lead to a biased sampling. BUT can this lead to a geographically biased training? This bias may not be apparent even in the validation using data from three cruises, since they too are in same bands as before. But when applied to data from FVCOM, there are biases reported in Figure 5. Is it possible that the model is not generalized enough for other regions?**

We agree with the referee. The model is not generalized enough for other regions. This model was trained using cruises datasets on the ECS shelf, can be applied to retrieve pH_T on the ECS shelf. But *this approach can be applied to other regions to predict pH by suitably adapting the input variables and network structure using local datasets*.

2. **Lines 100-109 explains the choice of variables for all but one variable "month". I assume the variable was added to capture the seasonality. However, a significant bias was still reported in August 2013, and July 2016. These biases are being attributed to sudden increase in the river discharge, but did that not affect July 2014, 2015, 2017? What is the role "month" is playing in the ANN model? Once trained is the expectation for the model to be able to interpolate between the month when the samples were not taken?**

In order to retrieve monthly pH_T, the monthly T, S, DO, N, P and Si from the Changjiang Biology Finite-Volume Coastal Ocean Model (FVCOM) (http://47.101.49.44/wms/demo) were fed into the ANN model as input variables. Here a significant bias was reported in August 2013 and July 2016. See Lines 229-232 of the revised manuscript:

Overall, the retrieved pH_T agrees well (within the ANN model accuracy: ANN±RMSE) with the observed values at the surface, except for three samples in summer (Fig. 10). There are relatively large deviations (greater than the RMSE of 0.04) in August 2013 at station A1-5 and A6-9, and in July 2016 at station A8-5.

The reduced performance in summer can be attributed in large part a reduced performance of the Changjiang Biology FVCOM in predicting summertime input variables S, DO, and nutrients (Fig. S2). See Lines 219-222 of the revised manuscript:

Comparisons of monthly-average FVCOM model variables with surface and bottom observations on the ECS shelf showed that simulated T was close to observed values (Fig. S2a), simulated S was also close to observed values except at the bottom in August 2013 and at the surface in July 2016 (Fig. S2b), simulated DO was higher than observed at the bottom (Fig. S2c), and simulated nutrients were higher than observed at the surface (Fig. S2d-S2f).

Figure S2: Comparison of monthly-average environmental variables from the Changjiang Biology FVCOM with the corresponding observations at the surface and bottom on the East China Sea shelf. Blue and green solid lines represent surface and bottom simulated data from the Changjiang Biology FVCOM, respectively; red and black points show surface and bottom observation data from 2013 to 2016, respectively. (a)-temperature; (b)-salinity; (c)-dissolved oxygen; (d)-nitrate; (e)-phosphate; (f)-silicate.

The variable "month" was added to capture the seasonality. The reliability of the ANN model was evaluated using independent observations from 3 cruises (March, July, and October) in 2018, and showed a root mean square error accuracy of 0.04. The cruise dataset during October was not used in ANN model development.

3. **ANN application to FVCOM: Inputs to the ANN models training, based on cruise observations, were instantaneous measurements. What was the spatial resolution, time step and temporal output frequency from FVCOM model to provide comparable outputs. If monthly averages were used, please comment on applicability and validity of applying model trained based on instantaneous measurements to monthly averages?**

We agree with the referee. We will compare monthly input variables from FVCOM model with instantaneous observations and add the following sentence to the paragraph in II. 217-222 of the revised manuscript:

The resolution of the Changjiang Biology FVCOM output is 1-10 km in the horizontal, 10 depth levels in the vertical, and day in the temporal (refered Ge et al., (2013) for detail information). Comparisons of monthly-average FVCOM model variables with surface and bottom observations on the ECS shelf showed that simulated T was close to observed values (Fig. S2a), simulated S was also close to observed values except at the bottom in August 2013 and at the surface in July 2016 (Fig. S2b), simulated DO was higher than observed at the bottom (Fig. S2c), and simulated nutrients were higher than observed at the surface (Fig. S2d-S2f).

4. Application to FVCOM, scales the model to extended space and time, which I think is a key strength and contribution of this work. Spatial bias has been discussed and reported in the manuscript, but it would be important to discuss the model performance in time. Cruise observations were only from select few months, but is the model able to fill in between the seasons reasonably? And if yes, why? If no, why not? Final output of ANN applied to FVCOM data would be a time series of full spatial data i.e. pixel-wise pH estimate for ECS. That product is a key contribution that should be included in the manuscript, and spatial and temporal patterns of the outputs should be discussed.

We agree with the referee, we trained the ANN model only using few months observations. In order to assess the ability of the ANN model to fill in other months not used in ANN model development, we applied it to October (not used) in 2018 with a root mean square error accuracy of 0.04; we also added spatial and temporal patterns of ANN-derived pH_T. See Section 3.4.2 of the revised manuscript, as follows:

3.4.2 Spatial and temporal patterns of ANN-derived pH_T

The temporal and spatial variations of monthly surface pH_T from 2000-2016 based on Changjiang Biology FVCOM output are shown in Figure 13. During the dry season (November to March of the next year), pH_T values vary from ~7.62 to ~8.24. Relatively higher pH_T values are found in the southeastern of the study area (Chou et al., 2011), whereas lower pH_T values are found in the northeastern of the study area. During the wet season (April to October), pH_T values vary from ~7.77 to ~8.35, water of higher pH_T corresponded well to the seasonal dispersion of the Changjiang Dilute Water (Chou et al., 2009, 2013). Water of higher pH_T is found in the center of the study area during April, spreads to the southwestern part of the study area (along the coast of China) during May and June, shifts to the northeastern part of the study area during August. In September and October, water of higher pH_T is found in the southeastern part of the study area, strongly influenced by the Taiwan Warm Current (Qu et al., 2015).

A clear seasonality is that surface pH_T gradually increases during spring (March to May), after which it gradually decreases during summer and fall (June to November) (Fig. 14). The surface pH_T displays its maximum in May and minimum in December, and the pH_T varies seasonally by up to ~0.3 unit. Larger changes in pH were also discovered in the Washington Shelf, the pH varied ~1.0 unit over the seasons and ~1.5 unit spanning 8 years (Wootton et al., 2008). Accordingly, seasonal dynamics of surface pH_T can be mainly attributed to temperature changes and strong biological activities (production and respiration processes) over the season. From March to June, a rapid increase in surface pH_T indicates that production increases faster than respiration, which can be reflected in the drop in surface phosphate (Fig. S5d) and apparent oxygen utilization (AOU) (Fig. S5c). It may be driven by the Changjiang discharge (Fig. S4), which carries large amount of nutrients, result in stronger primary production in warm seasons under the combined action of nutrients and suitable temperature (Gong et al., 2011). From July to October, although surface temperature remains at a high level (Fig. S5a), the rise in surface AOU
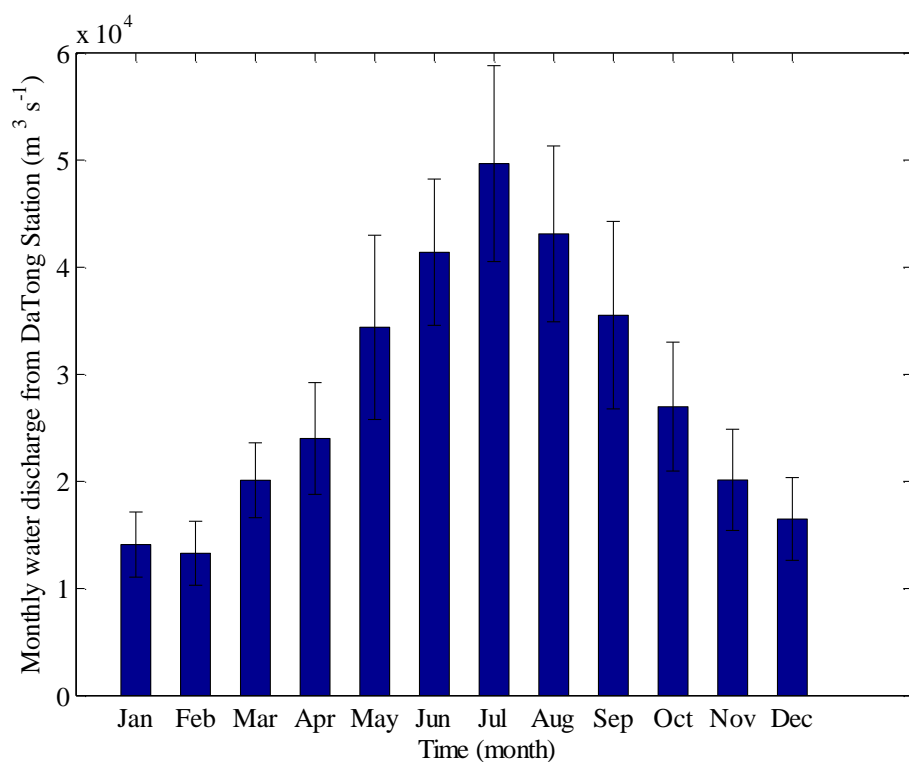
Figure 13: Spatial distribution of monthly average surface pH$_T$ retrieved by the ANN model using Changjiang Biology FVCOM output. (a)-January; (b)-February; (c)-March; (d)-April; (e)-May; (f)-June; (g)-July; (h)-August; (i)-September; (j)-October; (k)-November; (l)-December.

Figure 14: Seasonal cycles of surface pH$_T$ on the East China Sea shelf from 2000-2016. The green circles represent monthly regional average, the blue dashed represents mean value of each month.

Figure S4: Monthly average water discharge and its standard deviation (DaTong Station, data derived from the Hydrological Information Center of China, http://www.hydroinfo.gov.cn/).
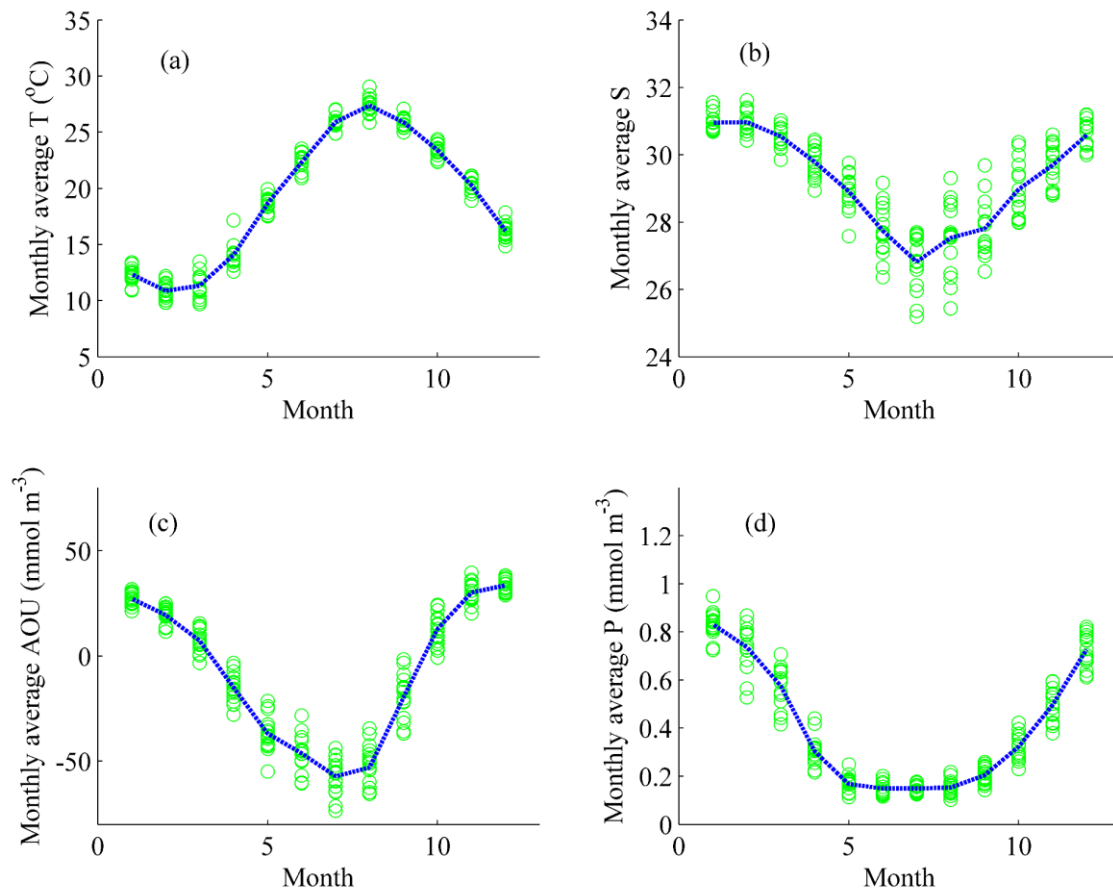
Figure S5: Seasonal cycles of surface T (a), S (b), AOU (c), and P (d) from Changjiang Biology FVCOM output on the East China Sea shelf from 2000-2016. The green circles represent monthly regional average, the blue dashed represents mean value of each month.

5. **Variable importance in the ANN model: The methodology here is not clear to me. What does adding 5% to environmental variable separately means? Is this a perturbation to the data to test its sensitivity? In either case, I am not at all convinced that this can be quantified as variable importance. There also is mention of "variable with greatest weight was DO, followed by S and T". What weights are looking at here, is this from the final trained model? From first layer, from second layer, or both? This section need additional detail and discussion to convey and convince the interpretation of variable importance.**

We agree with the referee. We only want to test the ANN model sensitivity to environmental input variables, and we re-wrote section 3.3. See Section 3.3 of the revised manuscript, as follows:

3.3 ANN model sensitivity to environmental input variables

To assess the ANN model sensitivity to different environmental input variables, we added 5% perturbation for each environmental variable separately. Statistically, with 5% T errors added, the ANN model showed slight overestimation in pH_T, with mean bias (MB) of 0.0059, RMSE of 0.0079, and R^2 of 0.9949 (Fig. 9a); with 5% DO errors added, the ANN model also showed slight pH_T overestimation, with MB of 0.0050, RMSE of 0.0090, and R^2 of 0.9934 (Fig. 9c); with 5% S errors added, the ANN model showed overestimation in pHT, with MB of -0.0111, RMSE of 0.0162, and R^2 of 0.9789 (Fig. 9b). These results suggested that the ANN model responded to T and DO errors in a positive way, S errors in a negative way. The positive response to increasing DO reflects positive correlation between pH_T and DO (Cai et al., 2011), which can be attributed to the processes of photosynthesis (generating DO and removing CO_2, hence increasing pH) and aerobic respiration (consuming

12

DO and generating CO₂, hence lowering pH); the negative response to increasing S reflects the influence of the (lower salinity) Changjiang discharge, carrying large amounts of nutrients that fuel increased primary production (uptake of nutrients and CO₂, hence raising the pH) in surface waters during warm seasons (Gong et al., 2011). It was found that the ANN model was insensitive to nutrients errors (Fig. 9d-9f) and most sensitive to S errors (Fig. 9b), followed by DO and T errors.
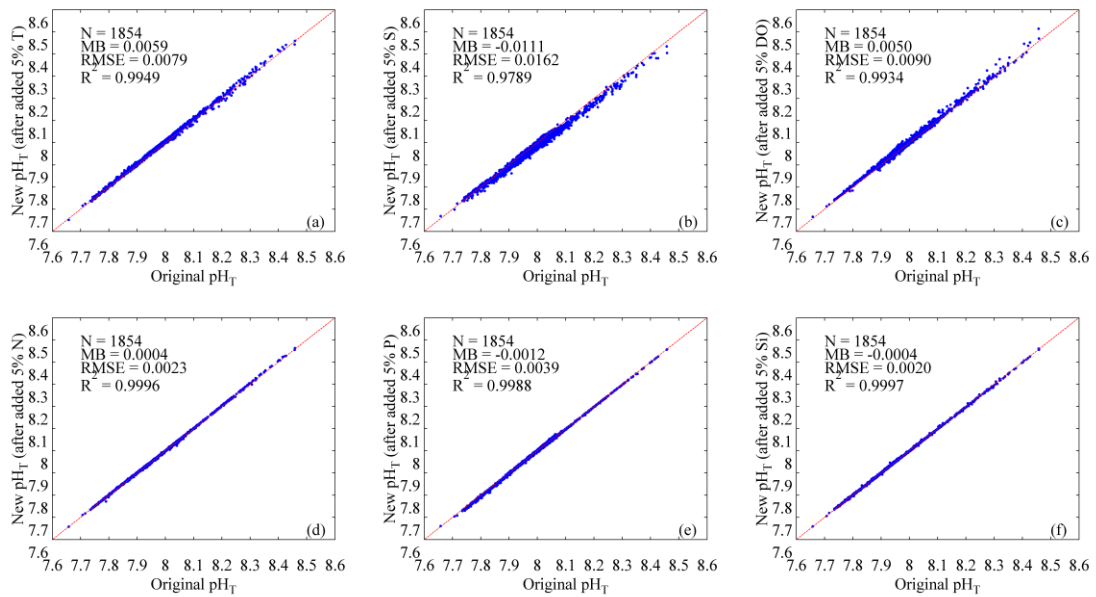

Figure 9: Sensitivity of the ANN model for environmental input variables. (a)-temperature (T); (b) salinity (S); (c)-dissolved oxygen (DO); (d)-nitrate (N); (e)-phosphate (P); (f)-silicate (Si). Three statistics approaches used are the mean bias (MB), the root mean squared error (RMSE), and the coefficient of determination (R^2). N represents the number of data points.

6. In Figures 9 and 11, I am unable to understand what "ANN Model - RMSE" and "ANN Model + RMSE", and thus the related discussions.

"ANN model + RMSE" and "ANN model - RMSE" represent upper and lower bounds of the ANN model accuracy. See Figure 10 (Line 533) and 12 (Line 541) of the revised manuscript.

Referee comment #1 and #2

There are problems with low resolution for all of the figures. Figure 1 is not really even readable. Figures need to be re-generated with much higher resolution, or using vector, rather than raster, formats.   From comment#1
Please improve the quality of figures 1 and 6, as they are difficult to read and follow.   From comment#2

We agree with the referees. Figures will be re-generated with higher resolution or using vector. As follows:
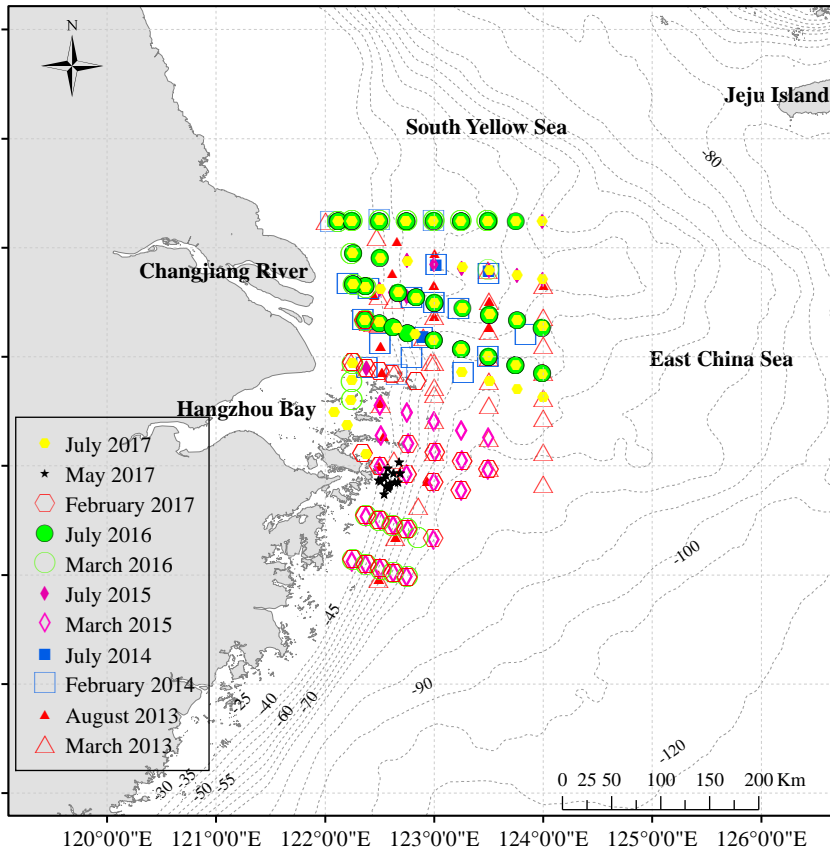
Figure 1 (revised): Sampling stations during 11 cruises (the confirmatory dataset) from 2013 to 2017 on the East China Sea shelf.
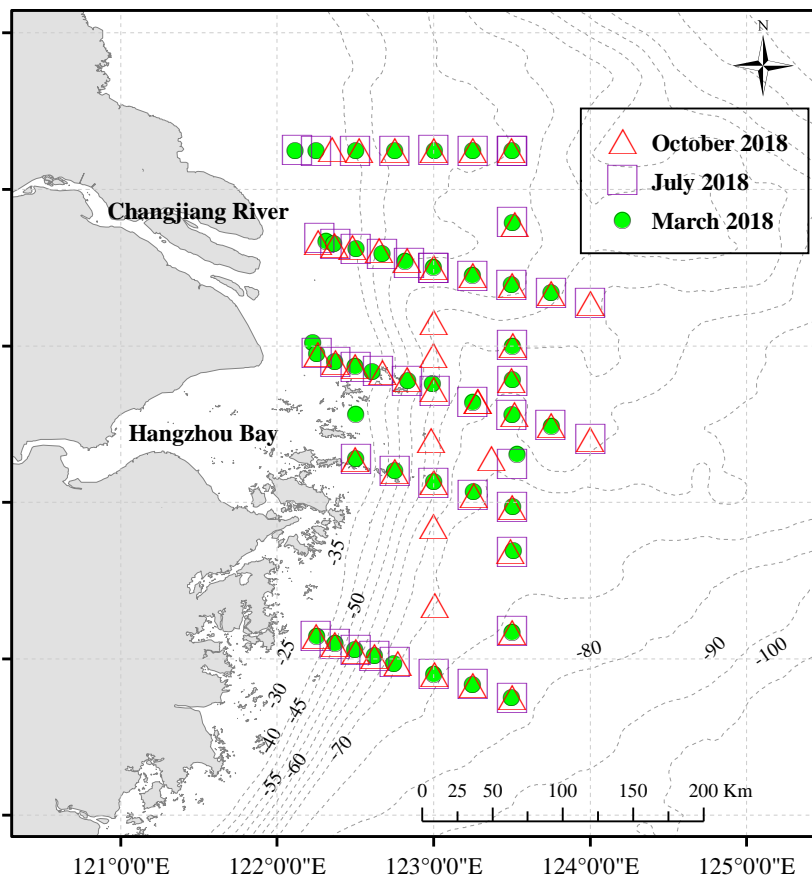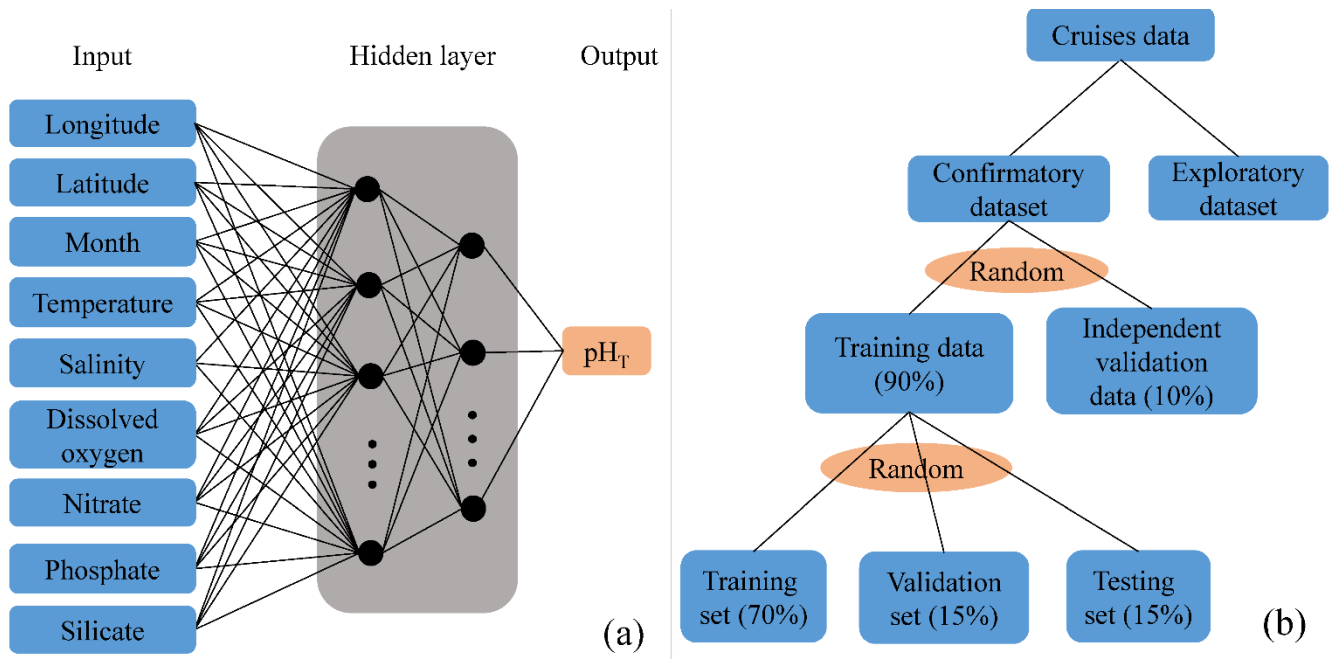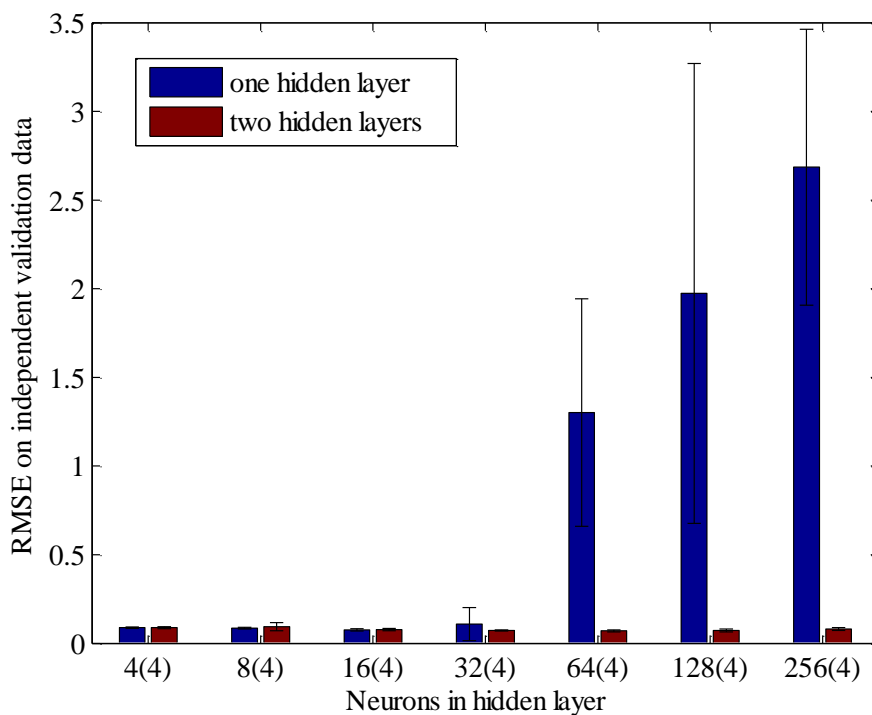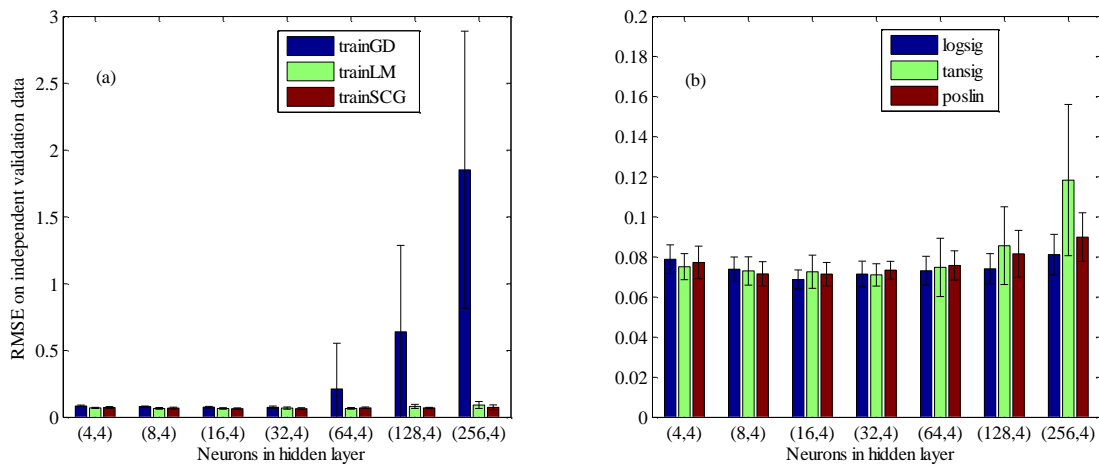

Figure 2 (revised): Sampling stations for three cruises (the exploratory dataset) used to extend the utility of the ANN model. The green circles represent March 2018, the purple squares represent July 2018, the red triangles represent October 2018.
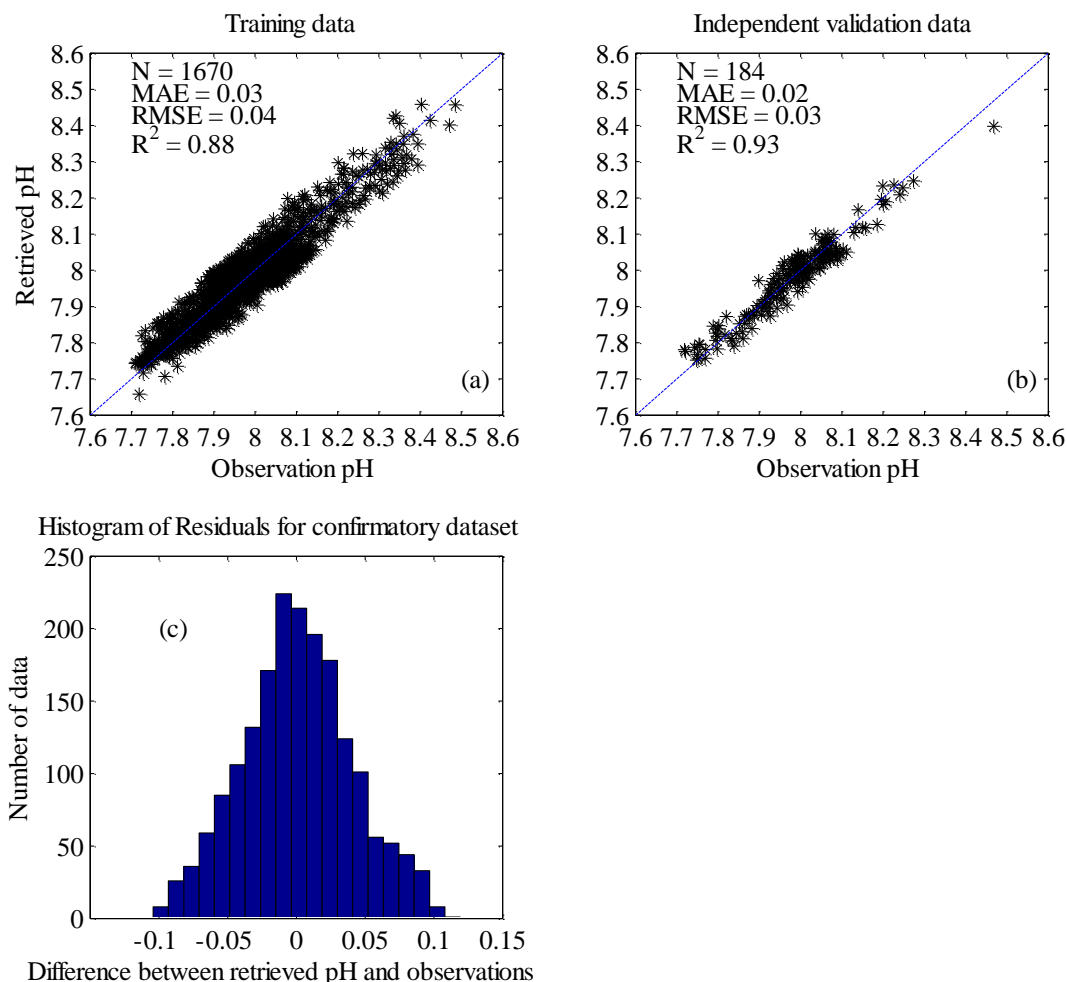
Figure 3 (revised): Schematic representation of the neural network algorithm to retrieve pH$_T$. (a)-the architecture of the ANN model. Input variables are observed temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate together with the geolocation (longitude and latitude) and time (month) of sampling; (b)-data distribution diagram for training and prediction.

Figure 4 (revised): Comparison of the performance of one hidden layer vs. two hidden layers in predicting independent validation data. The number of neurons in the first hidden layer was the same in the one hidden layer vs. two hidden layers model, numbers in parentheses show the number of neurons in the second hidden layer (for the two hidden layers model). Bars show the mean and standard deviation of the Root-Mean-Square-Error over a ten-fold cross-validation, for different numbers of neurons in the first hidden layer.

Figure 5 (revised): Comparison of the performance of different training functions and transfer functions on independent validation data. (a)-three training functions: Gradient descent backpropagation (trainGD), Levenberg-Marquardt backpropagation (trainLM), and Scaled conjugate gradient backpropagation (trainSCG); (b) three transfer functions: Log-sigmoid transfer function (logsig), Hyperbolic tangent sigmoid transfer function (tansig), and Positive linear transfer function (poslin). Bars show the mean and standard deviation of the Root-Mean-Square-Error over a ten-fold cross-validation, for different numbers of neurons in the first hidden layer.

Figure 6 (revised): Comparison of pH$_T$ retrieved by the ANN model with corresponding observations. (a)-Training data (90% of confirmatory dataset); (b)-Independent validation data (10% of confirmatory dataset); (c)-Histogram of residuals for confirmatory dataset. The 1:1 line is shown in each plot as visual reference. Three statistics are the mean absolute error (MAE), the coefficient of determination (R^2), and the root mean squared error (RMSE). N represents the number of data points.
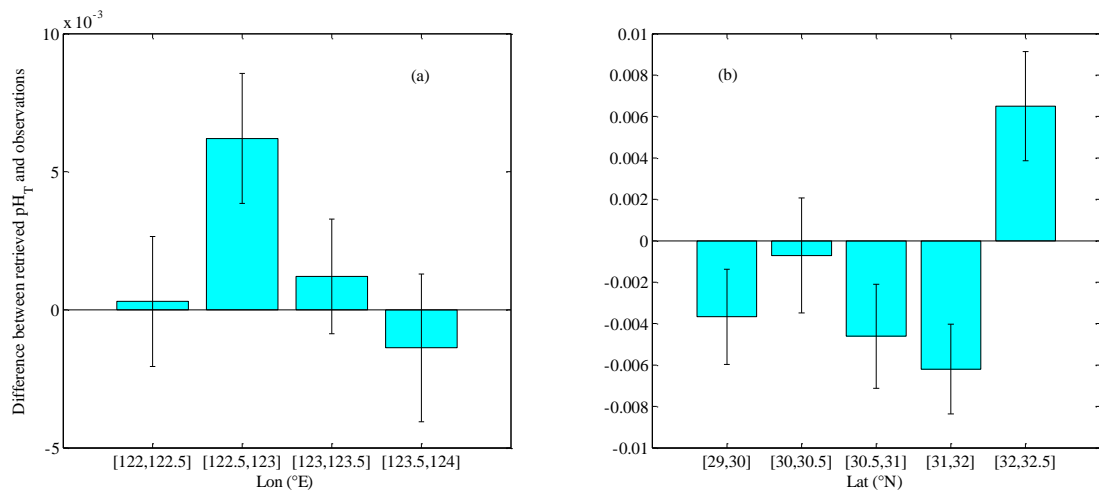
Figure 7 (revised): Box plots of the differences between retrieved pH$_T$ minus the observations. (a)-the differences vs longitude (Mean±SE); (b)-the differences vs latitude (Mean±SE). The height of each box represents the mean value of the differences, the whisker represents the standard error (SE) value of the differences.
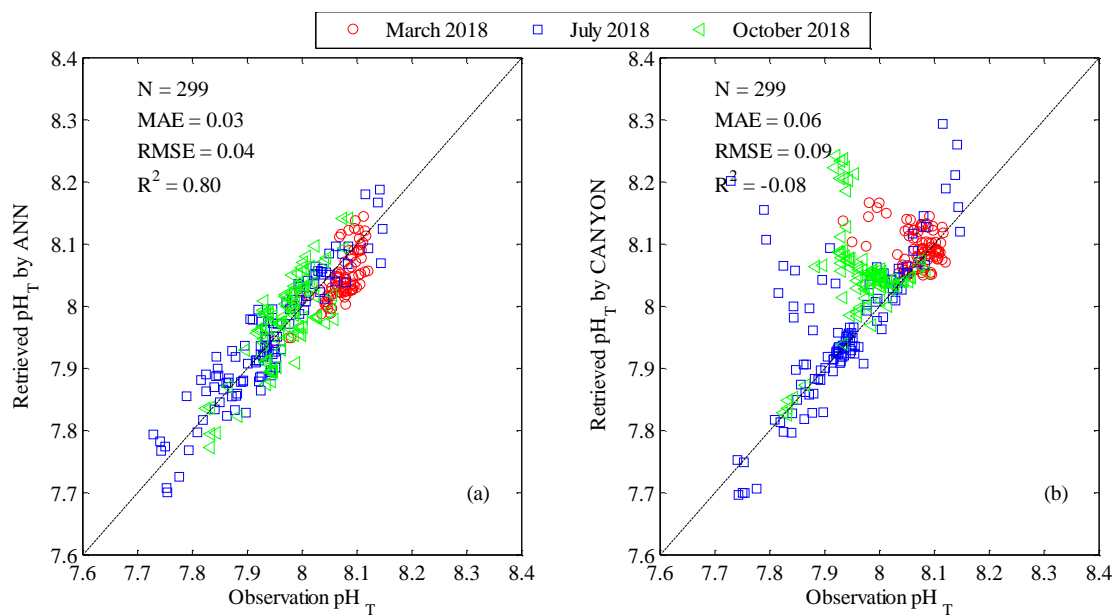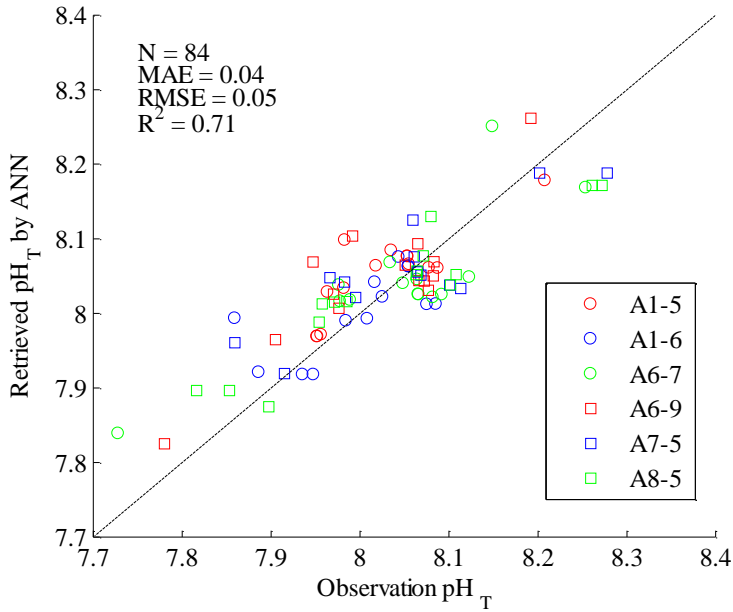

Figure 8 (revised): Comparison of retrieved pH$_T$ with corresponding observations for exploratory dataset. (a)-pH$_T$ retrieved by the ANN model vs observations; (b)-pH$_T$ retrieved by CANYON (Sauzède et al., 2017) vs observations. The red circles represent March 2018, the blue squares represent July 2018, the green triangles represent October 2018. The 1:1 line is shown in the plot as visual reference. Three statistics approaches used are the mean absolute error (MAE), the coefficient of determination (R^2), and the root mean squared error (RMSE). N represents the number of data points.

17
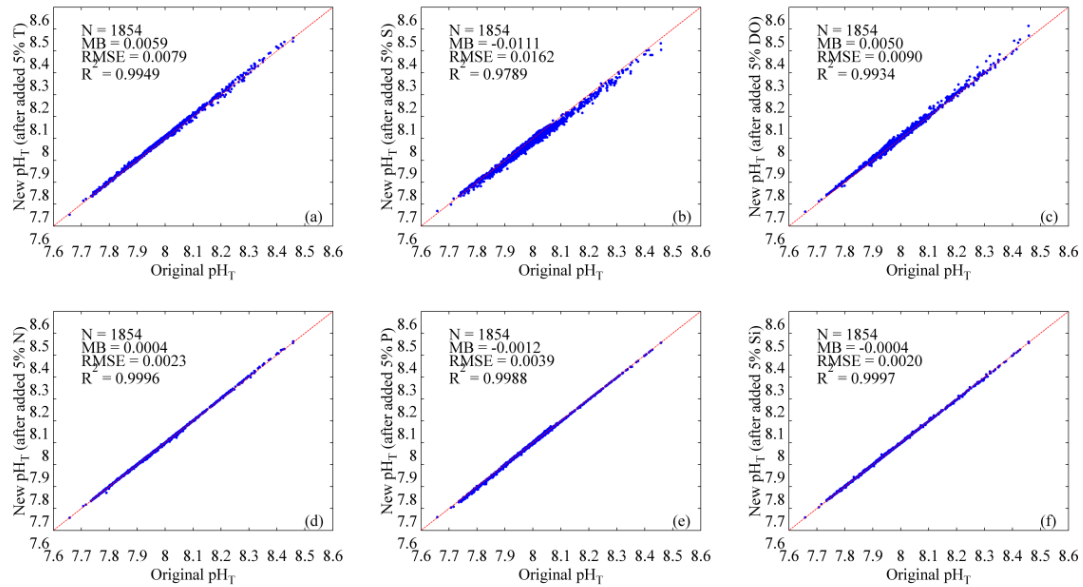
Figure 9 (revised): Sensitivity of the ANN model for environmental input variables. (a)-temperature (T); (b) salinity (S); (c)-dissolved oxygen (DO); (d)-nitrate (N); (e)-phosphate (P); (f)-silicate (Si). Three statistics approaches used are the mean bias (MB), the root mean squared error (RMSE), and the coefficient of determination (R^2). N represents the number of data points.
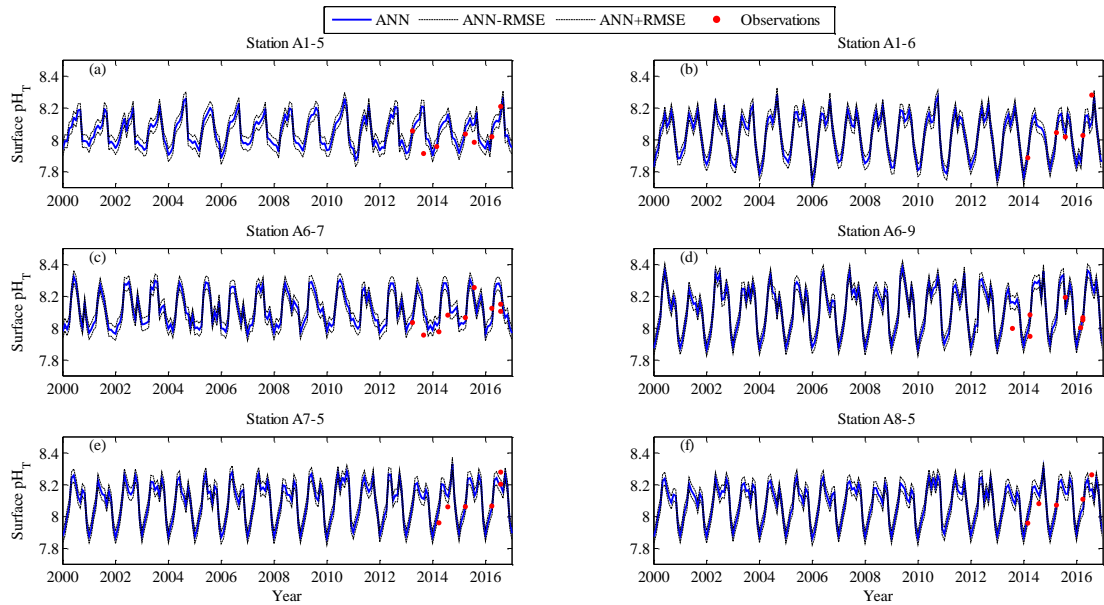

Figure 10 (revised): Comparison of surface pH$_T$ retrieved by the ANN model using Changjiang Biology FVCOM output with corresponding observations at six sites repeated sampling for 3 to 4 years. Red dots represent observations pH$_T$, blue solid line represents retrieved pH$_T$, black dotted lines represent upper and lower bounds of the ANN model accuracy (ANN ± RMSE). (a)-station A1-5; (b)-station A1-6; (c)-station A6-7; (d)-station A6-9; (e)-station A7-5; (f)-station A8-5.

Figure 11 (revised): Comparison of water column pH_T retrieved by the ANN model using Changjiang Biology FVCOM output with corresponding observations at six sites repeated sampling for 3 to 4 years. The 1:1 line is shown in the plot as a visual reference. Skill statistics include the mean absolute error (MAE), the coefficient of determination (R^2), and the root mean squared error (RMSE). N represents the number of data points.
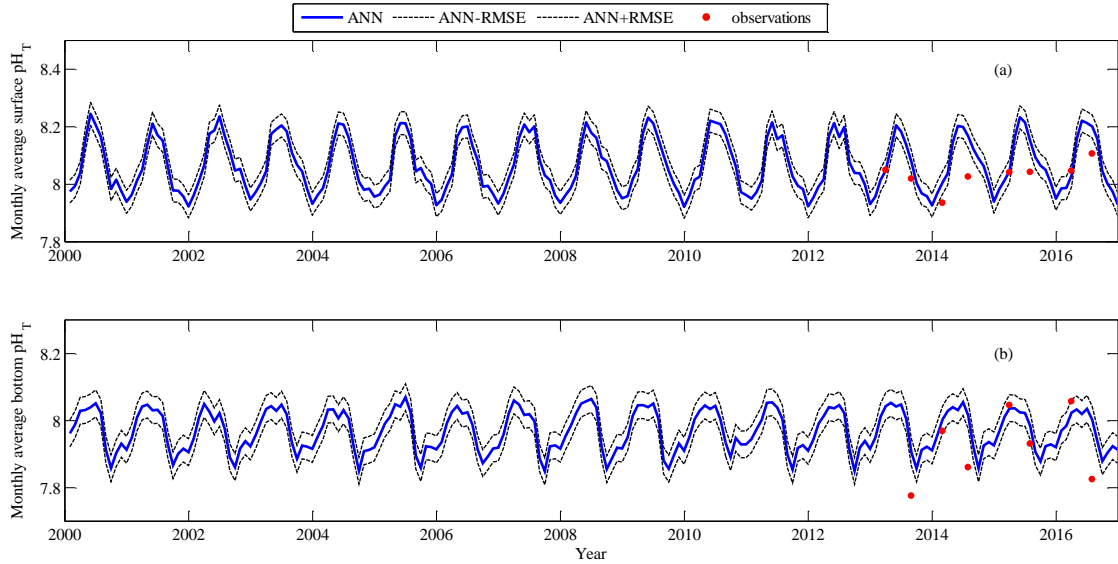
Figure 12 (revised): Comparison of monthly average pH_T on the East China Sea shelf. Blue solid line represents retrieved pH_T by the ANN model using Changjiang Biology FVCOM output; black dotted lines represent upper and lower bounds of the ANN model accuracy (ANN ± RMSE); red points show monthly-average pH_T observations from 2013 to 2016. (a)-surface; (b)-bottom.
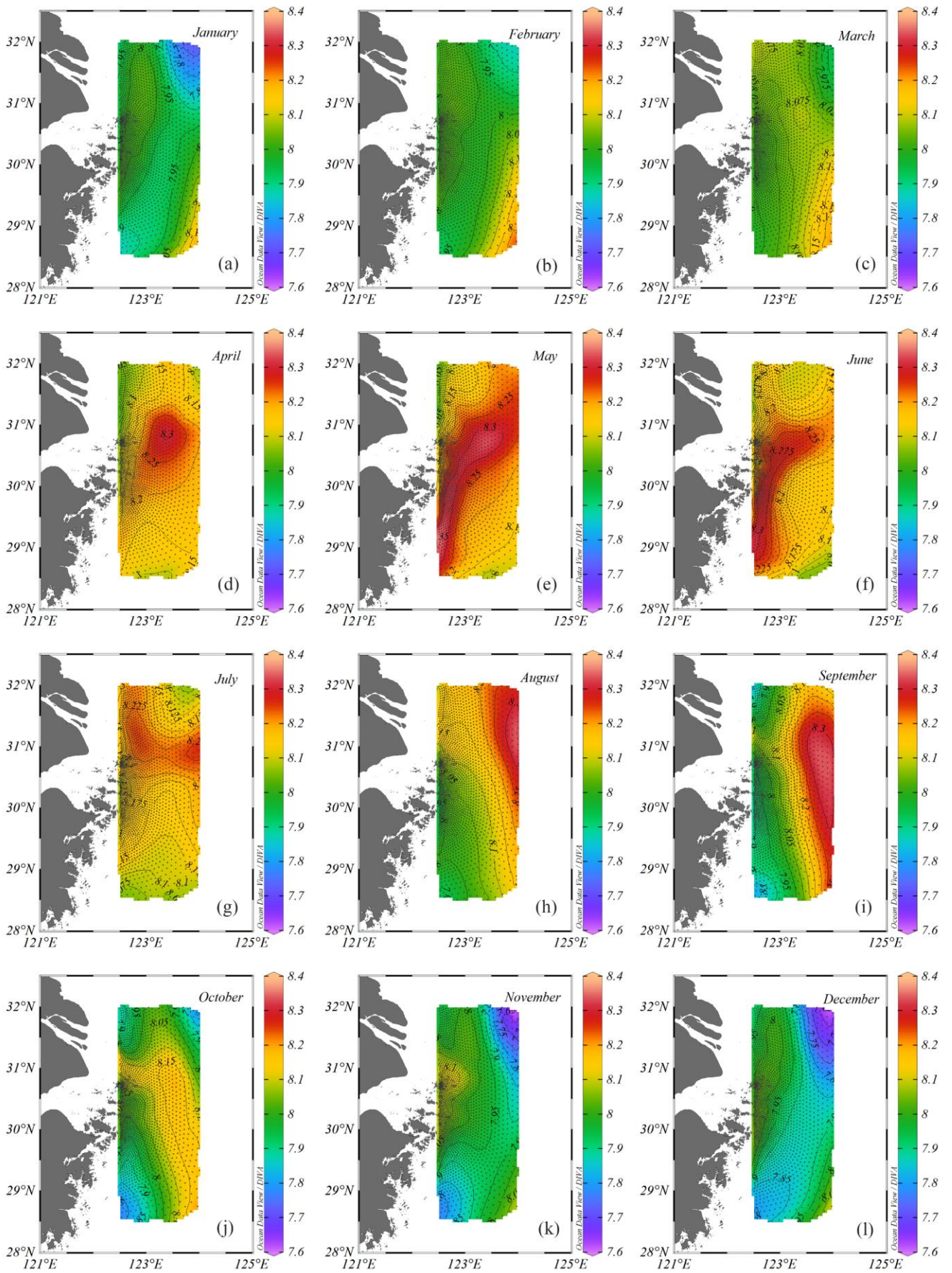
Figure 13 (revised): Spatial distribution of monthly average surface pH$_T$ retrieved by the ANN model using Changjiang Biology FVCOM output. (a)-January; (b)-February; (c)-March; (d)-April; (e)-May; (f)-June; (g)-July; (h)-August; (i)-September; (j)-October; (k)-November; (l)-December.
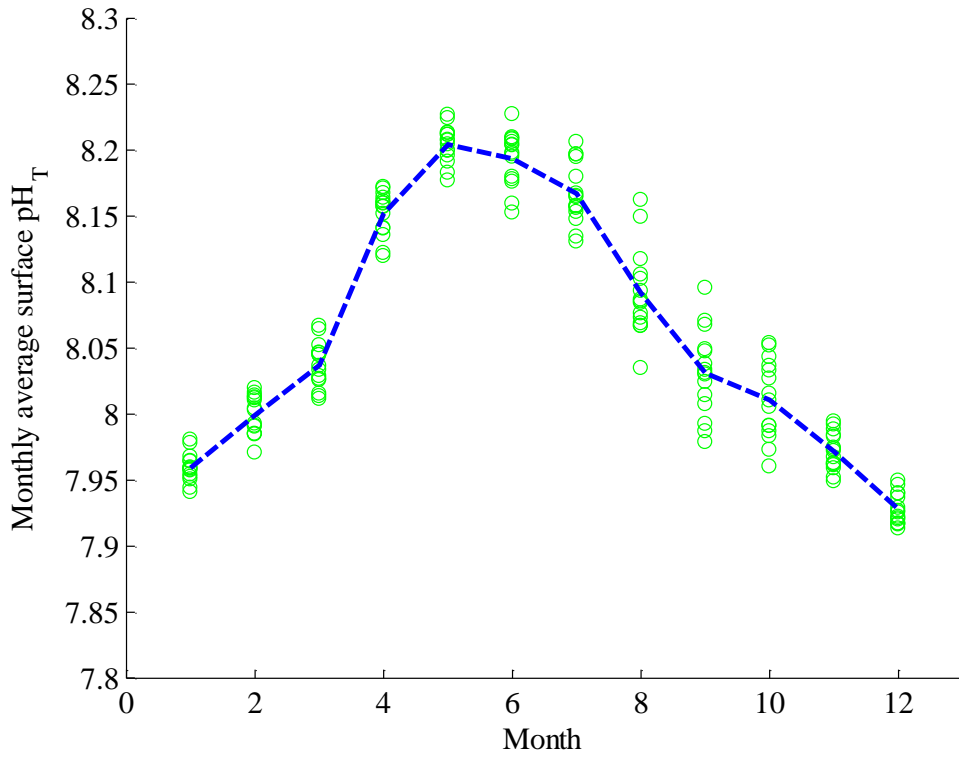
20

Figure 14 (revised): Seasonal cycles of surface pH$_T$ on the East China Sea shelf from 2000-2016. The green circles represent monthly regional average, the blue dashed represents mean value of each month.
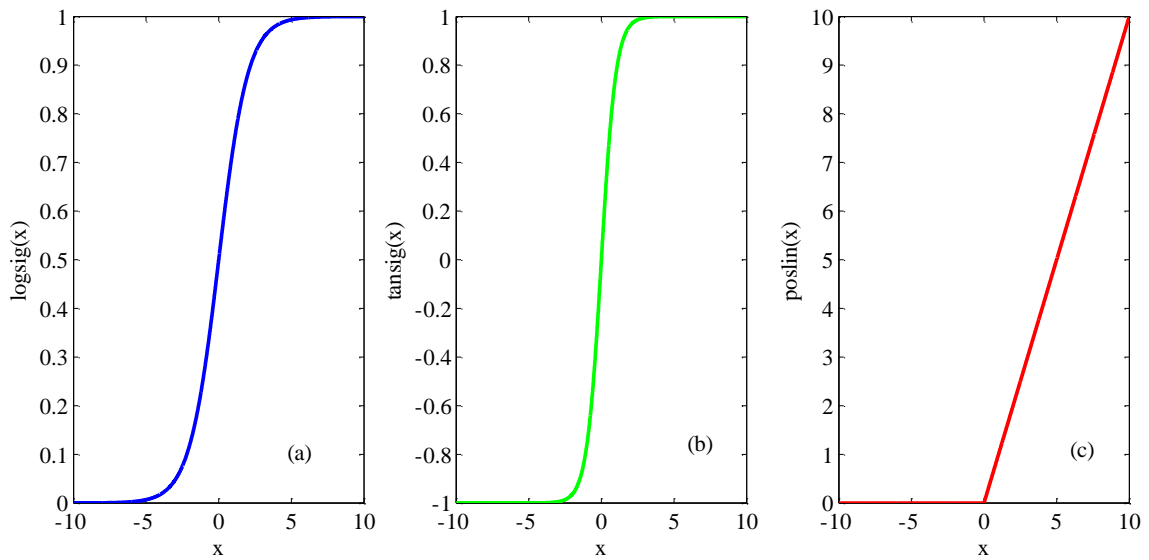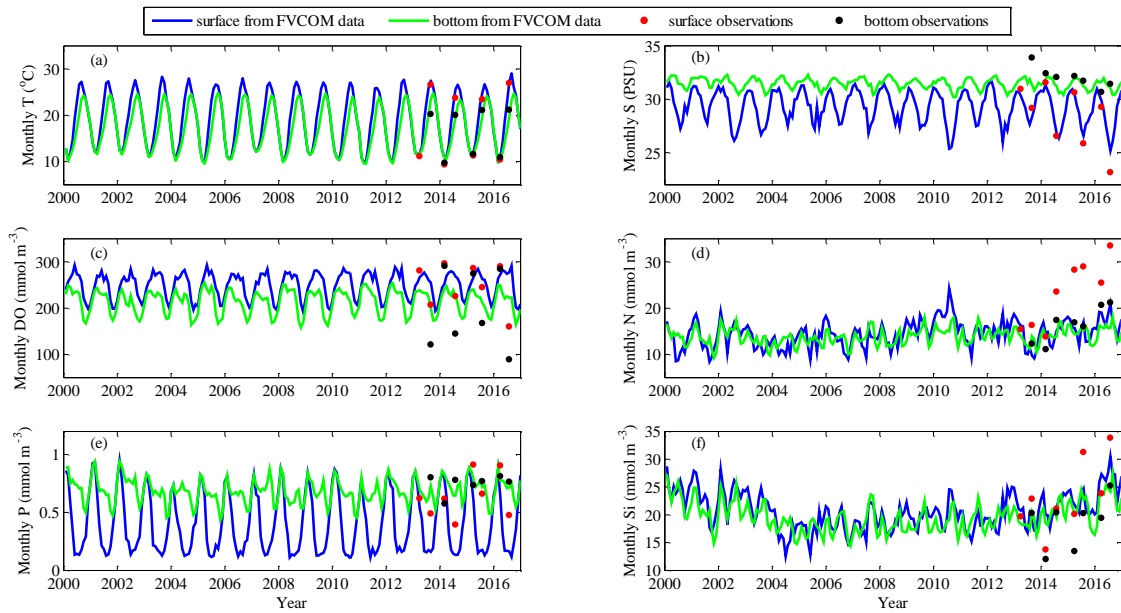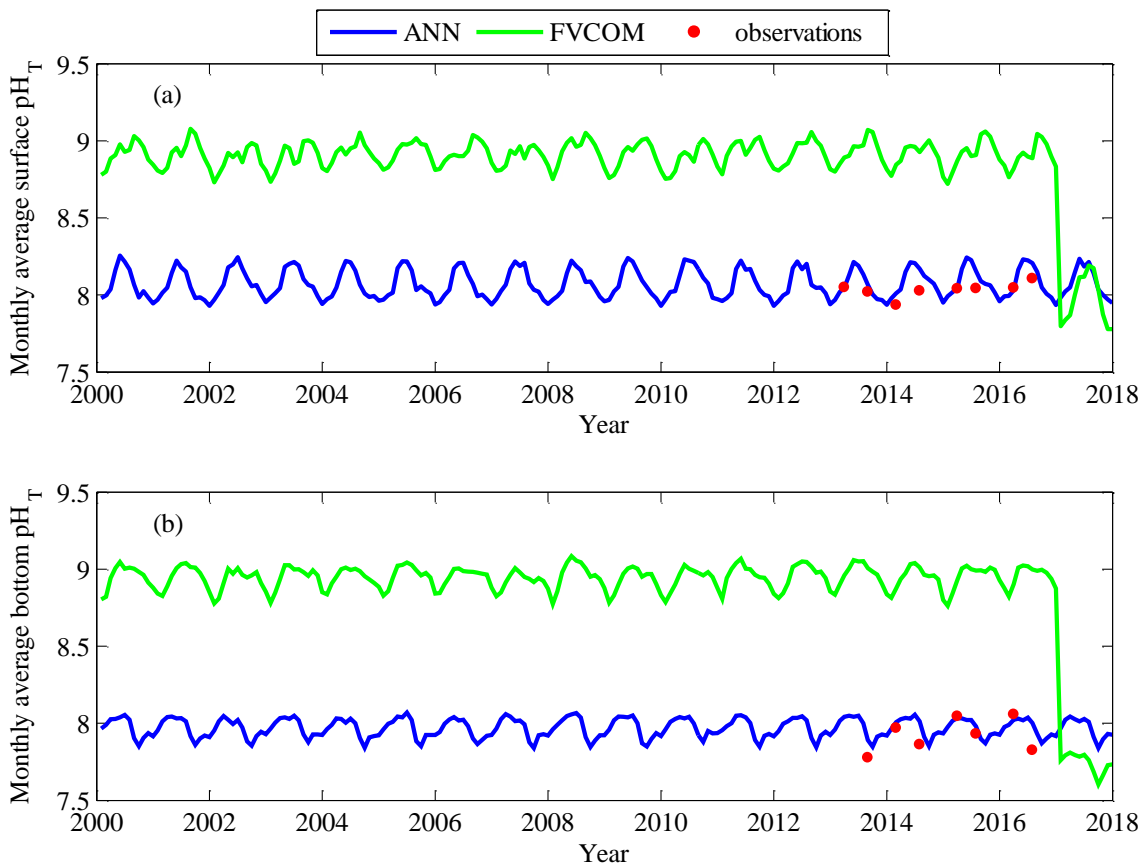

Figure S1 (revised): Comparison of three transfer functions. (a)-Log-sigmoid transfer function (logsig); (b) Hyperbolic tangent sigmoid transfer function (tansig); (c)-Positive linear transfer function (poslin).

Figure S2 (revised): Comparison of monthly-average environmental variables from the Changjiang Biology FVCOM with the corresponding observations at the surface and bottom on the East China Sea shelf. Blue and green solid lines represent surface and bottom simulated data from the Changjiang Biology FVCOM, respectively; red and black points show surface and bottom observation data from 2013 to 2016, respectively. (a)-temperature; (b)-salinity; (c)-dissolved oxygen; (d)-nitrate; (e)-phosphate; (f)-silicate.

Figure S3 (revised): Comparison of monthly average pH_T on the East China Sea shelf. Blue solid line represents retrieved pH_T by the ANN model using Changjiang Biology FVCOM output; green solid line represents simulated pH_T by the Changjiang Biology FVCOM; red points show monthly average pH_T observations from 2013-2016. (a)-surface; (b)-bottom.
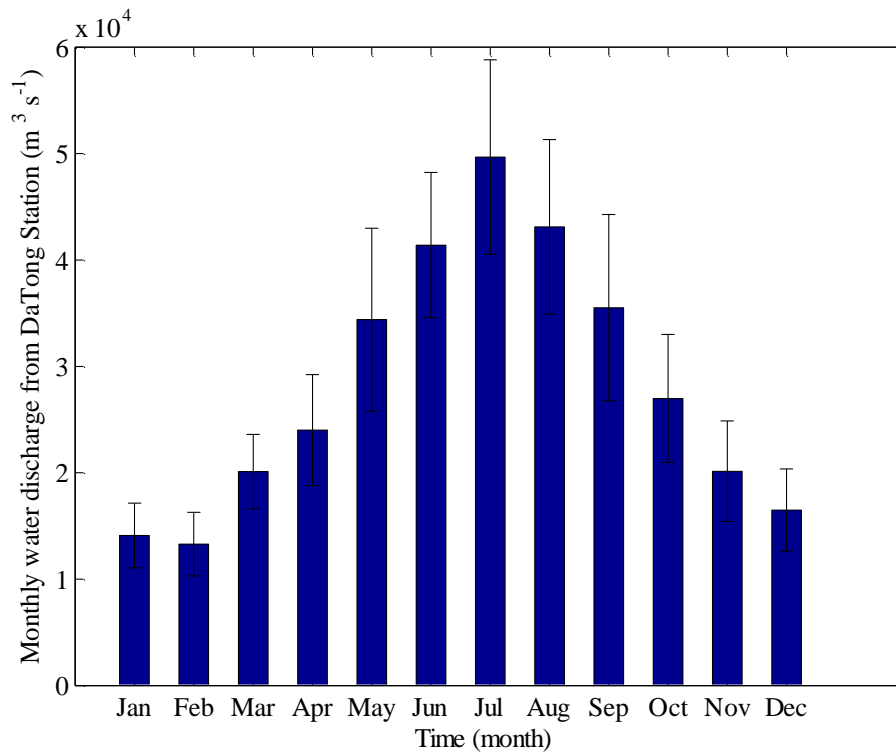
Figure S4 (revised): Monthly average water discharge and its standard deviation (DaTong Station, data derived from the Hydrological Information Center of China, http://www.hydroinfo.gov.cn/).

Figure S5 (revised): Seasonal cycles of surface T (a), S (b), AOU (c), and P (d) from Changjiang Biology FVCOM output on the East China Sea shelf from 2000-2016. The green circles represent monthly regional average, the blue dashed represents mean value of each month.

Retrieving monthly and interannual pH_T on the East China Sea shelf using an artificial neural network: ANN-pH_T-v1
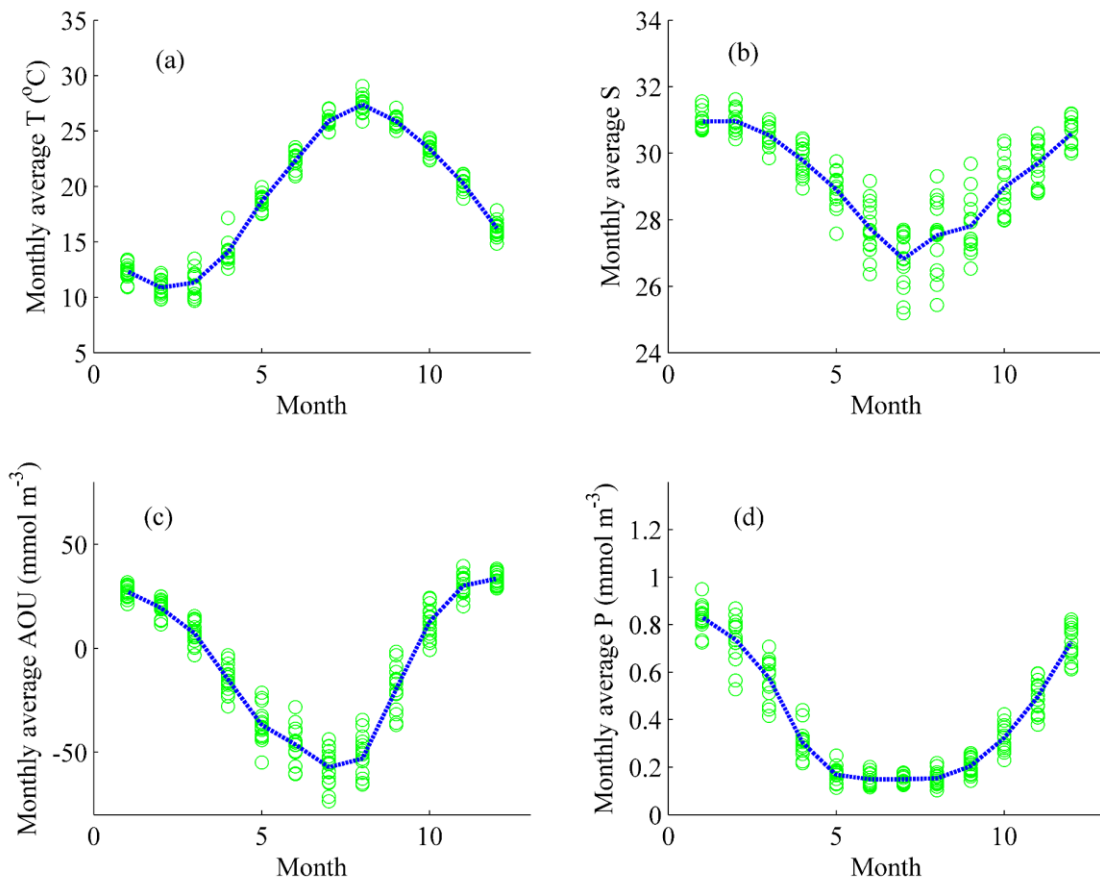
Xiaoshuang Li[1, 2], Richard Bellerby[1, 2], Jianzhong Ge[1], Philip Wallhead[2], Jing Liu[1], and Anqiang Yang[1]

[1]State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai, 200241, China

[2]Norwegian Institute for Water Research, Bergen, 5006, Norway

Correspondence to: Richard Bellerby (Richard.Bellerby@niva.no)

Abstract. While our understanding of pH dynamics has strongly progressed for open ocean regions, for marginal seas such as the East China Sea (ECS) shelf progress has been constrained by limited observations and complex interactions between biological, physical, and chemical processes. Seawater pH is a very valuable oceanographic variable but not always measured using high quality instrumentation and according to standard practices. In order to predict total scale pH (pH_T) and enhance our understanding of the seasonal variability of pH_T on the ECS shelf, an artificial neural network (ANN) model was developed using 11 cruise datasets from 2013 to 2017 with coincident observations of pH_T, temperature (T), salinity (S), dissolved oxygen (DO), nitrate (N), phosphate (P) and silicate (Si) together with sampling position and time. The reliability of the ANN model was evaluated using independent observations from 3 cruises in 2018, and showed a root mean square error accuracy of 0.04. A weight analysis of the ANN model variables suggested that DO, S, T were the most important predictor variables. The ANN model responded to T and DO errors in a positive way, S errors in a negative way, and the ANN model was most sensitive to S errors, followed by DO and T errors. Monthly water column pH_T for the period 2000-2016 was retrieved using T, S, DO, N, P, and Si from the Changjiang Biology Finite-Volume Coastal Ocean Model (FVCOM). The agreement is good here in winter, while the reduced performance in summer can be attributed in large part to limitations of the Changjiang Biology FVCOM in simulating summertime input variables.

1 Introduction

Atmospheric carbon dioxide (CO_2) levels have increased by nearly 46%, from approximately 278 ppm (parts per million) in 1750 (Ciais et al., 2013) to 405 ppm in 2017 (Le Quéré et al., 2018). The oceans have absorbed approximately 48% of the anthropogenic CO_2 emissions (Sabine et al., 2004), resulting in decreasing long-term pH trends of ~0.02 decade^{-1} in open ocean waters (e.g., Dore et al., 2009; González-Dávila et al., 2010; Bates et al., 2014; Lauvset et al., 2015). While a gradual decrease in pH is a predictable open ocean response to elevated anthropogenic CO_2 emissions, the seasonal changes and long-term trends in pH in coastal seas have not been fully understood due to the lack of long-term pH data and complexity of coastal systems. In this context, the development of approaches to predict carbonate chemistry parameters in coastal regions may assist both the management of local water quality and our wider understanding of the ocean carbon cycle.

Many attempts have been made to predict seawater pH by developing empirical relationships between pH and environmental variables, such as temperature (T) (Juranek et al., 2011), salinity (S) (Williams et al., 2016), dissolved oxygen (DO) (e.g., Juranek et al., 2011; Sauzède et al., 2017), nutrients (e.g., Williams et al., 2016; Carter et al., 2016, 2018), and longitude, latitude (Sauzède et al., 2017). Compared with traditional empirical methods, artificial neural networks (ANNs) have shown improved performance (Chen et al., 2017). ANNs have been proposed as powerful tools for modelling uncertain and complex systems such as ecosystems and environmental assessment (e.g., Olden and Jackson, 2002; Olden et al., 2004; Uusitalo, 2007; Raitsos et al., 2008; Chen et al., 2017). Their main advantage compared with multiple linear regression (MLR) models is that they do not require an a priori model but rather "learn" the model from existing data may be a greater flexibility and versatility in modelling complex nonlinear relationships. ANNs have been used for the retrieval of the partial pressure of carbon dioxide (pCO_2) (e.g., Friedrich and Oschlies, 2009; Laruelle et al., 2017), total alkalinity (e.g., Velo et al., 2013; Bostock et al., 2013; Sasse et al., 2013), total dissolved inorganic carbon (e.g., Bostock et al., 2013; Sasse et al., 2013), and phytoplankton functional

41 types (e.g., Raitsos et al., 2008; Palacz et al., 2013). However, these studies mainly focus on the open ocean; relatively few

42 studies have focused on coastal seas, perhaps because of the complexity and heterogeneity of the continental shelves. Alin et

43 al. (2012) developed an MLR model to reconstruct pH in the southern California Current System, while Moore-Maley et al.

44 (2016) evaluated the interannual variability of near-surface pH using a one dimensional, biophysical, mixing layer model in

45 the Strait of Georgia. To our knowledge, no empirical relationship for pH has yet been established for the ECS.

46 The ECS is the largest marginal sea in the western North Pacific Ocean and receives massive terrestrial inputs from the

47 Changjiang (Yangtze River) ~~River (Gong et al., 1996)~~. The shelf shallower than 200 m covers more than 70% of the entire

48 ECS (e.g., Ichikawa and Beardsley, 2002; Lie and Cho, 2016), where the dominant currents present seasonal circulation

49 patterns. The spatial and temporal distributions of the carbonate system have been investigated in the ECS (e.g., Chou et al.,

50 2009; Cao et al., 2011; Qu et al., 2015), and were found to largely reflect the distributions of various water masses. The pattern

51 of carbon sources and sinks exhibits substantial seasonal variation (Guo et al., 2015), and the ECS is generally considered as

52 a sink of atmospheric CO_2 throughout the year except in fall (e.g., Shim et al., 2007; Zhai and Dai, 2009). A mechanistic semi-

53 analytical algorithm (MeSAA) was developed to study pCO_2 variations in response to various controlling mechanisms during

54 summertime (Bai et al., 2015). However, the seasonal variability of pH has been very little studied in the ECS, mainly due to

55 the limited observational coverage and irregular variability caused by seasonal fluctuations of the Changjiang ~~River~~ discharge

56 and anthropogenic processes. Developing methods to extend the seasonal coverage of pH data may thus help to improve our

57 understanding of the ocean carbon cycle in the ECS.

58 This paper is structured as follows: section 2 describes the cruise data ~~used to build the~~ and ANN model building; section 3

59 shows the performance, sensitivity and application of the ANN model ~~performance and predictor variable importance, as well~~

60 ~~as an application to retrieve monthly pH for the period 2000-2016 on the ECS shelf using the monthly temperature, salinity,~~

61 ~~dissolved oxygen, nitrate, phosphate and silicate from the Changjiang Biology Finite-Volume Coastal Ocean Model (FVCOM)~~.

62 Summary and c~~C~~onclusions ~~and perspectives~~ are summarized in the last section.

63 **2 Data and method**

64 **2.1 Data**

65 Ten ~~Eleven~~ cruises were conducted on the ECS shelf ~~from 2013 to 2017 (Fig. 1). Ten of these were carried out~~ during the

66 "Shiptime Sharing Project of National Natural Science Foundation of China" from 2013 to 2017 (Fig. 1), the summer cruise

67 from 17 to 28 August 2013, 10 to 17 July 2014, 9 to 20 July 2015, 4 to 28 July 2016, 20 to 30 July 2017, the winter cruise

68 from 21 to 28 February 2014, 15 to 28 February 2017, the spring cruise from 4 to 20 March 2013, 11 to 21 March 2015, 7 to

69 19 March 2016;~~.~~ ~~the remaining cruise was carried out on the ECS shelf during 12-24 May 2017. Water samples were collected~~

70 ~~at three or four different depths during all cruises.~~ T and S profiles were obtained directly using a conductivity temperature-

71 depth/pressure (CTD) recorders (SBE 25plus or 911plus). Measurement of DO followed the Winkler procedure, as described

72 previously by Zhai et al. (2014). Nutrients samples were first filtered with 0.45 μm Whatman GF/F membrane, then stored in

73 250 mL HDPE bottles until chemical analysis. Nitrate (N), phosphate (P) and silicate (Si) were determined using a segmented

74 flow analyzer (Model: Skalar SAN[PLUS], Netherlands) with a precision < 3% (Zhang et al., 2007), the detection limits are 0.14

75 μM for N, 0.06 μM for P, and 0.07 μM for Si. ~~During the May 2017 survey, pH samples were stored in 500 mL high-quality~~

76 ~~borosilicate glass bottles without filtering and poisoned by addition of 200 μL saturated HgCl₂ solution until measured in the~~

77 ~~lab. The pH_T (total scale) was measured using an Automated Flow-through system for Embedded Spectrophotometry (AFtes)~~

78 ~~with a precision of 0.0005 pH_T unit and uncertainty of < 0.003 (Reggiani et al., 2016). During all other cruises,~~ pH samples

79 were stored in 140 mL brown borosilicate glass bottles and ~~poisoned~~ sterilized by addition of 50 μL saturated $HgCl_2$ solution.

80 Three traceable pH buffers were used including NIST (National Institute of Standards and Technology) buffers pH = 4.00,

81 7.02, 10.09. As described by Zhai et al. (2012, 2014), we converted it into total scale by subtracting 0.143 and the overall

82 accuracy of the pH$_T$ data was estimated as 0.01.

83 Three cruises were carried out on the ECS shelf in 2018 (Fig. 2) during the "National Natural Science Foundation Shared

84 Voyage Plan", from 10 to 19 March, 12 to 20 July, 12 to 21 October, and one cruise was carried out near the Changjiang

85 Estuary during May 2017 (Fig. 1). The measurement methods of T, S, DO, and nutrients are the same as that of the above ten

86 voyages. pH samples were stored in 500 mL high-quality borosilicate glass bottles without filtering and sterilized by addition

87 of 200 μL saturated HgCl$_2$ solution until measurement in the lab. The pH$_T$ was measured at the temperature in the flow cell

88 using an Automated Flow-through system for Embedded Spectrophotometry (AFtes) with a precision of 0.0005 pH unit and

89 uncertainty of < 0.003 (Reggiani et al., 2016). Water samples were collected at three or four different depths during all cruises.

90 We omitted data points where one or more other physical variables were missing. ~~The final number of data used by the ANN~~

91 ~~model was 1854, and the distribution of the sampling sites from the 11 cruises is shown in Fig. 1.~~ The three cruises during

92 2018 (Fig. 2) were used to estimate model predicted performance as an exploratory dataset, while the remaining eleven cruises

93 (Fig. 1) were used to train the model as a confirmatory dataset. The final number of observations in the confirmatory dataset

94 was 1854 (see Table 1 for more detailed information on the field survey).

95 **2.2 Artificial neural network development**

96 The ANN we used is a feed-forward multilayer perceptron (Tamura and Tateishi, 1997) with two hidden layers. The neurons

97 of each layer are connected with the neurons of the previous layer and the next layer by weights (Fig. ~~2~~3a). The coefficients

98 of the weight matrix are iteratively tuned in the training step. ~~Here we used the back-propagation conjugate-gradient technique~~

99 ~~(Hornik et al., 1989).~~ In order to avoid overfitting, a ten-fold cross-validation was used to assess model prediction accuracy

100 (Fig. 3b). Here, ~~the cruise data were~~ the confirmatory dataset was randomly divided into ten equal subsamples. One subsample

101 was used as the independent validation data (10% of ~~all cruise data~~ the confirmatory dataset)~~, which~~ and was always excluded

102 from training; ~~and~~ the remaining nine ~~remaining~~ subsamples were ~~together~~ used as training data (90% of ~~all cruise data~~ the

103 confirmatory dataset). The training data were further divided randomly into a training set (70% of the training data), validation

104 set (15% of the training data), and testing set (15% of the training data) during the training process. The training set was used

105 for computing the gradient and updating the network weights and biases~~, T~~; the validation set was used to monitor the error

106 and control model stop during the training process, and the testing set was used to monitor whether the model was over-

107 fitted~~matched~~ (Palacz et al., 2013). We compared performances in predicting the independent validation data from the ten-fold

108 cross-validation and selected the optimal model based on the lowest root mean square error (RMSE). Then we applied the

109 optimal model to the exploratory dataset (Fig. 2) and evaluated model performance by calculating error statistics. In our study,

110 calculations were done in the MathWorks Matlab environment, using the Deep Learning Toolbox.

111 First, we compared the performance of one hidden layer vs. two hidden layers in predicting independent validation data. The

112 number of neurons varied from 2^2 to 2^8 for the first hidden layer and was fixed at four in the second hidden layer for the two

113 hidden layers model; the number of neurons in the first layer was the same in the one hidden layer vs. two hidden layers model

114 (Fig. 4), The ten-fold cross-validation showed that the model with two hidden layers performed better as the number of neurons

115 increased. Second, in order to choose suitable training techniques and activation functions of the ANN model with two hidden

116 layers, we tested three training functions (Gradient descent backpropagation (trainGD), Levenberg-Marquardt

117 backpropagation (trainLM), and Scaled conjugate gradient backpropagation (trainSCG)), which differed in how the weights

118 are modified, and three transfer functions (Log-sigmoid transfer function (logsig), Hyperbolic tangent sigmoid transfer

119 function (tansig), and Positive linear transfer function (poslin)) (Fig. 5). The output values of logsig, tansig and poslin were

120 compressed onto [0, 1], [-1, 1], and [0, +∞], respectively (Fig. S1). As the number of neurons increased, the performances of

121 trainGD and tansig became poor. Although there was no obvious difference between trainLM and trainSCG, the training

122 technique trainSCG was selected and transfer function logsig was applied to two hidden layers considering the overall

123 performance (Fig. 5). Third, in the training phase of the ANN model, the number of neurons was tested, varying from 4 to 128

124 for two hidden layers (Table S1). Best performance for both training data and independent validation data was obtained with

125 40 neurons in the first hidden layer and 16 neurons in the second layer. Finally, d~~D~~ifferent combinations of input variables

126 were tested to choose the optimal architecture of the ANN model (Table 2); best performance was obtained using ~~As shown~~

127 ~~in Fig. 2, input variables include~~ longitude, latitude, month, T, S, DO, N, P and Si as input variables. ~~We selected these variables~~

128 ~~as principal inputs for the following reasons~~ The utility of these variables for predicting pH has a strong a priori basis: the

129 carbonate system thermodynamic relationships depend on T and S (Lueker et al., 2000); ~~DO was expected to vary with pH~~

130 ~~and there was a tight positive link~~ a positive correlation is expected between DO and pH (Wootton et al., 2012) because of the

131 role of photosynthesis and respiration in removing or generating CO_2 in the water; various nutrients influence phytoplankton

132 growth and abundance, ~~which might~~ thereby increas~~e~~ing organic carbon fixation~~, increasing inorganic carbon~~/uptake and

133 increasing pH (Wootton et al., 2008, 2012). We found geographical information to be a powerful addition in improving the

134 skill of the method (~~see~~ Table 2), allowing the network to learn spatio-temporal patterns that could not be explained by other

135 input variables (Sasse et al., 2013). ~~The number of neurons in the two hidden layers was tested, varying between 1 and 100 for~~

136 ~~the first hidden layer and between 1 and 50 for the second hidden layer. The optimal architecture was composed of two hidden~~

137 ~~layers with 40 neurons in the first and 16 neurons in the second.~~

138 In order to avoid bias towards high-value inputs/outputs and to eliminate the dimensional influence of the data, all data used

139 by the ANN model were normalized using the following equation (e.g., Sauzède et al., 2015, 2016):

140 $$x_{i,j} = \frac{2}{3} * \frac{x_{i,j} - mean(x_{i,j})}{\sigma(x_{i,j})} \qquad (1)$$

141 with σ the standard deviation of the considered input variables or output variable pH_T. Similar to the approach of Sauzède et

142 al. (2015, 2016), the longitude and month input variables were transformed as follows to account for the periodicity:

143 $slongitude = \sin\left(\frac{Lon*\pi}{180}\right)$, $clongitude = \cos\left(\frac{Lon*\pi}{180}\right)$ (2)

144 $smonth = \sin\left(\frac{month*\pi}{6}\right)$, $cmonth = \cos\left(\frac{month*\pi}{6}\right)$ (3)

145 The latitude variable was transformed into the range of the sigmoid function by dividing by 90 ~~(Sauzède et al., 2015)~~, then

146 normalized using (1).

3 Result and discussion

3.1 ~~The~~ ANN model performance

149 To evaluate the performance of the ANN model, we compared model simulated pH_T (pH_T^M) with corresponding observations

150 (pH_T^O) using several statistical indices, including the mean absolute error (MAE), the coefficient of determination (R^2), and

151 the root mean squared error (RMSE). The model simulated pH_T with a RMSE ~~=~~ of 0.04 and R^2 ~~=~~ of 0.88 for the training data

152 (90% of ~~all data, Fig 3a~~ confirmatory dataset, Fig. 6a), and predicted pH_T with a RMSE ~~=~~ of 0.03 and R^2 ~~=~~ of 0.93 for the

153 independent validation data (10% of ~~all data, Fig 3b~~ confirmatory dataset, Fig. 6b). ~~The distributions of the differences (pH_T^M~~

154 ~~- pH_T^O) were approximately normal with no obvious outliers (Fig. 4).~~ The histogram of residuals in confirmatory dataset (Fig.

155 6c) showed that 68% of the residuals were within the RMSE of 0.04.

156 In order to further explore where the ANN model may lead to large errors, we plotted distributions of differences (pH_T^M -

157 pH_T^O) with respect to the longitude and latitude (Fig. ~~5~~7). The points with large errors are mainly concentrated in the longitude

158 range [122.5°E, 123°E] and the latitude range [3~~2~~1°N, 32.5°N], in an area strongly influenced by the Changjiang Dilute Water

159 (CDW). The reduced performance of the ANN model here may be primarily due to the strong seasonal oscillations of the

160 Changjiang ~~River~~ discharge (Dai and Trenberth, 2002). As a reference, the performance of some other empirical approaches,

161 including MLR, multi-variate nonlinear regression (MNR), decision tree, random forest, and Support Vector Machine (SVM)

162 regression, is shown in Table 3. The selected ANN model (Table 2, Model#10) showed better performance than the other

163 tested approaches using the same input variables (Table 3). ~~Although the RMSE for pH_T we obtained here was higher than~~
164 ~~obtained in some previous studies (e.g., Juranek et al., 2011; Williams et al., 2016; Sauzède et al., 2017), their research regions~~
165 ~~were open ocean regions, not coastal seas. For example, Juranek et al. (2011) developed empirical algorithms to estimate pH~~
166 ~~with RMSE of 0.018 for data between 30-500 m in the NE subarctic Pacific; Williams et al. (2016) also developed empirical~~
167 ~~algorithms to predict pH with RMSE of 0.01 in the Southern Ocean; Sauzède et al. (2017) developed a neural network method~~
168 ~~to estimate pH with RMSE of 0.02 in the global ocean. However, coastal seas tend to show greater temporal and spatial~~
169 ~~variability than open oceans. Alin et al. (2012) developed a MLR approach to reconstruct pH with RMSE of 0.024 in the~~
170 ~~southern California Current System. Zhai et al. (2014) compared the field-measured pH with calculated pH from measured~~
171 ~~total alkalinity and dissolved inorganic carbon using the program CO2SYS.xls (Pelletier et al., 2011) and obtained~~
172 ~~discrepancies with standard deviation 0.05. Carbon chemistry parameters in this region are not only under the direct impact of~~
173 ~~Taiwan Warm Current and remote control of the Kuroshio water intrusion into the shelf but also significantly controlled by~~
174 ~~seasonal variations of the Changjiang River discharge (e.g., Isobe and Matsuno, 2008; Chen et al., 2008; Chou et al., 2009).~~
175 ~~Taking into account the highly complex hydrographic, biological and chemical conditions, the accuracy of pH_T presented is~~
176 ~~promising.~~

3.2 ~~Comparison with new field data~~ ANN model validation using the exploratory dataset

178 To further assess the ability of the ANN model to estimate pH_T on the ECS shelf, we applied the ANN model to an exploratory
179 dataset ~~data from three cruises~~ not used in ~~the~~ ANN model development ~~(Fig. 6):~~ and sampled during March, July, and October
180 2018 (Fig. 2). Scatterplots of retrieved pH_T vs observations (Fig. ~~7~~8a) show~~ed~~ ~~that the ANN model predicts pH_T with~~ an RMSE
181 of 0.04, MAE of 0.03 and R^2 of 0.80 ~~for these cruise data~~. ~~This result is~~ which is consistent with the ~~result~~ performance of the
182 training data (Fig. ~~3~~6a)~~, which further reflects the stability and reliability of the ANN model on the ECS shelf~~. Although the
183 RMSE for pH_T we obtained here was higher than obtained in some previous studies (e.g., Juranek et al., 2011; Williams et al.,
184 2016; Sauzède et al., 2017), these latter studies considered open ocean regions, not coastal seas. For example, Juranek et al.
185 (2011) developed empirical algorithms to estimate pH with RMSE of 0.018 for data between 30-500 m in the NE subarctic
186 Pacific; Williams et al. (2016) also developed empirical algorithms to predict pH with RMSE of 0.01 in the Southern Ocean;
187 Sauzède et al. (2017) developed a neural network method to estimate pH with RMSE of 0.02 in the global ocean. As a further
188 comparison we applied the CANYON model developed by Sauzède et al. (2017) to our coastal exploratory dataset (Fig. 8b),
189 and obtained an RMSE of 0.09 and MAE of 0.06. It is not surprising that the ANN model (developed here for the ECS shelf)
190 outperforms the CANYON model (developed for the global ocean) for predicting pH_T on the ECS shelf. The carbon chemistry
191 parameters in this region are not only under the direct impact of Taiwan Warm Current and remote control of the Kuroshio
192 water intrusion into the shelf, but are also significantly controlled by seasonal variations of the Changjiang discharge (e.g.,
193 Isobe and Matsuno, 2008; Chen et al., 2008; Chou et al., 2009). Taking into account the highly complex hydrographic,
194 biological and chemical conditions, the accuracy of pH_T presented is promising.

3.3 ~~Variable importance in the ANN model~~ ANN model sensitivity to environmental input variables

196 To assess the ANN model sensitivity to ~~relative importance of~~ different environmental input variables ~~in the ANN model~~, we
197 added 5% perturbation for each environmental variable separately. Statistically, with 5% T errors added, the ANN model
198 showed slight overestimation in pH_T, with mean bias (MB) of 0.0059, RMSE of 0.0079, and R^2 of 0.9949 (Fig. 9a); with 5%
199 DO errors added, the ANN model also showed slight pH_T overestimation, with MB of 0.0050, RMSE of 0.0090, and R^2 of
200 0.9934 (Fig. 9c); with 5% S errors added, the ANN model showed overestimation in pHT, with MB of -0.0111, RMSE of
201 0.0162, and R^2 of 0.9789 (Fig. 9b). These results suggested that the ANN model responded to T and DO errors in a positive
202 way, S errors in a negative way. ~~we used the following method: for each environmental variable separately, add 5% and~~
203 ~~calculate the resulting percentage change in the predicted pH_T. Predicted pH_T responded positively to (T, DO) and negatively~~

204 ~~to S (Fig. 8). The variable with the greatest weight was DO, followed by S and T, and the weights of nutrients were relatively~~

205 ~~small.This is consistent with (Cai et al., 2011) where~~ The positive response to increasing DO reflects positive correlations

206 between pH_T and DO (Cai et al., 2011), which ~~in the Gulf of Mexico and ECS were~~ can be attributed to the processes of

207 photosynthesis (generating DO and removing CO_2, hence increasing pH) and aerobic respiration (consuming DO and

208 generating CO_2, hence lowering pH). Th~~is~~e negative response to increasing S reflects the influence of the (lower salinity)

209 Changjiang ~~River~~ discharge, carrying large amounts of nutrients that fuel increased primary production (uptake of nutrients

210 and CO_2, hence raising the pH) in surface waters during warm seasons (Gong et al., 2011). It was found that the ANN model

211 was insensitive to nutrients errors (Fig. 9d-9f) and most sensitive to S errors (Fig. 9b), followed by DO and T errors.

3.4 ANN model application

3.4.1 Comparison

214 In order to retrieve monthly pH_T on the ECS shelf, the monthly T, S, DO, N, P and Si from the Changjiang Biology Finite-

215 Volume Coastal Ocean Model (FVCOM) (http://47.101.49.44/wms/demo) were fed into the ANN model as input variables.

216 ~~Monthly pH_T for the period 2000-2016 was obtained at the spatial resolution of the Changjiang Biology FVCOM output: 1-10~~

217 ~~km in the horizontal, 10 depth levels in the vertical, and 12 months.~~ The resolution of the Changjiang Biology FVCOM output

218 is 1-10 km in the horizontal, 10 depth levels in the vertical, and day in the temporal (refered Ge et al., (2013) for detail

219 information). Comparisons of monthly-average FVCOM model variables with surface and bottom observations on the ECS

220 shelf showed that simulated T was close to observed values (Fig. S2a), simulated S was also close to observed values except

221 at the bottom in August 2013 and at the surface in July 2016 (Fig. S2b), simulated DO was higher than observed DO at the

222 bottom (Fig. S2c), and simulated nutrients were higher than observed nutrients at the surface (Fig. S2d-S2f). Comparisons of

223 monthly average pH_T from the Changjiang Biology FVCOM model with pH_T retrieved by the ANN model suggested that the

224 ANN model can potentially provide a more accurate pH_T (Fig. S3). The possible reason was that the carbonate system from

225 the Changjiang Biology FVCOM was not optimized due to challenges obtaining sufficient boundary information.

226 Considering the discreteness and discontinuity of the sampling sites, we compared ~~retrieved~~ pH_T retrieved by the ANN model

227 using the Changjiang Biology FVCOM output with the corresponding observations at some sites with repeated sampling for

228 3 to 4 years. These sites were A1-5 (123.0140°E, 32.2145°N), A1-6 (123.2750°E, 32.2679°N), A6-7 (122.9880°E, 30.7050°N),

229 A6-9 (123.4990°E, 30.5723°N), A7-5 (123.4990°E, 30.2523°N), and A8-5 (123.4930°E, 29.9940°N). Overall, the retrieved

230 pH_T ~~from the Changjiang Biology FVCOM output~~ agrees well (within the ANN model accuracy: ANN±RMSE) with the

231 observed values at the surface, except for three samples in summer (Fig. ~~9~~10). There are relatively large deviations (greater

232 than the RMSE of 0.04) in August 2013 at station A1-5 and A6-9, and in July 2016 at station A8-5. ~~These may be primarily~~

233 ~~attributed to the sudden increase in the Changjiang River discharge (Dai and Trenberth, 2002).~~ To illustrate the application

234 performance in the water column, a scatterplot of retrieved pH_T vs observations at six sites with repeated sampling for 3 to 4

235 years (Fig. 1~~0~~1) showed that the ANN model predicted pH_T with a RMSE of 0.05 and R^2 of 0.71.

236 We further compared monthly pH_T retrieved by the ANN model using the Changjiang Biology FVCOM output with ~~retrieved~~

237 ~~pH_T using measured T, S and DO, and~~ in situ measured pH_T values (Fig. 1~~1~~2). The agreement is good (within the ANN model

238 accuracy: ANN±RMSE) here in winter, but large deviations (greater than the RMSE of 0.04) appear in summer. The reduced

239 performance in summer can be attributed in large part a reduced performance of the Changjiang Biology FVCOM model in

240 predicting summertime input variables S, DO, and nutrients (Fig. S~~1~~2)~~; using the observed values of DO, S, etc. as predictor~~

241 ~~variables, the skill of the ANN pH_T predictions is much improved (RMSE = 0.09 vs. RMSE = 0.02).~~

3.4.2 Spatial and temporal patterns of ANN-derived pH_T

243 The temporal and spatial variations of monthly surface pH_T from 2000-2016 based on Changjiang Biology FVCOM output

244 are shown in Figure 13. During the dry season (November to March of the next year), pH_T values vary from ~7.62 to ~8.24.

Relatively higher pH_T values are found in the southeastern of the study area (Chou et al., 2011), whereas lower pH_T values are found in the northeastern of the study area. During the wet season (April to October), pH_T values vary from ~7.77 to ~8.35, water of higher pH_T corresponded well to the seasonal dispersion of the Changjiang Dilute Water (Chou et al., 2009, 2013). Water of higher pH_T is found in the center of the study area during April, spreads to the southwestern part of the study area (along the coast of China) during May and June, shifts to the northeastern part of the study area during August. In September and October, water of higher pH_T is found in the southeastern part of the study area, strongly influenced by the Taiwan Warm Current (Qu et al., 2015).

A clear seasonality is that surface pH_T gradually increases during spring (March to May), after which it gradually decreases during summer and fall (June to November) (Fig. 14). The surface pH_T displays its maximum in May and minimum in December, and the pH_T varies seasonally by up to ~0.3 unit. Larger changes in pH were also discovered in the Washington Shelf, the pH varied ~1.0 unit over the seasons and ~1.5 unit spanning 8 years (Wootton et al., 2008). Accordingly, seasonal dynamics of surface pH_T can be mainly attributed to temperature changes and strong biological activities (production and respiration processes) over the season. From March to June, a rapid increase in surface pH_T indicates that production increases faster than respiration, which can be reflected in the drop in surface phosphate (Fig. S5d) and apparent oxygen utilization (AOU) (Fig. S5c). It may be driven by the Changjiang discharge (Fig. S4), which carries large amount of nutrients, result in stronger primary production in warm seasons under the combined action of nutrients and suitable temperature (Gong et al., 2011). From July to October, although surface temperature remains at a high level (Fig. S5a), the rise in surface AOU (Fig. S5c) suggest a decrease in primary production or increase of respiration, which leads to a gradual drop in surface pH_T (Wootton et al., 2012). It implies respiration processes dominate relative to primary production during summer and fall.

4 Summary and cConclusions and perspectives

We have developed an artificial neural network (ANN) model, demonstrated its reliability, and used it to retrieve monthly pH_T for the period 2000-2016 on the East China Sea shelf. This model predicts the water column pH_T using nine input components, and the three most important environmental input variables were dissolved oxygen, salinity and temperature. We trained this ANN model using 11 cruise datasets from 2013 to 2017. In order to choose the optimal architecture of the ANN model, we tested different training and transfer functions, the number of neurons in two hidden layers, and different combinations of input variables. We also validated the reliability of the ANN model with a root mean square error accuracy of 0.04 using three cruises in 2018 as exploratory dataset. The ANN model responded to temperature and dissolved oxygen errors in a positive way, salinity errors in a negative way, and was most sensitive to salinity errors, followed by dissolved oxygen and temperature errors. We also retrieved monthly-average pH_T using the ANN model in combination with input variables from the Changjiang Biology Finite-Volume Coastal Ocean Model (FVCOM).

The approach has several potential applications. First, it can provide estimates of seawater pH_T with known accuracies for the East China Sea shelf and the period 2013-2018. Within this region the model could be used as a cost-effective way to handle restrictions of marine observations conducted from ships, such as coarse resolution and under-sampling of carbonate system variables. Second, while the ANN model is not a replacement for direct measurements of the carbonate system, it may be a valuable tool for understanding the seasonal variation of pH_T in poorly observed regions. Third, this approach can be applied to other regions to predict pH by suitably adapting the input variables and network structure using local dataset. The MATLAB code used in this study to develop and apply the ANN model is freely available, and is accompanied by a README file providing detailed guidance on how to use and adapt the code.

Code and data availability

Matlab code of the ANN model for pH_T estimation and datasets are available:

285 http://doi.org/10.5281/zenodo.3519219

286 The monthly-average input variables (T, S, DO, N, P, Si) from the Changjiang Biology Finite-Volume Coastal Ocean Model

287 and retrieved pH_T values from 2000 to 2016 on the East China Sea shelf and three cruises data during 2018 used to evaluate

288 the ANN model are available:

289 http://doi.org/10.5281/zenodo.3519236

290 Requests to access the raw data should be directed to Richard Bellerby: Richard.Bellerby@niva.no

291 Six stations with repeated sampling for 3 to 4 years and corresponding retrieved pH values from the Changjiang Biology

292 FVCOM output are available: http://doi.org/10.5281/zenodo.3491747

Video supplement

294 Monthly distribution of surface pH_T on the East China Sea shelf from 2000 to 2016 year:

295 http://doi.org/10.5281/zenodo.2672943

296 Profile distribution of pH_T at 31°N on the East China Sea shelf from 2000 to 2016 year:

297 http://doi.org/10.5281/zenodo.2672929

Author contribution

299 Li, X. S. and Bellerby, R. contributed to the development of methodology and the design of the model. Ge, J. Z. provided ten

300 cruises dataset from 2013 to 2017 year and the input variables from the Changjiang Biology Finite-Volume Coastal Ocean

301 Model Data. Liu, J. and Yang, A. Q. provided four cruises dataset from 2017 to 2018 year. Li, X. S. developed the manuscript

302 with contributions from all co-authors.

References

312 Alin, S. R., Feely, R. A., Dickson, A. G., Hernández-Ayón, J. M., Juranek, L. W., Ohman, M. D., and Goericke, R.: Robust

313 empirical relationships for estimating the carbonate system in the southern California Current System and application to

314 CalCOFI hydrographic cruise data (2005–2011), J. Geophys. Res., 117, C05033, doi:10.1029/2011JC007511, 2012.

315 Bai, Y., Cai, W. J., He, X. Q., Zhai, W. D., Pan, D., Dai, M. H., and Yu, P. S.: A mechanistic semi-analytical method for

316 remotely sensing sea surface pCO_2 in river-dominated coastal oceans: A case study from the East China Sea, J. Geophys. Res.

317 Oceans, 120, 2331–2349, doi:10.1002/2014JC010632, 2015.

318 Bates, N. R., Astor, Y. M., Church, M. J., Currie, K., Dore, J. E., González-Dávila, M., Lorenzoni, L., Muller-Karger, F.,

319 Olafsson, J., and Santana-Casiano, J. M.: A time-series view of changing ocean chemistry due to ocean uptake of anthropogenic

320 CO_2 and ocean acidification, Oceanography, 27(1), 126–141, doi:10.5670/oceanog.2014.16, 2014.

Bostock, H. C., Mikaloff Fletcher, S. E., and Williams, M. J. M.: Estimating carbonate parameters from hydrographic data for the intermediate and deep waters of the Southern Hemisphere oceans, Biogeosciences, 10, 6199–6213, https://doi.org/10.5194/bg-10-6199-2013, 2013.

Cai, W. J., Hu, X. P., Huang W. J., Murrell, M. C., Lehrter, J. C., Lohrenz, S. E., Chou, W. C., Zhai, W. D., Hollibaugh, J. T., Wang, Y. C., Zhao, P. S., Guo, X. H., Gundersen, K., Dai, M. H., and Gong, G. C.: Acidification of subsurface coastal waters enhanced by eutrophication, Nature Geoscience, 4, 766-770, doi:10.1038/NGEO1297, 2011.

Cao, Z. M., Dai, M. H., Zheng, N., Wang, D., Li, Q., Zhai, W. D., Meng, F. F., and Gan, J. P.: Dynamics of the carbonate system in a large continental shelf system under the influence of both a river plume and coastal upwelling, J. Geophys. Res., 116, G02010, doi:10.1029/2010JG001596, 2011.

Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., and Takeshita, Y.: Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate, Limnol. Oceanogr. Methods, 16, 119-131, doi:10.1002/lom3.10232, 2018.

Carter, B. R., Williams, N. L., Gray, A. R., and Feely, R. A.: Locally interpolated alkalinity regression for global alkalinity estimation, Limnol. Oceanogr. Methods, 14, 268-277, doi:10.1002/lom3.10087, 2016.

Chen, C. S., Xue, P. F., Ding, P. X., Beardsley, R. C., Xu, Q. C., Mao, X. M., Gao, G. P., Qi, J. H., Li, C. Y., Lin, H. C., Cowles, G., and Shi, M. C.: Physical mechanisms for the offshore detachment of the Changjiang Diluted Water in the East China Sea, J. Geophys. Res., 113, C02002, doi:10.1029/2006JC003994, 2008.

Chen, S. L. and Hu, C. M.: Estimating sea surface salinity in the northern Gulf of Mexico from satellite ocean color measurements, Remote Sens Environ, 201, 115-132, https://doi.org/10.1016/j.rse.2017.09.004, 2017.

Chou, W. C., Gong, G. C., Sheu, D. D., Hung, C. C., and Tseng, T. F.: Surface distributions of carbon chemistry parameters in the East China Sea in summer 2007, J. Geophys. Res., 114, C07026, doi:10.1029/2008JC005128, 2009.

Chou, W. C., Gong, G. C., Tseng, C. M., Sheu, D. D., Hung, C. C., Chang, L. P., Wang, L. W.: The carbonate system in the East China Sea in winter, Mar. Chem., 123, 44-55, doi:10.1016/j.marchem.2010.09.004, 2011.

Chou, W. C., Gong, G. C., Hung, C. C., and Wu, Y. H.: Carbonate mineral saturation states in the East China Sea: present conditions and future scenarios, Biogeosciences, 10, 6453-6467, doi:10.5194/bg-10-6453-2013, 2013.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T. F., D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Dai, A. and Trenberth, K. E.: Estimates of freshwater discharge from continents: Latitudinal and seasonal variations, J. Hydrometeorol., 3, 660–687, https://doi.org/10.1175/1525-7541, 2002.

Dore, J., Lukas, R., Sadler, D., Church, M., and Karl, D.: Physical and biogeochemical modulation of ocean acidification in the central North Pacific, Proc. Natl. Acad. Sci. U. S. A., 106, 12235–12240, 2009.

Friedrich, T. and Oschlies, A.: Neural network-based estimates of North Atlantic surface pCO_2 from satellite data: A methodological study, J. Geophys. Res., 114, C03020, https://doi.org/10.1029/2007JC004646, 2009.

Gong, G. C., Chen, Y. L. L., and Liu, K. K.: Chemical hydrography and chlorophyll a distribution in the East China Sea in summer: implications in nutrient dynamics, Cont. Shelf Res, 16, 1561-1590, https://doi.org/10.1016/0278-4343(96)00005-2, 1996.

Ge, J. Z., Ding, P. X., Chen, C. S., Hu, S., Fu, G., and Wu, L. Y.: An integrated East China Sea–Changjiang Estuary model system with aim at resolving multi-scale regional–shelf–estuarine dynamics, Ocean Dyn., 63, 881-900, doi: 10.1007/s10236-013-0631-3, 2013.

363 Gong, G. C., Liu, K. K., Chiang, K. P., Hsiung, T. M., Chang, J., Chen, C. C., Hung, C. C., Chou, W. C., Chung, C. C., Chen,

364 H. Y., Shiah, F. K., Tsai, A. Y., Hsieh, C. H., Shiao, J. C., Tseng, C. M., Hsu, S. C., Lee, H. J., Lee, M. A., Lin, I. I., and Tsai,

365 F.: Yangtze River floods enhance coastal ocean phytoplankton biomass and potential fish production, Geophys. Res. Lett., 38,

366 L13603, doi:10.1029/2011GL047519, 2011.

367 González-Dávila, M., Santana-Casiano, J. M., Rueda, M. J., and Llinás, O.: The water column distribution of carbonate system

368 variables at the ESTOC site from 1995 to 2004, Biogeosciences, 7, 3067−3081, 2010.

369 Guo, X. H., Zhai, W. D., Dai, M. H., Zhang, C., Bai, Y., Xu, Y., Li, Q., and Wang, G. Z.: Air–sea CO_2 fluxes in the East China

370 Sea based on multiple-year underway observations, Biogeosciences, 12, 5495-5514, doi:10.5194/bg-12-5495-2015, 2015.

371 Ichikawa, H. and Beardsley, R. C.: The Current System in the Yellow and East China Seas, J. Oceanogr., 58, 77–92,

372 https://doi.org/10.1023/A:1015876701363, 2002.

373 Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, Neural Netw., 2,

374 359-366, https://doi.org/10.1016/0893-6080(89)90020-8, 1989.

375 Isobe, A. and Matsuno, T.: Long-distance nutrient-transport process in the Changjiang River plume on the East China Sea

376 shelf in summer, J. Geophys. Res.-Oceans, 113, C04006, doi:10.1029/2007JC004248, 2008.

377 Juranek, L. W., Feely, R. A., Gilbert, D., Freeland, H., and Miller, L. A.: Real-time estimation of pH and aragonite saturation

378 state from Argo profiling floats: Prospects for an autonomous carbon observing strategy, Geophys. Res. Lett., 38, L17603,

379 doi:10.1029/2011GL048580, 2011.

380 Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J. L., Delille, B., and Regnier, P.: Global high-resolution monthly pCO_2

381 climatology for the coastal ocean derived from neural network interpolation, Biogeosciences, 14, 4545-4561, doi:10.5194/bg-

382 14-4545-2017, 2017.

383 Lauvset, S. K., Gruber, N., Landschützer, P., Olsen, A., and Tjiputra, J.: Trends and drivers in global surface ocean pH over

384 the past 3 decades, Biogeosciences, 12(5), 1285–1298, doi:10.5194/bg-12-1285-2015, 2015.

385 Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., and Pickers, P. A., et al.: Global Carbon

386 Budget 2018, Earth Syst. Sci. Data, 10, 2141–2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

387 Lie, H. J. and Cho, C. H.: Seasonal circulation patterns of the Yellow and East China Seas derived from satellite-tracked drifter

388 trajectories and hydrographic observations, Prog. Oceanogr., 146, 121-141, http://dx.doi.org/10.1016/j.pocean.2016.06.004,

389 2016.

390 Lueker, T. J., Dickson, A. G., and Keeling, C. D.: Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and

391 equations for K_1 and K_2: Validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium, Mar. Chem.,

392 70, 105–119, doi:10.1016/S0304-4203(00)00022-0, 2000.

393 Moore-Maley, B. L., Allen, S. E., and Ianson, D.: Locally driven interannual variability of near-surface pH and Ω_A in the Strait

394 of Georgia, J. Geophys. Res. Oceans, 121, 1600–1625, doi:10.1002/2015JC011118, 2016.

395 Olden, J. D. and Jackson, D. A.: Illuminating the "black box": a randomization approach for understanding variable

396 contributions in artificial neural networks, Ecol. Model, 154, 135-150, https://doi.org/10.1016/S0304-3800(02)00064-9, 2002.

397 Olden, J. D., Joy, M. K., and Death, R. G.: An accurate comparison of methods for quantifying variable importance in artificial

398 neural networks using simulated data, Ecol. Model, 178, 389-397, https://doi.org/10.1016/j.ecolmodel.2004.03.013, 2004.

399 Palacz, A. P., John, M. A. S., Brewin, R. J. W., Hirata, T., and Gregg, W. W.: Distribution of phytoplankton functional types

400 in high-nitrate, low-chlorophyll waters in a new diagnostic ecological indicator model, Biogeosciences, 10, 8103-8157,

401 https://doi.org/10.5194/bgd-10-8103-2013, 2013.

402 Pelletier, G. J., Lewis, E., and Wallace, D. W. R.: CO2SYS.XLS: A calculator for the CO_2 system in seawater for Microsoft

403 Excel/VBA, Ver. 16, Washington State Department of Ecology, Olympia, Washington, 2011.

404 Qu, B. X., Song, J. M., Yuan, H. M., Li, X. G., Li, N., Duan, L. Q., Chen, X., and Lu, X.: Summer carbonate chemistry

405 dynamics in the Southern Yellow Sea and the East China Sea: Regional variations and controls, Cont. Shelf Res, 111, 250-

406 261, https://doi.org/10.1016/j.csr.2015.08.017, 2015.

407 Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralabous, J., Richardson, A. J., and Reid, P. C.: Identifying four

408 phytoplankton functional types from space: an ecological approach, Limnol. Oceanogr., 53, 605–613,

409 https://doi.org/10.4319/lo.2008.53.2.0605, 2008.

410 Reggiani, E. R., King, A. L., Norli, M., Jaccard, P., Sorensen, K., and Bellerby, R. G. J.: FerryBox-assisted monitoring of

411 mixed layer pH in the Norwegian Coastal Current, Journal of Marine Systems, 162, 29-36, doi:10.1016/j.jmarsys.2016.03.017,

412 2016.

413 Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R.,

414 Tilbrook, B., Millero, F. J., Peng, T. H., Kozyr, A., Ono, T., Rios, A. F.: The oceanic sink for anthropogenic CO_2, Science,

415 305, 367−371, 2004.

416 Sasse, T. P., McNeil, B. I., and Abramowitz, G.: A novel method for diagnosing seasonal to inter-annual surface ocean carbon

417 dynamics from bottle data using neural networks, Biogeosciences, 10, 4319–4340, doi:10.5194/bg-10-4319-2013, 2013.

418 Sauzède, R., Bittig, H. C., Claustre, H., de Fommervault, O. P., Gattuso, J. P., Legendre, L., and Johnson, K. S.: Estimates of

419 Water-Column Nutrient Concentrations and Carbonate System Parameters in the Global Ocean: A Novel Approach Based on

420 Neural Networks, Front. Mar. Sci., 4, 128, doi:10.3389/fmars.2017.00128, 2017.

421 Sauzède, R., Claustre, H., Jamet, C., Uitz, J., Ras, J., Mignot, A., and D'Ortenzio, F.: Retrieving the vertical distribution of

422 chlorophyll-a concentration and phytoplankton community composition from in situ fluorescence profiles: a method based on

423 a neural network with potential for global-scale applications, J. Geophys. Res. Ocean, 120, 451–470,

424 doi:10.1002/2014JC010355, 2015.

425 Sauzède, R., Claustre, H., Uitz, J., Jamet, C., Dall'Olmo, G., D'Ortenzio, F., Gentili, B., Poteau, A., and Schmechtig, C.: A

426 neural network-based method for merging ocean color and Argo data to extend surface bio-optical properties to depth: retrieval

427 of the particulate backscattering coefficient, J. Geophys. Res. Ocean, 121, 2552–2571, doi:10.1002/2015JC011408, 2016.

428 Shim, J. H., Kim, D., Kang, Y. C., Lee, J. H., Jang, S. T., and Kim, C. H.: Seasonal variations in pCO_2 and its controlling

429 factors in surface seawater of the northern East China Sea, Cont. Shelf Res., 27, 2623–2636,

430 https://doi.org/10.1016/j.csr.2007.07.005, 2007.

431 Tamura, S. and Tateishi, M.: Capabilities of a Four-Layered Feedforward Neural Network: Four Layers versus Three, IEEE

432 Transactions on Neural Networks, 8, 251–255, doi:10.1109/72.557662, 1997.

433 Uusitalo, L.: Advantages and challenges of Bayesian networks in environmental modelling, Ecol. Model, 203, 312-318,

434 https://doi.org/10.1016/j.ecolmodel.2006.11.033, 2007.

435 Velo, A., Pérez, F. F., Tanhua, T., Gilcoto, M., Ríos, A. F., and Key, R. M.: Total alkalinity estimation using MLR and neural

436 network techniques, J Marine Syst, 111-112, 11-18, https://doi.org/10.1016/j.jmarsys.2012.09.002, 2013.

437 Williams, N. L., Juranek, L. W., Johnson, K. S., Feely, R. A., Riser, S. C., Talley, L. D., Russell, J. L., Sarmiento, J. L., and

438 Wanninkhof, R.: Empirical algorithms to estimate water column pH in the Southern Ocean, Geophys. Res. Lett., 43, 3415–

439 3422, doi:10.1002/2016GL068539, 2016.

440 Wootton, J. T., Pfister, C. A., and Forester, J. D.: Dynamic patterns and ecological impacts of declining ocean pH in a high

441 resolution multi-year dataset, Proc. Natl. Acad. Sci. U. S. A., 105, 18848–18853, https://doi.org/10.1073/pnas.0810079105,

442 2008.

443 Wootton, J. T. and Pfister, C. A.: Carbon System Measurements and Potential Climatic Drivers at a Site of Rapidly Declining

444 Ocean pH, PLoS ONE, 7(12): e53396, https://doi.org/10.1371/journal.pone.0053396, 2012.

445 Zhai, W. D. and Dai, M. H.: On the seasonal variation of air-sea CO_2 fluxes in the outer Changjiang (Yangtze River) Estuary,

446 East China Sea, Mar. Chem., 117, 2–10, https://doi.org/10.1016/j.marchem.2009.02.008, 2009.

447 Zhai, W. D., Zhao, H. D., Zheng, N., and Xu, Y.: Coastal acidification in summer bottom oxygen-depleted waters in
448 northwestern-northern Bohai Sea from June to August in 2011, Chinese Science Bulletin, 57, 1062-1068, doi:10.1007/s11434-
449 011-4949-2, 2012.
450 Zhai, W. D., Zheng, N., Huo, C., Xu, Y., Zhao, H. D., Li, Y. W., Zang, K. P., Wang, J. Y., and Xu, X. M.: Subsurface pH and
451 carbonate saturation state of aragonite on the Chinese side of the North Yellow Sea: seasonal variations and controls,
452 Biogeosciences, 11, 1103-1123, https://doi.org/10.5194/bg-11-1103-2014, 2014.
453 Zhang, G., Zhang, J., and Liu, S. M.: Characterization of nutrients in the atmospheric wet and dry deposition observed at the
454 two monitoring sites over Yellow Sea and East China Sea, J Atmos Chem., 57(1), 41-57, doi:10.1007/s10874-007-9060-3,
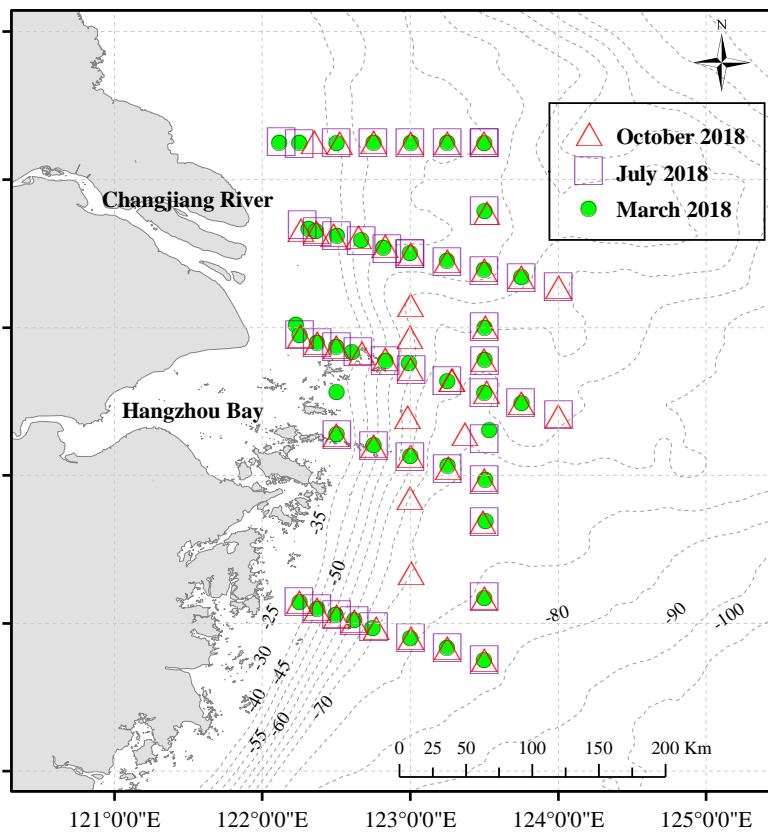455 2007.
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
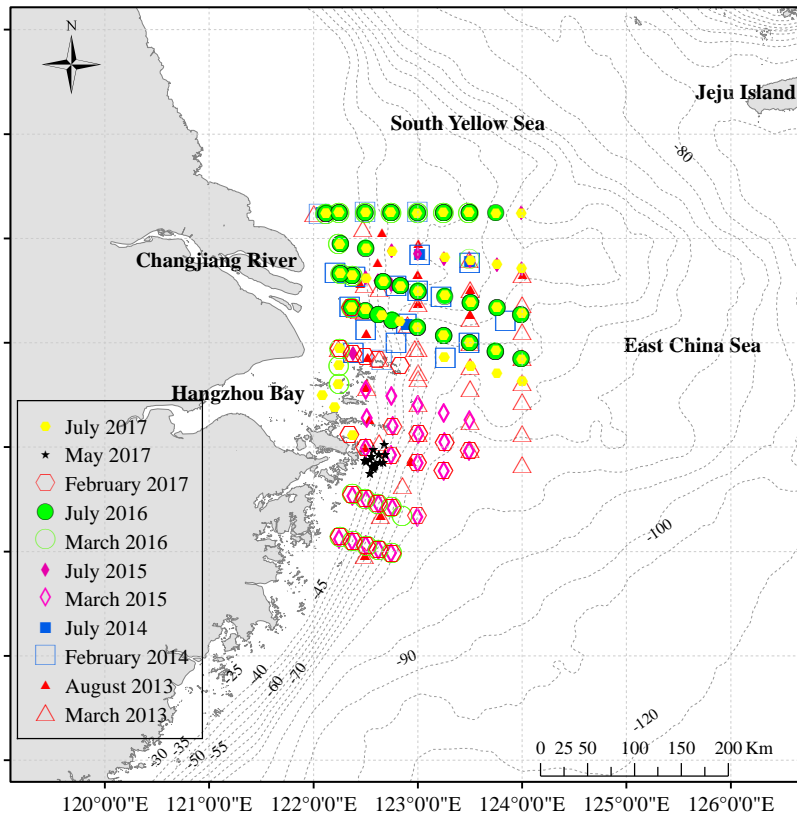
Figure 1: Sampling stations ~~from~~ during 11 cruises (the confirmatory dataset) from 2013 to 2017 on the East China Sea shelf.


Figure 2: Sampling stations for three cruises (the exploratory dataset) used to extend the utility of the ANN model. The green circles represent March 2018, the purple squares represent July 2018, the red triangles represent October 2018.

495
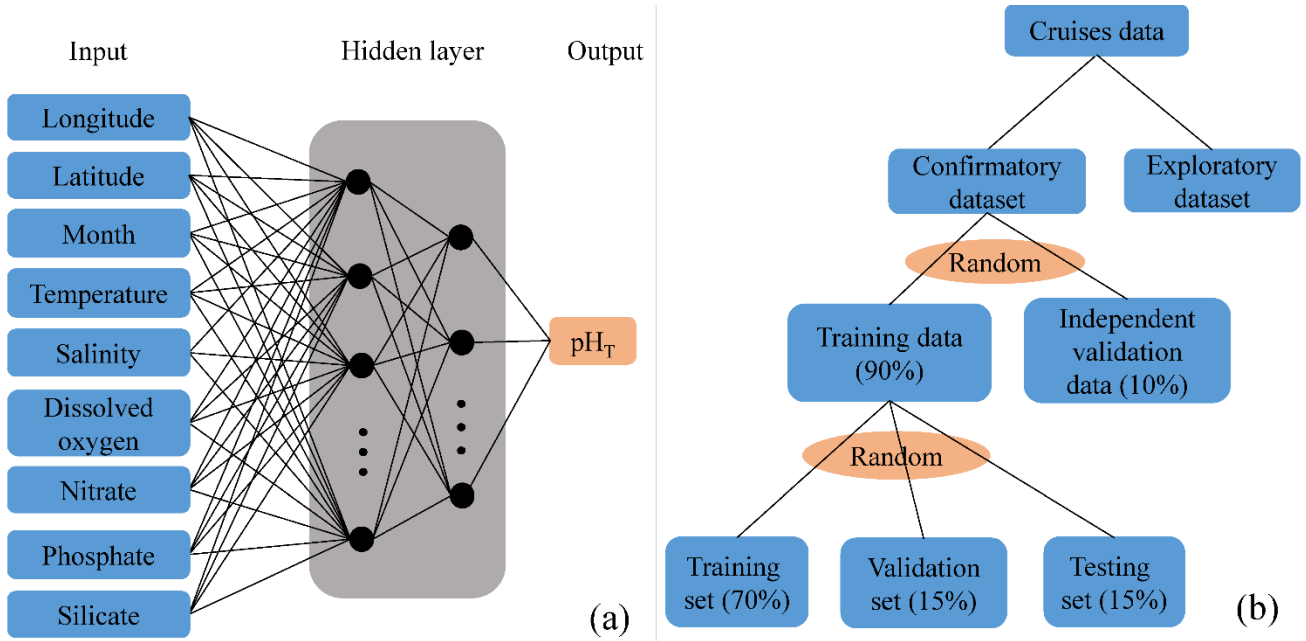
Figure 3: Schematic representation of the neural network algorithm to retrieve pH$_T$. (a)-the architecture of the ANN model. Input variables are observed temperature, salinity, dissolved oxygen, nitrate, phosphate, and silicate together with the geolocation (longitude and latitude) and time (month) of sampling; (b)-data distribution diagram for training and prediction.

499
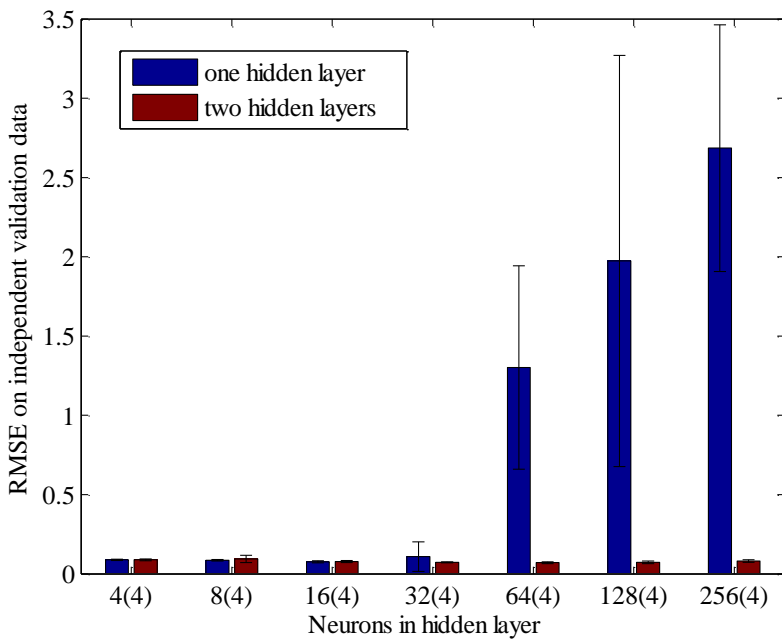
Figure 4: Comparison of the performance of one hidden layer vs. two hidden layers in predicting independent validation data. The number of neurons in the first hidden layer was the same in the one hidden layer vs. two hidden layers model, numbers in parentheses show the number of neurons in the second hidden layer (for the two hidden layers model). Bars show the mean and standard deviation of the Root-Mean-Square-Error over a ten-fold cross-validation, for different numbers of neurons in the first hidden layer.

504

505
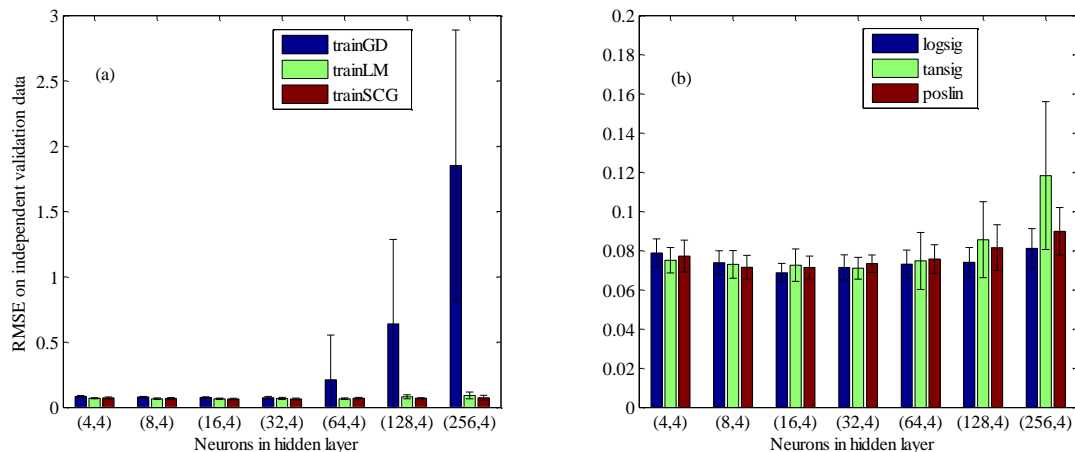
506 **Figure 5: Comparison of the performance of different training functions and transfer functions on independent validation data. (a)-**
507 **three training functions: Gradient descent backpropagation (trainGD), Levenberg-Marquardt backpropagation (trainLM), and**
508 **Scaled conjugate gradient backpropagation (trainSCG); (b) three transfer functions: Log-sigmoid transfer function (logsig),**
509 **Hyperbolic tangent sigmoid transfer function (tansig), and Positive linear transfer function (poslin). Bars show the mean and**
510 **standard deviation of the Root-Mean-Square-Error over a ten-fold cross-validation, for different numbers of neurons in the first**
511 **hidden layer.**
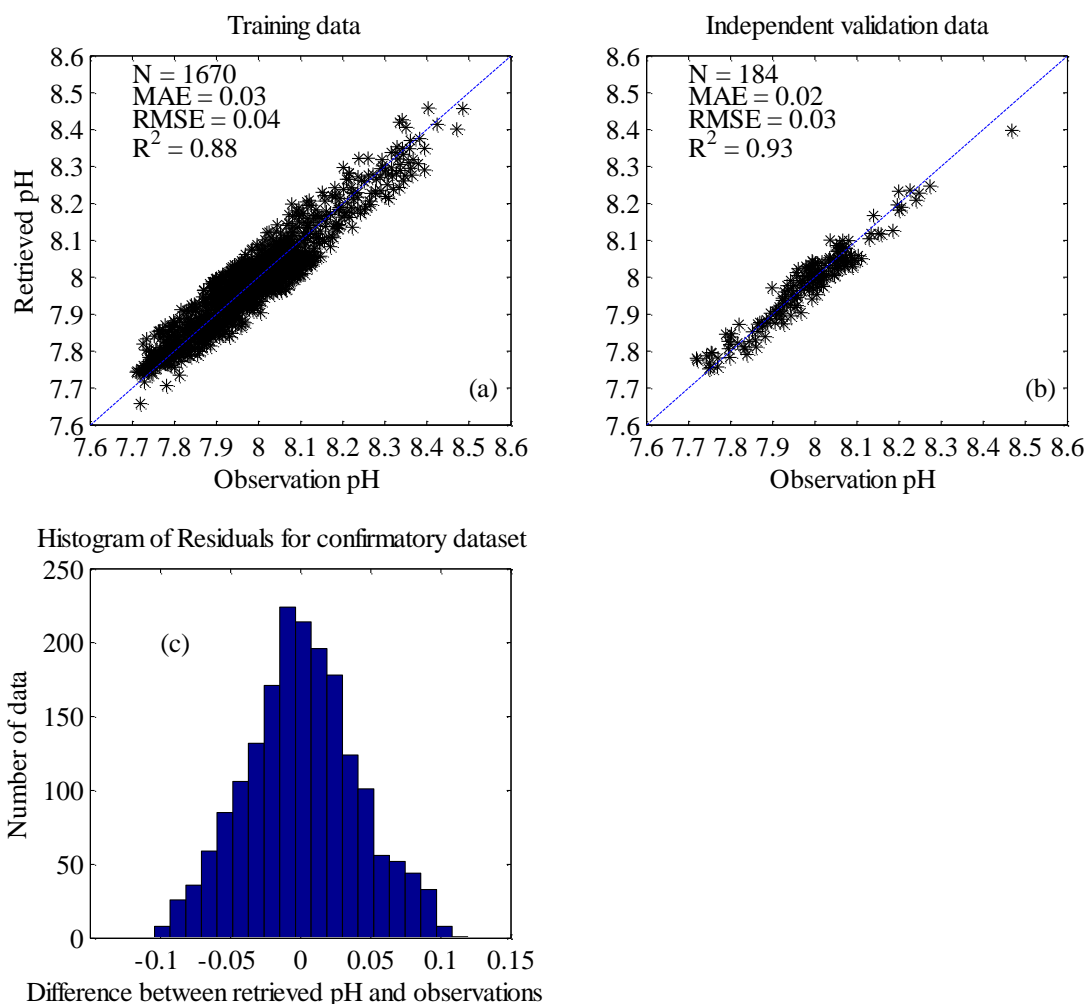
512 **Figure 6: Comparison of pH$_T$ retrieved by the ANN model with corresponding observations. (a)-Training data (90% of confirmatory**
513 **dataset); (b)-Independent validation data (10% of confirmatory dataset); (c)-Histogram of residuals for confirmatory dataset. The**
514 **1:1 line is shown in each plot as visual reference. Three statistics are the mean absolute error (MAE), the coefficient of determination**
515 **(R^2), and the root mean squared error (RMSE). N represents the number of data points.**

516

38
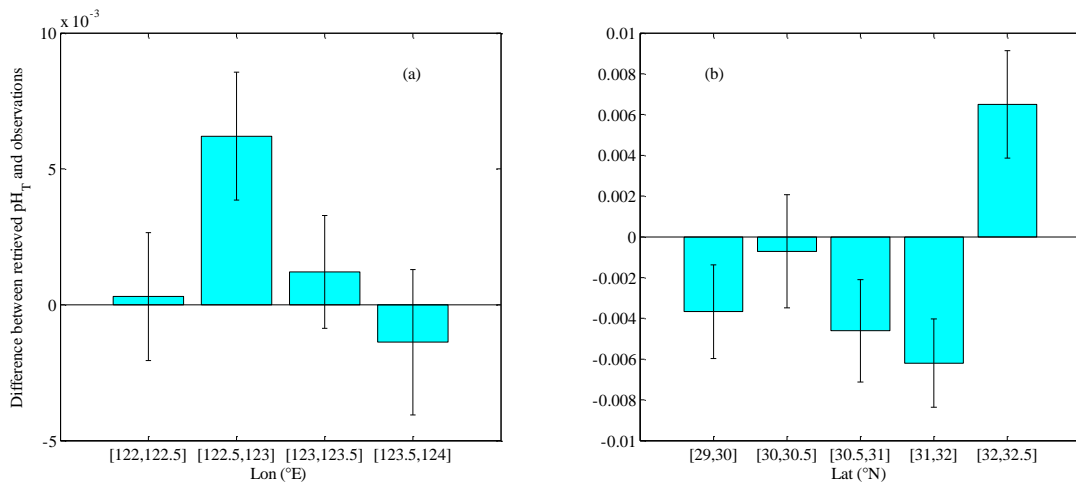
517 **Figure 7: Box plots of the differences between retrieved pH$_T$ minus the observations. (a)-the differences vs longitude (Mean±SE);**
518 **(b)-the differences vs latitude (Mean±SE). The height of each box represents the mean value of the differences, the whisker represents**
519 **the standard error (SE) value of the differences.**
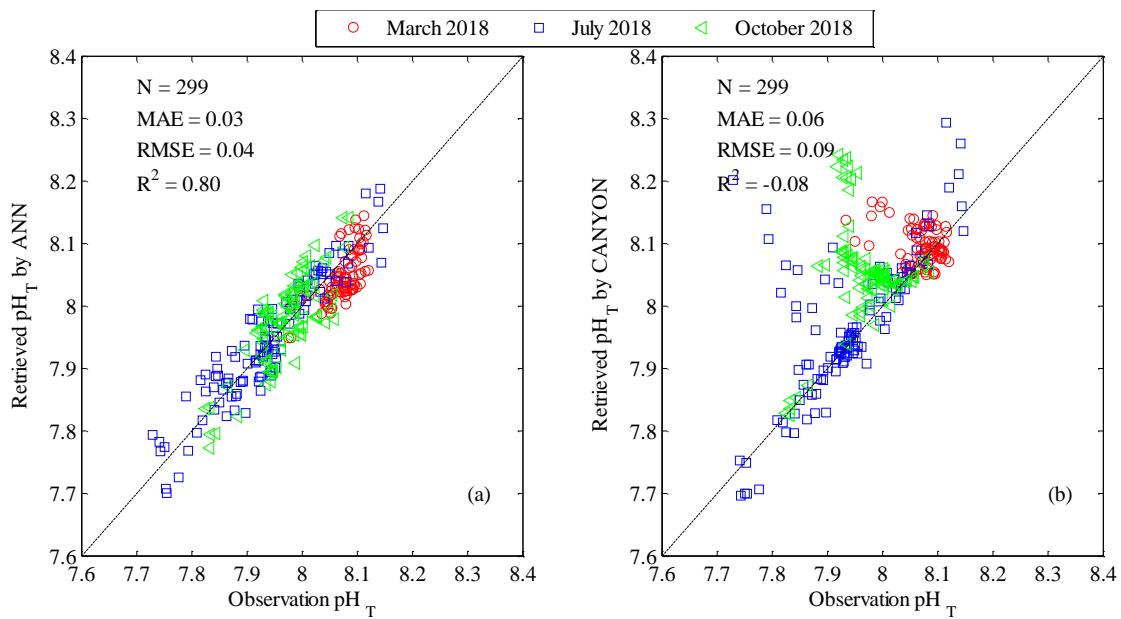
520 **Figure 8: Comparison of retrieved pH$_T$ with corresponding observations for exploratory dataset. (a)-pH$_T$ retrieved by the ANN**
521 **model vs observations; (b)-pH$_T$ retrieved by CANYON (Sauzède et al., 2017) vs observations.** The red circles represent March 2018,
522 **the blue squares represent July 2018, the green triangles represent October 2018. The 1:1 line is shown in the plot as visual reference.**
523 **Three statistics approaches used are the mean absolute error (MAE), the coefficient of determination (R^2), and the root mean**
524 **squared error (RMSE). N represents the number of data points.**

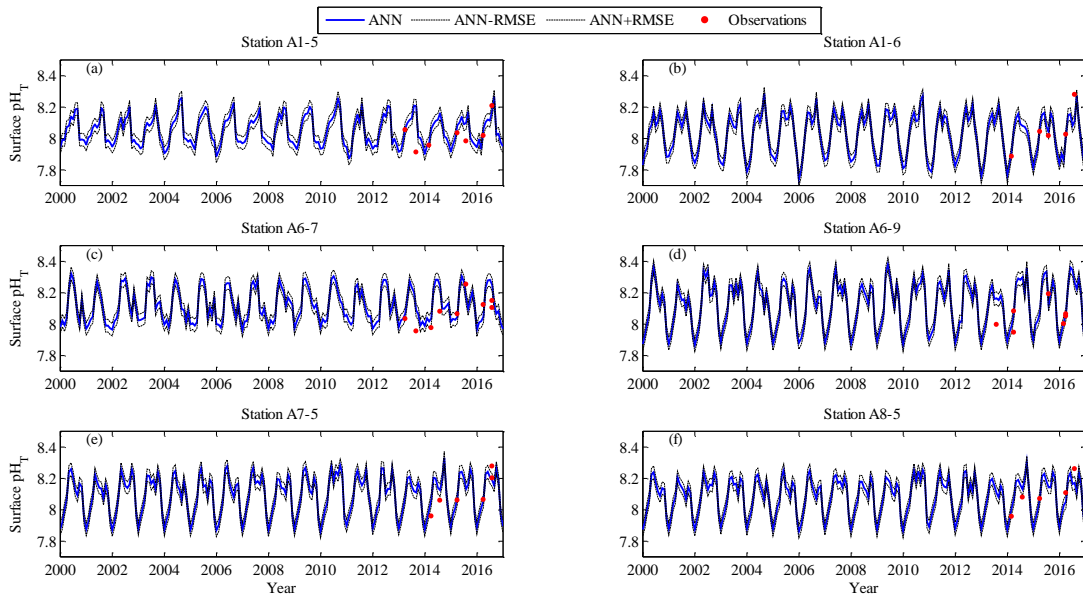525

526
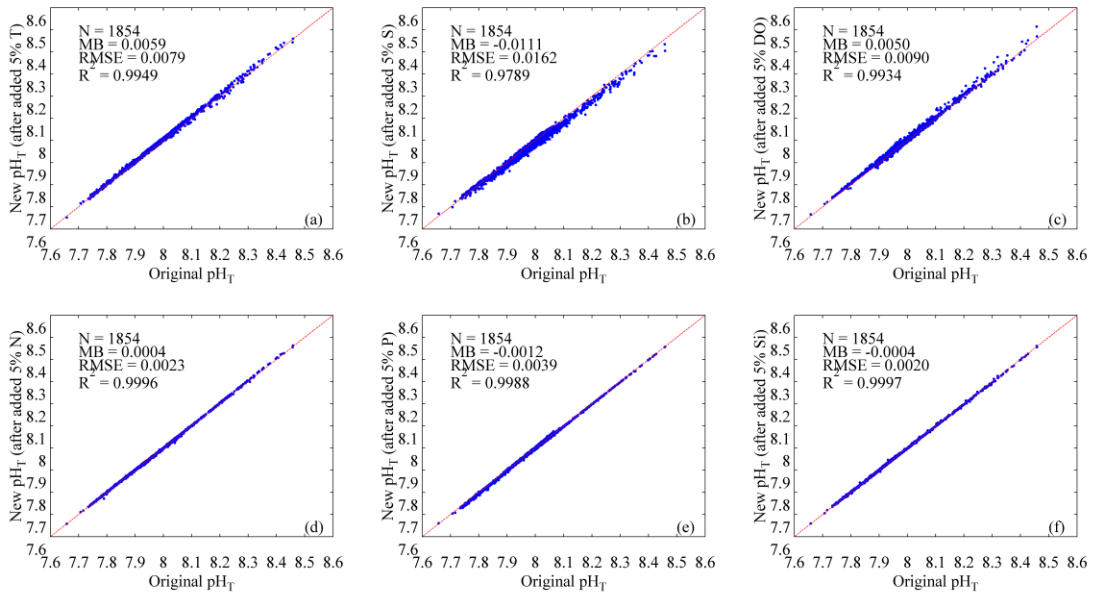
Figure 9: Sensitivity of the ANN model for environmental input variables. (a)-temperature (T); (b) salinity (S); (c)-dissolved oxygen (DO); (d)-nitrate (N); (e)-phosphate (P); (f)-silicate (Si). Three statistics approaches used are the mean bias (MB), the root mean squared error (RMSE), and the coefficient of determination (R^2). N represents the number of data points.

Figure 10: Comparison of surface pH_T retrieved by the ANN model using Changjiang Biology FVCOM output with corresponding observations at six sites repeated sampling for 3 to 4 years. Red dots represent observations pH_T, blue solid line represents retrieved pH_T, black dotted lines represent upper and lower bounds of the ANN model accuracy (ANN ± RMSE) ~~retrieved pH_T ± RMSE~~. (a)-station A1-5; (b)-station A1-6; (c)-station A6-7; (d)-station A6-9; (e)-station A7-5; (f)-station A8-5.
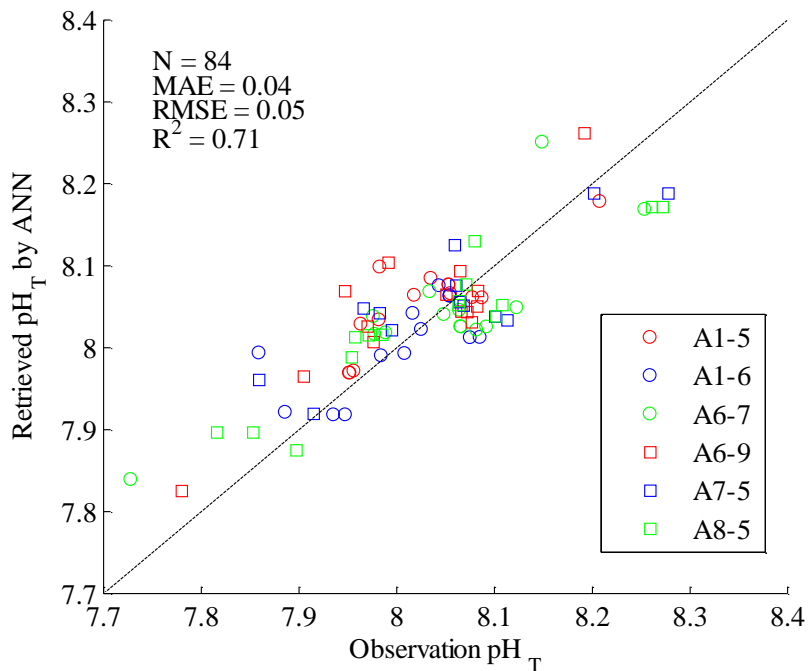
Figure 11: Comparison of water column pH_T retrieved by the ANN model using Changjiang Biology FVCOM output with corresponding observations at six sites repeated sampling for 3 to 4 years. The 1:1 line is shown in the plot as a visual reference. Skill statistics include the mean absolute error (MAE), the coefficient of determination (R^2), and the root mean squared error (RMSE). N represents the number of data points.
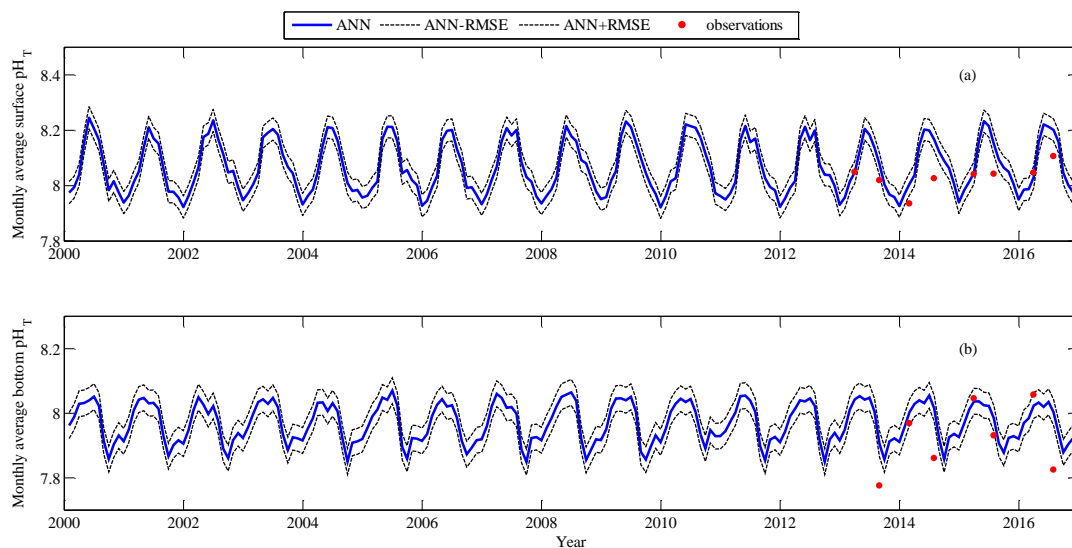
Figure 12: Comparison of monthly average pH_T on the East China Sea shelf. Blue solid line represents retrieved pH_T by the ANN model using Changjiang Biology FVCOM output; black dotted lines represent upper and lower bounds of the ANN model accuracy (ANN ± RMSE) retrieved pH_T ± RMSE; red points show monthly-average pH_T observations from 2013 to 2016. (a)-surface; (b)-bottom.
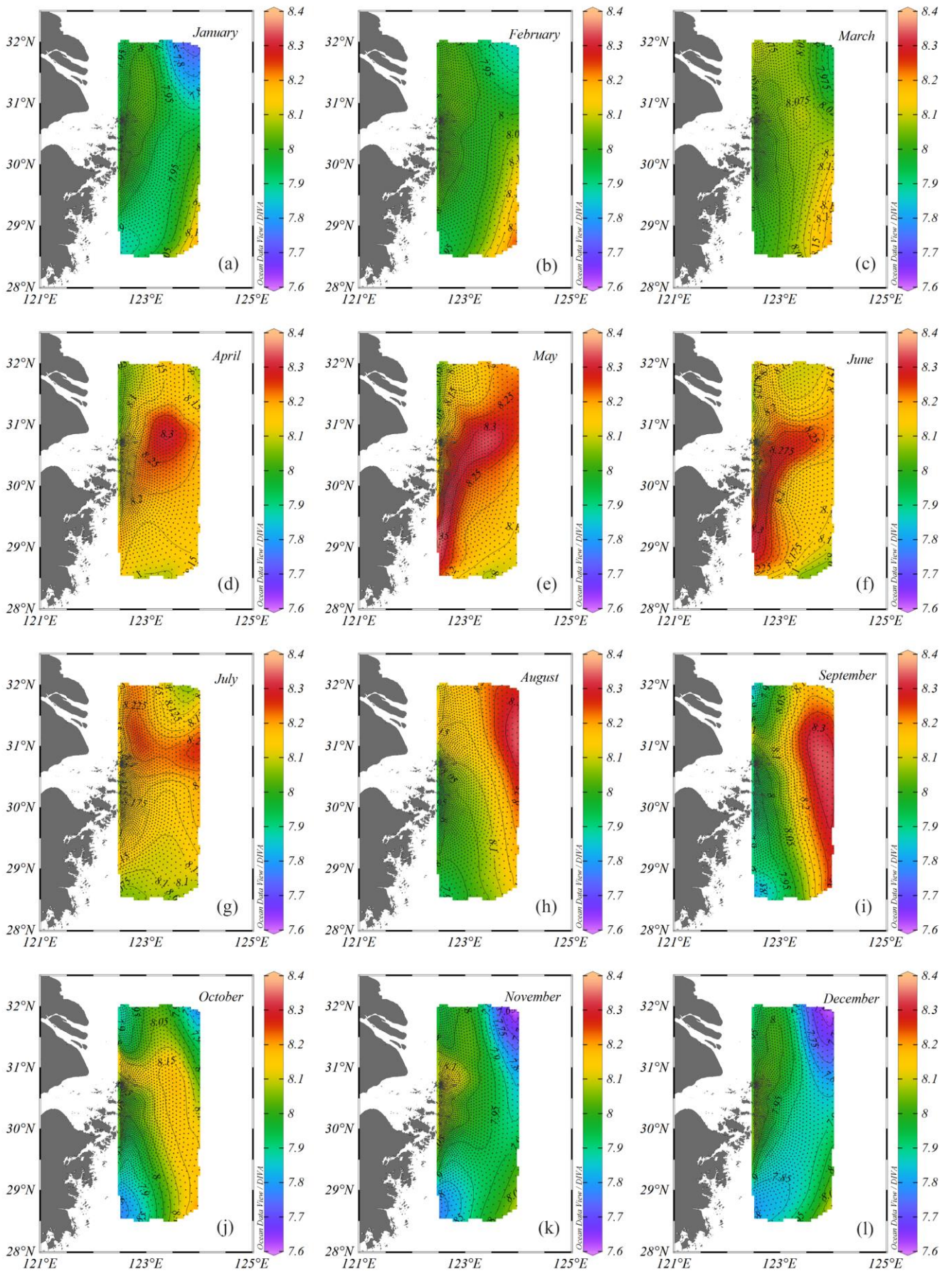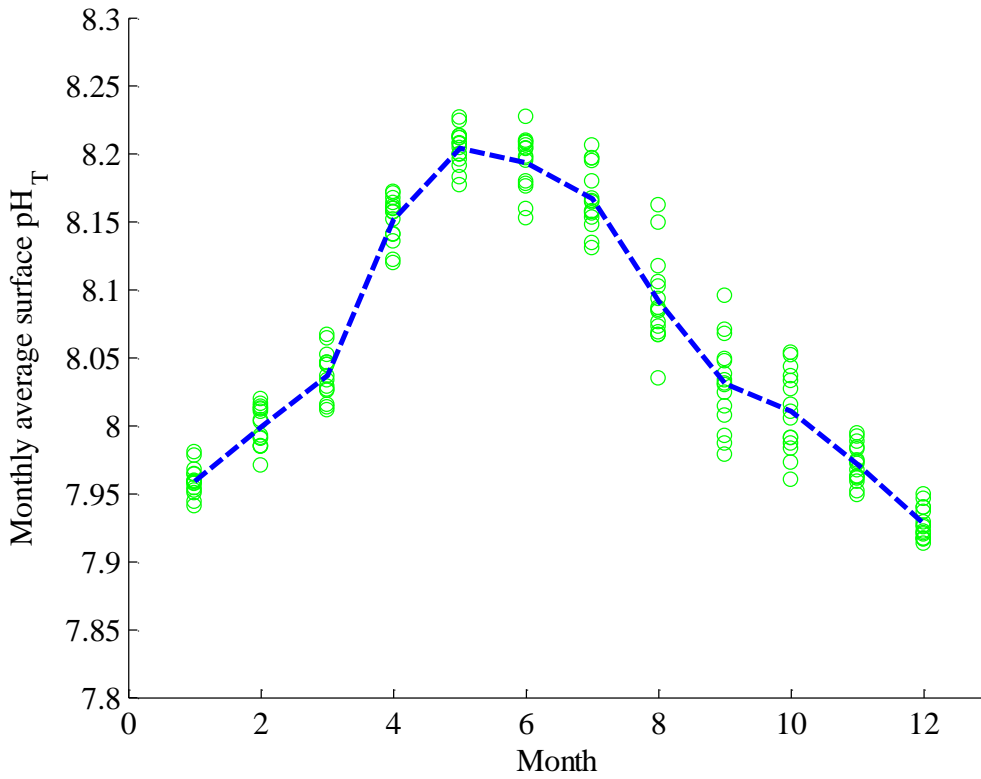
Figure 13: Spatial distribution of monthly average surface pH$_T$ retrieved by the ANN model using Changjiang Biology FVCOM output. (a)-January; (b)-February; (c)-March; (d)-April; (e)-May; (f)-June; (g)-July; (h)-August; (i)-September; (j)-October; (k)-November; (l)-December.

Figure 14: Seasonal cycles of surface pHT on the East China Sea shelf from 2000-2016. The green circles represent monthly regional
average, the blue dashed represents mean value of each month.

558

559

560 **Table 1: Field survey information and measurements of water temperature, salinity, dissolved oxygen, nitrate, phosphate, silicate**
561 **and pHT (Mean±SE).**

Sampling period	Temperature (℃)	Salinity	Dissolved oxygen (mmol m^{-3})	Nitrate (mmol m^{-3})	Phosphate (mmol m^{-3})	Silicate (mmol m^{-3})	pHT
March 4th-20th, 2013	11.54±1.34	32.04±2.26	275.28±19.30	12.25±8.25	0.58±0.17	17.54±7.65	8.19±0.04
August 17th-28th, 2013	23.45±3.17	32.32±2.91	142.22±63.45	12.16±8.05	0.55±0.32	16.47±12.18	8.04±0.18
February 21th-28th, 2014	9.56±2.38	32.14±1.78	293.07±19.52	11.92±9.17	0.59±0.18	12.52±6.50	8.10±0.04
July 10th-17th, 2014	21.66±2.13	29.50±5.10	186.44±43.29	21.57±22.10	0.57±0.46	21.45±17.76	8.07±0.11
March 11th-21th, 2015	11.42±1.44	31.57±2.60	279.72±15.29	22.04±18.88	0.81±0.35	16.48±11.64	8.19±0.03
July 9th-20th, 2015	22.14±1.55	29.73±4.71	207.32±56.12	19.73±18.62	0.60±0.42	20.87±17.48	8.13±0.09
March 7th-19th, 2016	10.77±2.02	30.85±2.92	284.00±31.40	20.26±12.80	0.82±0.25	19.17±11.62	8.20±0.05
July 4th-28th, 2016	23.19±3.19	28.17±6.67	122.90±49.97	25.77±23.60	0.63±0.46	28.56±25.03	8.06±0.16
February 15th-28th, 2017	11.03±2.57	32.00±2.43	296.21±21.27	12.30±9.13	0.56±0.18	13.09±7.45	8.13±0.05
May 12th-24th, 2017	17.71±1.54	29.62±2.79	171.58±49.52	12.60±4.83	0.29±0.24	10.95±4.29	8.08±0.13
July 20th-30th, 2017	24.85±3.41	27.70±6.31	192.11±76.55	20.57±23.23	0.42±0.34	19.28±18.92	8.09±0.18

562

563

564

565

566

567

568

569

570

Table 2: Different model structures and their performance in the training step. The variables (Lon (longitude), Lat (latitude), Month (month), T (temperature), S (salinity), DO (dissolved oxygen), N (nitrate), P (phosphate), Si (silicate)) marked with 1 represent the input variables. Skill statistics include the coefficient of determination (R^2), the root mean squared error (RMSE), and the mean absolute error (MAE).

Model	Lon	Lat	Month	T	S	DO	N	P	Si	Training data			Independent validation data		
										R^2	RMSE	MAE	R^2	RMSE	MAE
1						1				0.40	0.092	0.068	0.47	0.076	0.058
2					1	1				0.62	0.073	0.053	0.62	0.067	0.051
3				1	1	1				0.69	0.065	0.048	0.72	0.060	0.044
4				1	1	1	1			0.76	0.057	0.044	0.77	0.052	0.041
5				1	1	1		1		0.81	0.051	0.040	0.79	0.051	0.040
6				1	1	1			1	0.77	0.056	0.044	0.79	0.054	0.043
7				1	1	1	1	1		0.80	0.053	0.042	0.79	0.051	0.041
8				1	1	1		1	1	0.81	0.051	0.040	0.81	0.049	0.039
9				1	1	1	1		1	0.76	0.058	0.044	0.77	0.054	0.044
10				1	1	1	1	1	1	0.83	0.048	0.037	0.86	0.046	0.037
11			1	1	1	1	1	1		0.85	0.046	0.035	0.87	0.043	0.032
12			1	1	1	1		1	1	0.85	0.046	0.034	0.85	0.045	0.035
13			1	1	1	1	1		1	0.82	0.049	0.036	0.84	0.050	0.036
14			1	1	1	1	1	1	1	0.84	0.046	0.035	0.87	0.045	0.033
15	1	1	1	1	1	1	1			0.86	0.044	0.033	0.79	0.046	0.034
16	1	1	1	1	1	1		1		0.87	0.043	0.032	0.87	0.044	0.034
17	1	1	1	1	1	1			1	0.87	0.043	0.033	0.82	0.045	0.035
18	1	1	1	1	1	1	1	1		0.88	0.040	0.031	0.88	0.039	0.031
19	1	1	1	1	1	1		1	1	0.87	0.042	0.032	0.87	0.042	0.033
20	1	1	1	1	1	1	1		1	0.84	0.046	0.035	0.85	0.047	0.036
21	1	1	1	1	1	1	1	1	1	0.88	0.040	0.031	0.93	0.033	0.024

Table 3: Model comparison between traditional empirical methods (MLR and MNR) and mechine-learning based empirical methods (Decision tree, Random Forest, and SVM). The statistics was derived from confimatory dataset (training data independent validation data) using input variables: T, S, DO, N, P, and Si. Note R^2 statistics in our study was based on the calculation of coefficient of determination, therefore negative R^2 could be derived if there were strong bias.

| Model	Kernel Function	Input variables	RMSE	R^2	MAE
MLR	-	T, S, DO, N, P, Si	0.078	0.56	0.062
MNR	-	T, S, DO, N, P, Si	0.060	0.74	0.047
Decision Tree	Simple Tree	T, S, DO, N, P, Si	0.064	0.71	0.047
	Medium Tree	T, S, DO, N, P, Si	0.060	0.74	0.044
	Complex Tree	T, S, DO, N, P, Si	0.061	0.73	0.043
Random Forest	Boosted Trees	T, S, DO, N, P, Si	0.340	-7.51	0.339
	Bagged Trees	T, S, DO, N, P, Si	0.056	0.77	0.04
SVM	Linear	T, S, DO, N, P, Si	0.079	0.55	0.061
	Quadratic	T, S, DO, N, P, Si	0.061	0.73	0.046
	Cubic	T, S, DO, N, P, Si	0.060	0.74	0.043
	Fine Gaussian	T, S, DO, N, P, Si	0.064	0.70	0.042
	Medium Gaussian	T, S, DO, N, P, Si	0.054	0.79	0.041
	Coarse Gaussian	T, S, DO, N, P, Si	0.069	0.65	0.054

580

Supplementary material

Table S1: The performance of different number of neurons for two hidden layers in the training step. Three statistics are the coefficient of determination (R^2), the root mean squared error (RMSE), and the mean absolute error (MAE).

| Model	Number of neurons		Training data			Independent validation data		
	first hidden	second hidden	R^2	RMSE	MAE	R^2	RMSE	MAE
1	4	4	0.68	0.071	0.054	0.67	0.072	0.057
2	8	4	0.70	0.070	0.050	0.67	0.069	0.050
3	16	4	0.76	0.062	0.045	0.76	0.062	0.045
4	32	4	0.74	0.063	0.046	0.79	0.062	0.048
5	40	4	0.76	0.062	0.044	0.76	0.061	0.045
6	64	4	0.79	0.058	0.041	0.78	0.056	0.043
7	128	4	0.76	0.062	0.045	0.74	0.062	0.044
8	8	8	0.73	0.065	0.047	0.73	0.065	0.048
9	16	8	0.78	0.059	0.042	0.78	0.058	0.044
10	32	8	0.78	0.059	0.042	0.83	0.053	0.039
11	40	8	0.79	0.059	0.042	0.77	0.055	0.040
12	64	8	0.77	0.061	0.044	0.76	0.059	0.042
13	128	8	0.77	0.060	0.042	0.79	0.059	0.043
14	16	16	0.79	0.057	0.041	0.85	0.054	0.041
15	32	16	0.80	0.057	0.040	0.69	0.059	0.043
16	40	16	0.82	0.054	0.039	0.81	0.053	0.039
17	64	16	0.79	0.059	0.041	0.76	0.057	0.040
18	128	16	0.79	0.058	0.040	0.78	0.059	0.043
19	32	32	0.78	0.059	0.042	0.75	0.058	0.039
20	40	32	0.79	0.058	0.041	0.79	0.055	0.040
21	64	32	0.78	0.059	0.042	0.83	0.052	0.040
22	128	32	0.79	0.058	0.041	0.79	0.056	0.041
23	40	40	0.77	0.060	0.043	0.77	0.060	0.044
24	64	40	0.79	0.058	0.042	0.75	0.060	0.043
25	128	40	0.80	0.057	0.040	0.78	0.057	0.042
26	64	64	0.78	0.060	0.042	0.78	0.057	0.040
27	128	64	0.72	0.068	0.050	0.65	0.067	0.048
28	128	128	0.72	0.067	0.049	0.65	0.072	0.051

584

585
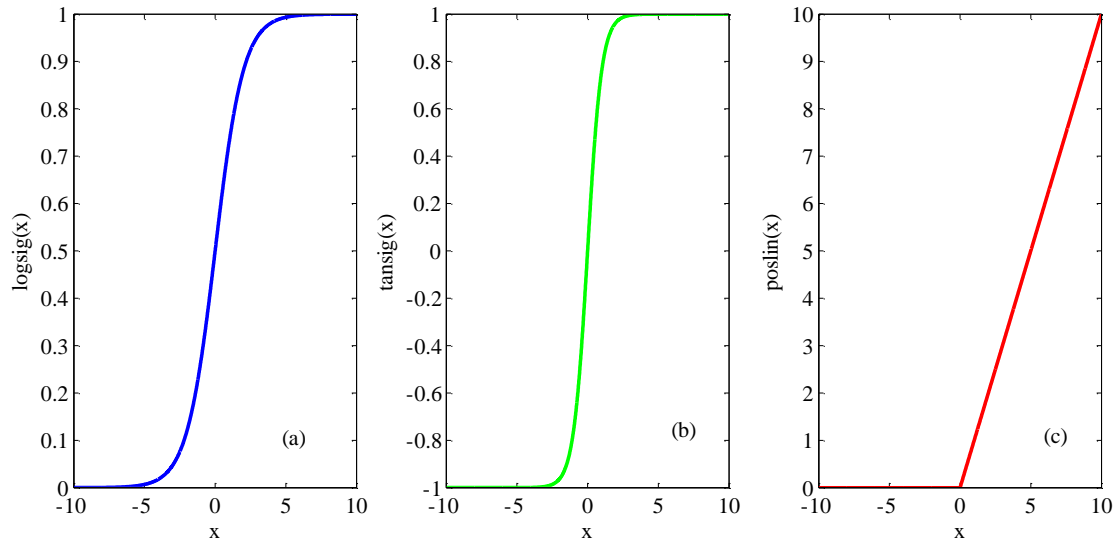
586 **Figure S1: Comparison of three transfer functions. (a)-Log-sigmoid transfer function (logsig); (b) Hyperbolic tangent sigmoid**
587 **transfer function (tansig); (c)-Positive linear transfer function (poslin).**

588
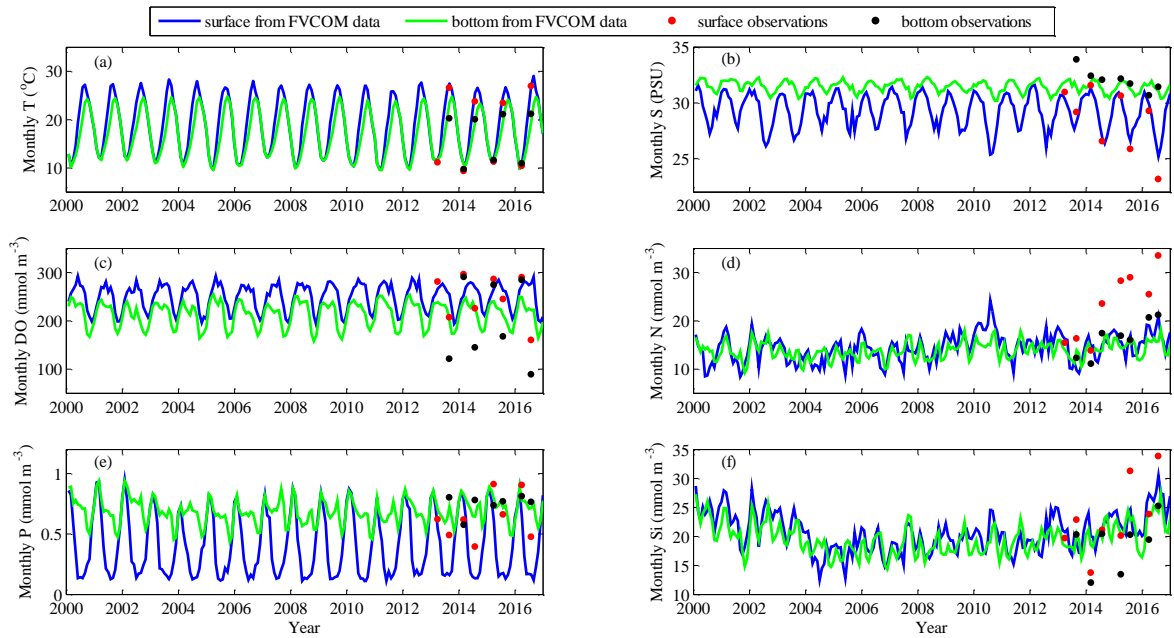
589 **Figure S2: Comparison of monthly-average environmental variables from the Changjiang Biology FVCOM with the corresponding**
590 **observations at the surface and bottom on the East China Sea shelf. Blue and green solid lines represent surface and bottom**
591 **simulated data from the Changjiang Biology FVCOM, respectively; red and black points show surface and bottom observation data**
592 **from 2013 to 2016, respectively. (a)-temperature; (b)-salinity; (c)-dissolved oxygen; (d)-nitrate; (e)-phosphate; (f)-silicate.**
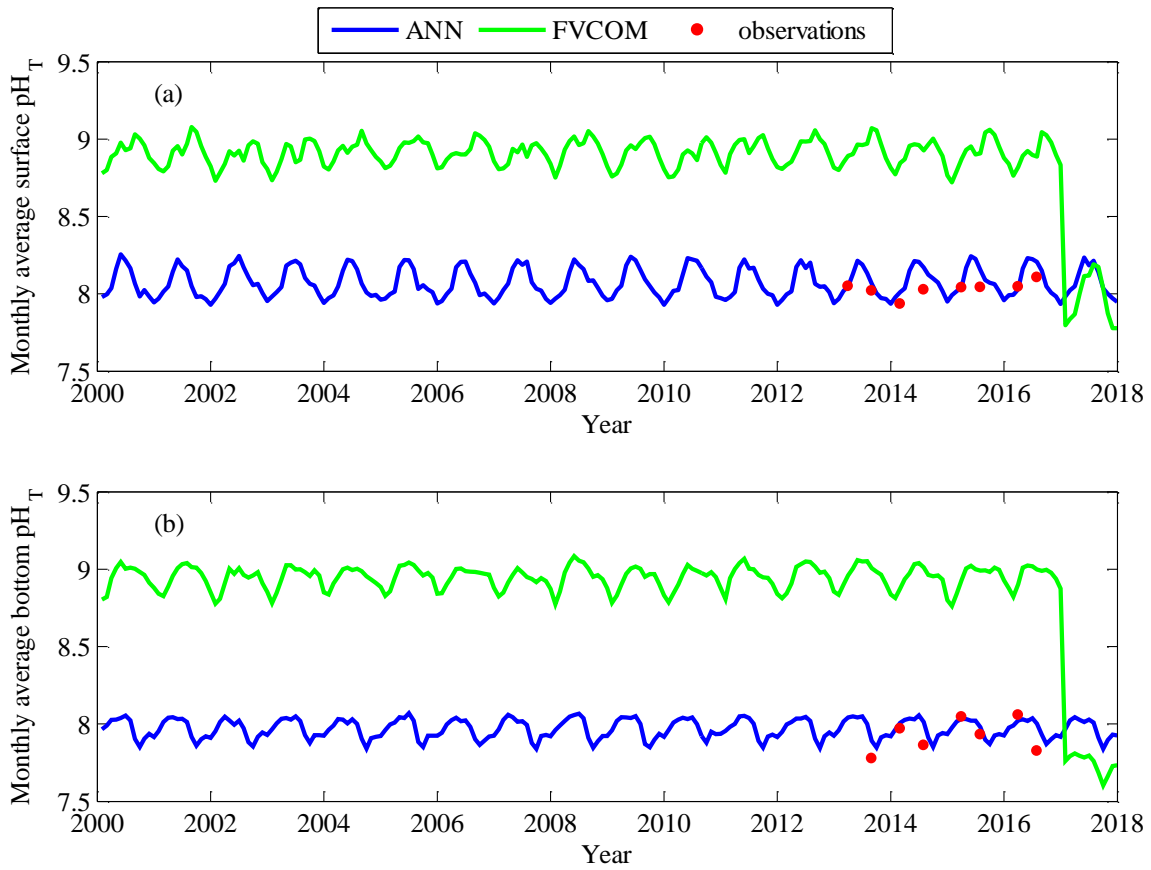
593   **Figure S3: Comparison of monthly average pH$_T$ on the East China Sea shelf. Blue solid line represents retrieved pH$_T$ by the ANN**
594   **model using Changjiang Biology FVCOM output; green solid line represents simulated pH$_T$ from the Changjiang Biology FVCOM;**
595   **red points show monthly average pH$_T$ observations from 2013-2016. (a)-surface; (b)-bottom.**
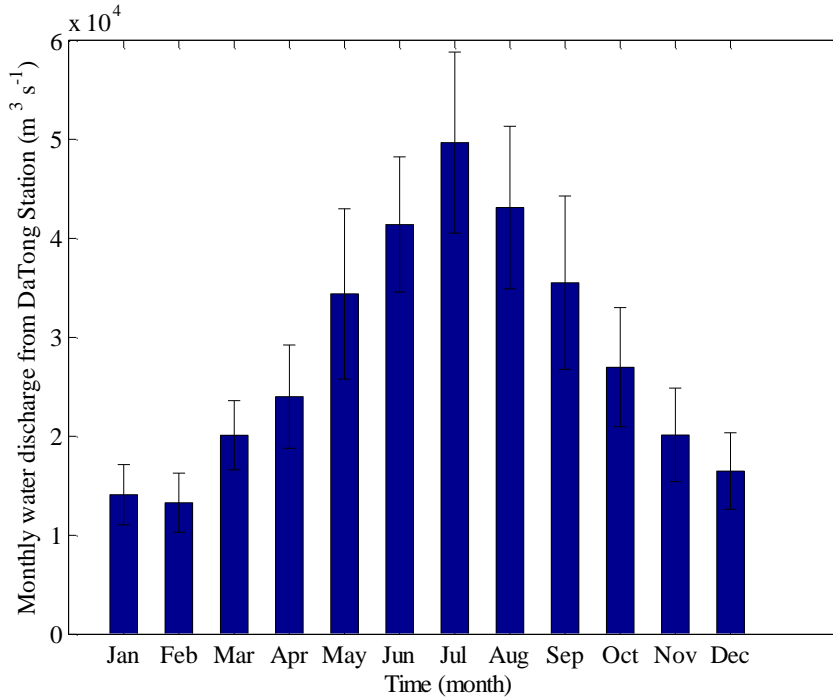
596   **Figure S4: Monthly average water discharge and its standard deviation (DaTong Station, data derived from the Hydrological**
597   **Information Center of China, http://www.hydroinfo.gov.cn/).**

598

599

600

47
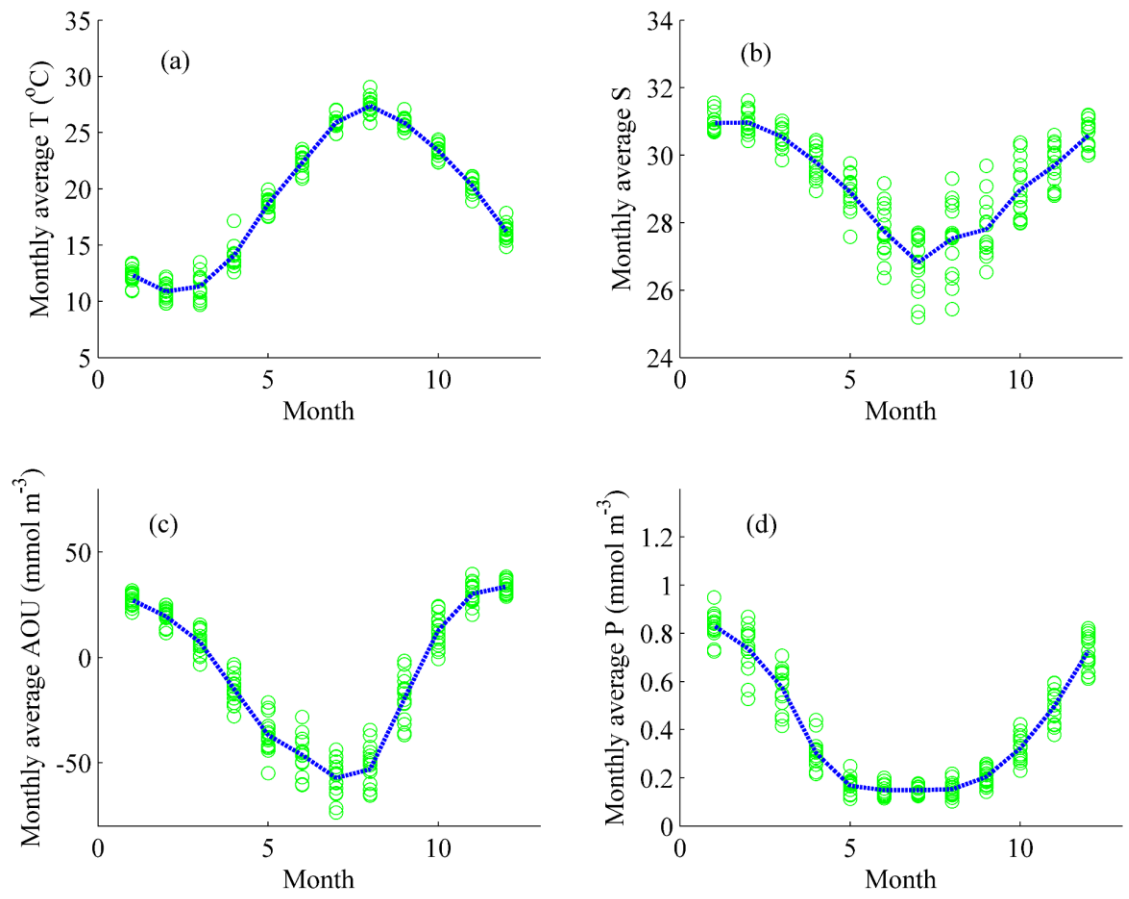
601 **Figure S5: Seasonal cycles of surface T (a), S (b), AOU (c), and P (d) from Changjiang Biology FVCOM output on the East China**
602 **Sea shelf from 2000-2016. The green circles represent monthly regional average, the blue dashed represents mean value of each**
603 **month.**