Original Research

Journal Home page : www.jeb.co.in ★ E-mail : editor@jeb.co.in
Journal of Environmental Biology

JEB™ p-ISSN: 0254-8704
e-ISSN: 2394-0379
CODEN: JEBIDP

DOI : http://doi.org/10.22438/jeb/41/2(SI)/JEB-23

Modeling of microbial contamination in the Marmara Sea, Bursa-Turkey

A. Katip*

Department of Environmental Engineering, Faculty of Engineering, Bursa Uludag University, Görükle Campus, Bursa-16059, Turkey

***Corresponding Author Email :** aballi@uludag.edu.tr

Paper received: 29.04.2019 *Revised received*: 26.09.2019 *Accepted*: 03.01.2020

Abstract

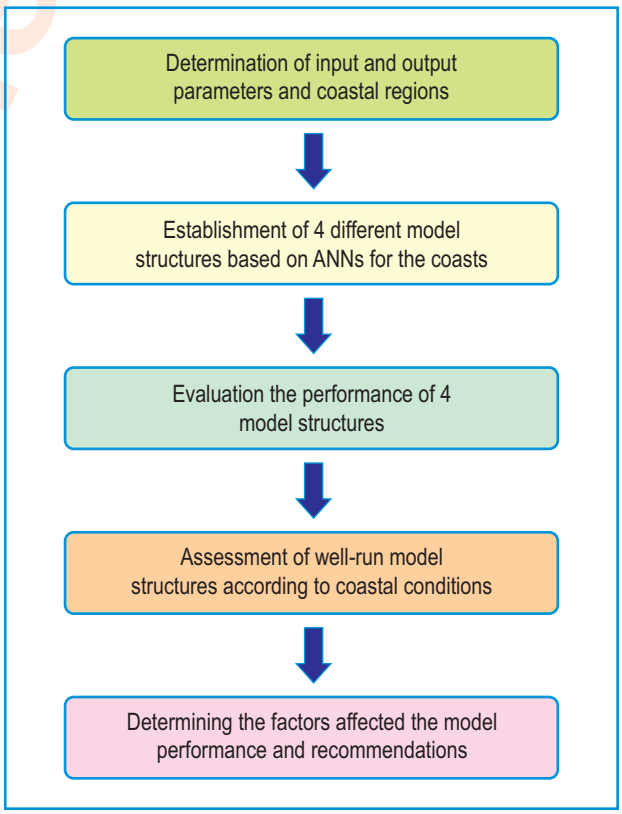
Aim: The main objective of this study was to design and develop the feed forward neural network (FNN) model structures for forecasting of faecal coliform concentrations and microbial water quality in Gemlik, Karacabey and Mudanya coastal areas alongside the Sea of Marmara, Turkey.

Methodology: Artificial neural networks (ANNs) are modeling tools for environmental parameters, especially water quality and provide working of inter-related multi parameters. In this study, 4 model structures were implemented to forecast the faecal coliform concentrations for the sea coasts of "Gemlik, Karacabey and Mudanya" alongside the Marmara Sea. Total coliform and faecal streptococci were input parameters. The Levenberg–Marquardt algorithm was applied for training the modeling studies. The results of the models were crosschecked with the real concentrations according to performance functions root mean squared error (RMSE).

Results: Comparison of the modeling results with the measured concentrations demonstrated that established model structures provided correct results. (R) Correlation coefficients were determined between 0.57 and 0.98. It was observed that during the trials enhancing the hidden layer counts in the model structures did not increase the model performance in each test. Kind and count of inputs affected the model productivity. The growing rates of the coliform group bacteria were dissimilar because, different types of contaminants in the seawater affect the metabolism. The error values of the forecasting results applied in Gemlik and Mudanya Coasts were larger because there were large quantities of pollution loads and pollutant diversities.

Interpretation: The developed model structures could predict the microbial contamination in the coastal environments and provided information on the more effective integrated sea coast management and protection of human health.

Key words: Faecal pollution, Feed forward neural network, Pathogenic microorganisms, Sea of Marmara, Water quality modeling



How to cite : Katip, A.: Modeling of microbial contamination in the Marmara Sea, Bursa-Turkey. *J. Environ. Biol.*, **41**, 432-438 (2020).

Introduction

Artificial neural networks have been used effectively for the last 60 years in the studies related to the environmental pollution, climatic change and planning of water resources (Morid *et al*., 2007; Rankovic *et al*., 2010, Kim and Valdes, 2003, Partal and Kişi, 2007, Reimer and Sodoudi, 2004, Belayneh and Adamowski, 2012, Mishra and Singh, 2010). Microbial pollution, limnological parameters, eutrophication state and other pollutants are the most studied topics of water resources (Brion and Lingireddy, 2003, Soyupak *et al*., 2003, Karul *et al.,* 2000). Sea which is a big source of carbon sink and support large number of biodiversity is being considered as a part of much discussion in the scientific and social environment due to continuous increase of micobial and macrobial contaminations because it supports us socially, economically as well as ecologically.

Fecal indicator bacteria (FIB) are used as proxies to measure the microbial water quality of aquatic ecosystems (Katip, 2018). Methods of modeling FIB has evolved in order to provide accurate and timely prediction to inform decisions by governing authorities to prevent risks to public health (Vijayashanthar *et al*., 2018).

Sea coasts are ever-mounting polluted and natural state of coastal zones is devastated, therefore sustainable use of coasts is difficult (Agúndez *et al*., 2014). This study is important in terms of environmental science and technology because ANN models provide completion of missing measurements and prediction of pollution.

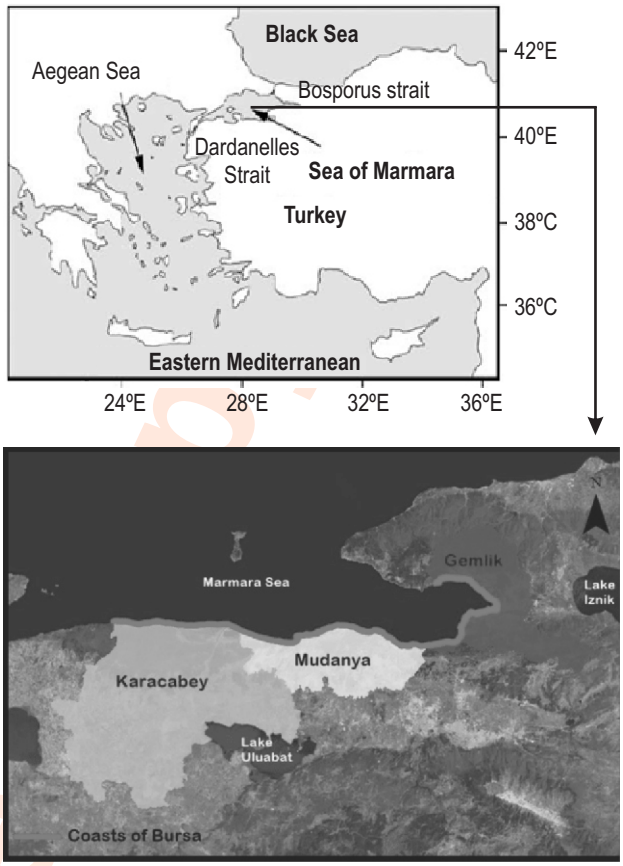


Fig. 1 : Site location map of the Sea of Marmara and Coasts of Bursa (Abrajano *et al.,* 2002; Anonymous, 2014).

Materials and Methods

Study area and Data: Sea coasts of Bursa province is 135 km long lies alongside the coasts of Marmara Sea and share 1.47% coast of Turkey (7816 km). These coasts are located within the boarders of sub-provinces in Bursa; mainly in Gemlik, Mudanya and Karacabey, with a very small part in Osmangazi. In terms of land use, agricultural areas and forestlands cover 45.5 and 42 %, respectively. Even though the rural areas form a ratio of 5 %, these coasts are subjected to domestic and industrial pollution, as being the most important region in Turkey, in terms of industry and population (Anonymous, 2014).

In this study, the faecal pollution concentrations of the coasts in Gemlik, Mudanya and Karacabey sub-provinces; with most of the seacoast in Bursa; were evaluated using the data procured from the Bursa Provincial Directorate of Public Health (Anonymous, 2015). Bacterial counts were performed by membrane filtration (MF) method and expressed as CFU100 ml^{-1} (APHA, 2012, TSWQR, 2006). Total coliform, faecal coliform and fecal streptococci parameters were modeled with ANN and evaluated.

ANN Model Application*:* Feed forward neural network structure and Levenberg Marquardt training algorithm were selected (Hagan and Menhaj, 1994). Two counts of particular hidden layer neuron for each input neuron were tested. In the recent studies, two different approaches were used for specifying the count of hidden layer neurons. The activation function was applied for the output layer as a linear function and the tangent sigmoid transfer function was preferred as the activation function for the hidden layer (Hornik, 1991, Rankovic *et al*., 2010, Okkan and Mollamahmutoğlu, 2010).

The productivity of the prediction from the data-driven model structures was judged. The correlation between measured and predicted values were calculated with Pearson correlation coefficient. Smaller values of MAE (average of the absolute error) and MSE (mean square error) were used for determining better model productivity (Belayneh and Adamowski, 2012).

The total coliform, faecal coliform and faecal streptococci concentrations were monitored between 2010-2014 and modeled. Total measurement count used in modeling for each parameter was 663; 293 of these numbers belonged to Gemlik, 250 to Mudanya and 120 to Karacabey. In water quality studies related to ANN modeling, the data size ranged from 60 to 442

(Brion and Lingireddy, 2003, Mas *et al.,* 2007, Radojevic *et al*., 2013). The ANN model structures were prepared with the MATLAB ANN toolbox (Matlab, R2015). 80 % of the data was used for training, 10% for testing and 10% for validation (Brion and Lingireddy, 2003, Mas *et al*., 2007, Radojevic *et al.,* 2013). The cross validation technique in all ANN structures was used for the partition of data sets. The data set for training was used for calculating the error gradient and update biases together with the network weights. The data set for testing was used to validate the model productivity.

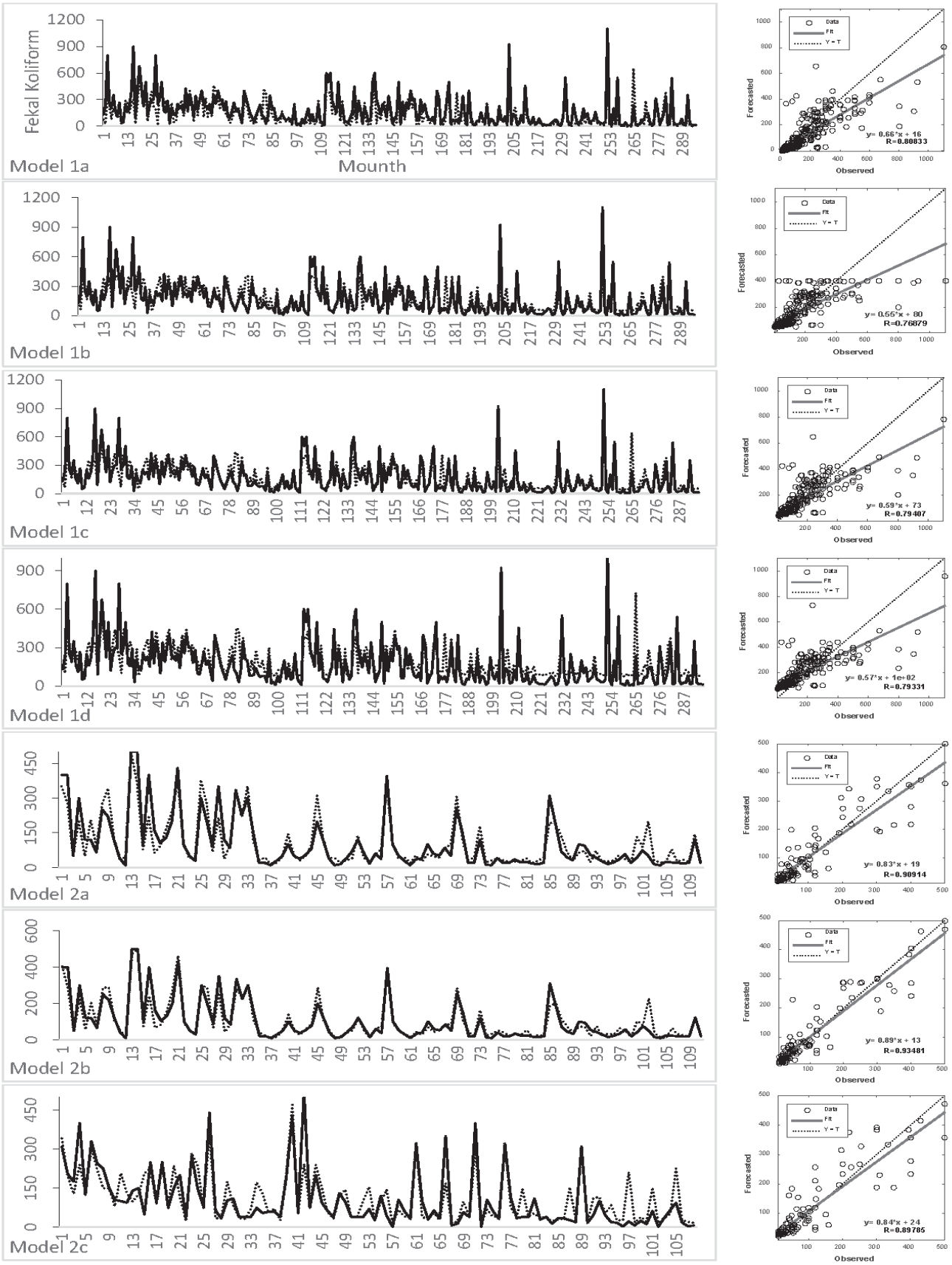
Monthly mean of concentrations were used in this study and 4 model structures (a, b, c, d) were applied for 3 coasts (Gemlik-Model 1, Karacabey-Model 2, Mudanya-Model 3). The results of different model structures with particular hidden layer count were compared. In a and b model structures, the input parameters were total coliform and faecal streptococci, in the c and d model structures total coliform was selected as input parameter. Faecal coliform was output parameter in all structures. Faecal coliforms were in coliform group bacteria and were indicator of faecal pollution. Total coliforms can be phytogenetic and soil born, but faecal coliforms arise from faecal pollution discharges (Casadevall and Pirofski, 2014, Alberts *et al*., 2002). For estimating the effects of sewage discharges, faecal coliforms were modeled as output parameter.

Results and Discussion

The productivity of the model structures was determined with different inputs and hidden neuron counts. The productivity

Table 1 : Created model structures and calculated RMSE and R values

Coast Models	Model No	Inputs	Outputs	ANN	RMSE			R			
					Training	Validation	Testing	Training	Validation	Testing	Whole Data Set
Gemlik (Model 1)	Model 1 a)	Total coliform Faecal streptococci	Faecal coliform	2-2-1	107.44	141.69	75.51	0.82	0.69	0.86	0.81
	Model 1 b)	Total coliform Faecal streptococci	Faecal coliform	2-5-1	117.53	67.36	73.01	0.75	0.81	0.91	0.77
	Model 1 c)	Total coliform	Faecal coliform	1-1-1	107.48	82.16	96.39	0.77	0.93	0.83	0.79
	Model 1 d)	Total coliform	Faecal coliform	1-3-1	102.46	173.23	65.11	0.82	0.57	0.92	0.79
Karacabey (Model 2)	Model 2 a)	Total coliform Faecal streptococci	Faecal coliform	2-2-1	53.06	25.21	19.07	0.91	0.86	0.98	0.91
	Model 2 b)	Total coliform Faecal streptococci	Faecal coliform	2-5-1	39.30	56.04	40.87	0.92	0.95	0.87	0.96
	Model 2c)	Total coliform	Faecal coliform	1-1-1	50.15	41.81	71.45	0.90	0.95	0.87	0.90
	Model 2d)	Total coliform	Faecal coliform	1-3-1	55.21	57.39	34.85	0.88	0.98	0.97	0.89
Mudanya (Model 3)	Model 3 a)	Total coliform Faecal streptococci	Faecal coliform	2-2-1	80.40	190.90	71.00	0.96	0.70	0.94	0.93
	Model 3 b)	Total coliform Faecal streptococci	Faecal coliform	2-5-1	105.65	54.84	59.14	0.93	0.95	0.93	0.93
	Model 3 c)	Total coliform	Faecal coliform	1-1-1	103.67	128.69	65.31	0.93	0.73	0.99	0.92
	Model 3 d)	Total coliform	Faecal coliform	1-3-1	83.22	180.15	138.78	0.93	0.91	0.98	0.93


Fekal Koliform
Mounth
Model 1a
Model 1b
Model 1c
Model 1d
Model 2a
Model 2b
Model 2c
Data
Fit
Y = T
Forecasted
Observed
y= 0.66*x + 16
R=0.80833
y= 0.55*x + 80
R=0.76879
y= 0.59*x + 73
R=0.79407
y= 0.57*x + 1e+02
R=0.79331
y= 0.83*x + 19
R=0.90914
y= 0.89*x + 13
R=0.93481
y= 0.84*x + 24
R=0.89785

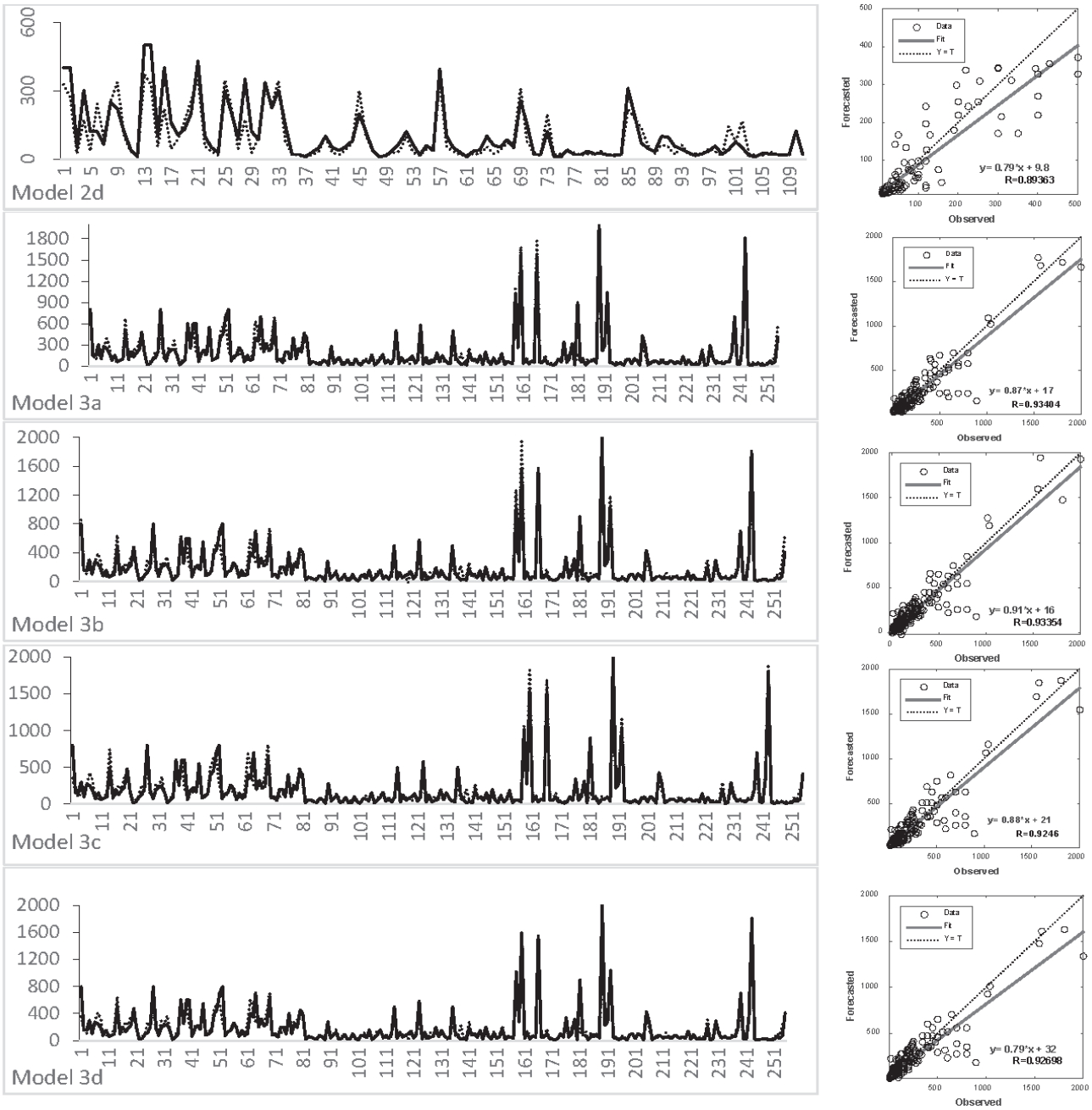


Fig. 2: Measured time series and predicted concentrations using created model structures and scatter plots.

function (RMSE) values and correlation coefficients (R) were calculated for training, validation, testing and whole data set. The values of productivity functions (RMSE) and correlation coefficients (R) belonging to the created model structures are given in Table 1.

Model 1 a and b structures predicted productively , because RMSE value of training of model 1-a was smaller than model 1-b, whole data set of R correlation number of model 1-a was greater than model 1-b. Enhancing of hidden neuron numbers in Model 1-b was not enhanced to whole data set of R correlation value. Model 1–c and d predicted productively. Hidden neuron numbers of model 1-d enhanced, therefore, RMSE values of training and testing data set reduced and R values of Model 1 c and d accrued; R value of whole data set remained unchanged. Gemlik Coasts "Model 1" had pretreatment and deep-sea discharge systems and domestic waste wasters were not treated adequately (Anonymous, 2014).

In Karacabey Coast, model 2- a and b structures ran more efficiently than model 1 a and b. RMSE calculated from model 2 -a and b training data were greater than model 1 -a and b, while other RMSE values were smaller than model 1. All R-values were greater than model 1.

The coasts of Karacabey were exposed to less contaminant source as compared to the Gulf of Gemlik (Anonymous, 2010). These coasts seem to be subject to lower pollution. They seem to have less pollutant diversity and, therefore, interaction of pollutants with each other may be less. For this reason, Model 2 run was more efficient. Increase in hidden neuron count in Model 2-b enhanced R-values. In model 2 c and d, which used total coliform as input parameters, run was more efficient than model 1 c and d. Increase in the hidden neuron numbers in Model, 2-d did not raise model R correlation values. The number of estimations in Karacabey coast were less than Gemlik, therefore RMSE values of Model 2 -a and b training data set were greater than Model 1 –a and b. Raising of hidden neuron numbers in Model 2-b enhanced R correlation values. In model 2 c and d, which used total coliform as input parameters, productivity was more than model 1 c and d.

In Mudanya coasts, model 3- a and b predicted productively. The calculation results of RMSE were smaller than model 1 a and b, greater than model 2 a and b. Model 3 a and b forecasted better than model 1 a and b such as model 2. Model 2 c and d were more productive than model 3 c and d because RMSE values of model 3 c and d were bigger than model 2 c and d. When model 3 c and d and model 1 c and d were compared, it was found that model 3 c and d were more efficient than model 1 c and d; because R values of model 3 c and d were bigger than model 1 c and d. The coasts of Mudanya as well as Karacabey are cleaner than the coasts of Gemlik. Model results of the coasts of Mudanya gave better results than Gemlik, since the amount and species of pollutant were less than Gemlik. It was found that all models of 3 coasts were forecasted productively. The calculated R and RMSE results showed appropriate values coinciding with other published literature (Soyupak *et al.,* 2003, Brion and Lingireddy, 2003, Yonar and Yalili Kilic, 2014, Mas and Ahlfeld, 2007, Okpokwasili and Nweke, 2005, Radojevic *et al.,* 2013). The measured time series and predicted concentrations using the created model structures and scatter plots are shown in Fig. 2.

This research demonstrated that the artificial neural network studies can be used to forecast coliform group bacteria in the sea coasts. The created model structures were run productively to predict the faecal coliform parameter in Gemlik, Karacabey and Mudanya coasts alongside the Sea of Marmara. Multilayer feedforward networks were used. All these trials indicated that to raise the count of the hidden layers has not always raised the model productivity. Kind and count of the inputs could be more influential. The usage of ANNs alongside the sea coastal areas can be used in a more beneficial way for the protection of natural habitats and human health, as well as sustainable coastal management studies.

References

Abrajano, T., A. Aksu, R. Hiscott and P. Mudie: Aspects of carbon isotope biogeochemistry of late Quaternary sediments from the Marmara Sea and Black Sea. *Marine Geology*, **190**, 151-164 (2002).

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter: Molecular Biology of the cell. 4th Edn., *Garland Science* (2002).

Anonymous: The Sea of Marmara Pollution Monitoring Project-Final Report. T.C. Ministry of Environment and Forestry, Environmental Management General Directorate, İstanbul (2010).

Anonymous: Bursa Metropolitan Municipality. Department of Coastal Services (2014).

Anonymous: Bursa Provincial Public Health Department. Retrieved from bhsm.gov.tr (Access date: 12.07.2018), (2015).

APHA: Standard Methods for the Examination of Water and Wastewater. 22nd Edn., APHA, AWWA, WPCF, Washington DC, USA (2012).

Belayneh, A. and J. Adamowski: Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *App. Comput. Intell. Soft Comp.*, **2012**, 1-13 (2012).

Brion, G.M. and S. Lingireddy: Artificial neural network modeling: A summary of successful applications relative to microbial water quality. *Water Sci. Technol.*, **47**, 235–240 (2003).

Casadevall, A. and L. Pirofski: Ditch the term pathogen. *Comment. Nature,* **516**, 165–166 (2014).

Cordier, M., J.P. Agúndez, W. Hecq and B.A. Hmaide: Guiding framework for ecosystem services monetization in ecological-economic modeling. *Ecosysl. Serv.*, **8**, 86-96 (2014).

Hagan, M. and M. Menhaj: Training feed forward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.*, **5**, 989–993 (1994).

Hornik, K.: Approximation capabilities of multilayer feedforward network. *Neural Networks ,* **4**, 251–257 (1991).

Karul, C., S. Soyupak, C.A., Akbay and N.E. Germen: Case studies on the use of neural networks in eutrophication modeling. *Ecological Modeling*, **134**, 145–152 (2000).

Katip, A.: The usage of artificial neural networks in microbial water quality Modeling: A case study from the Lake İznik. *Appl. Ecol. Environ. Res.*, **16**, 3897-3917 (2018).

Kim, T.W. and J.B. Valdes: Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *J. Hydrol. Engin.*, **8**, 319–328 (2003).

Mas, D. and D. Ahlfeld: Comparing artificial neural networks and regression models for predicting fecal coliform concentrations. *Hydrol. Sci.* J., **52**, 713-731 (2007).

MATLAB. Getting Started Guide R2015b. Retrieved from www.mathworks.com (Access date: 09.07.2015) (2015).

McHugh, D., C. Tanner, R. Mechoulam, R. Pertwee and R. Ross: Inhibition of human neutrophil chemotaxis by endogenous cannabinoids and phytocannabinoids: Evidence for a site distinct from CB1 and CB2. *Mol. Pharmacol.*, **73**, 441–50 (2008).

Mishra, A.K. and V.P. Singh: A review of drought concepts. *J. Hydrol.,* **391**, 202–216 (2010).

Morid, S., V. Smakhtin and K. Bagherzadeh: Drought forecasting using artificial neural networks and time series of drought indices. *Inter. J. Climatol.*, **27**, 2103–2111 (2007).

Okpokwasili, G.C. and C.O. Nweke: Microbial growth and substrate utilization kinetics. *Afri. J. Biotechnol.,* **5**, 305-307 (2005).

Okkan, U. and A. Mollamahmutoðlu: Daily runoff prediction of coruh river by artificial neural networks. Süleyman Demirel Univertsity *J. Natur. App. Sci*., **14**, 251-261 (2010).

Partal, T. and Ö. Kişi: Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *J. Hydrol.*, **342**, 199–212 (2007).

Radojevic, I.D., L. Čomić, V. Rankovic and M.D. Topuzovic: Applying neural networks for predicting the facultative oligotrophic bacteria in two reservoirs with different trophic state. *J. Environ. Protec. Ecol.*, **14**, 55–63 (2013).

Rankovic, V., J. Radulovic, I. Radojevic, A. Ostojic and L. Comic: Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia. *Ecol. Model*., **221**, 1239-1244 (2010).

Reimer, E. and S. Sodoudi: Climate model for the catchment area of the Elbe River with neuro-fuzzy method. EMS Annual Meeting Abstracts, European Meteorological Society, Nice, France (2004).

Soyupak, S., F. Karaer, H. Gurbuz, E. Kivrak, E. Senturk and A. Yazici: A neural network-based approach for calculating dissolved oxygen profiles in reservoirs. *Neural Computing and Applications*, **12**, 166–172 (2003).

TSWQR: Turkey Swimming Water Quality Regulation (76/160/EU), Ankara: Official Gazette No: 26048 and dated 09.01.2006 (2006).

Vijayashanthar, V., J. Qiao, Z. Zhu, A.M. ASCE, P. Entwistle and G. Yu: Modeling fecal indicator bacteria in urban waterways using artificial neural networks. *J. Environ. Enginee.,* **144**, 1-9 (2018).