Scientific Research Publishing

Open Journal of Statistics, 2016, 6, 789-804
http://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Improved the Prediction of Multiple Linear Regression Model Performance Using the Hybrid Approach: A Case Study of Chlorophyll-a at the Offshore Kuala Terengganu, Terengganu

Muhamad Safiih Lola[1], Mohd Noor Afiq Ramlee[1], G. Sugan Gunalan[1], Nurul Hila Zainuddin[1], Razak Zakariya[2], MdSuffian Idris[2], Idham Khalil[2]

[1]School of Informatics and Applied Mathematics, University Malaysia Terengganu, Kuala Terengganu, Malaysia
[2]School of Science Marin, University Malaysia Terengganu, Kuala Terengganu, Malaysia
Email: safiihmd@umt.edu.my, afiqramlee91@gmail.com, sugangunalan@yahoo.com, Hila.zainuddin@gmail, ajak@umt.edu.my, suffian@umt.edu.my, idham@umt.edu.my

How to cite this paper: Lola, M.S., Ramlee, M.N.A., Gunalan, G.S., Zainuddin, N.H., Zakariya, R., Idris, M. and Khalil, I. (2016) Improved the Prediction of Multiple Linear Regression Model Performance Using the Hybrid Approach: A Case Study of Chlorophyll-a at the Offshore Kuala Terengganu, Terengganu. *Open Journal of Statistics*, **6**, 789-804.
http://dx.doi.org/10.4236/ojs.2016.65065

Received: July 22, 2016
Accepted: October 14, 2016
Published: October 18, 2016

Abstract

Efficiency and precision in prediction of Chlorophyll-a using this model is still a pandemic among researchers, due to the natural conditions in ocean water systems itself, which involved chemical, biological and physical processes and interaction among them may affect the model performance drastically. Thus, to overcome this problem as well as to improve the strength of MLR, we proposed a hybrid approach, *i.e.*, an Artificial Neural Network to the MLR coins as Artificial Neural Network-Multiple Linear Regression (ANN-MLR). To investigate the performance of the proposed model, we compared Multiple Linear Regression (MLR), Artificial Neural Network (ANN) and proposed hybrid Artificial Neural Network and Multiple Linear Regression (ANN-MLR) in the prediction of chlorophyll-a (chl-a) concentration by statistical measurement which are MSE and MAE. Achieving our objectives of study, we used 4 parameters, *i.e.* temperature (˚C), pH, salinity (ppt), DO (ppm) at the Offshore Kuala Terengganu, Terengganu, Malaysia. The results showed that our proposed model can improve the performance of the model as compared to ANN and MLR due to small errors generated, error reduced, and increased the correlation coefficient for all parameters in both MSE and MAE, respectively. Thus, this result indicated that our proposed model is efficient, precise and almost perfect correlation as compared to ANN and MLR.

Keywords

Multi Linear Regression, Artificial Neural Network, ANN-MLR, Chlorophyll-a, Correlation

1. Introduction

Chlorophyll (chl) is a bio-indicator of the aquatic ecosystem, as it is common to all photosynthesis organisms. It's also widely used for estimation of phytoplankton in ecological studies and water quality [1]-[5]. The chlorophyll concentration is commonly used in satellite ocean colour products [6]-[10]. Chlorophyll can be found in algae, plants and phytoplankton. This molecule used as photoreceptors in photosynthesis. Chlorophyll appears to be green in plants and algae because it reflects the green wavelengths found in sunlight, while absorbing all other colours. There are 6 different chlorophylls (A, B, C, D, E, F) where each of them reflects different ranges of green wavelengths. In every single photosynthesis organism chlorophyll-a can be found, from algae to land plants and cyanobacteria. Phytoplankton has chlorophyll-a, where a chlorophyll sensor used to detect these organisms *in-situ*. It also provides immediate data and can be used for long-term recording and monitoring. However, as a chlorophyll sensor assumes that all cyanobacteria and algae have the same levels of chlorophyll-a, it can provide a rough estimate of biomass but cannot be used to identify specific species. Even with this limitation, *in-situ* chlorophyll measurements are recommended in Standard Methods for Examination of Water and Wastewater to estimate algal populations. Chlorophyll sensors used to determine in-situ method for the trophic state of an aquatic system.

The phytoplankton in the ocean water system involved chemical, biological and physical processes as well as interactions among the processes. In order to predict and study the strength of the dependent and independent variables among them, the Multiple Linear Regression (MLR) method is used. The Multiple Linear Regression (MLR) method is commonly used techniques to obtain a linear input output model for a given dataset [11] [12]. However, this model will face some difficulties, especially when the independent variables are following certain distribution. Thus, Artificial Neural Network (ANN) was adopted as an approach to extracting information, required no priori assumptions about the model in terms of mathematical relationships or distribution data and it is a well suited method with self-adaptive, self-organizing and error tolerance [13] [14].

2. Materials and Methods

2.1. Multiple Linear Regression (MLR)

Multiple Linear Regression expresses the relation between dependent variable y and more independent variable $\left(x_1, x_2, \cdots, x_p\right)$. Linear regression simply has one dependent variable which varies with one independent variable. However, when we need to explain about the dependent variable with two or more independent variables we need to use multiple linear regression. The multiple linear regression model as in Equation (1) is as follow:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \tag{1}$$

where, β is the coefficient of parameters, y and x are dependent and independent

variables respectively, while ε *i.e.* error term.

2.2. Artificial Neural Network (ANN)

Artificial Neural Networks is a field of Artificial Intelligent (AI) where we, by inspiration from the human brain, find data structures and algorithms for learning and classification of data. Many tasks that humans perform naturally fast, such as the recognition of familiar faces, proves to be very complicated task for a computer when conventional programming methods are used. By applying Neural Network techniques a program can learn by examples, and create an internal structure of rules to classify different inputs, such as recognizing images.

The hidden layer back propagation network is the most widely used models for modelling, forecasting and classification. This is model is characterized by a series of three-layer processing unit which connected with acyclic links. The relation between output (y) and input $\left(x_1, x_2, \cdots, x_p\right)$ which can be classified into this mathematical Equation (2) [15] is as follows:

$$y = w_0 + \sum_{j=1}^{q} w_j \cdot g\left(w_{0,j} + \sum_{i=1}^{p} w_{i,j} \cdot x_{t-i}\right) + \varepsilon_t \tag{2}$$

where, $w_{i,j}$ with $i = 0,1,2,\cdots,p$; $j = 1,2,\cdots,q$ is the parameter model as known as connection weight, p and q are number of inputs and number of hidden nodes, respectively, g is a sigmoid transfer function, w_0 and $w_{0,j}$ are weights of the arcs leaving from the bias terms, while ε_t is error.

The data which were inputted will pass through the input layer of the neural network. Then it will pass through hidden layer and exit through the output layer which was shown in Diagram 1. Each hidden layer and output layer node will collect the data from each node in there (either input layer or the hidden layer) and used as activation function as in Equation (3) [15]:

$$Sig\left(x\right) = \frac{1}{1+\exp\left(-x\right)} \tag{3}$$

The activation function can take many forms. The type of activation function is shown by the neurons in the network. Then, the ANN model as in Equation (4), do not have a linear mapping function from the past observation for the future y_t which is:

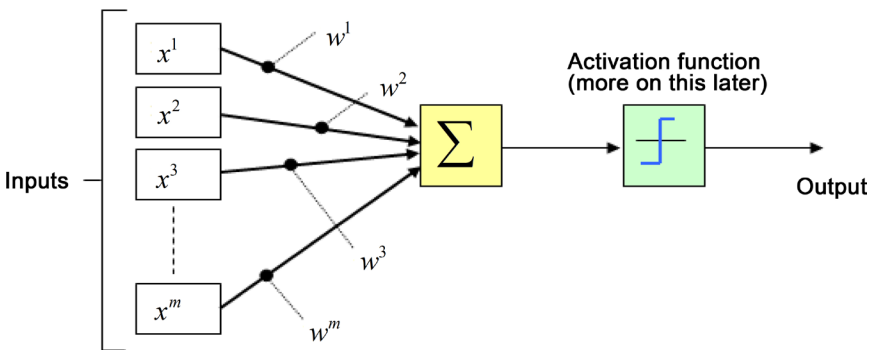

Diagram 1. The general design of neural network $\left(N^{p-q-1}\right)$.

$$y_t = f\left(y_{t-1}, \cdots, y_{t-p}, w\right) + \varepsilon_t \tag{4}$$

where, w is the vector for all parameters, $f(\cdot)$ and ε_t are the function for network structure and connection weights and error, respectively.

Therefore, it can be said that the neural network is equivalent to non-linear model. The simple network, which was stated in Equation (2) it is efficient until able to estimate the random function as the number of hidden nodes when q is large enough. In the research done, the structure of the network has small number of hidden nodes or slightly work will often to forecast beyond study sample due to the strong over fitting effect on the network which often can be found in the modelling process of neural network. The first has a good adaptability match to the sample used to build a model but it has the ability to make general weak conclusions against the outside of the sample data [15]. The q depends on the data and do not have fixed systematic rule for determining this parameter. In addition, to choose the suitable hidden layer nodes, another important task for ANN model to choose the number of missing observation, p, dimension of input vector.

There are many types of neural network such as multilayer [16], feed forward, and back propagation method. **Diagram 1** shows the artificial neurons. A neural network consisted of a system of interconnected nodes. The neural network is the combination of neurons to solve a certain problem. The first node layer (input layer) will bring the information to be processed in the neural network. The nodes in the layers are called neurons, because of the function which act as neurons. Each neuron consist of two parts. They are linear and nonlinear activation function.

The value of the input layer is distributed to both hidden layer nodes, where the addition and activation of the function are performed. The output value of the hidden layer of the input values from the output layer also performs the addition and activation function.

2.3. Hybrid Model (ANN-MLR)

In order to produce a more general model, linear hybrid model and more accurate non-linear model, the hybrid ANN-MLR model as in Equation (5) was introduced. In this model the time series is also considered as a function of linear and non-linear components as follows:

$$Y_t = L_t + N_t \tag{5}$$

where L_t and N_t are linear and non-linear components, respectively. In the first stage, the main purpose is to get the linear component which is the MLR model. Then the error in Equation (6) from the first stage contain non-linear relation whereas the linear model cannot be modelled and e_t is the error from time t in term of linear model:

$$e_t = Y_t - \hat{Y}_t \tag{6}$$

$\hat{Y}_t$ is the forecasted value of time t .The result of the predicable and linear modelling error is the result from the first stage and will be used in the next stage. Meanwhile, the linear trend augmented by the MLR model to be used in the second stage. In the second

stage, the main focus will be on non-linear model. The multi-layer perceptron is used to model the non-linear relation, the possibility of simultaneous linear model that still remain in linear model error and the linear and non-linear relationships in the original data. With n as input node, the ANN model of the error is:

$$e_t = f\left(e_{t-1}, e_{t-2}, \cdots, e_{t-n}\right) + \mu_t \tag{7}$$

where $f(\cdot)$ is non-linear function to determine neural network and μ_t is random error. When observe $\hat{Y}_t$ in Equation (6), the combine forecast model will be formed as:

$$YF_t = LF_t + NF_t \tag{8}$$

2.4. The Performance Criteria of Comparison

In this study, both linear and nonlinear models were used in the data sets. According to the error of estimation, the smaller the error, the higher the accuracy of the data. The performance criteria evaluation model to measure the error of data as well as error reduction is as follows:

$$MSE = \frac{\sum_{i=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}{n}, \tag{9}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|Y_t - \hat{Y}_t\right|}{n}, \tag{10}$$

$$\text{Error Reduction} = \frac{|a-b|}{a} \times 100. \tag{11}$$

2.5. Area and Scope of Study

This study was carried out around the coast of the South China Sea in the area of Kuala Terengganu, Kampung Marang, Kampung Setiu and Kuala Besut. The data that involved in this study are *in-situ* data in 2015 (30th April to 3rd May). This data contain 126 readings of the optimal parameters and the concentration of chlorophyll-a from different stations as shown in **Figure 1**. The predicted chlorophyll-a based on the four water data quality parameters which were temperature, pH, salinity, and Dissolved Oxygen (DO). Due to *in-situ* data taken from the field, this data was estimated to be 100% accurate as there are no constraints that prevent the error from reading.

3. Results and Discussion

The MLR, ANN and ANN-MLR models for optimal parameters predicted based on the five water data quality parameters which were temperature, pH, salinity, DO (ppm) and DO (% saturation).

3.1. Multi Linear Regression

The MLR for optimal parameters predicted based on the five water data quality parameters which were temperature, pH, salinity, DO (ppm) and DO (% saturation). According

to Table 1, the absolute value of b_i is greater and twice its standard error (*i.e.*, SE_{bi}), the ^{i}th variable regarded as a significant variable [17]. Here the four out of five parameters, *i.e.* temperature, pH, salinity and DO (ppm) were determined through a regression coefficient greater than twice their standard errors (see the boldface numbers in the constant b_i, in Table 1). These results suggest that four out of five parameters are important variables for explaining the chlorophyll-a levels in the China Sea.

The value of b_i in the correlated explanatory variables, however rely on other variables in MLR model. Consequently, the computed values of b_i for a certain explanatory variable strongly rely on the degree of its correlation with other variables in the

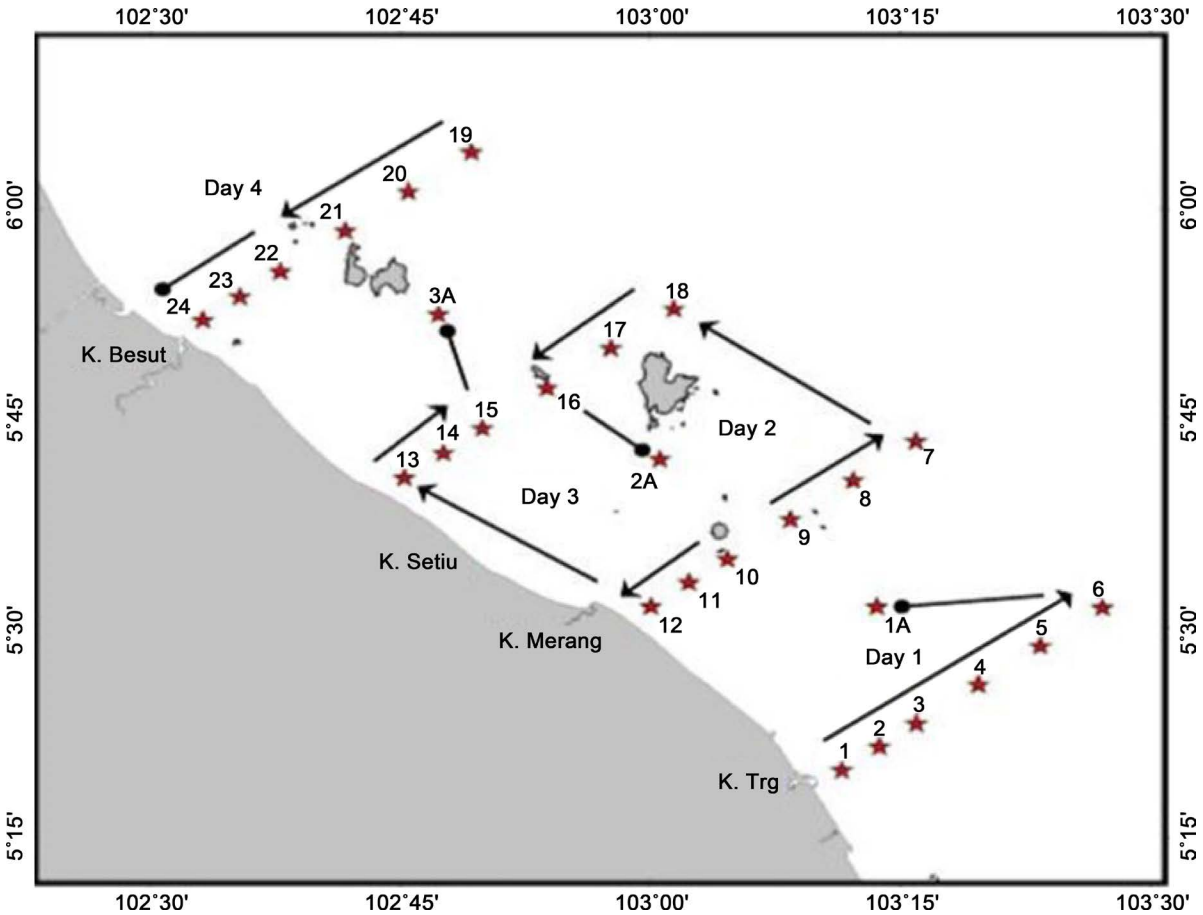

Figure 1. Sampling area in Kuala Terengganu offshore, Terengganu.

Table 1. Multiple linear regression (MLR) of chlorophyll-a using 5 water quality parameters.

		Temp(˚C)	pH	Salinity(ppt)	DO (% sat)	D0 (ppm)
Model	R	0.722	0.867	0.939	0.887	0.934
	R^2	0.521	0.752	0.882	0.787	0.872
	F	26.122	72.642	179.867	88.69	163.433
Constant	b_i	**29.991**	**7.145**	**69.309**	−72.059	**6.914**
	Std. Error	**15.868**	**3.371**	**35.602**	17.276	**2.256**

MLR model. The MLR graph of the parameters (testing and validation) which related to chlorophyll-a prediction as shown in Figures 2(a)-(e) shows that the prediction of the MLR is not satisfied with the original observation for the validation period and testing period. The correlation coefficient values between predicted models and observed data for temperature, pH, salinity, DO and chlorophyll-a are 0.722, 0.939, 0.867, 0.887 and 0.728 respectively (Figures 3(a)-(e)).

3.2. Artificial Neural Network (ANN)

For developing the prediction model for determining the optimal parameters for chlorophyll-a, data such as DO, temperature, salinity and pH was used as input. As for the output it was changed according to the parameters that we need to determine and predict. This model uses the method of adjustment tool (fitting tool). This method of customization tools used if the goal of the study want to map between numeric input data or variables with a set of target data. The neural network adjustment tool will help to select the data to create and train the networks. There are several types of samples used in ANN. The first was training. During the training exercise the inputs will be submitted

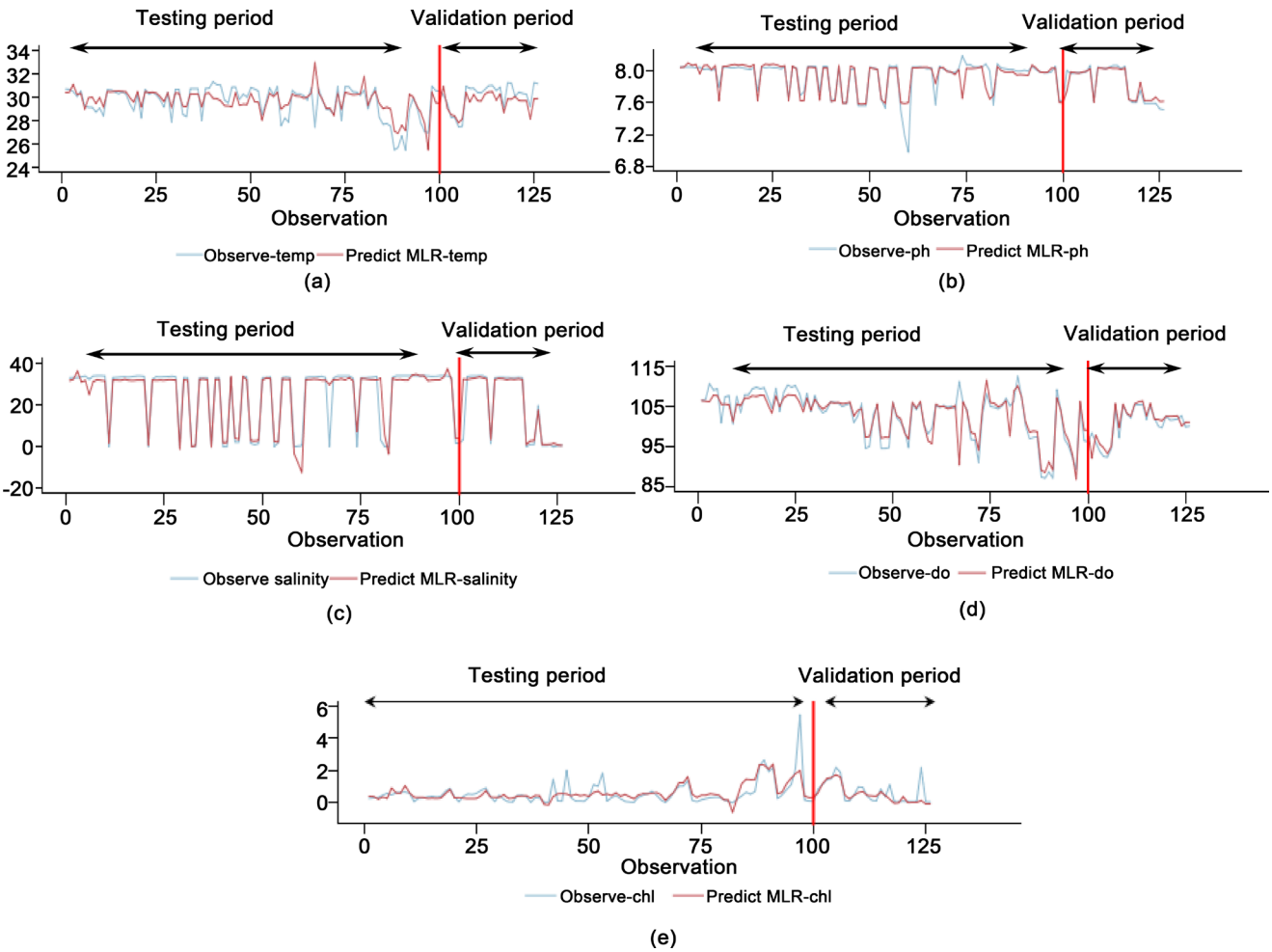

Figure 2. MLR model verification for each optimal parameters for chlorophyll-a.

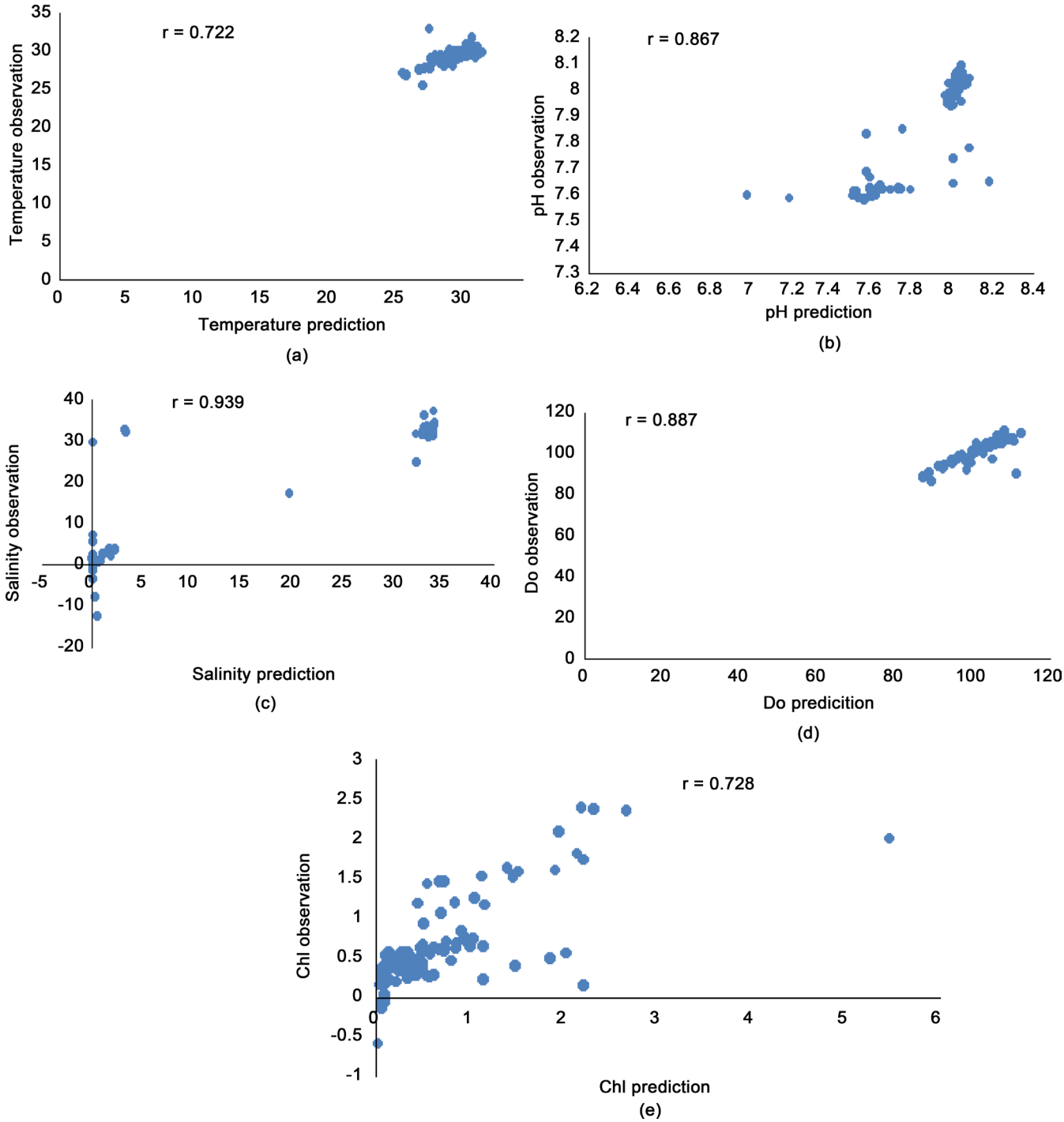

Figure 3. Observed versus MLR predicted data for each optimal parameters for chlorophyll-a.

to the network and the network will adjust based on the error in the model. The second sample was tested, which was not dependent on the gauge. Then the third sample was used to measure the network generalization and will stop the training when the generalization stopped increasing. A total of 126 data used and divided into 3 samples with 70% training, 20% and 10% verification process. The graph ANN model verification for

each optimal parameters for chlorophyll-a for testing and validation period as shown in Figures 4(a)-(e). In the ANN model verification, we can say that ANN is slightly match the prediction. These results indicate that the neural network model was able to recognize the pattern of determining the optimal parameter for the prediction of chlorophyll-a to provide better predictions, due to the predict the line of ANN- chlorophyll-a closer to the observed chlorophyll-a value.

The correlation coefficient values between predicted models and observed data for temperature, pH, salinity, DO and chlorophyll-a are 0.936, 0.920, 0.958, 0.928 and 0.738 respectively which are satisfactory in common model applications as shown in Figures 5(a)-(e).

3.3. Hybrid Model (ANN-MLR)

The algorithm of the hybrid ANN-MLR model has two steps. For the first step, we need to analyze the problem for the linear part, an MLR model was employed. In the second step, the residuals from the MLR model were modelled by using the ANN model. Since the MLR model cannot detect the non-linear structure of the optimal parameters for the prediction of chlorophyll-a, the residuals of the linear model will contain information about the non-linearity. The outputs of the neural network can be used as predictions

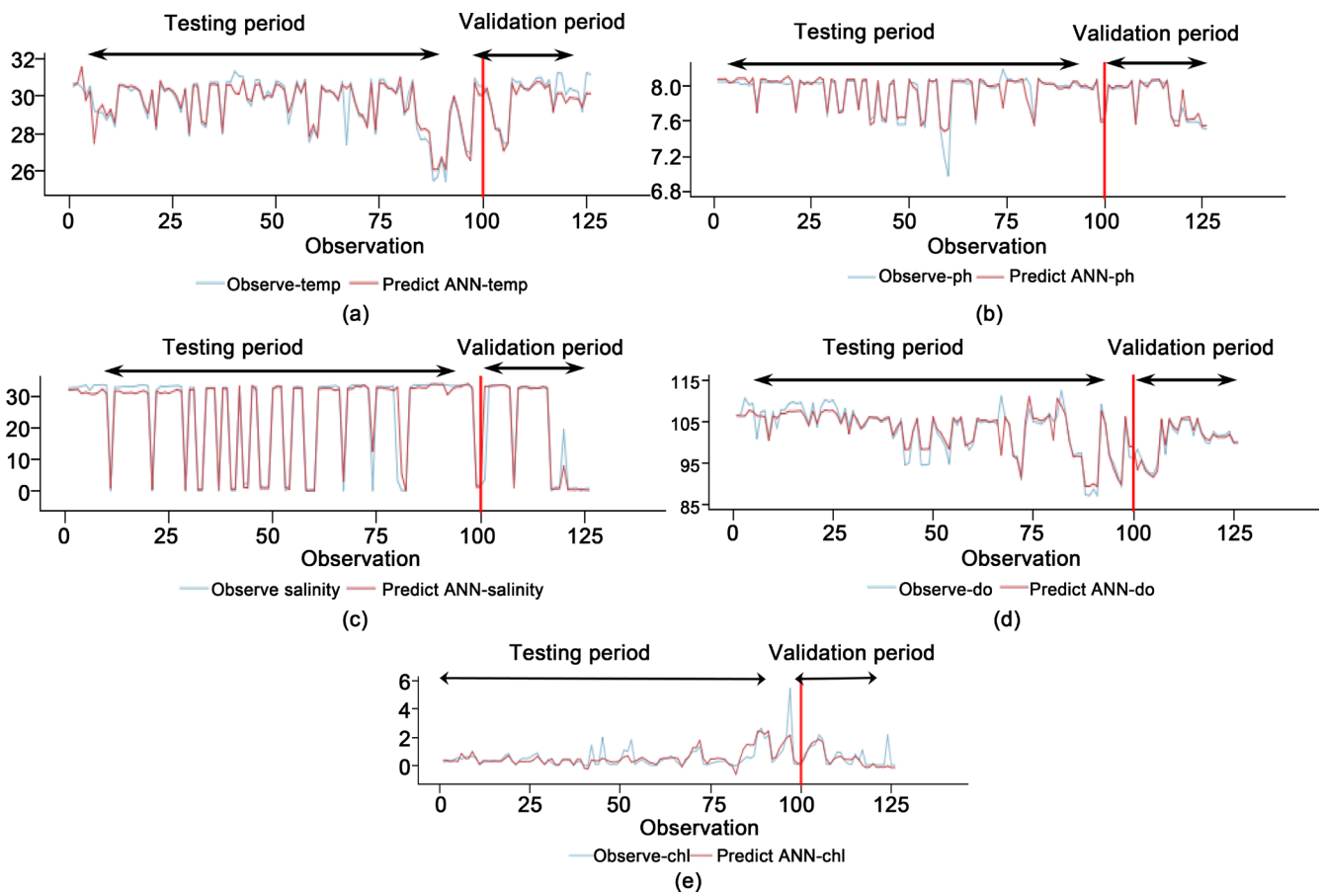

Figure 4. ANN model verification for each optimal parameters for chlorophyll-a.

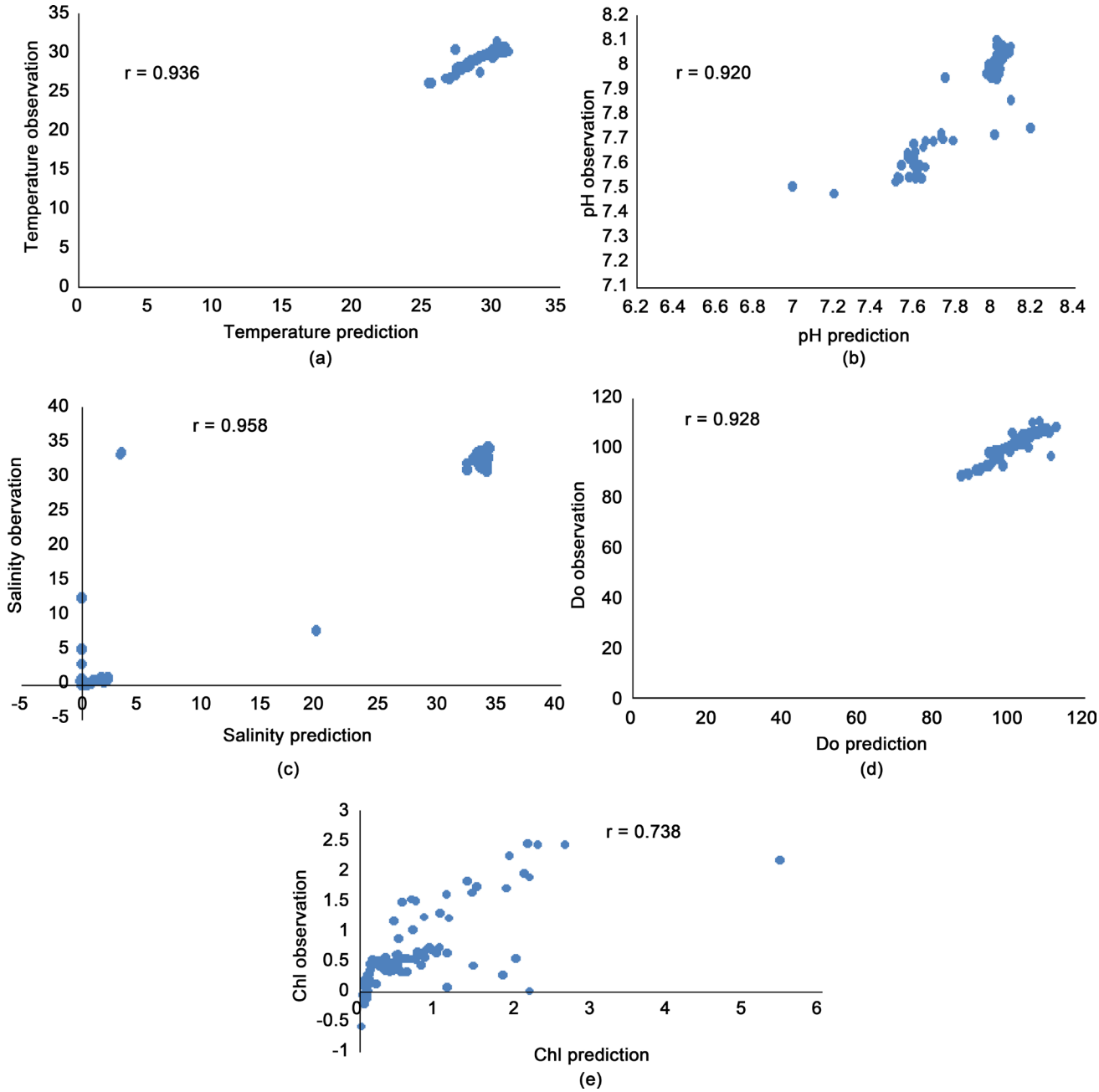

Figure 5. Observed versus ANN predicted data for each optimal parameters for chlorophyll-a.

of the error terms in the MLR model. The hybrid model utilizes the unique feature and strength of the MLR model as well as an ANN model in determining different patterns. Therefore, it may be favorable to model linear and non-linear patterns separately by using different models and then combine the predictions to improve the overall modelling and predicting performance.

When we compare the hybrid (ANN-MLR) model with the ANN model its shows more accuracy than the ANN model for the prediction of the parameters (see Table 2). The accuracy of each parameter shows over 91% for the prediction. In the hybrid model

algorithm, the input and output of the optimal parameter for chlorophyll-a, were normalized to [0.1]. In the modelling process, the hybrid was trained to adjust the model that the model predicted the parameters match well with observed data. Figures 6(a)-(e) show the predictions and observations of the models for the testing and validation period. The results indicated the model prediction reasonably match the observed parameters.

The correlation coefficient values between predicted models and observed data for

Table 2. Hybrid Model of chlorophyll-a using 5 water quality parameters.

Parameters	MSE			MAE		
	MLR	ANN	Hybrid	MLR	ANN	Hybrid
Temperature, ˚C	0.8198	0.2184	0.0320	0.6555	0.27803	0.1363
pH	0.0110	0.0069	0.0022	0.0503	0.0416	0.0177
Salinity (ppt)	25.5721	18.0246	0.4031	2.2258	1.4486	0.4025
DO (ppm)	6.6867	4.3905	0.0742	1.5440	1.3227	0.2095

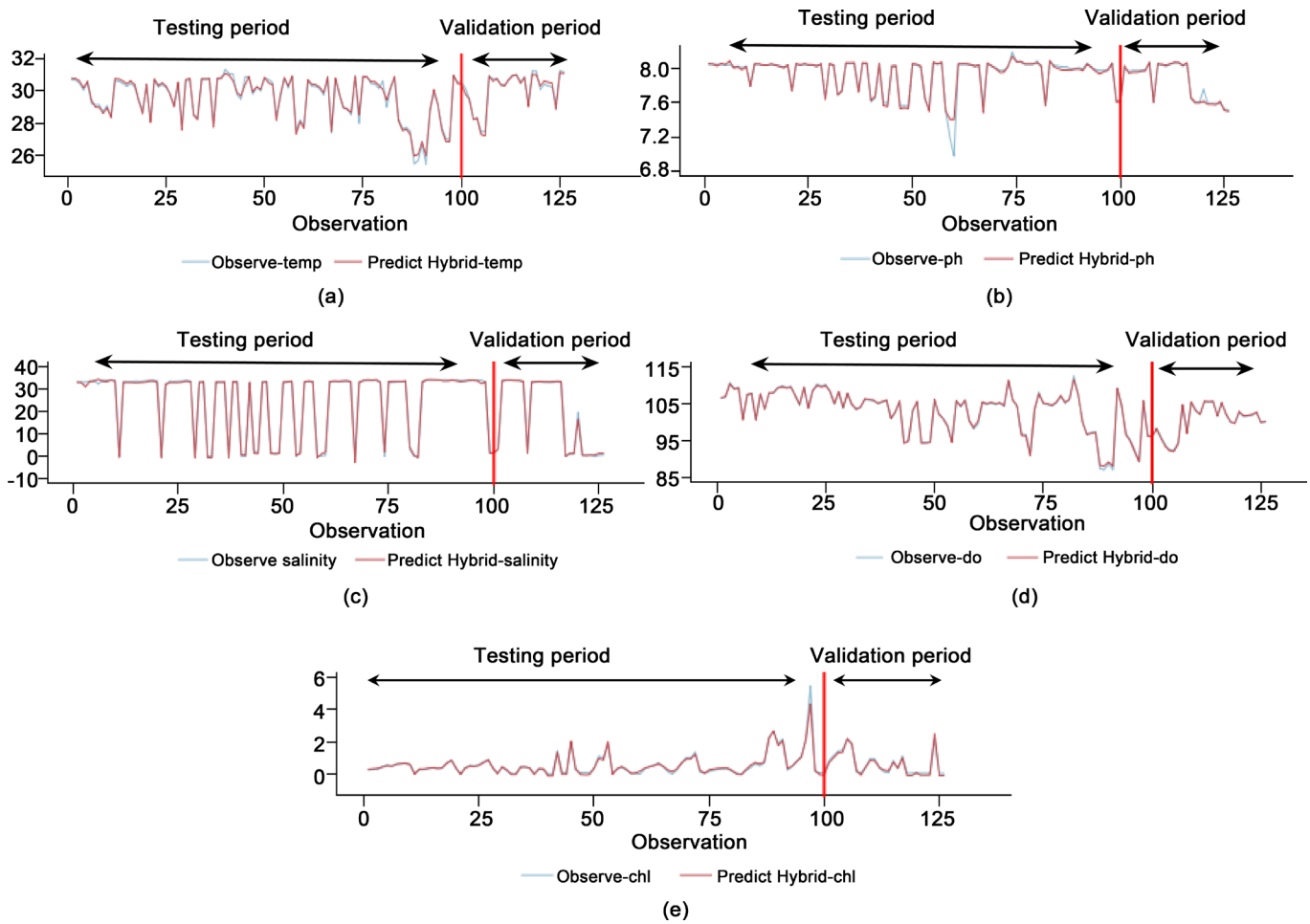

Figure 6. Hybrid model verification for each optimal parameters for chlorophyll-a.

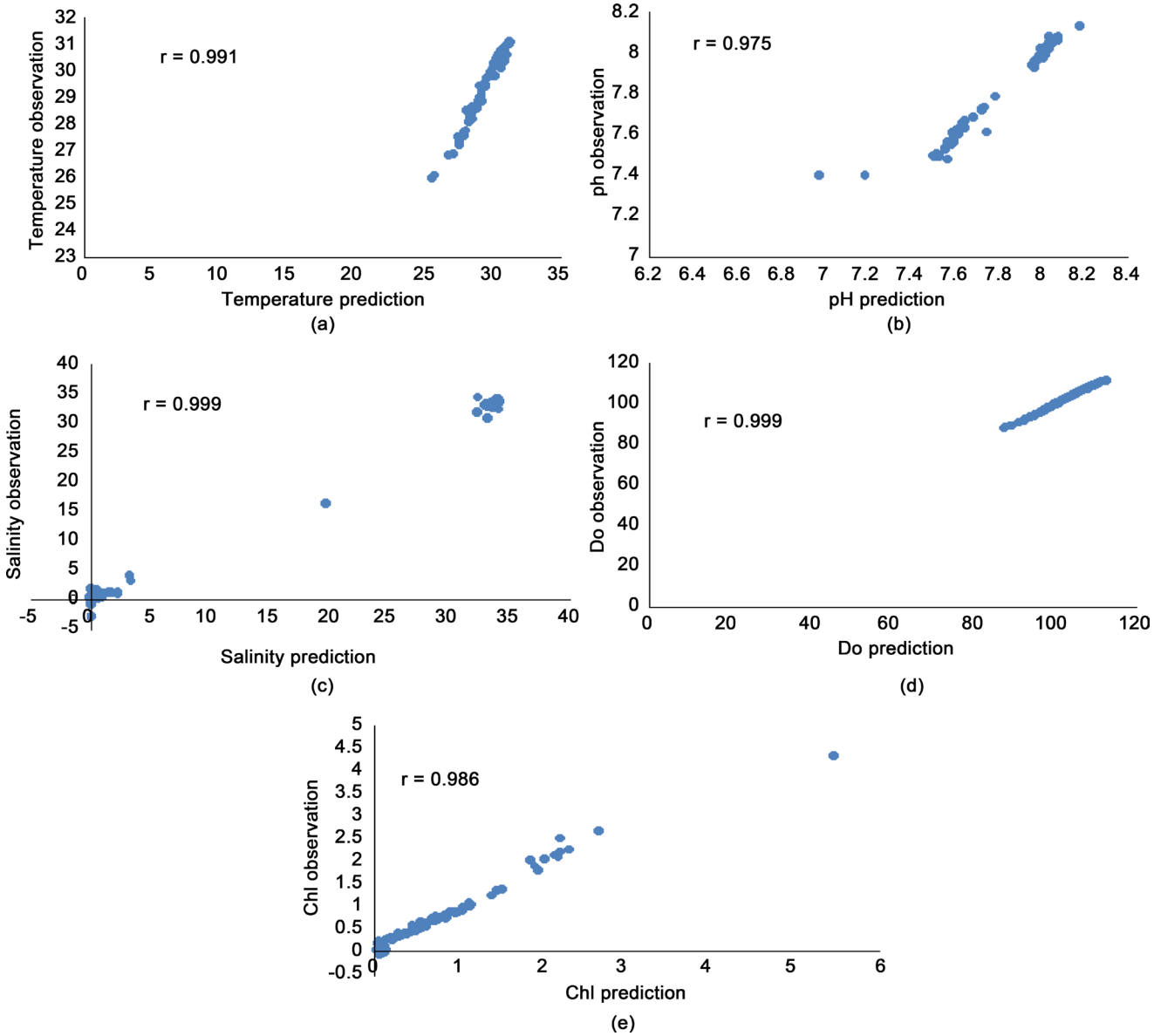

Figure 7. Observed versus and Hybrid predicted data for each optimal parameters for chlorophyll-a.

temperature, pH, salinity, DO and chlorophyll-a are 0.991, 0.939, 0.975, 0.999 and 0.999 respectively (**Figures 7(a)-(e)**) which are very strong correlation as compare to MLR and ANN.

3.4. Comparison of the Models of MLR, ANN and Hybrid (ANN-MLR)

To determine the best model of the predicted and observed data of the hybrid, ANN and MLR models for the period of four days is compared using MSE and MAE. The results revealed that the predicted optimal parameters for prediction chlorophyll-a using MLR was not found to be well matched with the observed parameters, the ANN model approximates closed to observed parameter. However the ANN-MLR seems closed and match precision with the observed for all parameters for chlorophyll-a.

Table 3 shows the estimation of errors of three different errors approaches used in the study for chlorophyll-a parameters. The MSE's between observed and predicted values were calculated in MLR models as 0.8198˚C, 0.0110, 25.5721 ppt, 6.6867 ppm and 0.2449 mg/m^3 for temperature, pH, salinity, DO and chl-a respectively. For ANN modelling approach, the MSEs between observed and predicted values were calculated as 0.2184˚C, 0.0069, 18.0246 ppt, 4.3905 ppm and 0.095463 mg/m^3 for temperature, pH, salinity, DO and chlorophyll-a respectively. In the hybrid method there was a decrease of 85.36%, 68.02%, 97.76%, 98.2%, and 99.84% in MSE values of ANN for temperature, pH, salinity, DO and chlorophyll-a respectively. Moreover, the MAEs between observed and predicted values for temperature, pH, salinity, do and chlorophyll-a were appeared to be slightly less for the ANN modelling approach. Error prediction for ANN model produced MAEs of 0.27803˚C, 0.0416, 1.4486 ppt, 1.3227 ppm, 1.3415 mg/m^3 for temperature, pH, salinity, DO and chlorophyll-a respectively. In the MAE values, the improvement of the hybrid model over ANN model were 50.99%, 57.46%, 72.21%, 84.16% and 98.08% for temperature, pH, salinity, DO and chlorophyll-a respectively. The results indicated that the hybrid model performed well for predicting of temperature, pH, salinity, DO and chlorophyll-a.

It is clearly known that the ANN-MLR is able to predict the parameters with a high degree of accuracy as compared to the ANN and MLR models. In conclusion, ANN-MLR approach can produce the best prediction of temperature, pH, salinity, DO and chlorophyll-a in the sea than the ANN and MLR modelling approach. Table 4 shows

Table 3. Statistical comparison for the MLR, ANN and ANN-MLR modelling.

Parameters	MSE			MAE		
	MLR	ANN	Hybrid	MLR	ANN	Hybrid
Temperature, ˚C	0.8198	0.2184	0.0320	0.6555	0.27803	0.1363
pH	0.0110	0.0069	0.0022	0.0503	0.0416	0.0177
Salinity (ppt)	25.5721	18.0246	0.4031	2.2258	1.4486	0.4025
DO (ppm)	6.6867	4.3905	0.0742	1.5440	1.3227	0.2095
Chlorophyll-a (mg/m^3)	0.2449	0.0955	1.50E-04	3.4502	1.3415	0.0257

Table 4. Comparison of error reduced for an MLR-ANN and ANN-MLR.

Parameters	Error Reduced in MSE (%)		Error Reduced in MAE (%)	
	MLR-ANN	MLR-Hybrid	MLR-ANN	MLR-Hybrid
Temperature, ˚C	73.36	96.10	57.58	79.21
pH	37.21	80.00	17.26	64.81
Salinity (ppt)	29.51	98.43	34.92	81.92
DO (PPM)	34.34	98.89	14.33	86.43
Chlorophyll-a (mg/m^3)	61.00	99.94	61.12	99.26

the reduced error in percentage for MSE and MAE by MLR to ANN and MLR to ANN-MLR. This results are based on the calculation in Equation (11) and Table 3. For example, the 73.36% error reduction for temperature (˚C) in MSE (%) of MLR-ANN is calculated based on Equation (11), *i.e.* [(0.8198 − 0.2184)/(0.8198) × 100]

Table 4 proved that the ANN-MLR model will reduce more error compare to other models. Figures 8-12 are the comparison for each optimal parameter between the observed and predicted values of MLR, ANN and ANN-MLR. From this figures shows that the ANN-MLR models are closed to the observed values as compared to the MLR and ANN models.

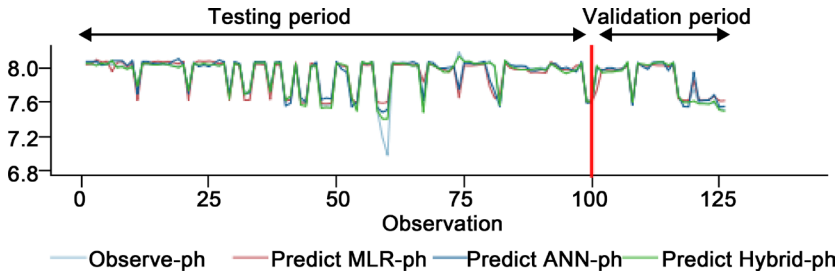

Figure 8. The observed and predicted values of MLR, ANN and ANN-MLR for pH.

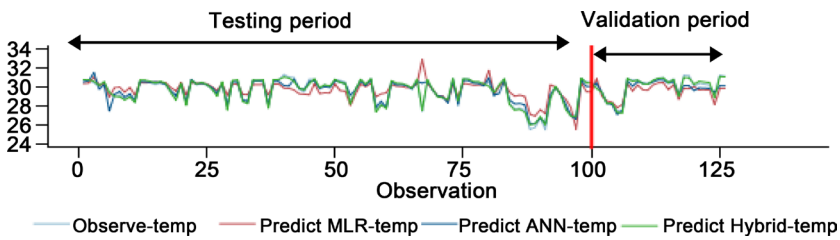

Figure 9. The observed and predicted values of MLR, ANN and ANN-MLR for temperature.

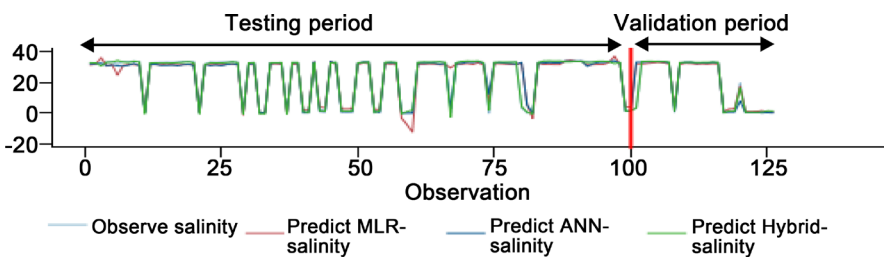

Figure 10. The observed and predicted values of MLR, ANN and ANN-MLR for salinity.

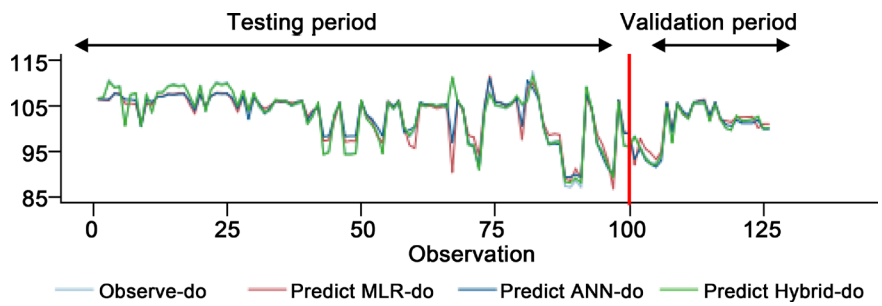

Figure 11. The observed and predicted values of MLR, ANN and ANN-MLR for DO.

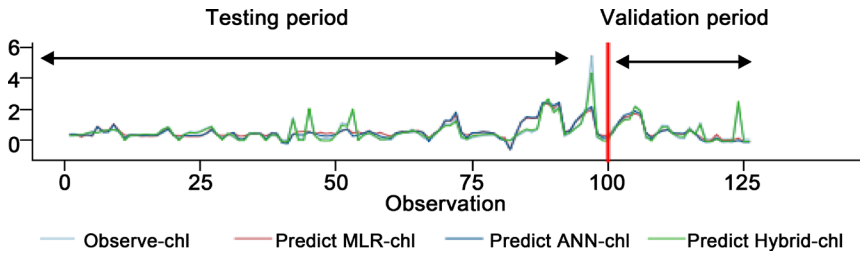

Figure 12. The observed and predicted values of MLR, ANN and ANN-MLR for chlorophyll-a.

4. Conclusion

A new approach of determining the optimal parameters for predicting of chlorophyll-a is studied. A calculative evaluation on the performance of the ANN-MLR modelling approach is proposed to predict the optimal parameters of chlorophyll-a prediction. To examine the ANN-MLR model performance compared to MLR and ANN, statistical measurements such as MSE and MAE are used. The results are as follows: The MLR model a shows a poor pattern. The ANN model we're able to provide more accurate prediction for parameters. However, ANN-MLR model revealed that this approach performs better in predicting the chlorophyll-a. In other words, ANN-MLR model is capable to recognize the patterns and the non-linearity characteristics. The accuracy measures MSE and MAE proved that the ANN-MLR provided much better accuracy over the ANNs and MLR methods for the prediction of parameters. The results of this study confirmed that the proposed model successfully improved the performance ability in determining and comparing the eutrophication.

Acknowledgements

A special gratitude for School of Informatics and Applied Mathematics (SIAM) and Research Management Centre (RMC), University Malaysia Terengganu for supporting this research paper.

References

[1] Cho, K.H., Kang, J.H., Ki, S.J., Park, Y., Cha, S.M. and Kim, J.H. (2009) Determination of the Optimal Parameters in Regression Models for the Prediction of Chlorophyll-*a*: A Case Study of the Yeongsan Reservoir, Korea. *Journal Science of the Total Environment*, **407**, 2536-2545. http://dx.doi.org/10.1016/j.scitotenv.2009.01.017

[2] Nas, B., Karabork, H., Ekercin, S. and Berktay, A. (2008) Mapping Chlorophyll-a through *in-Situ* Measurements and Terra ASTER Satellite Data. *Environmental Monitoring and Assessment*, **157**, 375-382. http://dx.doi.org/10.1007/s10661-008-0542-9

[3] Handan, C., Nilsun, D., Kanik, A. and Keskyn, S. (2004) Use of Principal Component Scores in Multiple Linear Regression Models for Prediction of Chlorophyll-a in Reservoirs. *Journal of Ecological Modelling*, **181**, 581-589.

[4] Oguz, T. and Ediger, D. (2006) Comparision of *in Situ* and Satellite-Derived Chlorophyll Pigment Concentrations, and Impact of Phytoplankton Bloom on the Suboxic Layer Structure in the Western Black Sea during May-June 2001. *Journal Deep-Sea Research II*, **53**, 1923-1933. http://dx.doi.org/10.1016/j.dsr2.2006.07.001

[5] Pereira, G.C., Evsukoff, A. and Ebecken, N.F.F. (2009) Fuzzy Modelling of Chlorophyll Production in a Brazilian Upwelling System. *Journal Ecological Modelling*, **220**, 1506-1512. http://dx.doi.org/10.1016/j.ecolmodel.2009.03.025

[6] Jouini, M., Lévy, M., Crépon, M. and Thiria, S. (2013) Reconstruction of Satellite Chlorophyll Images under Heavy Cloud Coverage Using a Neural Classification Method. *Journal Remote Sensing of Environment*, **131**, 232-246. http://dx.doi.org/10.1016/j.rse.2012.11.025

[7] Johnson, R.W. (1978) Mapping of Chlorophyll *a* Distributions in Coastal Zones. *Photogrammetric Engineering and Remote Sensing*, **44**, 617-624.

[8] Ritchie, J.C., Schiebe, F.R. and McHenry, J.R. (1976) Remote Sensing of Suspended Sediments in Surface Waters. *Photogrammetric Engineering and Remote Sensing*, **42**, 1539-1545.

[9] Robinson, I.S. (2004) Measuring the Oceans from Space: The Principle and Methods of Satellite Oceanography, Springer-Praxis, Chichester, UK.

[10] Spyrakos, E., Vilas Gonzalez, L., Torres Palenzuela, J. and Barton, E.D. (2011) Remote sensing Chlorophyll-a of Optically Complex Waters (riasBaixas, NW Spain): Application of a Regionally Specific Chlorophyll a Algorithm for MERIS Full Resolution Data during Upwelling Cycle. Remote Sensing Environment, **115**, 2471-2485. http://dx.doi.org/10.1016/j.rse.2011.05.008

[11] Sahoo, G., Schladow, S. and Reuter, J. (2009) Forecasting Stream Water Temperature Using Regression Analysis, Artificial Neural Network, and Chaotic Non-Linear Dynamic Models. *Journal of Hydrology*, **378**, 325-342. http://dx.doi.org/10.1016/j.jhydrol.2009.09.037

[12] Torres-Palenzuela, J.M., Vilas-González, L. and Mosqura-Giménez, Á. (2005) Correlation between MERIS and *in-Situ* Data for Study of Pseudo-nitzschia spp. Toxic Blooms in Galician Coastal Area. Millpress, Rotterdam, 497-507.

[13] González Vilas, L., Spyrakos, E. and Torres Palenzuela, J.M. (2011) Neural Network Estimation of Chlorophyll *a* from MERIS Full Resolution Data for the Coastal Waters of Galician *rias* (NW Spain). *Remote Sensing of Environment*, **115**, 524-535. http://dx.doi.org/10.1016/j.rse.2010.09.021

[14] Vilas, L. G., Spyrakos, E. and Palenzuela, J.M.T. (2010) Neural Network Estimation of Chlorophyll-a from MERIS Full Resolution Data for the Coastal Waters of Galician *rias* (NW Spain). *Journal Remote Sensing of Environment*, **115**, 524-535.

[15] Khashei, M. and Bijari, M. (2011) A Novel Hybridization of Artificial Neural Networks and ARIMA Models for Time Series Forecasting. *Applied Soft Computing*, **11**, 2664-2675. http://dx.doi.org/10.1016/j.asoc.2010.10.015

[16] Silva, L., Marques, J. and Alexandre, L.A. (2008) Data Classification with Multilayer Perceptrons Using a Generalized Error Function. *Neural Networks*, **21**, 1302-1310. http://dx.doi.org/10.1016/j.neunet.2008.04.004

[17] Rawlings, J. Pantula, S.G. and Dickey, D.S. (1998) Applied Regression Analysis. 2nd Edition, Springer-Verlag, New York. http://dx.doi.org/10.1007/b98890