An ANN Application for Water Quality Forecasting

Article *in* Marine Pollution Bulletin · September 2008

DOI: 10.1016/j.marpolbul.2008.05.021 · Source: PubMed

3 authors:

Sundarambal Palani
National University of Singapore

Shie-Yui Liong
National University of Singapore

Pavel Tkalich
National University of Singapore

Marine Pollution Bulletin 56 (2008) 1586–1597

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: www.elsevier.com/locate/marpolbul

An ANN application for water quality forecasting

Sundarambal Palani *, Shie-Yui Liong, Pavel Tkalich

Tropical Marine Science Institute, National University of Singapore, 14 Kent Ridge Road, Singapore 119223, Singapore

ARTICLE INFO

Keywords:
Singapore seawater
Data driven technique
Neural network
Water quality
Prediction
Forecasting
Southeast Asia

ABSTRACT

Rapid urban and coastal developments often witness deterioration of regional seawater quality. As part of the management process, it is important to assess the baseline characteristics of the marine environment so that sustainable development can be pursued. In this study, artificial neural networks (ANNs) were used to predict and forecast quantitative characteristics of water bodies. The true power and advantage of this method lie in its ability to (1) represent both linear and non-linear relationships and (2) learn these relationships directly from the data being modeled. The study focuses on Singapore coastal waters. The ANN model is built for quick assessment and forecasting of selected water quality variables at any location in the domain of interest. Respective variables measured at other locations serve as the input parameters. The variables of interest are salinity, temperature, dissolved oxygen, and chlorophyll-*a*. A time lag up to $2\Delta t$ appeared to suffice to yield good simulation results. To validate the performance of the trained ANN, it was applied to an unseen data set from a station in the region. The results show the ANN's great potential to simulate water quality variables. Simulation accuracy, measured in the Nash–Sutcliffe coefficient of efficiency (R^2), ranged from 0.8 to 0.9 for the training and overfitting test data. Thus, a trained ANN model may potentially provide simulated values for desired locations at which measured data are unavailable yet required for water quality models.

1. Introduction

Seawater is a primary natural resource for many coastal development sectors. It is also one of the most sensitive and vulnerable resources, because it is negatively impacted by a variety of anthropogenic activities. Deterioration of coastal water quality has triggered the initiation of serious management efforts in many countries. Tourism, recreation, and fishing require at least an acceptable level of seawater quality. Singapore, an island state, has stringent enforcement and continuously upgraded waste handling facilities to guarantee seawater of a certain quality.

Most acceptable ecological and social decisions are difficult to make without careful modeling, prediction/forecasting, and analysis of seawater quality for typical development scenarios. Water quality prediction enables a manager to choose an option that satisfies a large number of identified conditions. For instance, water quality variables, such as salinity, temperature, nutrients, dissolved oxygen (DO), and chlorophyll-*a* (Chl-*a*), in coastal water describe a complex process governed by a considerable number of hydrologic, hydrodynamic, and ecological controls that operate at a wide range of spatiotemporal scales. Sources of the admixtures often cannot be clearly identified, and the locally influenced complex mass exchange between the variables may not be known. Due to the correlations and interactions between water quality variables, it is interesting to investigate whether a domain-specific mechanism governing observed patterns exists to prove the predictability of these variables. The identification of such forecast models is particularly useful for ecologists and environmentalists, since they will be able to predict seawater pollution levels and take necessary precaution measures in advance.

Classical process-based modeling approaches can provide good estimations of water quality variables, but they usually are too general to be applied directly without a lengthy data calibration process. They often require approximations of various processes, and these approximations may overlook some important factors affecting the processes in seawater. A process-based model requires a lot of input data and model parameters that are often unknown, while data-driven techniques provide an effective alternative to conventional process-based modeling. Models developed by data-driven techniques are computationally very fast and require fewer input parameters than process-based models. Data-driven modeling techniques have gained popularity in the last 20 years. The scientific and engineering communities have acquired already extensive experience in the development and usage of data-driven techniques. ANNs are, however, still not widely used tools in the fields of water quality prediction and forecasting. ANNs are able to approximate accurately complicated non-linear input–output relationships. Like their physics-based numerical model counterparts, ANNs require training or calibration. After training,

* Corresponding author. Tel.: +65 6516 3105; fax: +65 6872 4067.
E-mail address: tmssp@nus.edu.sg (S. Palani).

0025-326X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.marpolbul.2008.05.021

each application of the trained ANN is an estimation of a simple algebraic expression with known coefficients and is executed practically instantaneously. The ANN technique is flexible enough to accommodate additional constraints that may arise in the application.

This paper demonstrates the application of ANNs to model the values of selected seawater quality variables, having the dynamic and complex processes hidden in the monitored data itself. The ANN model can reveal hidden relationships in the historical data, thus facilitating the prediction and forecasting of seawater quality. The steps followed in the development of such models include the choice of performance criteria, division and pre-processing of the available data, determination of appropriate model inputs and network architecture, optimization of the connection weights (training), and model validation. In this paper, a study of ANN modeling to predict and forecast temperature, salinity, DO, and Chl-*a* in Singapore coastal waters is presented. These water quality parameters were measured weekly at various locations. These models could be used as a prediction tool, which complements the process-based model and ongoing field monitoring program in the region. The results of the ANN prediction and forecast model in the East Johor Strait are discussed in this paper.

2. Materials and methods

2.1. Study area and water quality data

Singapore is a tropical country located between latitudes 1°06′N and 1°24′N and longitudes 103°24′E and 104°24′E, 137 km north of the equator. Singapore is located in Southeast Asia at the southern tip of the Malaysian Peninsula between Malaysia and Indonesia. The coastal water of Singapore is bounded by the Johor Strait in the north and the Singapore Strait in the south. Singapore is a small island having a total land area of 699 km^2, with air temperature ranging from 21.1 to 35.1 °C, and annual average rainfall of 2136 mm (Singapore Department of Statistics, 2005). Because of its geographical location, its climate is characterized by uniform temperature and pressure, high humidity, and abundant rainfall, and it is influenced strongly by the Asian monsoon. Singapore has two main monsoon seasons: the Northeast Monsoon (NEM) season lasts from November to March, and the Southwest Monsoon (SWM) season lasts from June to September. These seasons are separated by two shorter inter-monsoon (IM) periods that last from April to May and October to November. The tides in Singapore are the result of tidal generation in the South China Sea and Northern Indian Ocean. The hydrodynamic conditions of Singapore are dominated by semi-diurnal tides, having residual currents mainly in the eastern (SWM) and western (NEM) directions. In the Singapore Strait, the monsoon-driven currents and tidal fluctuation are significant. Spring tides and neap tides occur two days after the new/full moon or first/last quarter of the moon, respectively. Tidal velocities vary spatially from 0.5–1 m/s (in the open waters of the Singapore Strait) to 1.5–2 m/s (in constricted channels between the islands), but they decrease to less than 0.5 m/s in the eastern half of the Johor Strait (Zhang, 2000; Pang and Tkalich, 2003). During flood tides, the tidal current transports water from the Kuala Johor into the East Johor Strait; during ebb tides, the tidal current reverses its direction of flow. In southern Kuala Johor, the tidal current moves westwards during flood water. This direction is opposite to that at ebb tide. Both the surface and near-bottom tidal currents in Kuala Johor are stronger than those in the East Johor Strait. In both areas, ebb flows are stronger than flood flows and surface currents are stronger than near-bottom ones (Lim, 1983; Pang and Tkalich, 2003). Temperature and salinity in the East Johor Strait are influenced mainly by daily and seasonal changes as well as heavy rainfalls. The daily variation patterns of temperature and salinity in the East Johor Strait were found to be quite similar to those of the Singapore Strait, and both patterns are semi-diurnal. Due to freshwater flow from rivers, increased phytoplankton densities, higher water temperatures, and lower salinity values were observed during low tide.

The data used in this study were derived from the Singapore seawater quality survey conducted by Chuah (1998) between December 1996 and June 1997 at the entrance to the East Johor Strait (Fig. 1). The data are used to investigate both, water quality and the potential for red tide occurrences. The selected four sites (Fig. 1) lie in the area at which the East Johor Strait merges with Kuala Johor before entering the main Singapore Strait. This area is strongly influenced by tidal flow in and out of the East Johor Strait. Station 3 is close to the Squance buoy and several marine fish farms. Station 1 is located in deeper waters near the Changi Jetty. Station 2 may depict the transition between Station 3 in the Strait and Station 1 offshore. During ebb tide, the Serangoon Harbor receives a high volume of water flowing eastward out of the East Johor Strait (Lim, 1983; Pang and Tkalich, 2003). Weekly collected surface seawater samples were analyzed by Chuah (1998) for the estimation of Chl-*a*, water temperature, salinity, pH, secchi depth (SD), DO, and nutrients like ammonium (NH_4), nitrite (NO_2), nitrate (NO_3), total nitrogen (TN), phosphate (PO_4), and total phosphorus (TP). Other water quality variables, including nitrate nitrogen ($NO_2 + NO_3$), dissolved inorganic nitrogen ($DIN = NH_4 + NO_2 + NO_3$), organic nitrogen ($ON = TN - DIN$), and organic phosphorous ($OP = TP - PO_4$), were also derived. The concentration range of the water quality variables measured between January 1997 and May 1997 in the East Johor Strait in the model domain is given in Table 1.

2.2. Structure of neural network model

2.2.1. ANN architecture

The concept of artificial neurons was first introduced in 1943 (McCulloch and Pitts, 1943), and applications of ANNs in research areas began with the introduction of the back-propagation training (BP) algorithm for feedforward ANNs in 1986 (Rumelhart et al., 1986). An ANN is an information processing system that roughly replicates the behavior of a human brain by emulating the opera-

Fig. 1. Map showing the geographical setting of the present survey area with four field monitoring Stations [1. Paku (∼20 m water depth), 2. Malang Papan (∼13 m water depth), 2alt Sajahat (∼15 m water depth), and 3. Squance (∼12 m water depth)], Pasir Gudang port, the Causeway, and the major rivers in the East Johor Strait.

Table 1
Water quality properties in the ANN model domain measured between January 1997 and May 1997 in the East Johor Strait

Parameters	Unit	Minimum	Maximum	Mean	SD
Temperature	°C	27.60	31.80	29.36	1.10
Salinity	ppt	19.00	31.00	27.63	2.58
pH		7.50	8.70	8.26	0.27
SD	m	0.70	3.20	1.35	0.42
DO	mg/l	4.31	14.40	7.55	2.08
NH_4	mg/l	0.34	0.73	0.54	0.09
$NO_2 + NO_2$	mg/l	0.06	0.20	0.11	0.02
DIN	mg/l	0.44	0.86	0.65	0.10
TN	mg/l	1.97	5.77	3.22	0.92
ON	mg/l	1.30	5.11	2.57	0.88
PO_4	mg/l	0.01	0.08	0.03	0.02
TP	mg/l	0.02	0.12	0.06	0.02
OP	mg/l	0.01	0.07	0.03	0.01
Chl-*a*	mg/m^3	0.85	71.70	11.58	13.78

tions and connectivity of biological neurons. ANNs represent complex, non-linear functions with many parameters that are adjusted (calibrated or trained) in such a way that the ANN's output becomes similar to measured output on a known data set. ANNs need a considerable amount of historical data to be trained; upon satisfactory training, an ANN should be able to provide output for previously "unseen" inputs. The main differences between the various types of ANNs involve network architecture and the method for determining the weights and functions for inputs and neurodes (training) (Caudill and Butler, 1992). The multilayer perceptron (MLP) neural network has been designed to function well in modeling non-linear phenomena. A feed forward MLP network consists of an input layer and output layer with one or more hidden layers in between. Each layer contains a certain number of artificial neurons. An artificial neuron in a typical ANN architecture (Fig. 2) receives a set of inputs (signals (x) with weight (w)), calculates their weighted average ((z), using the summation function), and then uses some activation function to produce an output ($o = f(z)$, where $z = \sum_{i=1}^{n} x_i w_i$).

The connections between the input layer and the middle or hidden layer contain weights, which usually are determined through training the system. The hidden layer sums the weighted inputs and uses the transfer function to create an output value. The transfer function (local memory) is a relationship between the internal activation level of the neuron (called the activation function) and the outputs. A typical transfer function is the sigmoid function $f(z)$, which varies from 0 to 1 for a range of inputs (Caudill and Butler, 1992) where $f(z) = \frac{1}{1+e^{-z}}$. In time series prediction, supervised training is used to train the ANN in such a way as to minimize the difference between the network output and the measured target. Therefore, training is a process of weight adjustment that attempts to obtain a desirable outcome with least squares residuals. The most common training algorithm used in the ANN literature is called BP.

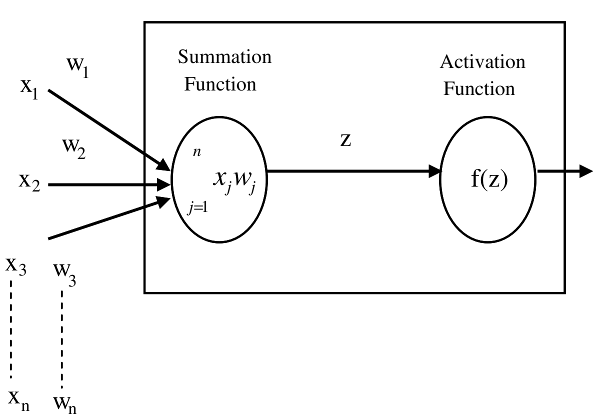

Fig. 2. A typical multilayer perceptron ANN architecture.

General regression neural networks (GRNNs) were invented by Specht (1991) and are also used in this paper. The GRNN predicts continuous outputs. GRNN nodes require two main functions to calculate the difference between all pairs of input pattern vectors and estimate the probability density function of the input variables. The difference between input vectors is calculated using the simple Euclidean distance between data values in attribute space. Weighting the calculated distance of any point by the probability of other points occurring in that area yields a predicted output value (Eqs. (1) and (2)).

$$E_{Y/X}(X) = \int y^* f_{XY}(x,y)\mathrm{d}y / \int f_{XY}(x,y)\mathrm{d}y; \tag{1}$$

$$\bar{y} = \sum_{i=1}^{n} y_i^* \exp(-D(x,x_i)) / \sum_{i=1}^{n} \exp(-D(x,x_i)), \tag{2}$$

where y_i is the ith case actual output value, $D(x, x_i)$ is calculated from Eq. (3), and n is the total number of cases in the dataset (Masters, 1993).

$$D(x,x_i) = \sum_{j=1}^{p} (x_j - x_{ij}/\sigma_j)^2, \tag{3}$$

where x is the input vector, x_i is the ith case vector, x_j is the jth data value in the input vector, x_{ij} is the jth data value in the ith case vector, and σ_j is the smoothing factor (Parzen's window) for the jth variable (Masters, 1993). The error is measured by the means of the mean square error (MSE). The MSE measures the average of the square of the amount by which the estimator differs from the quantity to be estimated. After determining the error (and depending on the optimization technique used), the above calculation may be run numerous times with different smoothing factors (sigmas). Training stops when either a threshold minimum square error value is reached or the test set square error begins to rise. A typical GRNN architecture is shown in Fig. 3.

2.2.2. ANN parameter selection

2.2.2.1. Hidden layers and nodes. Determining the number of hidden layers and nodes is a usually a trial and error task in ANN modeling. A rule of thumb for selecting the number of hidden nodes relies on the fact that the number of samples in the training set should at least be greater than the number of synaptic weights (Tarassenko, 1998). A one-hidden-layer network is commonly adopted by most ANN modelers; the number of hidden nodes M in this model is between I and $2I + 1$ (Hecht-Nielsen, 1987), where I is the number of input nodes. As a guide, M should not be less than the maximum of $I/3$ and the number of output nodes O. The optimum value of M is determined by trial and error. Networks with fewer hidden nodes are generally preferable, because they usually have better generalization capabilities and fewer overfitting problems. If the number of nodes is not large enough to capture the underlying behavior of the data, however, then the performance of the network may be impaired. In this study, a trial and error procedure for hidden node selection was carried out by gradually varying the number of nodes in the hidden layer.

2.2.2.2. Learning rate (η) and momentum (α). The function of these parameters is to speed up the training process while maintaining error reduction. There is no specific rule for the selection of values for these parameters. However, the training process is started by adopting one set of values (i.e. $\eta = 0.2$, range 0 to 1, and $\alpha = 0.5$, range 0 to 0.9) and then adjusting these values as necessary (i.e.

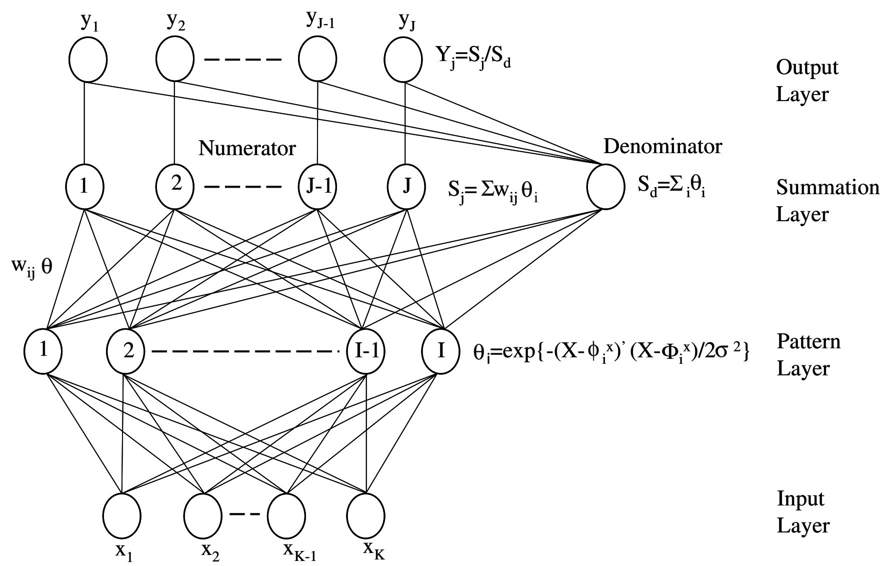

Fig. 3. General GRNN architecture. A typical GRNN architecture consists of four layers of processing units (input, pattern, summation, and output neurons). The input layer receives the input vector X and distributes the data to the pattern layer. Each neuron in the pattern layer then generates an output θ and presents the result to the summation layer. The numerator and denominator neurons in the subsequent summation layer compute the weighted and simple arithmetic sums based on the values of θ and w_{ij} learned during supervised training. The actual mathematical operations performed by these neurons in the hidden layers are illustrated in the diagram. The neurons in the output layer then carry out the division of the sums computed by the neurons in the summation layer. (Adopted from Leung et al. (2000)).

if the error reduction of the network is too slow or begins oscillating). The learning process in this study was controlled by the method of internal validation. Roughly 20% of calibration data was withheld and used to test the error at the end of each epoch (Tsoukalas and Uhrig, 1997). The weights were updated at the end of each epoch. The number of epochs with the smallest internal validation error indicates which weights to select. The ANN with the best performance when applied to the validation set was selected.

2.2.2.3. Initial weights. The weights of a network that is to be trained by BP must be initialized to some small, non-zero random values. If a network stops training before reaching an acceptable solution because a local minimum has been found, a change in the number of hidden nodes or learning parameters will often fix the problem; alternatively, we can simply start over with a different set of initial weights. When a network reaches an acceptable solution, there is no guarantee that it has reached the global minimum. The synaptic weights of the networks were initialized with normally-distributed random numbers in the range of -1 to 1.

2.2.2.4. Stopping criteria. Two commonly used stopping criteria involve (i) stopping after a certain number of runs through all of the training data (note that each run through all of the training data is called an epoch) and (ii) stopping when the total sum-squared (target) error reaches some low level. The value of the target error is actually problem dependent; due to computational time considerations, however, the training process can typically be stopped whenever the number of good patterns (i.e. those data with normalized errors less than the target error) is larger than 98%. In addition, a maximum number of iterations is usually specified.

2.2.3. Selection of input variables

In an ANN, one of main tasks is to determine the model input variables that significantly affect the output variable(s). In addition to input variables ($I1, I2, \ldots, In$), individual immediate past historical records (e.g. Ij, $t-1$; Ij, t-2,…,Ij, $t-N$; $j = 1, 2, \ldots, n$) may also influence the output variables. The choice of input variables is generally based on *a priori* knowledge of causal variables, inspections of time series plots, and statistical analysis of potential inputs and outputs. The choice of input variables for the present neural network modeling is based on a statistical correlation analysis of the field data, the prediction accuracy of water quality variables, and the domain knowledge. For example, Chl-*a* was significantly correlated with DO ($r^2 = 0.624$, P-value = 0.007) and phosphate ($r^2 = -0.516$, P-value = 0.034) at time t. Pearson indices calculated for Chl-*a* at each sampling station in the ANN model domain are shown in Table 2. It can be seen that the Chl-*a* level at Station 2 is significantly correlated ($r^2 > 0.6$, P-value $\leqslant 0.01$) with the Chl-*a* levels at Stations 1 and 3 at the same time lag. Further, the Chl-*a* level at Station 3 is highly correlated ($r^2 > 0.85$, P-value $\leqslant 0.01$) with the Chl-*a* level at Station 1. This suggests that the use of Chl-*a* with a lag time may be sufficient to predict Chl-*a* at different times and locations. The water quality at one location significantly influences that in surrounding locations. Statistical analysis of spatiotemporal water quality data can be used in combination with a data mining technique to identify important parameters and locations in the selected study area. Chau and Muttil (2007) applied data mining techniques to an ecological system in coastal waters. A stepwise approach, in which separate networks are trained for each input variable, can also be used. We have experimented with the water quality variables included in the parameters listed above in several models to both, identify the optimal predictive model, and reduce the monitoring cost by including fewer input parameters.

Having chosen appropriate input variables, the next step involves determining appropriate lags for each of these variables. This is particularly important for complex problems, in which the number of potential inputs is large and no *a priori* knowledge is available to suggest possible lags at which strong relationships exist between the output and the input time series. The network performing best is retained, and the effect of adding each of the remaining inputs in turn is assessed. This process is repeated for each combination of input variables until the addition of extra variables does not result in a significant improvement in the model's performance (Masters, 1993). The correlations between the input variables and output variable are computed separately for each lagged input variable. The significance of each of the variables in

Table 2
Pearson indices calculated for Chl-*a* at each sampling station in the ANN model domain

Parameter	Chl-*a* (t) at Station 1	Chl-*a* ($t-1$) at Station 1	Chl-*a* ($t-2$) at Station 1	Chl-*a* (t) at Station 3	Chl-*a* ($t-1$) at Station 3	Chl-*a* ($t-2$) at Station 3	Chl-*a* (t) at Station 2	Chl-*a* ($t-1$) at Station 2
Chl-*a* ($t-1$) at Station 1	-0.50^{*}							
Chl-*a* ($t-2$) at Station 1	0.44^{*}	-0.53^{*}						
Chl-*a* (t) at Station 3	0.86^{*}	-0.33^{*}	0.27^{**}					
Chl-*a* ($t-1$) at Station 3	-0.44^{*}	0.88^{*}	-0.33^{*}	-0.19^{***}				
Chl-*a* ($t-2$) at Station 3	0.22^{***}	-0.47^{*}	0.87^{*}	0.15^{***}	-0.20^{***}			
Chl-*a* (t) at Station 2	0.61^{*}	-0.28^{**}	0.24^{***}	0.61^{*}	-0.21^{***}	0.13^{***}		
Chl-*a* ($t-1$) at Station 2	-0.31^{*}	0.61^{*}	-0.28^{**}	-0.17^{***}	0.62^{*}	-0.21^{***}	0.05^{***}	
Chl-*a* ($t-2$) at Station 2	0.24^{***}	-0.31^{*}	0.61^{*}	0.15^{***}	-0.15^{***}	0.62^{*}	0.43^{*}	0.06^{***}

**P*-Value $\leqslant 0.01$; $^{**}0.01 < P\text{-value} \leqslant 0.05$; $^{***}P\text{-value} > 0.05$.

ANN modeling was also investigated using a "network growing" technique. Initially, water quality variables at time t, $t \pm 1$, and $t \pm 2$ at Stations 1 and 3 were considered as input variables to predict variables at time t. Optimal networks were obtained with these time-lagged variables. In the next stage, each of the remaining variables with the same time lag was considered along with variables to be predicted as input variables. Optimal networks for each of these combinations were obtained, and the results were compared.

2.2.4. Data partition

The data in neural networks are categorized into three sets: training or learning sets, test or overfitting test sets, and production sets. The learning set is used to determine the adjusted weights and biases of a network. The test set is used for calibration, which prevents overtraining networks. The general approach for selecting a good training set from available data series involves including all of the extreme events (i.e. all possible minimum and maximum values in the training set). The overfitting test set should consist of a representative data set. The order in which training samples were presented to the network was also randomized from iteration to iteration. Such measures are taken to improve the performance of the back-propagation algorithm. It is important to divide the data set in such a way that both training and overfitting test data sets are statistically comparable. The test set should be approximately 10–40% of the size of the training set of data (Neuroshell 2™, 2000). In the current case, the water quality data (5 months, total of 32 numbers) from Stations 1 and 3 were divided into two sets. The first set contained 80% of the records and was used as a training set; the second test contained 20% of the records and was used as an overfitting test set. The data for Stations 2 (11 numbers) and 2alt (5 numbers), which the network had never seen before, were used as the validation set. A vector pair with one or more missing measurements cannot be used for training a model, and an input vector with a missing measurement cannot be evaluated by a developed model. Therefore, the efficacy of implementing a neural network model is highly dependent on the quantity, historical range, and quality of the data used for its development.

2.2.5. Model performance evaluation

A model trained on the training set can be evaluated by comparing its predictions to the measured values in the overfitting test set. These values are calibrated by systematically adjusting various model parameters. The performances of the models are evaluated using the root mean square error (RMSE) [Eq. (4)], the mean absolute error (MAE) [Eq. (5)], the Nash–Sutcliffe coefficient of efficiency (R^2) [(Eq. (6)] (Nash and Sutcliffe, 1970), and the correlation coefficient (r). Scatter plots and time series plots are used for visual comparison of the observed and predicted values. R^2 values of zero, one, and negative indicate that the observed mean is as good a predictor as the model, a perfect fit, and a better predictor than the model (Wilcox et al., 1990), respectively. Depending on sensitivity of water quality parameters and the mismatch between the forecasted water quality variable and that measured, an expert can decide whether the predictability of the ANN model is accurate enough to make important decisions regarding data usage.

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum(X_{\text{observed}} - X_{\text{predicted}})^2} \tag{4}$$

$$\mathrm{MAE} = \frac{1}{N}\sum|X_{\text{predicted}} - X_{\text{observed}}| \tag{5}$$

$$\begin{aligned} R^2 &= 1 - \left(\frac{F}{F_0}\right) \\ F &= \sum(X_{\text{observed}} - X_{\text{predicted}})^2 \\ F_0 &= \sum(X_{\text{observed}} - X_{\text{mean observed}})^2, \end{aligned} \tag{6}$$

where N is the total number of observations in the data set.

2.3. ANN and water resources

Applications of ANNs in the areas of water engineering, ecological sciences, and environmental sciences have been reported since the beginning of the 1990s. In recent years, ANNs have been used intensively for prediction and forecasting in a number of water-related areas, including water resource study (Liong et al., 1999, 2001; Muttil and Chau, 2006), oceanography (Lee et al., 2002, 2004; Makarynskyy, 2004), and environmental science (Grubert, 2003). The use of data-driven techniques for modeling the quality of both freshwater (Karul et al., 2000; Chen and Mynett, 2003) and seawater (Barciela et al., 1999; Lee et al., 2000, 2003) has met with success in the past decade. Whitehead et al. (1997), Yabunaka et al. (1997), and French et al. (1998) have modeled the temporal changes of particular algal species in freshwater systems. Reckhow (1999) studied Bayesian probability network models for guiding decision making regarding water quality in the Neuse River in North Carolina. Blockeel et al. (1999) studied the predictive value of multiple physico-chemical properties of river water. Dzeroski et al. (2000) used regression trees with biological and chemical data for predicting water quality in Slovenian rivers. Maier and Dandy (2000) reviewed recent papers dealing with the use of neural network models for the prediction and forecasting of water resources variables. Feedforward networks with sigmoidal-type transfer functions have been used almost exclusively for the prediction and forecasting of water resources variables (Maren et al., 1990; Masters, 1993; Flood and Kartam, 1994; Hassoun, 1995; Rojas, 1996). Chau (2006) has reviewed the development and current progress of the integration of artificial intelligence (AI) into water quality modeling.

Only a limited ANN water quality prediction system for coastal water was available elsewhere. Hatzikos et al. (2005) utilized neural networks with active neurons as a modeling tool for the prediction of seawater quality indicators like water temperature, pH, DO, and turbidity. Hatzikos et al. (2007) developed an expert system using fuzzy logic that determines when certain environmental parameters exceed certain "pollution" limits, which are specified either by the authorities or environmental scientists, and sends out appropriate alerts. The earliest experimental efforts to understand the ecosystem in Singapore seawater were conducted by Tham (1953), Gu (1998), Chuah (1998), and Gin et al. (2000). Hui (2000) and Xiaohua (2000) continued to study water quality and eutrophication in Singapore Strait. To understand complex, highly non-linear water quality dynamics that vary in space and time, one can use either a process-based, three-dimensional eutrophication model (which is often complex in nature and computationally demanding) or data-driven models. Gin and Tkalich (1998), Zhang (2000), Cheong et al. (2000), Sundarambal and Tkalich (2003) carried numerical water quality modeling studies of tropical coastal waters in Singapore. The present study attempted to model Singapore seawater quality variables using ANN modeling for the first time.

Limited water quality data and the high cost of water quality monitoring often pose serious problems for process-based modeling approaches. ANNs provide a particularly good option, because they are computationally very fast and require many fewer input parameters and input conditions than deterministic models. ANNs do, however, require a large pool of representative data for training. The objective of this study is to investigate whether it is possible to predict the values of water quality variables measured by a seawater quality monitoring program; this task is quite important for enabling selective monitoring of seawater quality variables. The primary sources of inflow to Singapore that produce water quality changes arise from rivers from Singapore and Malaysia, abundant rainfall, tidal circulations, and adjacent seas.

A commercial neural net software package Neuroshell 2^{TM} (2000) Release 4.0 (Ward Systems Group) was used to develop the neural network prediction and forecast model. A set of inputs and outputs were defined, and a suitable training set was selected. The modeling steps involved in the present study were explained in Section 2.2. To determine the non-linear relationships between the water quality and eutrophication indicators, ANN models based on BP (wardnet) and the GRNN training algorithm were chosen in the present investigation. A typical GRNN architecture consists of four layers of processing units (input, pattern, summation, and output neurons) with genetic adaptive calibration. The Wardnet (WN) ANN architecture consisted of BP with three hidden layers with different Gaussian, hyperbolic-tangent, and Gaussian-complement activation functions. BP is especially good at fitting high dimensional, continuously-valued functional approximations to data. A requirement for training and evaluating neural networks is that model data be arranged into input/output vector pairs representing the variables of interest. Salinity, temperature, DO, and Chl-*a* are represented in the output layer of their respective ANN models. The parameters to be modeled, such as temperature, salinity, DO, and Chl-*a* at t and $t+1$, were the only output variables considered for the ANN's prediction and forecast modeling of water quality variables. When the trained model was applied for temperature, salinity, DO, and Chl-*a* prediction or forecasting, only the selected water quality variables at Stations 1 and 3 were used as input. Data from only two Stations (1 and 3) out of the four available (see Fig. 1) were selected for the ANN training and overfitting tests, and data from Stations 2 and 2alt were selected for ANN validation to model the water quality variables. Once the validations were sufficiently accurate, the field monitoring stations were spatially optimally selected; this in turn reduced monitoring cost. Details of the input variables, output variable, and ANN architecture with respect to the temperature, salinity, DO, and Chl-*a* ANN prediction and forecast model are given in Table 3.

3. Results and discussion

3.1. Temperature model results

The ANN model was developed to simulate weekly seawater temperature in the East Johor Strait of Singapore with respect to time and space. It used an ANN architecture BP with three hidden layers with different activation functions (wardnet) and an initial weight of 0.3. The optimum learning rate of 0.1 and momentum of 0.1 were selected as explained in Section 2.2.2. The sensitivities of the above parameters for the temperature prediction are smaller for the training and test data sets than they are for the validation data set. The parameters that produce the "best results" for validation data set (Stations 2 and 2alt) were retained for the final temperature prediction. Temperature is largely determined by the amount of solar energy absorbed by the water. Temperature is one of the more important measurements to be considered when examining water quality. Many biological, physical, and chemical parameters are dependent on temperature, and temperature can dramatically affect the rates of chemical and biological reactions. These include the solubility of chemical compounds in water, the

Table 3
Different tested ANNs for prediction/forecasting of selected water quality (WQ) variables in the domain of interest

Modeled WQ variable	Inputs	Output	Architectures
Prediction model at t			
Temperature	[Location (longitude, latitude), $(\text{temperature})_{t,t-1,\ t-2}$]at Stations 1 and 3	Temperature	Ward Net
Salinity	[Location (longitude, latitude), $(\text{temperature and Salinity})_{t,t-1,\ t-2}$]at Stations 1 and 3	Salinity	Ward Net
DO	[Location (longitude, latitude), $(\text{temperature, Salinity and DO})_{t,t-1,t-2}$]at Stations 1 and 3	DO	GRNN
Chl-*a*	[Location (longitude latitude), $(\text{DO, } PO_4\text{, temperature and Chl-}a)_t$] at Stations 1 and 3	Chl-*a*	GRNN
Forecast model at t + 1, week ahead			
Temperature	[Location (longitude, latitude), $(\text{temperature})_{t,t+1,t+2}$] at Stations 1 and 3	Temperature	Ward Net
Salinity	[Location (longitude, latitude), $(\text{temperature and salinity})_{t,t+1,t+2}$] at Stations 1 and 3	Salinity	Ward Net
DO	[Location (longitude, latitude), $(\text{temperature, salinity and DO})_{t,t+1,t+2}$] at Stations 1 and 3	DO	GRNN
Chl-*a*	[Location (longitude latitude), $(\text{DO, temperature and Chl-}a)_{t,t-1,t-2}$]at Stations 1 and 3	Chl-*a*	GRNN

distribution and abundance of organisms, the rate of growth of biological organisms, water density, mixing of different water densities, and current movements. Fig. 4a and b show the results of temperature model training and overfitting, validation Station 2, and validation Station 2alt for the prediction and forecasting model respectively. The results show that adequate temperature prediction/forecasting was obtained with temperature data as the only input. The neural network is able to simulate the water temperature with an accuracy of a degree or less (MSE < 0.5 °C and $R^2 > 0.7$).

3.2. Salinity model results

The ANN model was developed to simulate weekly seawater salinity in the East Johor Strait of Singapore with respect to time and space. It used an ANN architecture BP with three hidden layers with different activation functions (Wardnet) and initial weights of 0.3. The optimum learning rate (0.1) and momentum (0.1) were selected as explained in Section 2.2.2. The sensitivities of the above parameters for the salinity prediction are smaller for the training and test data sets than they are for the validation data set. The parameters that produce the "best results" for the validation data set (Station 2 and 2alt) were retained for the final salinity prediction. Fig. 5a and b show the salinity prediction and forecasting model results respectively. Except for the validation Station 2alt (MSE < 6 ppt), the neural network is able to simulate the water salinity within 2 ppt (MSE < 1.3 ppt and $R^2 > 0.66$) MSE for training, overfitting and validation Station 2. The results are acceptable, and the accuracy of the model can be improved by adding either more data for the validation of Station 2alt or input variables related to freshwater input.

3.3. DO model results

The ANN model was developed to simulate weekly DO concentrations in the East Johor Strait of Singapore with respect to time and space using GRNN architecture with genetic adaptive calibration. For GRNN modeling, a genetic breeding size (GBS) of 100 and genetic adaptive calibration were used. The optimum GBS was selected based on the "best results" for training, overfitting test and validation data sets from attempted GBS of 75, 100, and 200. The sensitivity of the GBS on the DO model prediction and forecasting results was marginally small. The general findings of our study include (1) with salinity and temperature as input at measurement stations, DO can be correlated to DO at nearby locations through ANN modeling, (2) sensitivities between these variables vary greatly with changing tidal and ambient conditions, and (3) a physical interpretation of the system's process physics can be readily made by examining ANN response surfaces.

The developed ANN models accurately simulated the DO concentrations of Singapore seawater. Typical ANN DO prediction model results are shown as a scatter diagram of computed and measured DO concentrations for the training and overfitting test sets ($R^2 = 0.95$; MSE = 0.28) in Fig. 6a. Using temperature and salinity as inputs, the ANN DO prediction model accurately simulates the range of DO values at Stations 2 ($R^2 = 0.47$; MSE = 0.61) and 2alt ($R^2 = -0.26$; MSE = 0.49). The model simulates DO concentrations with a good accuracy, producing MSEs of 0.28, 0.61, and 0.49 mg/l at the four field measurement stations. Typical DO forecasting results are shown in Fig. 6b for the training, overfitting test, and validation data sets. There was a small amount of uncertainty in the DO prediction and forecasting for some field measurements and ANN-predicted values, but this was likely

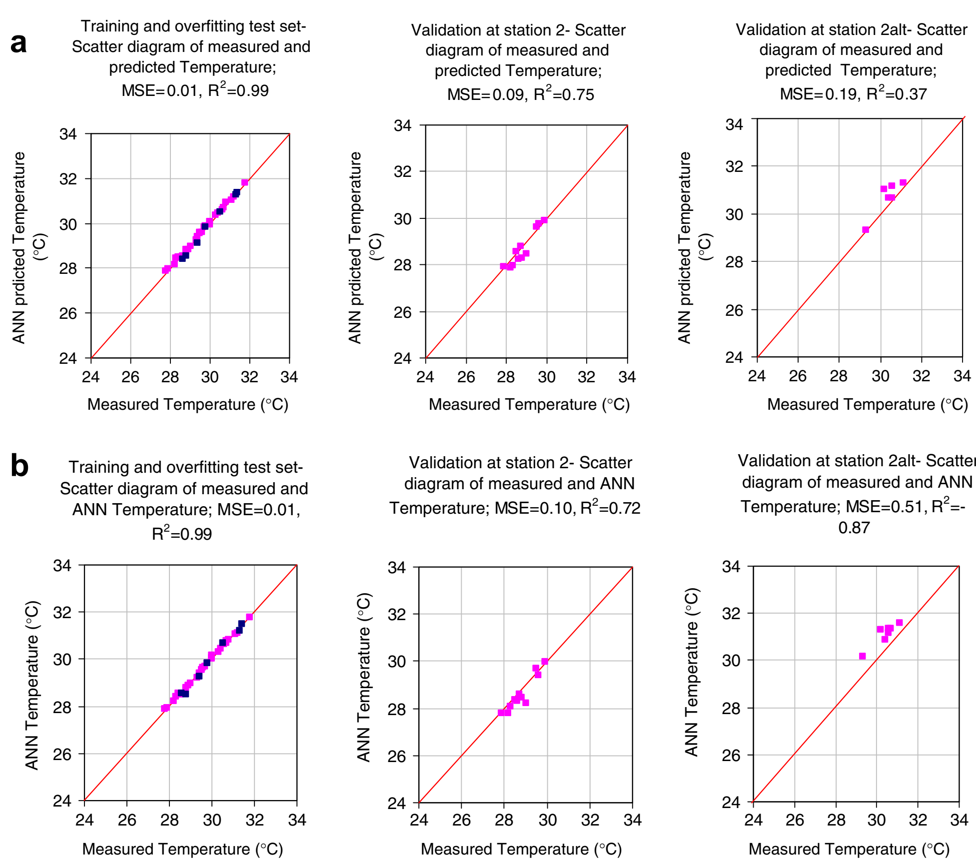

Fig. 4. Scatter diagram of predicted versus observed temperature for training (pink dot), overfitting (blue dot) tests, validations at Stations 2 and 2alt. Results for (a) temperature prediction model and (b) temperature forecast model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

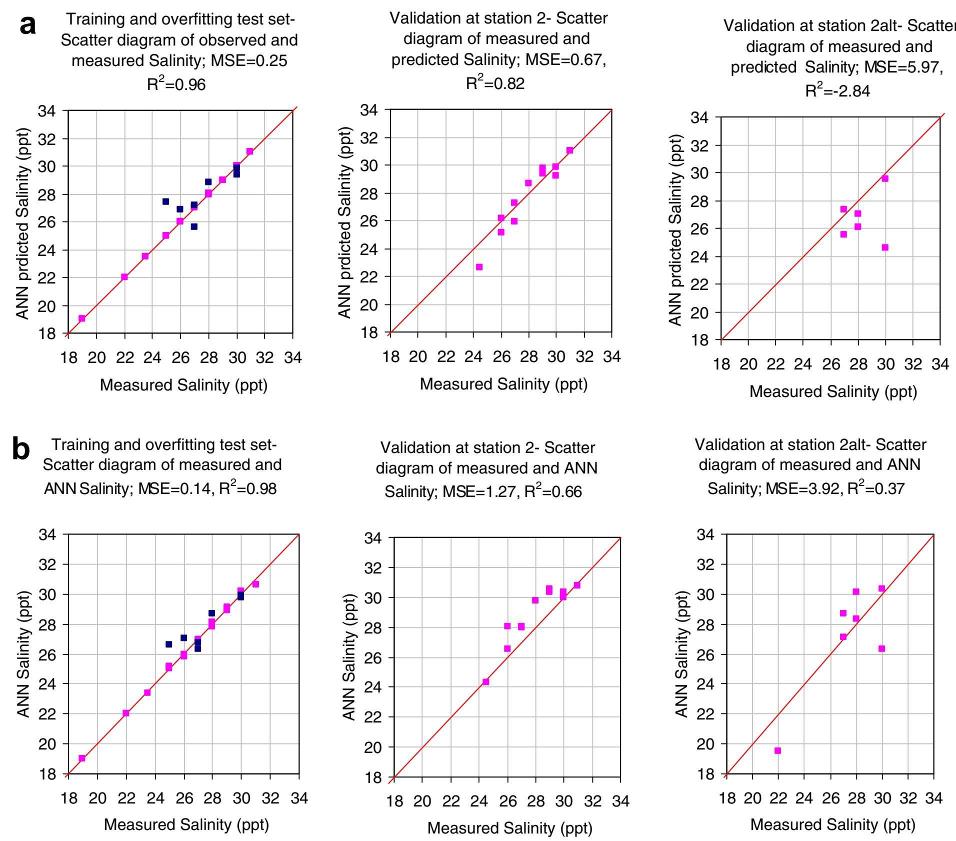

Fig. 5. Scatter diagram of predicted versus observed salinity for training (pink dot), overfitting (blue dot) tests, validations at Stations 2 and 2alt. Results for (a) salinity prediction model and (b) salinity forecast model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

due to unknown local factors. Even though there was uncertainty in the training and overfitting test sets during model construction, the performance accuracy of the developed ANN DO prediction and forecasting model are shown well in Fig. 6 when it was applied to "unseen" validation sets for Stations 2 and 2alt. The trained model can be applied to forecast DO concentrations of Singapore seawater at any location or time in the domain of interest (circled area in Fig. 1).

The ANN model of DO in Singapore seawater was successful in simulating the magnitude and patterns in measured DO concentrations that result from seasonal temperature variations, periodic blooms of phytoplankton (production of DO), and point-source discharges of oxygen-consuming substances like ammonia and BOD. In tropical seawaters where there are frequent phytoplankton blooms to increase DO levels, DO concentrations fluctuate and supersaturate during these blooms (Fig. 7). The Pearson correlation between the measured Chl-*a* and measured DO levels is positive ($R^2 = 0.624$, $P = 0.007$). The DO models simulate the dynamics of the measured concentration and are within the range of the measured values (MSE <0.61 mg/l for the prediction model and 1.13 mg/l for the forecasting model). ANN models were successful in predicting patterns in the weekly DO data, and they can be applied on hourly, daily, monthly, and seasonal time scales. The ANN model's performance was good, with a mean squared error less than 1.5 mg/l. High levels of accuracy were observed in the DO estimates (R^2 ranging from 0.9 to 1) obtained for training and overfitting at "any" location or time in the domain containing the training stations. We have demonstrated that DO levels in seawater can be forecasted with an acceptable accuracy from a small set of measurements.

3.4. Chl-a model results

Chl-*a* is one of the most important indicators of the existence and degree of eutrophication in water bodies. In order to develop an ANN model to predict/forecast phytoplankton dynamics, one can utilize scientific knowledge derived from either a process-based model or available monitoring data in the study area. Chl-*a* levels have a close, significant correlation with both nutrient species and water quality variables like salinity, temperature, and secchi depth. Measured phytoplankton blooms tended to coincide with neap tide and close values of DO and PO_4 (Fig. 7a and b). We observed that algal biomass underwent short-term variations throughout the year; further, the DO levels exhibited corresponding fluctuations. Preliminary analysis of the data indicated that the most significant bloom events fell during 17th March 1997 and 17th April 1997 at Station 3. A similar trend was followed by the other two stations, but their magnitude was comparatively less (Fig. 7).

The existence of correlations between Chl-*a* at all monitoring stations and time lags t, $t-1$, and $t-3$ showed that Chl-*a* data itself might be sufficient for developing a Chl-*a* model with respect to time and space. Chl-*a* at time (t) was highly correlated with DO and PO_4 levels at the Station 1, salinity, temperature, and DO and PO_4 levels at Station 3, and $NO_2 + NO_3$ and DO levels at Station 2. These correlations indicate a freshwater influence at Stations 1 and 3 but a seawater influence at Station 2. The eutrophication dynamics in the seawater of Singapore with respect to time and space were modeled in terms of the Chl-*a* concentration. The ANN models were trained to predict the algal biomass from the input of 15 water quality variables, including Chl-*a* at different

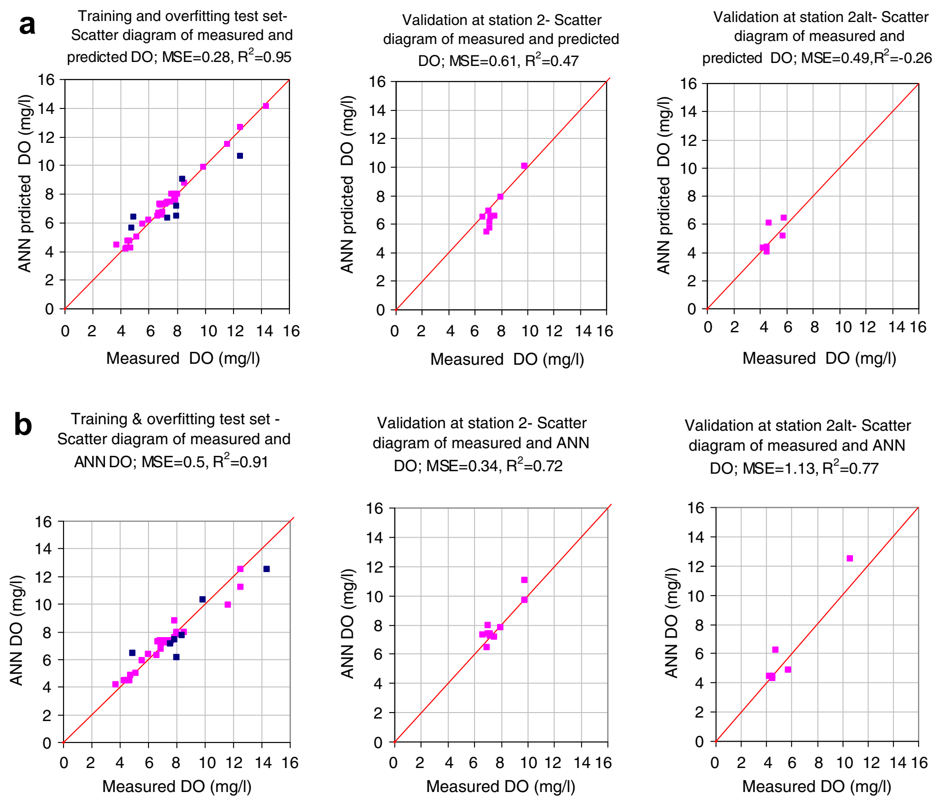

Fig. 6. (a) Prediction of DO concentration (mg/l) in Singapore seawater. Scatter diagram of the measured and ANN-predicted DO levels ($R^2 = 0.95$; MSE = 0.28) for training (pink dot) and overfitting (blue dot) tests; Validations at Stations 2 ($R^2 = 0.47$; MSE = 0.61) and 2alt ($R^2 = -0.26$; MSE = 0.49); (b) One week ahead forecast of DO concentration (mg/l) in Singapore seawater. Scatter diagram of the measured and ANN-predicted DO levels ($R^2 = 0.91$; MSE = 0.5) for training (pink dot) and overfitting (blue dot) tests; Validations at Stations 2 ($R^2 = 0.72$; MSE = 0.34) and 2alt ($R^2 = 0.77$; MSE = 1.13). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lagged times. Various model scenarios with different inputs parameters and different ANN architectures were tested for the prediction of Chl-*a*(t). Optimization of the input variables was performed by the removal/addition of one parameter at a time. In this study, however, the model scenario that used location (longitude latitude), DO, PO_4, Temp, and Chl-*a* at lagged time t from Stations 1 and 3 only as input parameters and the general regression neural network architecture performed best for the prediction of Chl-*a*(t) at any location within modeled domain. The addition of other variables did not significantly improve the performance of the network. The developed ANN model predicted the Chl-*a* and DO dynamics at Stations 2 and 2alt using the water quality observations from Stations 1 and 3. Fig. 8a shows the measured and predicted values of Chl-*a* (mg/m^3) for the best ANN model. The scatter diagram for the best Chl-*a* (mg/m^3) prediction model and forecasting model for the training, overfitting test, and validation data sets are shown in Fig. 8b and c, respectively. The results of ANN modeling (Fig. 8) suggest that it is possible to forecast water quality variables within the study domain with data from just two monitoring Stations (1 and 3). This approach reduces the monitoring cost by allowing researchers to choose an appropriate number of sampling stations for field monitoring. We suggest that the collection of field data and further improvement the ANN should continue, however, because the existing cases may not sufficiently cover the required sampling space. Performance indicators for the best Chl-*a* prediction and forecasting ANN model are shown in Table 4.

The ANN models developed were able to predict phytoplankton concentration dynamics (in terms of Chl-*a*) fairly well using minimal input parameters even though there were unknown factors affecting seawater quality and limited data sets. The results show that complex behavior in the eutrophication process could be modeled using the ANN technique, and we successfully estimated some extreme values from the validation data sets that were not used in training the neural network. The developed model can be used to (1) estimate the Chl-*a* concentration when the real value cannot be obtained, (2) estimate interpolated data between two consecutive samples, and (3) simulate different water quality scenarios for extreme ranges of input and output parameters. This eutrophication modeling approach is helpful for predicting Chl-*a* levels at any location or time (t) in the domain of interest where there are training stations. The uncertainty in our results may be due to dynamic characteristics of seawater quality around East Johor Strait, hydrodynamic forcing and unknown sources of pollution, or freshwater addition in the region of interest within a short temporal duration. It should be noted that data were measured in weekly intervals at one instance of the tidal cycle. The ANN was able to satisfactorily model the Chl-*a* dynamics and DO levels in Singapore seawater in a very short computational time. We recognize that a limited number of sampling data records were used in this study. More data sampling is thus required, and future work should recalibrate and revalidate the models to generalize our conclusions. The Chl-*a* model for seawater will be further fine-tuned for higher accuracy by including the tidal currents, wave data, and rain data as input parameters. ANN modeling can optimize monitoring network by removing stations that do not introduce new useful information, might help to analyze the factors that control the occurrence and development of HAB, provides fast predictions or forecasting, and minimizes the number of significant parameters required for modeling.

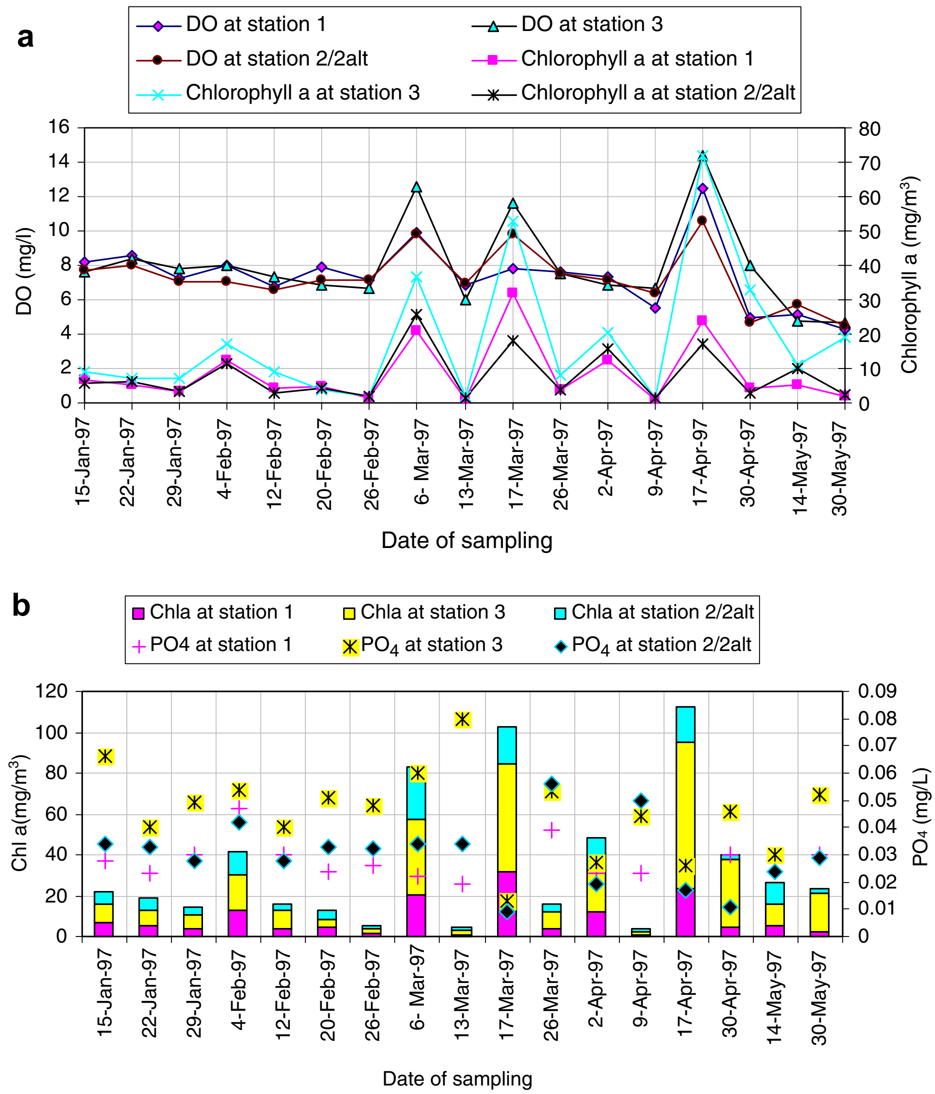

Fig. 7. Temporal variation of the measured concentrations of Chl-*a* and PO_4 at four monitoring stations in the East Johor Strait.

4. Summary and conclusions

ANN models were developed to predict salinity, water temperature, dissolved oxygen and Chl-*a* concentrations in Singapore coastal waters both temporally and spatially using continuous weekly measurements of water quality variables at different stations. In spite of largely unknown factors controlling seawater quality variation and the limited data set size, a relatively good correlation was observed between the measured and predicted values. The ANN modeling technique's application for dynamic seawater quality prediction was presented in this study. The ANN model has tremendous potential as a forecasting tool. An ANN's prediction capability was tested and found to be faster than that from a process-based model with minimal input requirements. In particular, we observed that the GRNN and MLP methods were able to "learn" the mechanism of convective transport of water quality variables quite well. Based on the "best performance" in water quality parameter forecasting, we observed that Ward Net is the better architecture for the temperature and salinity models, but the GRNN is better for DO and Chl-*a* models. This model could be used in parallel with physics-based models as a new prediction tool. It can identify important parameters for enabling both selective physical/chemical monitoring and quick water quality assessment of Singapore seawater. The limitations of this study include its limited data set. The lack of fit between the observed and estimated data indicates that new patterns must be incorporated into the model, and thus the model should be recalibrated and revalidated as more data are collected. Even though the available data size is relatively small, reasonably good results were obtained for the water quality prediction of unseen validation dataset at locations separate from the training dataset stations. If more data become available, the proposed approach should provide better predictions. The accuracy of the developed models should be compared using n-fold cross-validation, because the available data set is limited. ANN modeling is a promising and useful tool that optimizes monitoring networks by identifying essential monitoring stations (thereby permitting cost reduction) and forecasts seawater quality variables with acceptable accuracy. Further studies that apply multivariate models and incorporate new key input variables are necessary in order to arrive at better results after the pruning of non-useful connections. Additionally, future work should use new types of algorithms that are more appropriate for time series forecasting.

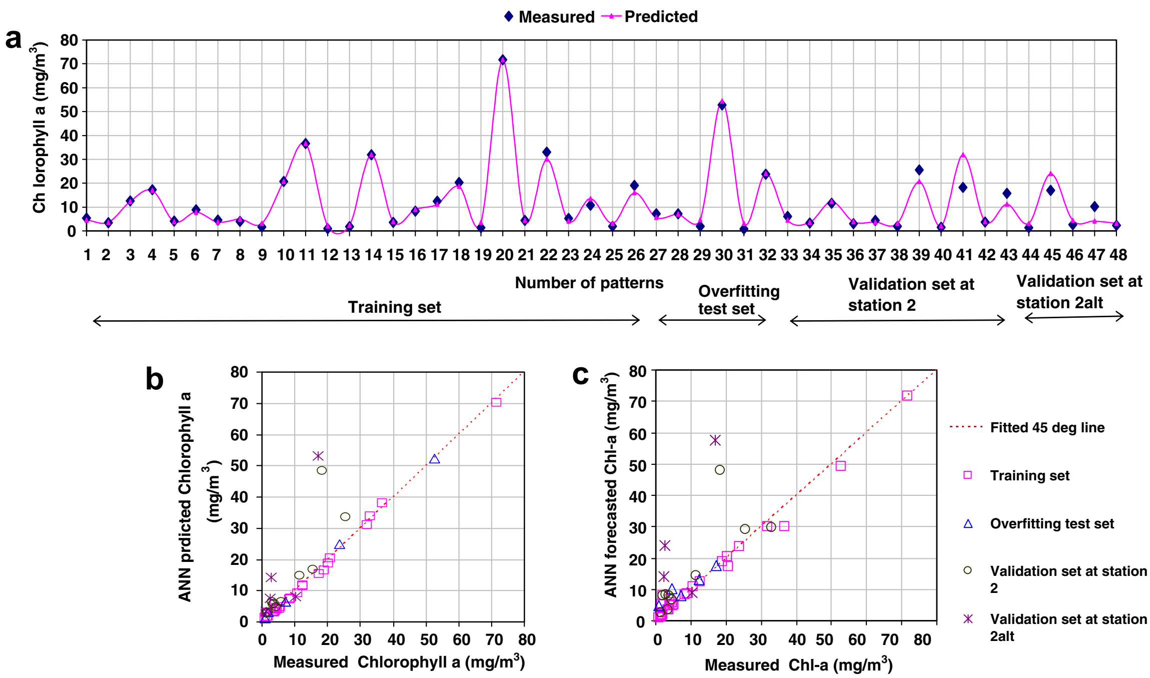

Fig. 8. (a) Measured and predicted values of Chl-*a* (mg/m^3); (b) Scatter diagram of the predicted and measured values of Chl-*a* (mg/m^3) for training, overfitting test, and validation data sets; (c) Scatter diagram of the forecasted and measured values of Chl-*a* (mg/m^3) for training, overfitting test, and validation data sets.

Table 4
Performance indicators for best Chl-*a* prediction/forecast ANN model

ANN model type	Dataset	R^2	RMSE	r^2	Network neurons
GRNN-prediction model	Training	0.99	1.33	1.00	*I–H–O*
	Overfitting test	0.99	1.74	1.00	10–32–1
	Validation 1	0.63	4.64	0.85	
	Validation 2	0.50	4.33	0.86	
GRNN-forecast model	Training	0.98	2.17	0.99	*I–H–O*
	Overfitting test	0.97	3.19	0.98	20–34–1
	Validation 1	0.70	10.4	0.84	
	Validation 2	0.51	23.6	0.71	

Note: validations 1 and 2: data at Stations 2 and 2alt, respectively; *I*-input, *H*-hidden, *O*-output layers.

Acknowledgments

The authors thank the authors A.L Chuah and K.Y.H. Gin for the published data in "Water quality study of the East Johor Strait" (1998). We also thank the Tropical Marine Science Institute for financial support and provided technical facilities.

References

Barciela, R.M., Garcia, E.E., Fernandez, E., 1999. Modelling primary production in a coastal embayment affected by upwelling using dynamic ecosystem models and artificial neural networks. Ecological Modelling 120, 199–211.
Blockeel, H., Dzeroski, S., Grbovic, J., 1999. Simultaneous prediction of multiple chemical parameters of river water quality with tilde. In: Proceedings of the Third European Conference on Principles of Data Mining and Knowledge Discovery. LNAI, 1704. Springer-Verlag, p. 16.
Caudill, M., Butler, C., 1992. Understanding Neural Networks. Basic Networks, vol. 1. MIT Press, Cambridge, MA.
Chau, K.W., Muttil, N., 2007. Data mining and multivariate statistical analysis to ecological system in coastal waters. Journal of Hydroinformatics 9 (4), 305–317.
Chau, K.W., 2006. A review on integration of artificial intelligence into water quality modeling. Marine Pollution Bulletin 52 (7), 726–733.
Chen, Q., Mynett, A.E., 2003. Integration of data mining techniques and heuristic knowledge in fuzzy logic modelling of eutrophication in Taihu Lake. Ecological Modelling 162 (1/2), 55–67.
Cheong, H.F, Shankar, N.J., Ong, C.E., Huda, M.K., 2000. Baseline study of water quality in Singapore coastal waters. In: Proceedings of XXIX IAHR Congress, Beijing, China. web site: http://www.iahr.org/e-library/beijing_proceedings/Theme_B/BASELINESTUDYOFWATERQUALITY.html.
Chuah, A.L., 1998. Water quality study of the East Johor Strait. Thesis (M.Sc.). Department of Chemical Engineering, National University of Singapore, Singapore, 129 pp.
Dzeroski, S., Demsar, D., Grbovic, J., 2000. Predicting chemical parameters of river water quality from bioindicator data. Applied Intelligence 13 (7–17).
Flood, I., Kartam, N., 1994. Neural networks in civil engineering I: principles and understanding. Journal of Computing in Civil Engineering 8 (2), 131–148.
French, M., Recknagel, F., Jarrett, G.L., 1998. Scaling issues in artificial neural network modelling and forecasting of algal bloom dynamics. In: Abt, S.R., Young-Pezeshk, J., Watson, C.C. (Eds.), Proceedings of the International Water Resources Engineering Conference, ASCE, August 3/7, 1998, Memphis, Tennessee, 1, pp. 891/896.
Gin, K.Y.H., Tkalich, P., 1998. A three-dimensional eutrophication model for Singapore costal water. In: Conference Proceedings of Environmental Strategies for the 21st century, Singapore, pp. 280–286.
Gin, K.Y.H., Lin, X., Zhang, S., 2000. Dynamics and size structure of phytoplankton in the coastal waters of Singapore. Journal of Plankton research 22 (8), 1465–1484.
Grubert, J.P., 2003. Acid deposition in the eastern United States and neural network predictions for the future. Journal of Environmental Engineering and Science 2 (2), 99–109.
Gu, G., 1998. Phytoplankton dynamics in Singapore's coastal waters. Thesis (M.E). Department of Civil Engineering, National University of Singapore, 243 pp.
Hassoun, M.H., 1995. Fundamentals of Artificial Neural Networks. MIT Press, Cambridge.
Hatzikos, E., Anastasakis, L., Bassiliades, N., Vlahavas, I., 2005. Simultaneous prediction of multiple chemical parameters of river water quality with tide. In: Proceedings of the Second International Scientific Conference on Computer Science. IEEE Computer Society, Bulgarian Section.

Hatzikos, E., Bassiliades, N., Asmanis, L., Vlahavas, I., 2007. Monitoring water quality through a telematic sensor network and a fuzzy expert system. Expert Systems 24 (3), 143–161. 19.
Hecht-Nielsen, R., 1987. Kolmogorov's mapping neural network existence theorem. Proceedings of 1st IEEE International Joint Conference of Neural Networks. Institute of Electrical and Electronics Engineers, New York, NY.
Hui, N.S., 2000. Eutrophication studies of estuarine and marine waters. Thesis (M.E). Department of Civil Engineering, National University of Singapore.
Karul, C., Soyupak, S., Cilesiz, A.F., Akbay, N., Germen, E., 2000. Case studies on the use of neural networks in eutrophication modeling. Ecological Modelling 134, 145–152.
Lee, J.H.W., Wong, K.T.M., Huang, Y., Jayawardena, A.W., 2000. A real time early warning and modeling system for red tides in Hong Kong. In: Wang, Z.Y. (Ed.), Proceedings of the Eighth International Symposium on Stochastic Hydraulics. Balkema, Beijing, pp. 659–669.
Lee, J.H.W., Huang, Y., Dickmen, M., Jayawardena, A.W., 2003. Neural network modelling of coastal algal blooms. Ecological Modelling 159, 179–201.
Lee, T.L., Jeng, D.S., 2002. Application of artificial neural networks in tide forecasting. Ocean Engineering 29, 1003–1022.
Lee, T.L., 2004. Back-propagation neural network for long-term tidal predictions. Ocean Engineering 31, 225–238.
Leung, M.T., Chen, A.S., Daouk, H., 2000. Forecasting exchange rates using general regression neural networks. Computers and Operations Research 27 (11), 1093–1110. 18.
Lim, L.C., 1983. Coastal fisheries oceanographic studies in Johore Strait, Singapore. I. Current movement in the East Johore Strait and its adjacent waters. Singapore Journal of Primary Industries 2 (2), 83–97.
Liong, S.Y., Lim, W.H., Paudyal, G., 1999. Real time river stage forecasting for flood stricken Bangladesh: neural network approach. Journal of Computing in Civil Engineering ASCE 14 (1), 1–8.
Liong, S.Y., Khu, S.T., Chan, W.T., 2001. Derivation of pareto front with genetic algorithm and neural network. Journal of Hydrologic Engineering ASCE 6 (1), 56–61.
Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental Modelling and Software 15, 101–124.
Makarynskyy, O., 2004. Improving wave predictions with artificial neural networks. Ocean Engineering 31, 709–724.
Maren, A.J., Harston, C.T., Pap, R., 1990. Handbook of Neural Computing Applications. Academic Press, San Diego, CA.
Masters, T., 1993. Advanced Algorithms for Neural Networks. A C++ Sourcebook. John Wiley and Sons, Inc, New York.
McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas imminent in nervous activity. Bulletin and Mathematical Biophysics 5, 115–133.
Muttil, N., Chau, K.W., 2006. Neural network and genetic programming for modelling coastal algal blooms. International Journal of Environment and Pollution 28 (3/4), 223–238.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models; part I – a discussion of principles. Journal of Hydrology 10, 282–290.
Neuroshell 2™, 2000. Neuroshell Tutorial. Ward Systems Group, Inc., Frederick, MD, USA.
Pang, W.C., Tkalich, P., 2003. Modeling tidal and monsoon driven currents in the Singapore Strait. Singapore Maritime and Port Journal, 151–162.
Reckhow, K.H., 1999. Water quality prediction and probability network models. Canadian Journal of Fisheries and Aquatic Sciences 56, 1150–1158.
Rojas, R., 1996. Neural Networks: A Systematic Introduction. Springer-Verlag, Berlin.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. (Eds.), Parallel Distributed Processing. MIT Press, Cambridge.
Specht, D.F., 1991. A general regression neural network. IEEE Transactions on Neural Networks 2 (6), 568–576.
Sundarambal, P., Tkalich, P., 2003. Eutrophication modelling for the Singapore waters. In: Proceedings International Conference on Port and Maritime R&D and Technology, 10–12 September 2003, Singapore, vol. 2, pp. 51–58.
Tarassenko, L., 1998. A Guide to Neural Computing Applications. Arnold Publishers, London.
Tham, A.K., 1953. A preliminary study of the physical, chemical and biological characteristics of Singapore. Fishery Publication 1 (4), 6–39.
Tsoukalas, L.H., Uhrig, R.E., 1997. Fuzzy and Neural Approaches in Engineering. Wiley Interscience, New York. 587 pp.
Whitehead, P.G., Howard, A., Arulmani, C., 1997. Modelling algal growth and transport in rivers – a comparison of time series analysis, dynamic mass balance and neural network techniques. Hydrobiologia 349, 39–46.
Wilcox, B.P., Rawls, W.J., Brakensiek, D.L., Wight, J.R., 1990. Predicting runoff from rangeland catchments: a comparison of two models. Water Resources Research 26, 2401–2410.
Xiaohua, L., 2000. Eutrophication studies of Singapore's coastal waters. Thesis (M.E). Department of Civil Engineering, National University of Singapore.
Yabunaka, K.I., Hosomi, M., Murakami, A., 1997. Novel application of back-propagation artificial neural network model formulated to predict algal bloom. Water Science and Technology 36 (5), 89–97.
Zhang, Q.Y., 2000. A three dimensional eutrophication model for Singapore coastal water. Thesis (PhD). Dept. of Civil Engineering, National University of Singapore, 192 pp.