ASSESSMENT AND PREDICTION OF WATER QUALITY PARAMETERS IN LAKE KÖYCEĞİZ USING ARTIFICIAL NEURAL NETWORK APPROACH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

OYA ÖZÇELİK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ENVIRONMENTAL ENGINEERING

NOVEMBER 2015

Approval of the thesis:

ASSESSMENT AND PREDICTION OF WATER QUALITY PARAMETERS IN LAKE KÖYCEĞİZ USING ARTIFICIAL NEURAL NETWORK APPROACH

submitted by **OYA ÖZÇELİK** in partial fulfillment of the requirements for the degree of **Master of Science in Environmental Engineering Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver
Dean, Graduate School of **Natural and Applied Sciences** _______________

Prof. Dr. Faika Dilek Sanin
Head of Department, **Environmental Engineering** _______________

Assoc. Prof. Dr. Emre Alp
Supervisor, **Environmental Engineering Dept., METU** _______________

Prof. Dr. Ayşegül Aksoy _______________
Co-Supervisor, **Environmental Engineering Dept., METU**

Examining Committee Members:

Prof. Dr. Kahraman Ünlü
Environmental Engineering Dept., METU _______________

Assoc. Prof. Dr. Emre Alp
Environmental Engineering Dept., METU _______________

Prof. Dr. Ayşegül Aksoy
Environmental Engineering Dept., METU _______________

Assist. Prof. Dr. Barış Kaymak
Environmental Engineering Dept., METU _______________

Assoc. Prof. Dr. Merih Aydınalp Köksal
Environmental Engineering Dept., Hacettepe University _______________

Date: Nov 27, 2015

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Oya ÖZÇELİK

Signature:

ABSTRACT

ASSESSMENT AND PREDICTION OF WATER QUALITY PARAMETERS IN LAKE KÖYCEĞİZ USING ARTIFICIAL NEURAL NETWORK APPROACH

Özçelik, Oya
M.S., Department of Environmental Engineering
Supervisor: Assoc. Prof. Dr. Emre Alp
Co-Supervisor: Prof. Dr. Ayşegül Aksoy

November 2015, 148 pages

Water quality monitoring plays a significant role on water resources management and planning. European Union (EU) Water Framework Directive aims to achieve "good status" for all waters.. Within the adaptation period to EU, Turkey aims to improve its water quality monitoring network; however, this will lead to time, budget and human resources problems. Purpose of this thesis is the application of a method that will provide water quality assessment (WQA) under limited budget and data conditions. The method was applied to Lake Köycegiz that is located in Muğla at junction point of Mediterranean and Aegean regions and has nearly 5500 ha surface area

WQA based on multivariate statistical analysis (hypothesis testing and principal component analysis (PCA)) was conducted and water quality status of Lake Köyceğiz and its tributaries were determined based on Surface Water Quality Management Regulation. The results showed that the lake is eutrophic and although there are seasonal differences for water quality parameters, there is no spatial difference between different locations of Lake Köycegiz. In addition, PCA explains the main pollution causes to the lake as fertilizer use in the area or wastewater discharge.

Artificial neural network (ANN) approach was performed to predict water quality parameters in the lake using monthly measured water quality parameters of tributaries as input. Different input combinations and performance criteria were tried to find the best predictions. Results revealed low error and high correlation values between measured and estimated parameters. These results indicate great potential of ANNs to predict water quality parameters.

Keywords: Lake Köyceğiz, ANN, water quality, hypothesis testing, principal component analysis

ÖZ

KÖYCEĞİZ GÖLÜNÜN SU KALİTESİ PARAMETRELERİNİN DEĞERLENDİRİLMESİ VE YAPAY SİNİR AĞI YÖNTEMİYLE TAHMİN EDİLMESİ

Özçelik, Oya
Yüksek Lisans, Çevre Mühendisliği Bölümü
Tez Yöneticisi: Doç. Dr. Emre Alp
Ortak Tez Yöneticisi: Prof. Dr. Ayşegül Aksoy

Kasım 2015, 148 sayfa

Su kalitesinin izlenmesi, su kaynaklarının yönetilmesi ve planlanmasında önemli bir rol oynamaktadır. Avrupa Birliği (AB) SU Çerçeve Direktifi bütün suların “iyi durumda” olmasını hedeflemektedir. AB ile uyum sürecinde, Türkiye’nin su kalitesi izleme ağını geliştirmeyi amaçlaması zaman, bütçe ve insan kaynağı problemlerini de beraberinde getirmektedir. Bu tezin amacı kısıtlı bütçe ve veri koşullarında Su Kalitesi Değerlendirmesi (SDK) sağlayan bir metodun uygulanmasıdır. Metot, Ege ve Akdeniz bölgelerinin kesişim noktasındaki Muğla’da yer alan ve yaklaşık 5500 ha yüzey alanına sahip Köyceğiz Göl’üne uygulanmıştır.

Çok değişkenli istatistiksel analizler (hipotez testi ve temel bileşenler analizi (TBA)) ile SKD yapılmış ve Köyceğiz Gölü ve gölü besleyen nehirlerin su kalitesi durumu Yüzeysel Su Kalitesi Yönetimi Yönetmeliği’ne göre tespit edilmiştir. Sonuçlar, gölün ötrofik olduğunu ve su kalitesi parametrelerinde mevsimsel farklılıklar gözlenmesine rağmen Köyceğiz Gölü’nün farklı noktalarında mekansal farklılıkların olmadığını göstermiştir. Ayrıca TBA, göle etki eden ana kirliliğin bölgedeki gübre kullanımından veya atıksu deşarjından kaynaklandığını açıklamaktadır.

Göldeki su kalitesi parametrelerinin tahmin edilmesi için, gölün nehir kollarının aylık ölçülen su kalitesi parametrelerini girdi olarak alan, yapay sinir ağları (YSA) yaklaşımı kullanılmıştır. En iyi tahminleri elde etmek amacıyla farklı girdi kombinasyonları ve performans kriterleri denenmiştir. Hedeflenen ve tahmin edilen parametreler arasında düşük hata ve yüksek korelasyon değerleri elde edilmiştir. Bu sonuçlar, YSA'nın su kalitesi parametrelerini tahmin etmekteki büyük potansiyelini ortaya koymaktadır.

Anahtar Kelimeler: Köyceğiz Gölü, YSA, su kalitesi, hipotez testi, temel bileşenler analizi

To my family, especially to my husband and our daughter…

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to my supervisor Assoc. Prof. Dr. Emre Alp for his guidance and support throughout my study.

I would also like to give my appreciations and thanks to my co-supervisor Prof. Dr. Ayşegül Aksoy for her guidance, encouragement, and unlimited support throughout my thesis study.

I would also thank to my committee members Prof. Dr. Kahraman Ünlü, Assist. Prof. Dr. Barış Kaymak, Assoc. Prof. Dr. Merih Aydınalp Köksal for their time spent in evaluating my thesis.

Special thanks to Ezgi Tanrıkul Demir and Gökçe Somun Odabaşı for their patience, friendship, encouragements and contributions.

I would express my deep appreciation to my parents Fahriye and Mehmet Ali Özçelik, for their love, support and endless patience through all my life. In addition, I am thankful to my brothers Oğuz and Kemal, Sevim Gedik, my sisters, my cousins, and other members of my family for their helps whenever I needed.

I grateful my dear husband Dr. Osman Serdar Gedik for his love, endless patience, moral and technical support. Without him, this success was impossible. Final thank to our little daughter Ayşe Sare for her waiting patiently through my study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AME	Absolute Maximum Error
ANN	Artificial Neural Network
ART	Adaptive Resonance Theory
BOD	Biological Oxygen Demand
COD	Chemical Oxygen Demand
DO	Dissolved Oxygen
FC	Fecal Coliforms
GA	Genetic Algorithm
GDPNH	General Directorate for Preservation of Natural Heritage
LB	Lake Beach
LC	Lake Center
MAE	Mean of Absolute Error
MSE	Mean Square Error
NC	Namnam Creek
NMAE	Normalized Mean Absolute Error
OECD	Organisation for Economic Co-operation and Development
PC	Principal Component
PCA	Principal Component Analysis
SOM	Self Organizing Map
SPSS	Statistical Package for the Social Sciences
SWQMR	Surface Water Quality Management Regulation
T	Temperature
TC	Total Coliforms
TN	Total Nitrogen
TP	Total Phosphorus
WFD	Water Framework Directive
WQA	Water Quality Assessment
YC	Yuvarlakçay Creek

CHAPTER 1

INTRODUCTION

Turkey has been in the proses of harmonizing the requirements of European Union (EU) to becoming a member state. To achieve this goal, the EU regulations required to be adopted and implemented. In Turkey's program for alignment with the EU Acquis, one of the priority areas is "environment" area. Within the adjustment program of Turkey to EU, one of the most important legal developments in this area is the issue of a Framework Water Law, based on the adoption of EU's Water Framework Directive (WFD) (2000/60/EU). WFD is the most comprehensive and important directive on water quality that entered into force in 2000. The main aim of the WFD is forming a directive for the protection of surface waters, transition waters, coastal waters and groundwater, and provision of "good status" for all waters. WFD describes how the monitoring should be done during the river basin management plan (Dalkılıç, et al., 2008).

In Turkey, for prevention of the pollution, considerable amount sub-directives of the WFD were implemented to the Turkish Legislation within the scope of the EU works. These are the Protection of Waters against Pollution caused by Nitrates from Agricultural Sources (18.02.2004 - O.J. No: 25377), the Quality of Surface Water Intended for the Abstraction of Drinking Water Regulation (17.02.2005 - O.J. No: 25730), Control of Pollution by Dangerous Substances in Water and its Environment Regulation (26.11.2005 - O.J. No: 26005), Urban Wastewater Treatment Regulation (08.01.2006 – O.J. No: 26047), the Quality of Surface Water Intended for the Abstraction of Drinking Water Regulation (17.02.2005 - O.J. No: 25730) and Bathing Water Quality Regulation (09.02.2006 – O.J. No: 26048) (OECD, 2008).

The first project for implementation of WFD to Turkey is MATRA project. In this project, WFD methodology was examined and it was decided to establish a roadmap for future applications. For MATRA project, Big Menderes Watershed was selected as pilot a watershed. During the project, “characterization”, “cause and effect”, “ecology” and “precautions” titles were developed. In addition to MATRA project, two separate applications that are SMART (Sustainable Management of Scarce Resources in the Coastal Zone) and OPTIMA (Optimization for Sustainable Water Management) projects were conducted at Gediz Watershed. SMART project models the water resources in terms of quality and quantity; examines the different development scenarios; demographic, socioeconomic and technologic progress indicators. Due to the examination of politics scenarios, SMART is contributing to the creation of management decision. Second project OPTIMA in Gediz Waterhed aims to find conciliating solutions to adverse demands. Moreover, OPTIMA target to make a new approach for water resources management then aims to examine, to assess and to develop this approach to increase productivity. In this way, the project provides a strong method scientifically (Dalkılıç, et al., 2008).

In addition, some projects about monitoring and sustainable management of water resources completed in the last years (2014, 2015) in the scope of the Ministry of Forestry and Water Affairs are indicated below:

- River Basin Management Plans (EU)
- Basin Monitoring and Determining Reference Points Project (EU)
- The National Water Information System
- Identification and Improvement of Water Resources Quality Criteria and Project
- Automatic Continuous Measurement Stations Setup Project
- Gördes Watershed Catchment Special Provision Identification Project
- Dry Spell Management and Action Plans Preparation Project
- Watershed Management Projects
- Atatürk Reservoir Protection Research and Development Project

- Climate Change Effects on Water Resources Project
- Assessment of Drinking Water Resources and Treatment Plants Project (EU)
- Integrated Watershed Management Project (The Ministry of Forestry and Water Affairs, 2015).

Monitoring of water quality has great importance in determining the status of water bodies, for information about realization of specified measures or for effectiveness of realized measures. As described in the previous paragraphs, Turkey has made a significant progress on this subject and many projects are still ongoing. The biggest challenge for Turkey in this subject is development of monitoring network and in this process, performing of water quality assessment correctly with limiting number of data. Hence, the purpose of the thesis is to assess and predict the water quality parameters in Lake Köyceğiz using limited data. Firstly, some of the important multivariate statistical methods (hypothesis testing, correlation analysis and principal component analysis (PCA)) were used to assess the water quality and later Artificial Neural Network (ANN) approach was applied to estimate the parameters of Lake Köyceğiz with using limited data.

Lake Köyceğiz is located in city of Köyceğiz in Muğla where the junction point of Mediterranean and Aegean regions. Lake Köyceğiz, which is the most important lake of the region and the first scientifically defined meromictic lake of Turkey, is at south part of the city. Because of its location of being in a very important touristic area with rich ecological value, Lake Köyceğiz was declared as Köyceğiz–Dalyan Specially Protected Environmental Area (SPEA) in the year 1988. It is fed by a lot of creeks with different sizes and flows to Mediterranean Sea through Dalyan Channel. Namnam and Yuvarlakçay Creeks are main creeks feeding the lake. In addition, there are many medium scaled creeks, seasonal creeks and 10 drainage channels contributing to the lake. Flowrate of Namnam and Yuvarlakçay Creeks decreases in summer months due to the Mediterranean climate effect and irrigation. In Lake Köyceğiz Watershed, General Directorate for Preservation of Natural Heritage (GDPNH) has carried out a water quality project since the year 2006. GDPNH has taken monthly samples from creek, lake and sea areas for identification of water quality parameters. According to the project, high concentrations of nitrite nitrogen and coliform have been observed which indicates an eutrophication problem (Environmental Protection Agency for Special Areas, 2007).

Eutrophication is a complex process that accelerates the growth of algae on higher forms of plant life due to the enrichment of water nutrients (especially compounds of nitrogen and phosphorus) and inducing an undesirable disturbance to the balance of living organisms in the water and to the quality of the water. Large input of nutrients to a water body is the main cause of eutrophication of a water body and high levels of phytoplankton biomass leading to algal blooms due to the imbalance in the food web is the main effect of it (WHO, et al., 2002). Enrichment of water bodies is often dramatically increased by nutrients from human activities although natural origin nutrients can enrich water bodies. Nutrients come to water bodies from point and non-point sources. Control of non-point sources of nutrients is more difficult than point sources because they come many different locations and sources.

The thesis consists of five chapters. In Chapter 1, aim of the thesis and information about methodology in general terms are presented.

In Chapter 2, information about study area and available data for the study are given. Firstly, water quality parameters of Namnam and Yuvarlakçay creeks are classified based on Surface Water Quality Management Regulation (SWQMR). Then, trophic status of Lake Köyceğiz is identified based on SWQMR and Organisation for Economic Co-operation and Development (OECD). Secondly, water quality parameters of the lake are compared with respect to seasons and sampling location points by using hypothesis testing. Thirdly, correlation analysis is conducted to the water quality parameters of the lake and the creeks. As last analysis of the assessment chapter, principal component analysis (PCA) is applied to the lake water quality parameters to assess the contributions to the lake. At the final of the chapter, some studies about Köyceğiz and water quality assessment analysis are mentioned.

In Chapter 3, detailed infromation is given about ANNs. ANNs are powerful tools, which have been widely used in many fields such as classification, forecasting and modeling of areas, for encountered problems in real life. ANN is a biologically inspired computing methodology that has the learning skill by imitating the learning method of human brain. ANNs are black boxes in which the unknown model parameters (connection weights of an ANN structure) are adjusted to obtain the best matching between input data set and corresponding output data set (Gümrah, et al., 2000). In addition, application of ANN methodology in Lake Köyceğiz is presented in Chapter 3. For the ANN model, different data sets (input scenarios) are used to predict water quality parameters in Lake Köyceğiz for a period of 2008-2014. Six different input scenarios are created to find best prediction result. Chapter 4 includes conclusions and discussion of the study. In addition, recommendations of the study are presented in Chapter 5.

CHAPTER 2

ASSESSMENT OF WATER QUALITY PAREMETERS IN LAKE KÖYCEĞİZ

In this chapter, water quality assessment is conducted for Lake Köyceğiz; Namnam and Yuvarlakçay Creeks feeding the lake. Information about study area and available data is detailed at the beginning of the chapter. Next, the creeks and the lake are classified based on water quality criteria in Surface Water Quality Management Regulation (SWQMR). Then, available parameters for the sampling points of the lake and creeks are examined seasonally and spatially. Later, correlation analysis is conducted for available data to understand relationship between parameters and locations of them. Finally, principal component analysis is applied to the data. With this analysis, related parameters are collected under principal components, which are unrelated with others, and pollutant factor sources for the creeks and the lake are tried to explain based on these components.

2.1 Study Area

City of Köyceğiz is located at the junction point of Mediterranean and Aegean regions, 75 km away from Muğla (Figure 2.1). Köyceğiz having 1,758 km^2 surface area is a touristic area with rich and natural beauties. The region is formed by different landforms that are around Lake Köyceğiz. Köyceğiz has a population with 34,027 as of 2104. 65% of the population lives at villages and they live off agriculture, animal husbandry, forestry and tourism. County's climate and topography is suitable for polyculture agriculture. Another source of income is migratory beekeeping at region. Grey Mullet fish farming is done at Dalyan Channel connecting Lake Köyceğiz to Mediterranean Sea. (Köyceğiz Municipality, 2015)

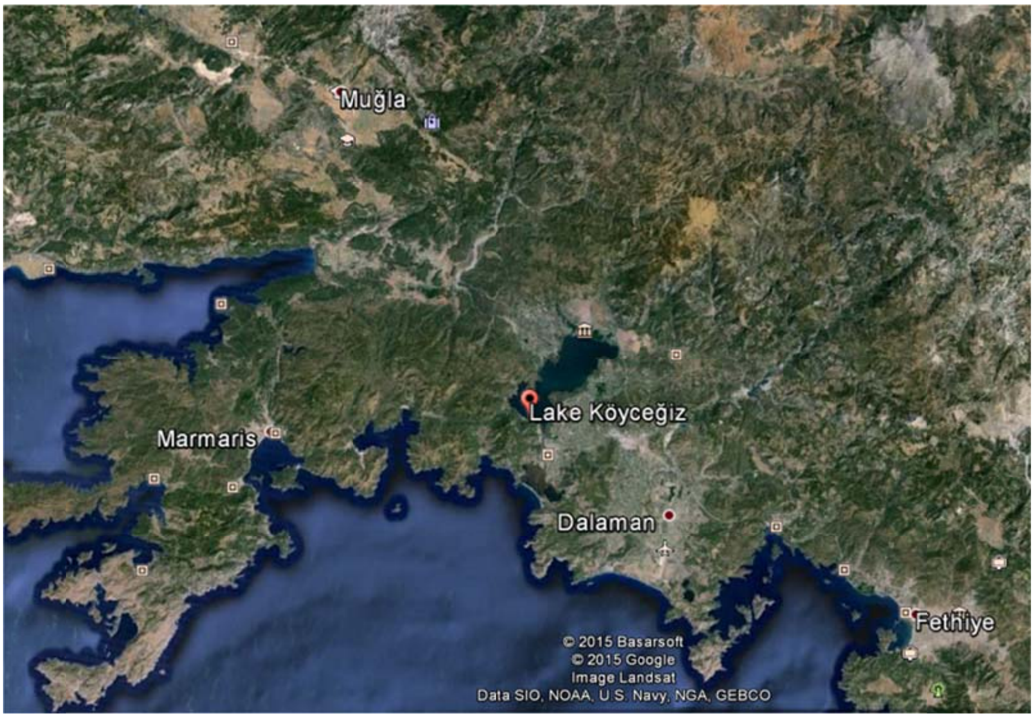


Figure 2.1: Location of Köyceğiz (Google Earth, 2013)

There are different tourism activities in the region. Hot-cold thermal springs in Sultaniye Village and mud baths at Çandır Village are important centers for health tourism. Also, safari tourism at Yayla Village and Gökçeova attracts attention. Dalaman Channel is important for trekking and rafting. Ekincik located at Mediterrenean seaside of Köyceğiz is another place of interest for water surfing, water ski and swimming with its long beach, brilliant bay and marina. Besides, around the lake, infrastructure works are continuing for bike tracks between Köyceğiz and Dalyan (Köyceğiz Municipality, 2015).

Lake Köyceğiz located at south part of the city is the most important lake in the region. In 1988, Lake Köyceğiz was declared as Köyceğiz–Dalyan Specially Protected Environmental Area (SPEA) because of its location in a very important touristic area with rich ecological value.

Lake Köyceğiz has nearly 5500 ha surface area and drainage area with 1072.70 km^2. Length and width of the lake is 12-13 km and 5-6 km, respectively. Lake is formed by two basins as Köyceğiz Basin (north side) and Sultaniye Basin (south side). Köyceğiz Basin is bigger but shallower than Sultaniye Basin. It has a 24 m average depth while Sultaniye Basin has a 28 - 30 m. Two basins are connected with a quite narrow and shallow pass having 4 m average depth (Gürel, et al., 2002). Bathymetric map of Köyceğiz Lake is given in Figure 2.2.

.

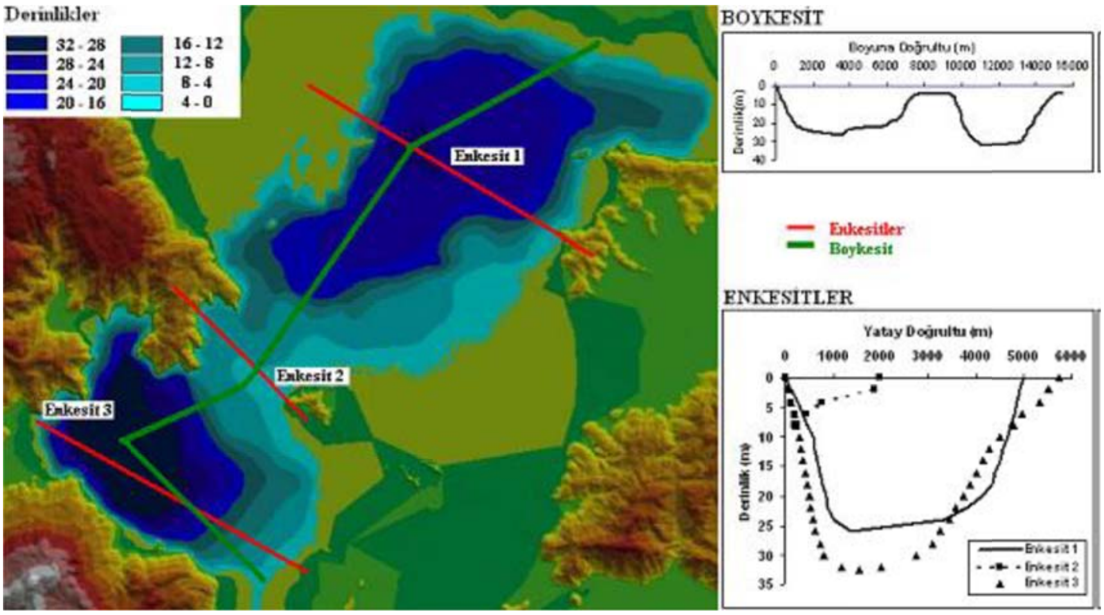


Figure 2.2: Bathymetric map of Lake Köyceğiz (Gürel, et al., 2002)

Lake Köyceğiz is the first scientifically defined meromictic lake of Turkey. There are two separate water layers in the lake. Upper layer is consist of fresh water while bottom one is brackish water (Kazancı, et al., 1992).

Lake Köyceğiz is fed by many creeks with different sizes and flows to Mediterranean Sea through Dalyan Channel. Namnam and Yuvarlakçay are the two biggest creeks contributing the lake. Besides, Asardere, Araplar, Kargıcak Creeks are medium scaled ones and Değirmendere, Çakmakdere, Cehennembendi are three seasonal creeks feeding Lake Köyceğiz (Özdemir, et al., 2003) (Türedi, 2006). Moreover, there are ten drainage channels as Kapnıç, Maden iskelesi, Sarıöz, Arıtma, Hamitköy, Kargıcak, Asartepe, Kocaöz, Mera and Gavurbeli drainage channels in the region. Lake is surrounded by Hamitköy (Namnam), Köyceğiz, Yuvarlak and Dalyan Delta Flats (Buhan, 1996).

One of the two big creeks Namnam Creek has nearly 13 km mainstream length and 25-30 m average depth. Drainage area of Namnam Creek is 607 km^2 which is the largest drainage area of the Köyceğiz Watershed (Özdemir, et al., 2003). Yuvarlakçay, which is other big creek in the watershed area, has 21 km length and 8-14 m average depth. Yuvarlakçay Creek is mainly fed by ground waters. There are fish farms on Yuvarlakçay Creek. Flowrate of these creeks decreases in summer months due to the irrigation and Mediterranean climate effect. Dalyan Channel, which is a bridge between Mediterranean Sea and Lake Köyceğiz, has 10 m length. Its width changes between 5 and 70 m. Furthermore, its depth again changes between 1 and 6 m. Dalyan Channel has a delta area with approximately 1,150 ha and it mixes to sea with a 100 m width and nearly 4-5 km length beach (Türedi, 2006).

Köyceğiz Wastewater Treatment Plant is operated at the north side of the lake since 2003 (GDPNH, 2015). There two important fault lines at Lake Köyceğiz. First of them is at south of the lake from northwest to northeast. There are thermal sources on this line that is Sultaniye Thermal Spring. This makes the lake popular for hot thermal activities. Second line passes through the center of lake in north-south direction. 45-70 m depths were observed on this fault line in the lake. Two meteorological stations in the area are Köyceğiz and Dalaman stations. According to these stations, 56% of annual precipitation is observed in winter months. Fall, spring and summer traces the winter with 23%, 17% and 4% of annual precipitation, respectively (Türedi, 2006).

Considering the land use of Köyceğiz watershed, three main vegetation is dominant. 70% of the watershed is formed by forest, remain is covered by reedy fields and scrubs. (Köyceğiz Municipality, 2015)

Table 2.1: Land use of Köyceğiz watershed (Gönenç, et al., 2002)

Land Use Type	Area (ha)	Ratio (%)
Agricultural Area	13,080	8.6
Scrubs	887	0.6
Forest	107,413	70.5
Barren land	30,994	20.3

Köyceğiz is the first region of Aegean Region based on Citrus farming. Lemon, orange, grapefruit, mandarin, pomegranate and tomato are main farm products of the region (Gönenç, et al., 2002). In addition, sweet gum tree, which is a special tree, grows only in Köyceğiz and Honduras all over the World (Köyceğiz Municipality, 2015).

In the literature, there are many studies using Köyceğiz-Dalyan Watershed, Lake Köyceğiz and its tributaries as study area.

Kazancı et al. (2000) studied the longitudinal and seasonal distribution of macro invertebrates and the physical and chemical variables of Yuvarlakçay Creek between April 1992 and April 1993. The diversity, frequency, dominance, abundance and similarity records of macro invertebrates are obtained in the study. Moreover, for water quality assessment, Belgian Biotic Index is used for the first time in Turkey. Minimum, maximum and mean NH_4-N concentration for mouth point of Yuvarlakçay is found as 0.00 mg/L, 0.27 mg/L and 0.04 mg/L, respectively. It is concluded that there are continuous, slight and moderate organic pollution in

Yuvarlakçay Creek. In addition, necessity of the water quality monitoring for the creek is emphasized for protection the water quality of Lake Köyceğiz.

Ahmet et al. (2009) assessed the environmental impact of the flow-through rainbow trout farm on Yuvarlakçay creek. In the study, bimonthly sampled data is used between February and December 2006 from the inlet and outlet of the farm. The study is pointed that changes in T, TP, TN, BOD_5 and COD concentrations between the inlet and outlet of the fish farm are insignificant ($p>0.05$) while, the pH and DO values of the inlet are significantly decreased in the effluent water ($p<0.05$). In addition, applied drum filters used for effluent treatment method help to reduce organic matter amount in water. Because of the use of high quality extruded feed and using a solid waste removal technique, the sampled farm can be regarded as a good example for the rest of the trout farms in Turkey.

Ekdal et al. (2011) studied for coastal ecological analysis of Köyceğiz – Dalyan Lagoon by using an integrated modeling system consisting water quality and watershed models. The Soil and Water Assessment Tool (SWAT) was used as the watershed model for prediction of the non-point pollutant loads arising from the study area, and Water Quality Analysis Simulation Program (WASP) was conducted for the water quality modeling studies. Estimated flow and pollutant loads obtained from SWAT are used as the input to WASP. The study shows that obtained simulation results are able to catch the general trend of the monitoring data. In addition, in most cases, the results for the boundaries fit at higher level than other regions of the system.

Nedim et al. (2012) investigates the water quality of Namnam Creek which feeds Lake Köyceğiz. Data between May 2010 and April 2011 is used to find physico-chemical parameters. The results are obtained as temperature (12.04-25.25 °C), pH (7.23-8.68), EC (348-675 µS/cm), DO (4.53-7.43 mg/L), BOD (BDL-4.65 mg/L), NO_2-N (BDL-0.48 mg/L), NO_3-N (BDL-26.10 mg/L), NH_4-N (BDL-2.54 mg/L), andortho-phosphate (BDL-3.85 mg/L). Results were assessed according to the Water

Pollution Control Regulation. The assessment revealed that agricultural pollutants and anthropogenic factors are mostly reasons of the pollution in Namnam Creek.

2.2 Available Data

In this section, data utilized for estimating the water quality parameters and the meteorological data are introduced.

2.2.1 Water Quality Data

In Lake Köyceğiz Watershed, General Directorate for Preservation of Natural Heritage (GDPNH) has carried out a water quality project since the year 2006. Within the concept of the project, GDPNH has taken monthly samples from creeks, lake and sea areas. Creek sampling points are Namnam Creek before Lake Köyceğiz, Namnam Creek Döğüşbelen Site, Dalyan Channel before Mediterranean Sea, Yuvarlakçay before and after the fish farm, Yuvarlakçay before Lake Köyceğiz. Lake Köyceğiz sampling points are Lake Center, Lake Beach (north point of the lake) and Dalyan Entrance (south point of the lake). Other lakes sampling points are Alagöl Lake center and Sülüngür Lake center points. Sea area sampling points are Dalyan Ağzı Beach, Ekincik Bay and İztuzu Beach points. Drainage channels sampling points are Kapnıç, Maden iskelesi, Sarıöz, Arıtma, Hamitköy, Kargıcak, Asartepe, Kocaöz, Mera and Gavurbeli drainage channels points. Locations of the water quality sampling stations are given in Figure 2.3.

.

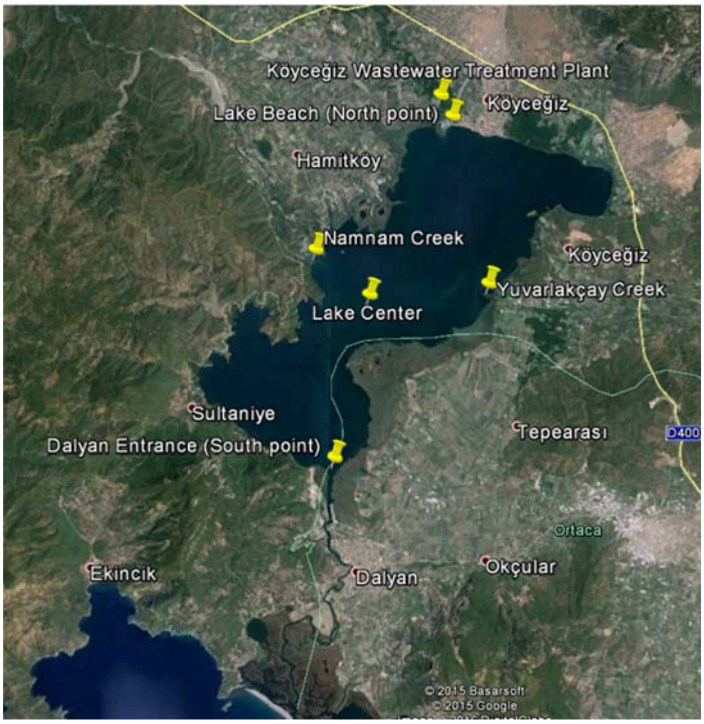


Figure 2.3: Water quality sampling stations (Google Earth, 2013)

In this study, Lake Köyceğiz and Namnam and Yuvarlakçay Creeks that feed the Lake Köyceğiz are assessed. Other lakes and sea areas are after the Lake Köyceğiz therefore they cannot influence the water quality of the Lake Köyceğiz. Drainage channels have data from the year 2010, and hence; number of data samples is insufficient for the study. Because of these reasons, parameters obtained from Namnam and Yuvarlakçay Creeks, which feed and therefore affect the water quality of the Lake Köyceğiz, are used as input data. Parameters obtained from Lake Köyceğiz sampling points (Lake center, Dalyan Channel entrance and Lake beach) are used as output data to assess the water quality of the Lake Köyceğiz.

Within the carried project, samples taken from creeks are analyzed for identification of T, pH, dissolved oxygen, ammonium nitrogen, nitrite nitrogen, nitrate nitrogen, total phosphorus, COD, fecal coliforms and total coliforms parameters since the year 2006. However, nitrate nitrogen and COD data are missing for the year 2008. In addition, electrical conductivity, salinity, flowrate and total suspended solids parameters are recorded since 2008; total kjeldahl nitrogen parameter data is available since 2009. Furthermore, total nitrogen, total dissolved solids, dissolved reactive phosphorus, orthophosphate, organic phosphorus and BOD parameters are analyzed for very short periods and flowrate date has a lot of missing values. Samples obtained from Lake Köyceğiz sampling points are analyzed for determination of pH, dissolved oxygen, total nitrogen, total suspended solids, total phosphorus, COD and total coliforms parameters. However, there are missing values in the COD parameter data for nearly two years. In addition, temperature, electrical conductivity, salinity and chlorophyll-a parameters are measured since 2008. Also, ammonium nitrogen, nitrite nitrogen, nitrate nitrogen total kjeldahl nitrogen, orthophosphate, organic phosphorus, BOD, fecal coliforms, secchi disc and light transmittance are analyzed for very short periods. For Dalyan Channel entrance sampling point, there are data for temperature, pH, electrical conductivity, salinity, dissolved oxygen, ammonium nitrogen, nitrite nitrogen and total phosphorus, fecal and total coliforms since 2008.

Table 2.2: Used sampling points, all available and used parameters in the study

Sampling points	All available parameters	Used parameters in the study
Namnam and Yuvarlakçay Creeks	T, pH, EC, salinity, Q, DO, NH_4-N, NO_2-N, NO_3-N, TKN, TN, TDS, TSS, DRP, PO_4^{-3}, OP, TP, BOD, COD, FC, TC	T, pH, EC,DO, NH_4_N, TP (as input data of ANN)
Dalyan Channel Entrance (south point of the lake)	T, pH, EC, salinity, Q, DO, NH_4-N, NO_2-N, NO_3-N, TKN, TN, TDS, TSS, DRP, PO_4^{-3}, OP, TP, BOD, COD, FC, TC	DO, NH_4_N, TP (as output data of ANN)
Lake center	T, pH, EC, salinity, DO, NH_4-N, NO_2-N, NO_3-N, TKN, TN, TDS, TSS, DRP, PO_4^{-3}, OP, TP, BOD, COD, FC, TC, Chl-a, SD, LT	DO, TN, TP, Chl-a (as output data of ANN)

Sampling points	All available parameters	Used parameters in the study
Lake beach (north point of the lake)	T, pH, EC, salinity, DO, NH_4-N, NO_2-N, NO_3-N, TKN, TN, TDS, TSS, DRP, PO_4^{-3}, OP, TP, BOD, COD, FC, TC, Chl-a, SD, LT	DO, TN, TP, Chl-a, TC (as output data of ANN)

(T: temperature; EC: electrical conductivity, Q: flowrate; DO: dissolved oxygen; NH_4-N: ammonium nitrogen; NO_2-N: nitrite nitrogen; NO_3-N: nitrate nitrogen; TKN: total kjeldahl nitrogen; TN: total nitrogen; TDS: total dissolved solids; TSS: total suspended solids; DRP: dissolved reactive phosphorus; PO_4^{-3}: orthophosphate; OP: organic phosphorus; TP: total phosphorus; BOD: biological oxygen demand; COD: chemical oxygen demand; FC: fecal coliforms; TC: total coliforms; Chl-a: Chlorophyll-a; SD: secchi disc; LT: Light transmittance)

In this thesis study, monthly data between the years 2008 and 2014 are used. In this way, study can be conducted with more parameter. There are 63 data for each used parameter. In Table 2.2, all available parameters and used parameters in the study based on completeness of them are summarized. The detailed descriptive statistics of the used water quality data are given in APPENDIX A and the means values of the long term water quality measurements are given in Table and Table 2.4.

Table 2.3: Mean values of the long term water quality measurements for the creeks

Parameter	Unit	Yuvarlakçay Creek	Namnam Creek
Temperature	°C	18.84	20.03
pH	-	8.25	8.44
Electrical Conductivity	µS/cm	560.41	789.73
Dissolved Oxygen	mg/L	8.51	8.20
Ammonium Nitrogen	mg/L	0.05	0.04
Total Phosphorus	mg/L	0.04	0.05

Table 2.4: Mean values of the long term water quality measurements for the lake sampling points

Parameter	Unit	Dalyan Entrance	Lake Center	Lake Beach
Temperature	°C	21.77	21.42	22.22
pH	-	8.50	8.52	8.48
Electrical Conductivity	µS/cm	5,459	5,080	-
Dissolved Oxygen	mg/L	7.94	7.97	7.84
Ammonium Nitrogen		0.05		
Total Nitrogen	mg/L		0.72	0.69
Total Phosphorus	mg/L	0.07	0.05	0.03
Chlorophyll-a	µg/L		1.97	1.80
Total Coliforms	CFU/100 mL		-	1,196.13

As can be seen from the tables, means of EC for lake points are higher than the creeks' due to the meromictic property of the lake.

Figure 2.4 shows temporal variations of temperature for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points. As can be seen from graph, all sampling points have similar trend for temperature parameter. Temperature of Yuvarlakçay Creek does not exceed 25.90 °C while; other points reach the higher values than 30 °C.

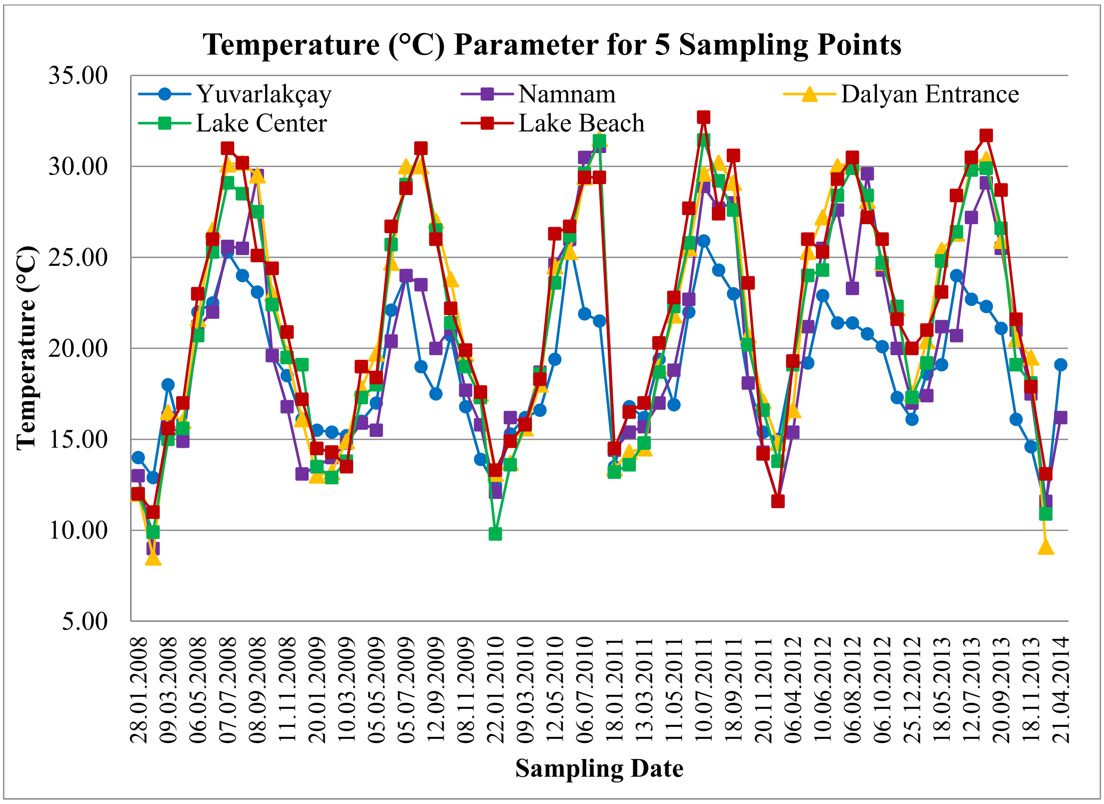


Figure 2.4: Temperature parameter for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points

Figure 2.5 illustrates the pH parameter measurements for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points. All pH values are between 6.6 and 9.6. In January 2008, all locations except Namnam and Lake Beach have least pH values. PH parameter of Namnam drops to the least value in November 2008. Lake Beach has the least value that 7.0 in September 2012. In addition, Namnam has highest pH values in May and August 2012.

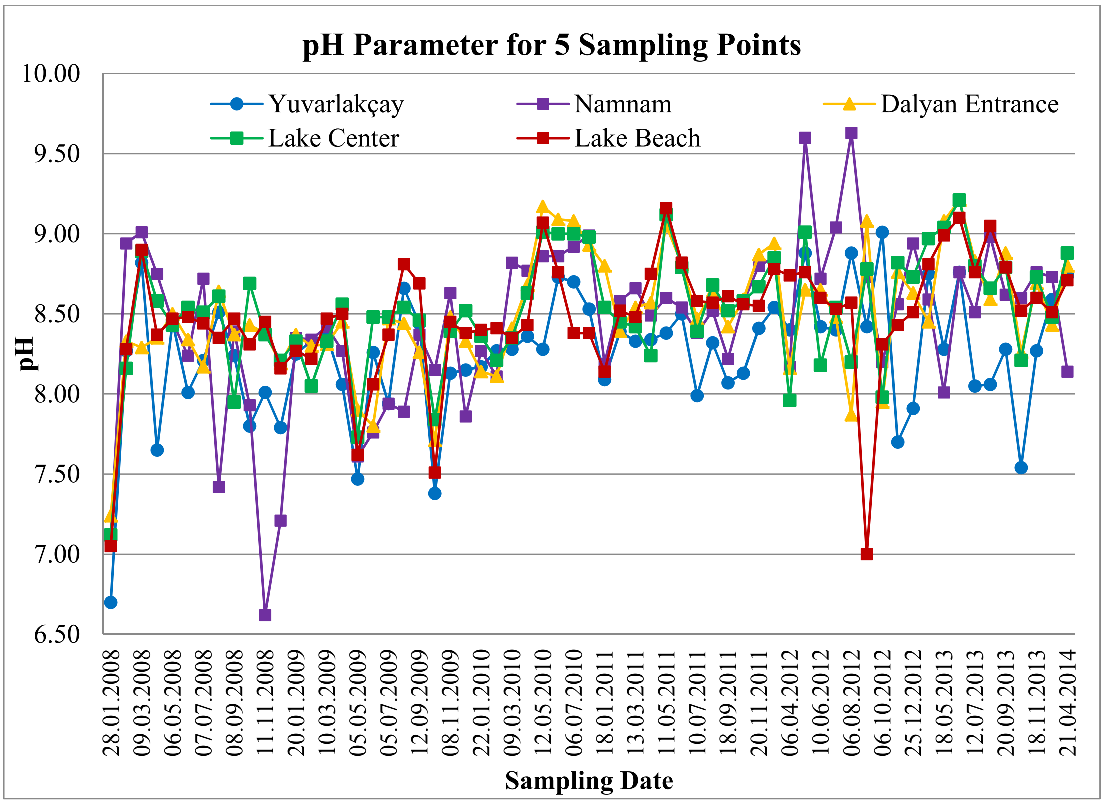


Figure 2.5: pH parameter for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points

Figure 2.6 shows the EC variations for Lake Center, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points. EC values are higher at lake sampling point than the creeks. This can be resulted from the meromictic property of the lake. Existence of salty water in the lake increases the EC values. Dalyan Entrance has the highest EC value in October 2008. This shows the powerful effect of sea water on the lake. In addition, Namnam has it's the highest EC value with 8,970 µS/cm in September 2008. In August 2013, Yuvarlakçay Creek has it's the highest EC value with 5,470 µS/cm.

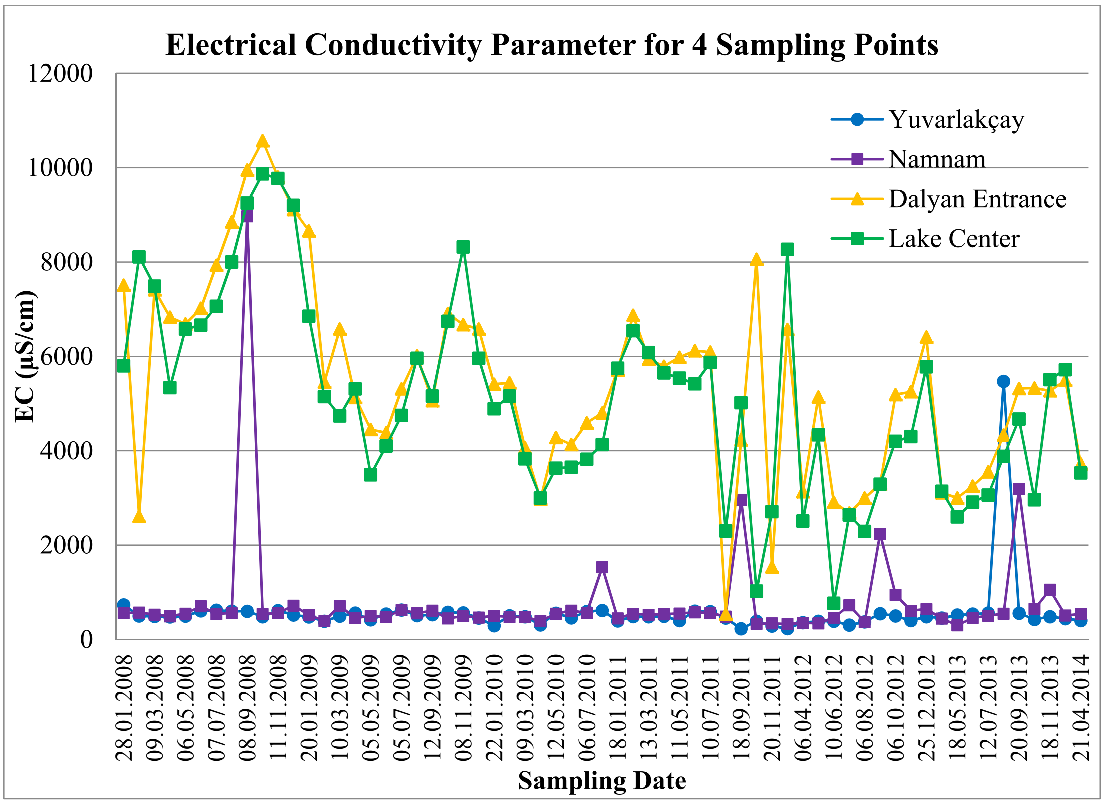


Figure 2.6: Electrical conductivity parameter for Lake Center, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points

Figure 2.7 shows temporal DO variations for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points. All sampling points show similar trends for DO parameter while; Lake Beach has the highest DO concentration with 15.52 mg/L. DO concentrations above the saturation value indicates the algal blooms in this region.

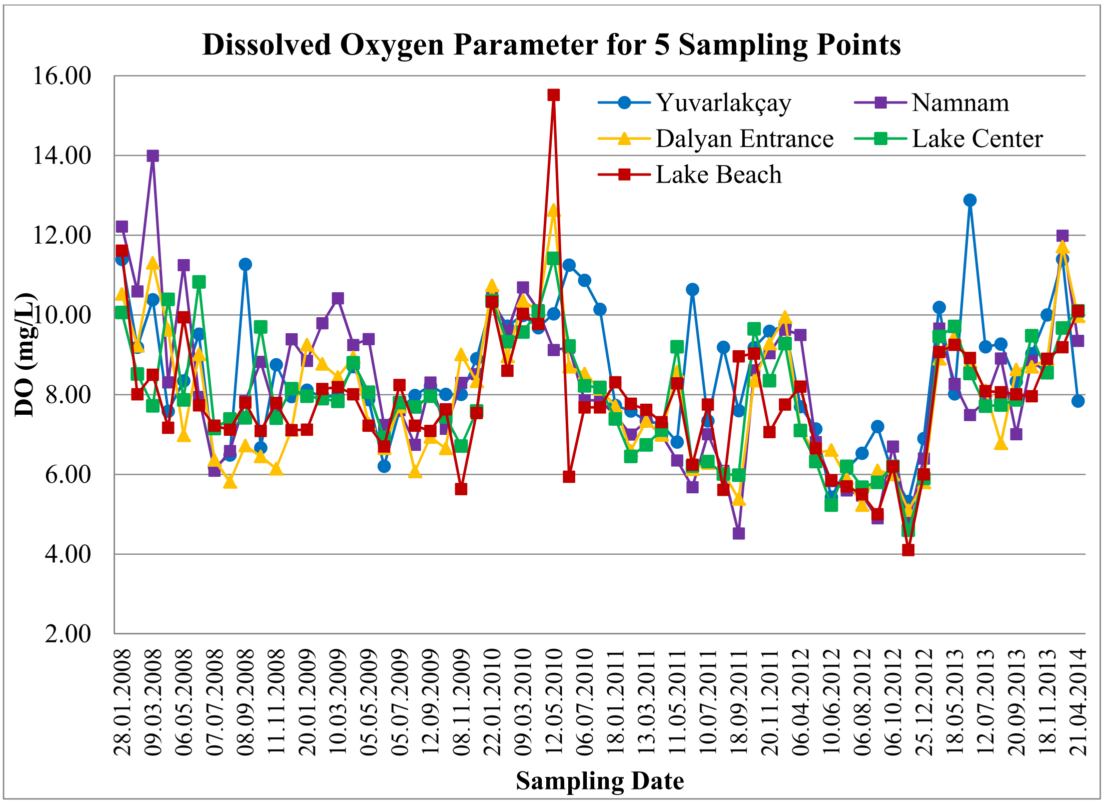


Figure 2.7: Dissolved oxygen parameter for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points

Figure 2.8 depicts NH_4-N measurements for Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points. In February 2008, Dalyan Entrance sampling point has the highest NH_4-N concentration with a 0.300 mg/L. In natural waters, NH_4-N is rapidly oxidized to nitrate and nitrite nitrogen with requirement of adequate DO concentration. High DO concentration of 9.23 mg/L in the sampling date for Dalyan Entrance might be because of algal activities that are trigger by high NH_4-N concentrations. The high NH_4-N concentration can be caused by a sewage discharge near the sampling point. In addition, NH_4-N concentration of same location is 0.2 mg/L in July 2013. Namnam Creek has nearly 0.180 mg/L NH_4-N concentration in February 2008, April and October 2011. NH_4-N concentration of Yuvarlakçay Creek exceeds the 0.200 mg/L in January 2008. Moreover, measurements are recorded as

"<0.010 mg/L" in the years 2009, 2010, 2013 and recorded as "<0.100 mg/L" in the year 2012.

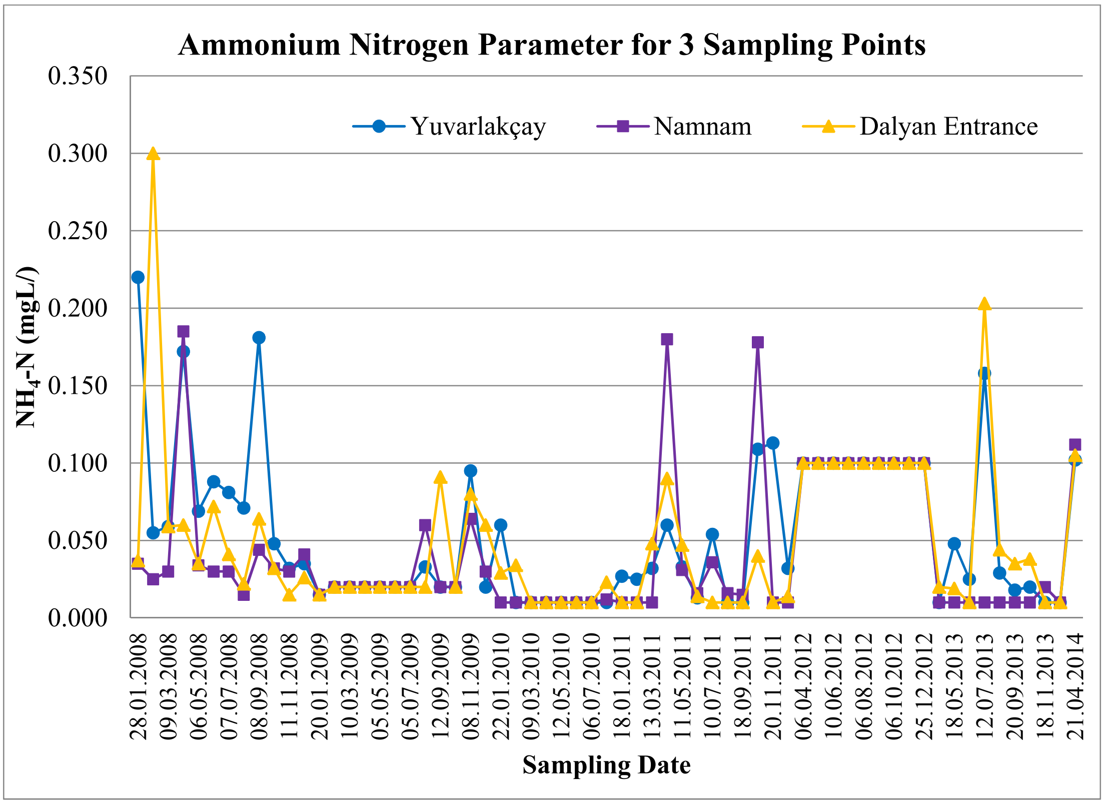


Figure 2.8: Ammonium nitrogen parameter for Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points

Figure 2.9 shows the graph of TN variations for Lake Center and Lake Beach. In general, two sampling points of the lake have similar trend for TN parameter. However, TN concentration of Lake Center reaches to 3.67 mg/L in June 2008. This peak might be due to surface runoff results from rainy spring months which leads to nonpoint source pollution includes the fertilizer etc. from agricultural activities.

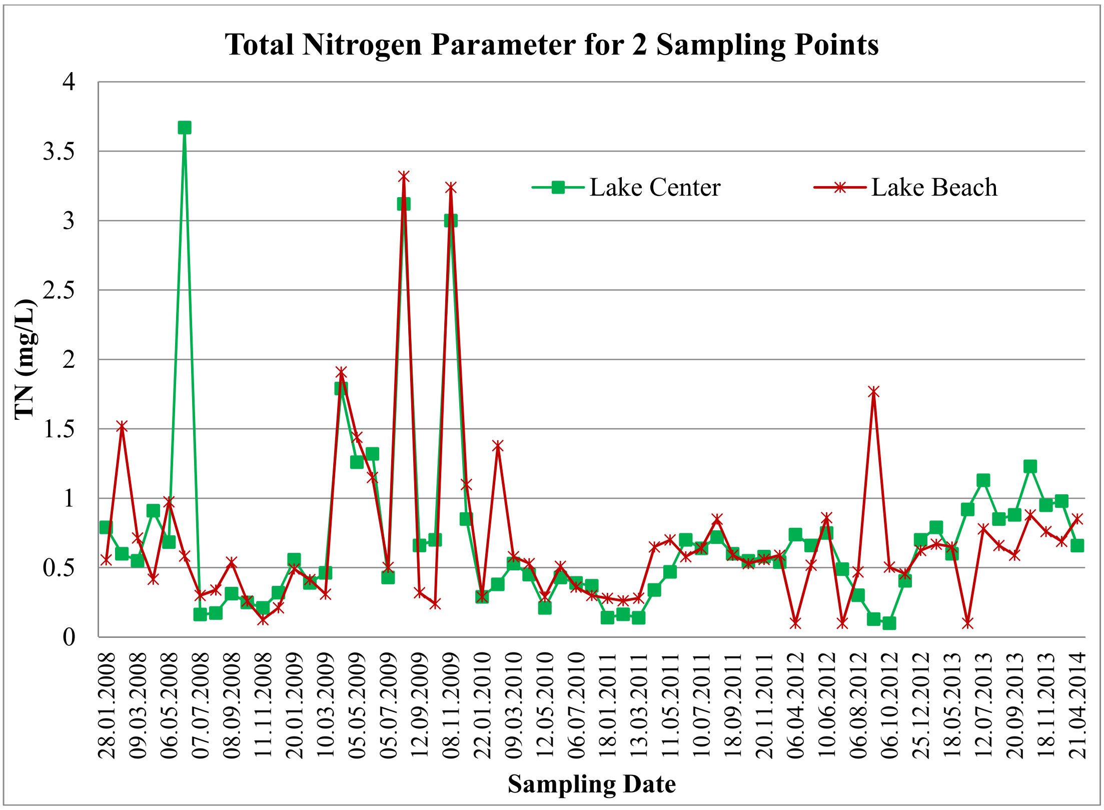


Figure 2.9: Total nitrogen parameter for Lake Center and Lake Beach sampling points

Figure 2.10 shows TP parameter values for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points. All locations except Lake Beach have peaks in the year 2012. The straight lines at 0.010 mg/L value in the graph are due to the measurement records as “<0.010 mg/L”.

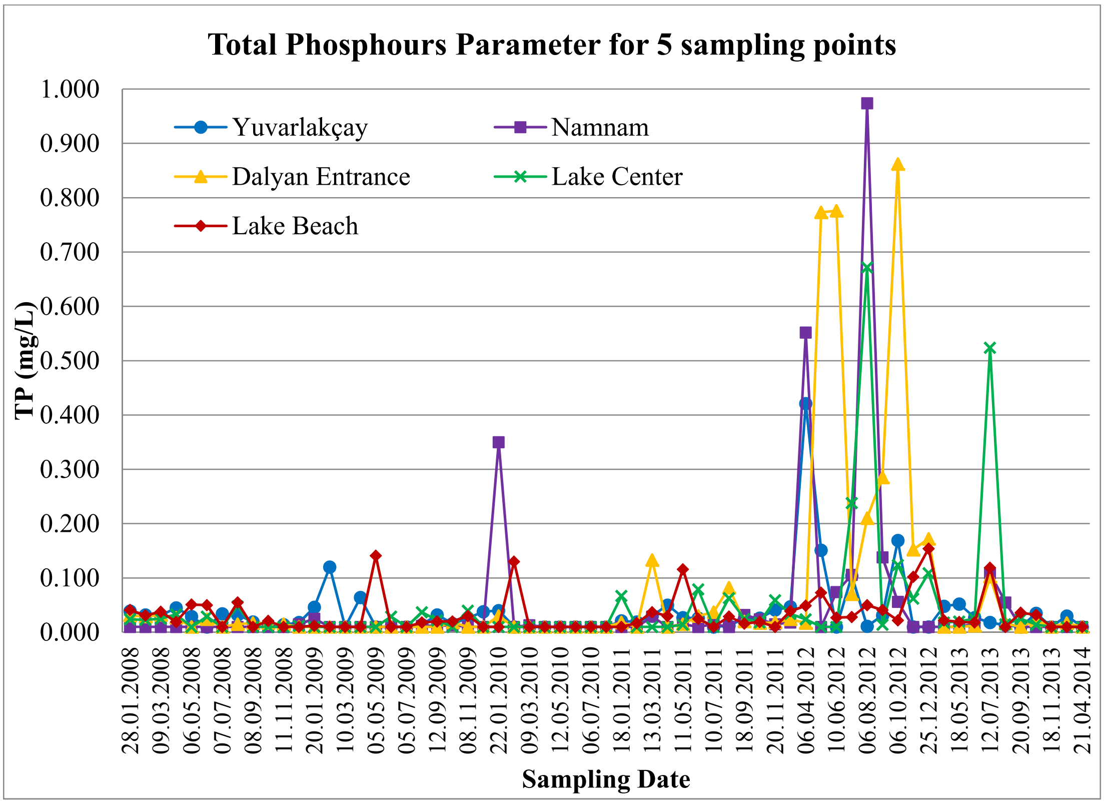


Figure 2.10: Total phosphorus parameter for Lake Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks sampling points

Figure 2.11 shows the chlorophyll-a concentrations in Lake Center and Lake Beach. In general, chlorophyll-a concentrations are high in the year 2008 for both of the lake sampling points. In February 2009, chlorophyll-a concentration of Lake Center reaches to 9.60 μg/L. It can be indicator of algal bloom due to the low DO concentration in this month. After June 2012, chlorophyll-a concentrations decreases to very low concentrations. The straight lines in the graph are due to the measurement records as "<0.10 μg/L" and "<1.00 μg/L".

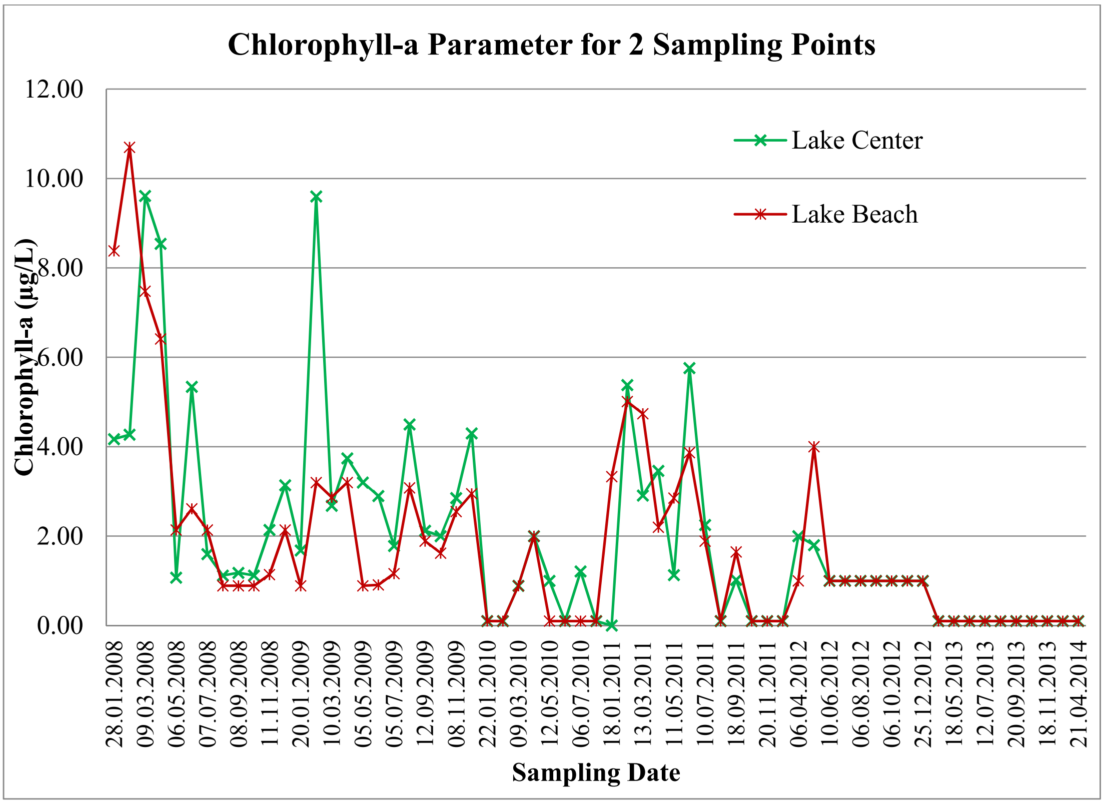


Figure 2.11: Chlorophyll-a concentrations for Lake Center and Lake Beach sampling points

Finally, Figure 2.12 shows the total coliforms measurements in Lake Beach. TC concentrations are lower than the given compulsory value (10,000/100 mL) in the Bathing Water Quality Directive (76/160/EC) until July 2013. However, TC concentration of the beach is 15,000/100 mL in July 2013 and 22,000/100 mL in April 2014. These peak values can be resulted from failure of the Köyceğiz Waste Water Treatment Plant near the beach.

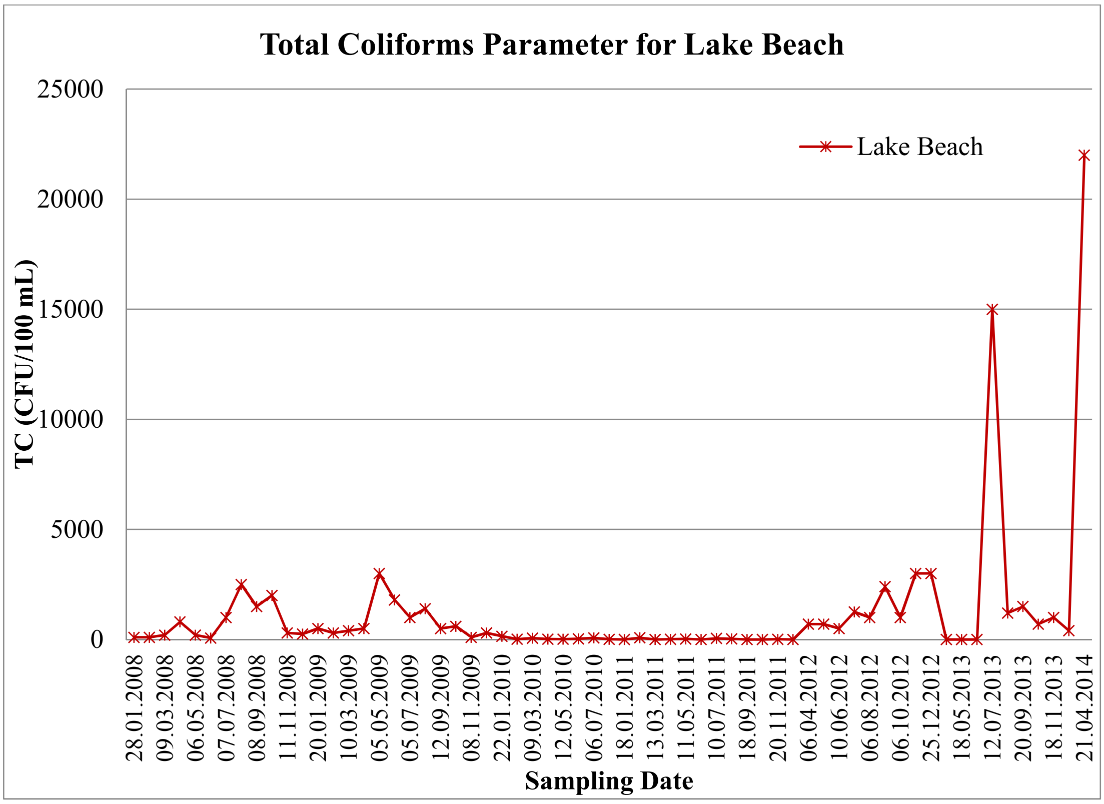


Figure 2.12: Total coliforms parameter for Lake Beach sampling point

2.2.2 Meteorological Data

Hourly and daily wind speed data, daily average moisture data, hourly and daily temperature data, daily total global solar radiation data, daily and monthly total precipitation data are available from two meteorological stations in the area that are at Köyceğiz and Dalaman stations. Temperature data is not used due to existence of water temperature data. Precipitation data is used as meteorological data due to its effects on the water quality. Precipitation amounts of the last fifteen days including sampling date are summed and used as precipitation data in the study. Consequently, 63 precipitation data samples are obtained. Precipitation and temperature data of Köyceğiz are given in Figure 2.13 and Figure 2.14, respectively. Peaks in precipitation data are in fall and winter months. There are missing temperature measurements at the zero values.

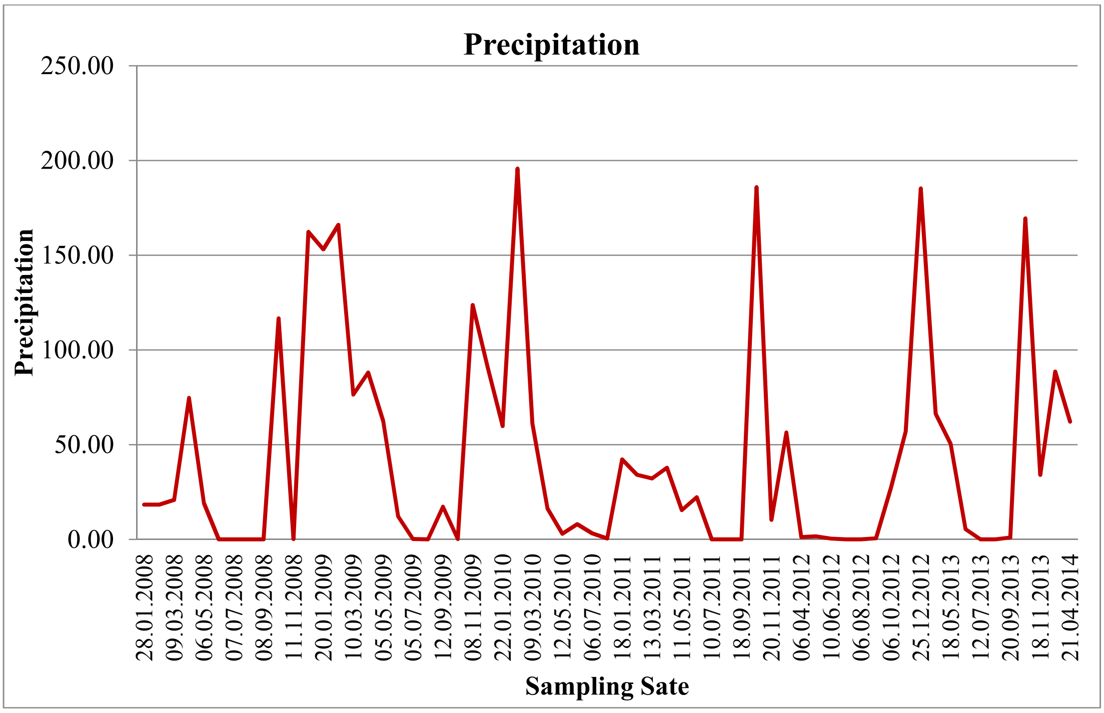


Figure 2.13: Precipitation amounts of Köyceğiz

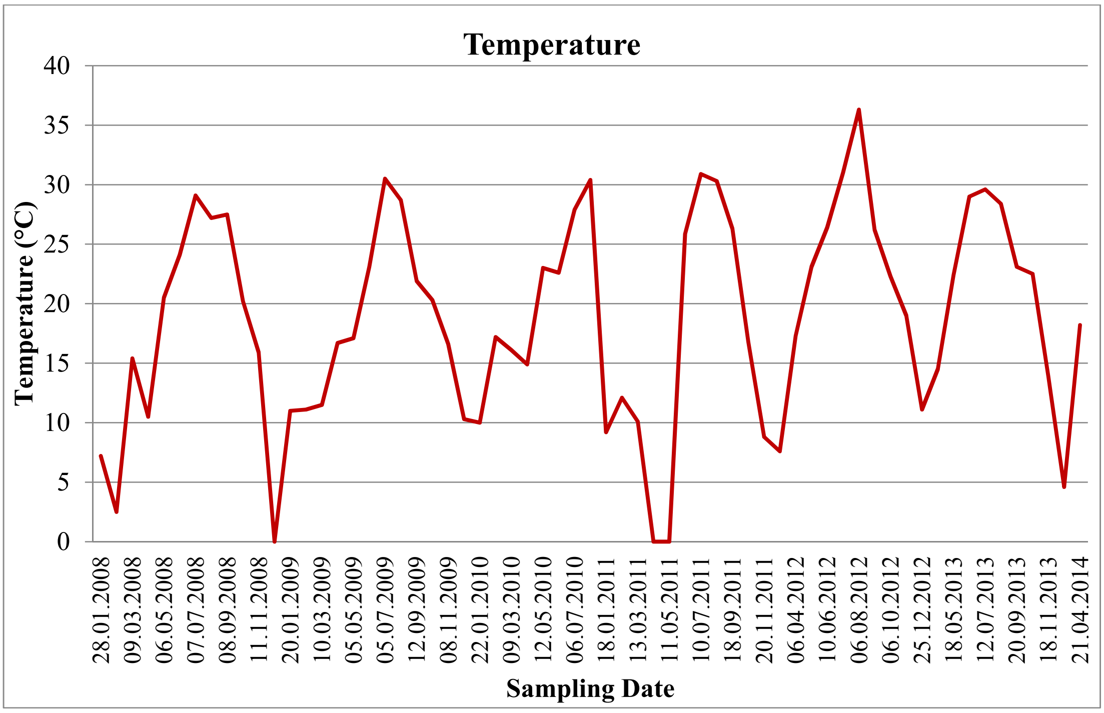


Figure 2.14: Temperature variations of Köyceğiz

2.3 Water Quality Classification

In this part of the water quality assessment chapter, measured water quality parameters of creeks (Yuvarlakçay and Namnam) and lake points are classified based on Surface Water Quality Management Regulation (SWQMR). For classification calculations, statistical guidelines in SWQMR are followed. The followed statistical methods are described under subtitles in related parameter.

2.3.1 Example calculation for creek classification

In this part, water quality parameters of Namnam and Yuvarlakçay Creeks are classified according to the 95% values of related parameters given in SWQMR as mentioned in Appendix 5, Table 5 of this regulation. 95% values are calculated by determining the types of statistical probability distributions. Statistical Package for the Social Sciences (SPSS) program is used for this purpose. For all parameters measured in all points, normal and log-normal distributions are tested. Then, for each point of each parameter, the optimum distribution is determined and 95% values are calculated. For identification of best probability distribution, statistical parameters are obtained then, results are confirmed graphically.

To explain briefly, the normal distribution is the most important and the most widely used example of a continuous random variable. Normal distribution has a bell shaped curve with the center of the bell located at the arithmetic mean (μ). Standard deviation (σ) controls the depth of this bell. A normal distribution is completely defined by arithmetic mean (μ) and its variance (σ^2) also; the distribution is expressed as N (μ, σ^2). Standard normal distribution has a mean of zero and a variance of one.

For a random variable x, when log(x) probability distribution is normal, probability distribution of x is defined as logarithmic normal distribution (log-normal distribution). Probability density function for lognormal distribution is given in (3.1)

$$f(y) = \frac{1}{y}\frac{1}{\sqrt{2\pi\sigma}}\exp(-\frac{1}{2}(\frac{\ln(y)-\mu}{\sigma})^2 \quad (3.1)$$

where μ is mean value; σ is standard deviation (Webb, et al., 2005).

If a parameter has normal distribution, its mean and median values should be very close to each other and skewness value of this parameter should be close to zero. Skewness value indicates the distribution side and closeness of it to zero means symmetric distribution of parameter. In the other case, the distribution deviates from the symmetry.

For calculation of 95% values, equation (3.2) is used. In standardized Z normal distribution table, Z_{95} value means the z value corresponding to 0.95 probability value (=1.65).

$$\mathbf{z} = \frac{\mathbf{x}-\boldsymbol{\mu}}{\boldsymbol{\sigma}} \quad (3.2)$$

where *x* is x_{95} value; μ is mean value; σ is standard deviation and z is z_{95} value.

Example: Yuvarlakçay Creek – Total Phosphorus

Total nitrogen parameter for Yuvarlakçay Creek is available from year 2006 to 2014. Obtained statistical values, which are mean, minimum, maximum, median, deviation and skewness coefficient values for normal and log-normal distributions, are given in Table 2.5.

Table 2.5: Statistical values for Yuvarlakçay Creek Total Phosphorus parameter

Statistical Values	**Normal Distribution**	**Log-normal Distribution**
Mean value	0.057	-1.579
Median value	0.025	-1.602
Standard deviation	0.148	0.419
Variance	0.022	0.176
Skewness coefficient	6.106	1.588
X_{95} value	2.001	0.130

According to the given information about distributions above, calculated statistical values for total phosphorus parameter of Yuvarlakçay Creek shows that the distribution is closer to log-normal distribution.

For confirmation of this result, probability plots are obtained via SPSS program. Probability plots are for variable's cumulative proportions against the cumulative proportions of any of a number of test distributions. They are generally used for determination of whether the distribution of a variable matches a given distribution. The points cluster around a straight line if variable matches the test distribution. P-P graphs plot the cumulative probabilities (values range from 0 to 1), with observed probabilities (cumulative proportion of cases). In normal P-P plots, normal distribution of data set is on y-axis (expected cumulative probabilities); in log-normal P-P plots, log-normal distribution of data set is on y-axis (expected cumulative probabilities). Based on these normal and log-normal P-P plots, data set are confirmed graphically related their distribution type. Normal and log-normal P-P plots for total phosphorus parameter of Yuvarlakçay Creek obtained by using SPSS program are given in Figure 2.15.

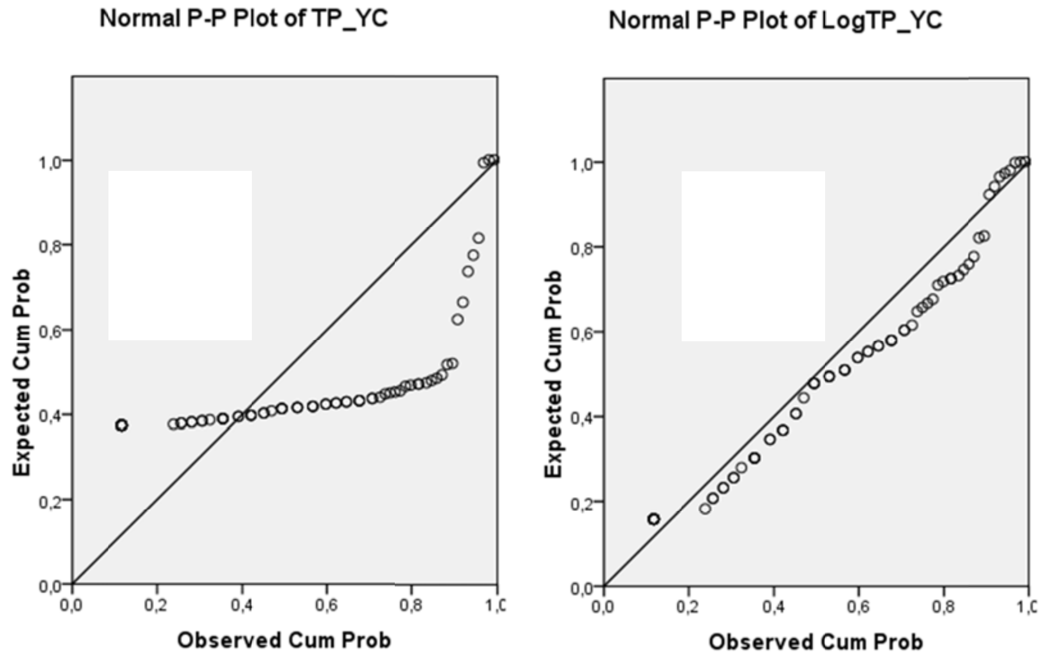


Figure 2.15: Normal and log-normal distribution plots for Yuvarlakçay Creek total phosphorus parameter

As it is revealed by the Figure 2.15, observed and expected probability values show a straighter distribution in log-normal P-P plot. Thus, 95% values were calculated based on log-normal distribution with equation 3.2.

$$z = \frac{x-\mu}{\sigma} \tag{3.2.}$$

where x is x_{95} value; μ is mean value; σ is standard deviation and z is z_{95} value.

In standardized z normal distribution table, z_{95} value means the z value corresponding to 0.95 probability value (=1.65).

μ: mean value = -1579 (for log-normal distribution)
σ: standard deviation = 0.419 (for log-normal distribution)
According to the equation (3.2):

1.65 = (log“95 percent value”(-1.579))/0.419

95 percent value = $10^{1.65x0.419-(1.579)}$ = 0.130 mg/L

According to the regulation, values higher than this calculated 95% value of 0.130 are removed from the data set, then arithmetic mean of remaining data is calculated. After calculation, 0.027 value is found. 0.027 mg/L TP value is in Class I according to SWQMR as mentioned in Appendix 5, Table 5 of this regulation, which is given in Table 2.6.

Table 2.6: Quality criteria based on intra-continental surface water resources (Ministry of Forestry and Water Affairs, 2012)

Water Quality Parameters	**Water Quality Classes**			
	I	**II**	**III**	**IV**
General Conditions				
Temperature (^{o}C)	≤ 25	≤ 25	≤ 30	> 30
pH	6.5-8.5	6.5-8.5	6.0-9.0	other than 6.0-9.0
Electrical Conductivity (µS/cm)	< 400	400-1000	1001-3000	> 3000
Color	**TCU 436 nm:** 1.5 **TCU 525 nm:** 1.2 **TCU 620 nm:** 0.8	**TCU 436 nm:** 3 **TCU 525 nm:** 2.4 **TCU 620 nm:** 1.7	**TCU 436 nm:** 4.3 **TCU 525 nm:** 3.7 **TCU 620 nm:** 2.5	**TCU 436 nm:** 5 **TCU 525 nm:** 4.2 **TCU 620 nm:** 2.8
(A) Oxygenation Parameters				
Dissolved Oxygen (mg O_2/L)[a]	> 8	6-8	3-6	< 3
Oxygen Saturation (%)[a]	90	70-90	40-70	< 40
Chemical Oxygen Demand (COD) (mg/L)	< 25	25-50	50-70	> 70
Biological Oxygen Demand (BOD_5) (mg/L)	< 4	4-8	8-20	> 20
B) Nutrient Parameters				
Ammonium nitrogen (mg NH_4^+-N/L)	< 0.2[b]	0.2-1[b]	1-2[b]	> 2
Nitrite nitrogen (mg NO_2^--N/L)	< 0.002	0.002-0.01	0.01-0.05	> 0.05

Water Quality Parameters	Water Quality Classes			
	I	II	III	IV
Nitrate nitrogen (mg NO_3^--N/L)	< 5	5-10	10-20	> 20
Total Kjeldahl-nitrogen (mg/L)	0.5	1.5	5	> 5
Total phosphorus (mg P/L)	< 0.03	0.03-0.16	0.16-0.65	> 0.65
C) Trace Elements (Metals)				
Mercury (μg Hg/L)	< 0.1	0.1-0.5	0.5-2	> 2
Cadmium (μg Cd/L)	≤ 2	2-5	5-7	> 7
Lead (μg Pb/L)	≤10	10-20	20-50	> 50
Copper (μg Cu/L)	≤20	20-50	50-200	> 200
Nickel (μg Ni/L)	≤20	20-50	50-200	> 200
Zink (μg Zn/L)	≤200	200-500	500-2000	> 2000
D) Bacteriological Parameters				
Fecal coliforms (CFU/100 mL)	≤10	10-200	200-2000	> 2000
Total coliforms (CFU/100 mL)	≤100	100-20000	20000-100000	> 100000

2.3.2 Water Quality Classification for Creeks

As detailed in the previous subsection, statistical values and distribution of each water quality parameter of Yuvarlakçay and Namnam Creeks are calculated. Detailed statistical values and classification results are given in APPENDIX B. In Table 2.7 and Table 2.8, distribution types, 95% values and classification results are given for Yuvarlakçay and Namnam Creeks, respectively.

Table 2.7: Distribution types, 95% values and classification results based on SWQMR for Yuvarlakçay Creek

Parameter	Distribution	95% value	Quality Class
Temperature (^oC)	Normal	18.87	I
pH	Normal	8.23	I
Electrical Conductivity (µS/cm)	Log-normal	481.23	II
Dissolved Oxygen (mg O_2/L)	Log-normal	8.36	I
Oxygen Saturation (%)	Log-normal	91.21	I
Ammonium Nitrogen(mg NH_4^+-N/L)	Normal	0.045	I
Nitrite Nitrogen (mg NO_2^--N/L)	Normal	0.016	III
Nitrate Nitrogen (mg NO_3^--N/L)	Log-normal	1.16	I
Total Phosphorus (mg P/L)	Log-Normal	0.027	I
Chemical Oxygen Demand (mg/L)	Log-normal	9.88	I
Fecal Coliforms (CFU/100 mL)	Log-normal	294.04	III
Total Coliforms (CFU/100 mL)	Log-normal	1301.20	II

Table 2.8: Distribution types, 95% values and classification results based on SWQMR for Namnam Creek

Parameter	Distribution	95% Value	Quality Class
Temperature (^oC)	Normal	19.77	I
pH	Normal	8.43	I
Electrical Conductivity (µS/cm)	Log-normal	532.03	II
Dissolved Oxygen (mg O_2/L)	Log-normal	7.88	II
Oxygen Saturation (%)	Log-normal	89.87	II
Ammonium Nitrogen (mg NH_4^+-N/L)	Normal	0.032	I
Nitrite Nitrogen (mg NO_2^--N/L)	Normal	0.008	II

Parameter	Distribution	95% Value	Quality Class
Nitrate Nitrogen (mg NO_3^--N/L)	Log-normal	0.23	I
Total Phosphorus (mg P/L)	Normal	0.021	I
Fecal Coliform (CFU/100 mL)	Log-normal	70.04	II
Total Coliform (CFU/100 mL)	Log-normal	836.64	II

The following points are ensued after analysis of tables:

- In Yuvarlakçay Creek, eight parameters (T, pH, DO, oxygen saturation, NH_4^+-N, NO_3-N, TP and COD) are in Class I, two parameters (EC and TC) are in Class II and again two parameters (NO_2-N and FC) are in Class III.
- In Namnam Creek, five parameters (T, pH, NH_4^+-N, NO_3^-N and TP) are in Class I and six parameters (EC, DO, oxygen saturation, NO_2-N, FC and TC) are in Class II.
- T, pH, NH_4^+-N, NO_3-N and TP parameters are in Class I for each two creeks.
- EC and TC parameters are in Class II for each two creeks.
- There is no parameter of Namnam Creek in Class III, all parameters are in Class I or II.
- DO and oxygen saturation parameters are in Class I for Yuvarlakçay Creek while in Class II for Namnam Creek.
- NO_2-N and FC parameters are in Class II for Namnam Creek while in Class III for Yuvarlakçay Creek.
- Due to the existence of missing values in COD data for Namnam Creek, COD classification could not be done for this creek.
- There are no criteria for salinity (%o), flow rate (m^3/s), total suspended solids (mg/L) and total nitrogen parameters in SWQMR.

2.3.3 Identification of Lake Trophic Status

For lake trophic status identification, arithmetic mean of data set is used. Arithmetic mean of lake parameters are classified according to SWQMR as mentioned in Appendix 7, Table 9 of this regulation and Organization for Economic Co-operation and Development (OECD), which are given in Table 2.9 and Table 2.10.

Table 2.9: Limiting values used in trophic status classification system for lakes, lagoons, reservoirs Surface Water Quality Management Regulation (SWQMR) (Ministry of Forestry and Water Affairs, 2012)

Trophic status	**Total P (µg/L)**	**Total N (µg/L)**	**Chlorophyll *a* (µg/L)**	**Depth of Secchi disk (m)**
Oligotrophic	≤ 10	≤ 350	< 3.5	> 4
Mesotrophic	10 > TP ≥ 30	350 > TN ≥ 650	3.5 - 9.0	4 - 2
Eutrophic	30 > TP ≥ 100	650 > TN ≥ 1200	9.1 - 25.0	1.9 - 1
Hipertrophic	> 100	> 1200	> 25.0	< 1

Table 2.10: Categories and ranges of trophic status considered (OECD, 1982)

Trophic Status	**Total Phosphorus (µg/L)**	**Chlorophyll-*a* (µg/L)**	**Depth of Secchi disk (m)**
Ultraoligotrophic	< 4	< 1	> 12
Oligotrophic	4 - 10	1 - 2,5	12 - 6
Mesotrophic	10 - 35	2.5 - 8	6 - 3
Eutrophic	35 - 100	8 - 25	3 – 1.5
Hipertrophic	> 100	> 25	< 1.5

For total phosphorus, total nitrogen and chlorophyll a parameters, arithmetic mean calculations were done for three sample points of the lake. Depth of secchi disc parameter is not available so, for classification this term could not be used.

Calculated arithmetic means and related trophic status based on SWQMR and OECD are given in Table 2.11.

Table 2.11: Calculated arithmetic means and related trophic status based on SWQMR and OECD

Parameters	Lake Köyceğiz			SWQMR	OECD
	South Point (Dalyan Entrance)	Lake Center	North Point (Lake Beach)		
Total Phosphorus (μg/L)	58	39	30	Eutrophic	Eutrophic
Total Nitrogen (μg/L)	-	698	754	Eutrophic	-
Chlorophyll-a (μg/L)	-	1.972	1.801	Eutrophic	Oligotrophic
Depth of Secchi Disc (m)	-	-	-	-	-

As can be seen from the Table 2.11, Lake Köyceğiz is in eutrophic state based on total phosphorus parameter according to SWQMR and OECD. There is no total nitrogen data for Dalyan Entrance sampling point and total nitrogen criteria in OECD. Thus, the lake is again in eutrophic state based on total nitrogen parameter according to only SWQMR. There is no chlorophyll-a data for Dalyan Entrance sampling point. The lake is eutrophic state based on SWQMR while in oligotrophic state based on OECD. For classification, the worst case is selected thus; the lake is in eutrophic state according to SWQMR. There is no depth of secchi disc data for all sampling points of the lake. Hence, classification could not be done with using this parameter.

2.4 Spatial and Seasonal Comparison by using Hypothesis Testing – One Way ANOVA Test

In this section, hypothesis test is applied to data for identification the differences of water quality parameters based on sampling point locations of the lake and four seasons. Means of total phosphorus (TP) and dissolved oxygen concentration (DO) parameters are compared spatially and seasonally for the lake sampling points that are located at south (Dalyan Entrance), center (Lake Center) and north (Lake Beach) points. Moreover, means of total nitrogen (TN) and chlorophyll-a (Chl-a) parameters are tested for only Lake Beach and Center.

Statistical hypothesis compares the means of samples taken from two separate bodies. In this way, two testing involves two hypotheses that are null (H_0) and alternate hypothesis (H_1). The rejection of null hypothesis (H_0) results the acceptance of alternate hypothesis (H_1). Acceptance of a null hypothesis means that there is no enough proof for rejection of it (Papoulis, 1991). In our case:

H_0: there is no significant difference between the means of given location or seasons.
H_0: $\mu = \mu_0$
H_1: there is a significant difference between the means of given location or seasons.
H_1: $\mu <> \mu_0$

Mean values of TP and DO parameters are compared based on south, center and north sampling points of the lake (Dalyan Entrance, Lake Center and Beach). Due to absence of data for Dalyan Entrance, mean values of TN and Chl-a parameters are compared according to only Lake Beach and Lake Center. Obtained comparison results are given in Table 2.12.

Table 2.12: Comparison of TP, DO, TN and Chl-a parameters according to location

Parameter	Location	N	P (Sig)	Decision
TP (mg/L)	South, center, north	82 * 3 = 246	0.228	Not reject Ho
DO (mg/L)	South, center, north	82 * 3 = 246	0.803	Not reject Ho
TN (mg/L)	Center, north	82*2 = 164	0.638	Not reject Ho
Chl-a (µg/L)	Center, north	82*2 = 164	0.662	Not reject Ho

As can be seen from the table, there is no a significant difference between the sampling points for TP, DO, TN and Chlorophyll-a parameters.

In Table 2.13, seasonal comparison of the water quality parameters are given.

Table 2.13: Seasonal comparison of TP, DO, TN and Chl-a parameters

Parameter	N	P (Sig)	Decision
TP (mg/L)	246	0.563	Not reject Ho
DO (mg/L)	246	0.000	**Reject Ho***
TN (mg/L)	164	0.359	Not reject Ho
Chl-a (µg/L)	126	0.003	**Reject Ho***

Examining Table 2.13, we can conclude that, there is no significant difference between four seasons for TP and TN parameters however; there is a significant difference ($P<0.05$) for DO and Chlorophyll-a parameters. For DO and Chlorophyll-parameters, homogeneity of variances test is applied. Significance value from this test for Chlorophyll-a is obtained as 0.000 thus, variances are not homogenous, not equal. Therefore, Temhane's T2 Test, which is one of the most used one of the Post

Hoc Test when equal variances not assumed case, is applied for Chlorophyll-a parameter.
Table 2.14 summarizes the significance values for comparison of Chlorophyll-a parameters seasonally.

Table 2.14: Post Hoc Test with Temhane's T2 for Chlorophyll-a parameter

Chlorophyll-a		Winter	Spring	Summer	Fall
Winter	**Mean**	2.767	-	-	-
	Mean Difference	-	-0.255	-1.370	-1.765
	Sig. Value	-	1.000	0.236	0.048*
Spring	**Mean**	-	2.512	-	-
	Mean Difference	0.255	-	-1.114	-1.510
	Sig. Value	1.000	-	0.151	0.009*
Summer	**Mean**	-	-	1.398	-
	Mean Difference	1.369	1.114	-	-0.396
	Sig. Value	0.236	0.151	-	0.709
Fall	**Mean**	-	-	-	1.887
	Mean Difference	1.765	1.510	0.396	-
	Sig. Value	0.048*	0.009*	0.709	-

* values below 0.5

There is a significant difference between winter and fall months and spring and fall months for Chlorophyll- a parameter while there is no important difference between other months.

Significance value from homogeneity of variances test for DO is obtained as 0.406 so, variances are homogenous, are equal. Therefore, Tukey Test, which is one of the most used one of the Post Hoc Test when equal variances assumed case, is applied for DO parameter. The significance values for comparison of DO parameter seasonally are given in Table 2.15.

Table 2.15: Post Hoc Test with Tukey for DO parameter

Dissolved Oxygen		Winter	Spring	Summer	Fall
Winter	**Mean**	8.473	-	-	-
	Mean Difference	-	0.081	-1.266	-1.259
	Sig. Value	-	0.081	0.000*	0.000*
Spring	**Mean**	-	8.555	-	-
	Mean Difference	-0.081	-	-1.347	-1.341
	Sig. Value	0.990	-	0.000*	0.000*
Summer	**Mean**	-	-	7.207	-
	Mean Difference	1.266	1.347	-	0.007
	Sig. Value	0.000*	0.000*	-	1.000
Fall	**Mean**	-	-	-	7.214
	Mean Difference	1.259	1.341	-0.007	-
	Sig. Value	0.000*	0.000*	1.000	-

* values below 0.5

As shown in the table, DO parameter changes between seasons. DO concentrations obtained from samples in winter months are significantly different from summer and fall months. In addition, DO concentrations measured in spring months are quite varied from obtained measurements in summer and fall months. However, there is no significance difference between DO amounts in summer and fall months. The changes for DO parameter should be resulted from temperature variations between the seasons.

2.5 Correlation Analysis

In this section, correlations between different parameters sampled from different locations of lake and creeks are analyzed by using Statistical Package for the Social Sciences (SPSS) program. Correlation analysis is conducted to understand relationship between water quality parameters and precipitation data. For Lake

Center, Lake Beach, Dalyan Entrance, Yuvarlakçay and Namnam Creeks, correlation analyses are conducted with available water quality parameters and precipitation data. A brief table and outcomes are given in this part. Detailed table obtained by SPSS is given in APPENDIX B.

Results are interpreted based on the rating as correlation coefficient between 0.5 and 0.7 means moderately correlated while coefficient with higher than 0.7 value means highly correlated. Correlation coefficients under 0.5 value are removed due to the weak relationship.

Correlation analysis is conducted to find correlation coefficients between water quality parameters of Yuvarlakçay and Namnam Creeks. They are given in Table 2.16. To conclude, T parameters of the creeks are highly related and moderately negative correlated with DO parameter of Namnam. In addition, their T parameter is moderately negative correlated with precipitation data. EC parameter of Yuvarlakçay is highly correlated with only NO_2-N of Namnam creek. DO parameters of the creeks are again moderately correlated. FC parameter of Yuvarlakçay is moderately correlated with FC and TC parameters of Namnam. Finally, NH_4-N parameters of the creeks are moderately positive related with each other.

Table 2.16: Correlation coefficients for Yuvarlakçay and Namnam creeks

T_YC		**DO_YC**		**FC_YC**		**EC_YC**	
T_NC	0.847	DO_NC	0.555	TC_NC	0.690	NO_2-NC	0.846
Prec.	-0.551	**NH_4_YC**		FC_NC	0.664		
DO_NC	-0.505	NH_4_NC	0.577				
T_NC		**DO_NC**		**FC_NC**		**TC_NC**	
T_YC	0.847	T_NC	-0.575	TC_NC	0.990	FC_NC	0.990
DO_NC	-0.575	DO_YC	0.555	FC_YC	0.664	FC_YC	0.690
NO_2_NC		**NH_4_NC**					
EC_YC	0.846	NH_4_YC	0.577				

(YC: Yuvarlakçay Creek; NC: Namnam Creek; T: Temperature; Prec.: Precipitation; EC: Electrical Conductivity; DO: Dissolved Oxygen; TC: Total Coliform; FC: Fecal Coliform)

Based on found correlation results for Yuvarlakçay and Namnam Creeks, an input scenario (scenario 5) is formed for the developed ANN model. In this scenario, correlated parameters are removed from all parameters of the creeks used before. Therefore, the requirement of correlation analysis before ANN procedure can be assessed based on the obtained results.

In Table 2.17, correlation coefficients for parameters of Dalyan Entrance point of the lake between parameters of the creeks are given. According to the table, T parameter of Dalyan Entrance point is highly correlated with same parameter of the creeks; moderately negative correlated with DO parameter of Namnam Creek and precipitation data. DO parameter is highly related with same parameter of Namnam Creek; moderately correlated with DO parameter of Yuvarlakçay Creek. In addition, FC parameter is highly correlated with FC and TC parameters of Namnam; moderately related with FC of Yuvarlakçay. Finally, TC parameter of this point is again highly correlated with same parameter of Yuvarlakçay Creek.

Table 2.17: Correlation coefficients for Dalyan Entrance

T_DE		**DO_DE**		**FC_DE**		**TC_DE**	
T_NC	0.933	DO_NC	0.757	TC_NC	0.781	TC_YC	0.712
T_YC	0.875	DO_YC	0.616	FC_NC	0.778		
DO_NC	-0.627			FC_YC	0.605		
Prec.	-0.549						

(YC: Yuvarlakçay Creek; NC: Namnam Creek; DE: Dalyan Entrance; T.: Temperature; Prec.: Precipitation; DO: Dissolved Oxygen; TC: Total Coliforms; FC: Fecal Coliforms)

Correlation coefficients between Lake Center and creeks' parameters are given in Table 2.18. As it can be seen from the table, T parameter of Lake Center is highly correlated with same parameter of the creeks; moderately negative related with DO parameter of Namnam and precipitation data. PH parameter is moderately related with pH of Yuvarlakçay Creek. DO is moderate related with Namnam and Yuvarlakçay Creeks' DO parameters. Finally, TC parameter of Lake Center is highly correlated with FC and TC parameters of Namnam; moderately related with FC parameter of Yuvarlakçay.

Table 2.18: Correlation coefficients for Lake Center

T_LC		**DO_ LC**		**TC_ LC**	
T_NC	0.922	DO_NC	0.647	TC_NC	0.844
T_YC	0.874	DO_YC	0.552	FC_NC	0.834
DO_NC	-0.623	**pH_ LC**		FC_YC	0.594
Prec.	-0.541	pH_YC	0.520		

(YC: Yuvarlakçay Creek; NC: Namnam Creek; LC: Lake Center; T: Temperature; Prec.: Precipitation; DO: Dissolved Oxygen; TC: Total Coliforms; FC: Fecal Coliforms)

In Table 2.19, correlation coefficients between Lake Beach and the creeks parameters are given. To conclude the table, T parameter of Lake Beach is highly correlated with same parameter of the creeks; moderately negative related with DO parameter of Namnam Creek. PH of the beach is moderately correlated with same parameter of Yuvarlakçay. DO parameter is moderately related with same parameter of Namnam Creek. Finally, TC parameter of the Lake Beach is highly correlated with TC and FC parameters of Namnam; moderately related with FC parameter of Yuvarlakçay.

Table 2.19: Correlation coefficients for Lake Beach

T_LB		**DO_ LB**		**TC_ LB**	
T_NC	0.911	DO_NC	0.562	TC_NC	0.804
T_YC	0.879	**pH_ LB**		FC_NC	0.804
DO_NC	-0.639	pH_YC	0.534	FC_YC	0.565

(YC: Yuvarlakçay Creek; NC: Namnam Creek; LB: Lake Beach; T: Temperature; Prec.: Precipitation; DO: Dissolved Oxygen; TC: Total Coliforms; FC: Fecal Coliforms)

According to correlation analysis results of three sampling point of the lake (Dalyan Entrance, Lake Center, Lake Beach), an input scenario (scenario 6) is created for the

ANN process. In this scenario, only correlated parameters are used for prediction of target output. In other words, for Dalyan Entrance and Lake Center, DO parameters of them is set as output while, DO parameters of the creeks are set as input which are correlated with output. Moreover, for Lake Beach, TC parameter is selected as output parameter while FC and TC parameters of Namnam and Yuvarlakçay's FC parameter are chosen as input parameters. By this way, input scenarios can be compared for DO and TC parameters and sufficiency of correlation analysis for these parameters can be identified for ANN process.

2.6 Principal Component Analysis

In this part of the chapter, Principal component analysis (PCA) is conducted for available data Lake Köyceğiz based on their correlation and covariance matrix. PCA is a powerful technique used to analyze a data set that consists of several inter-correlated quantitative dependent variables. The aim of this technique is the extraction of the important information from this data set and representation of it with a set of new orthogonal variables that are called as principal components (PCs) (Abdi, et al., 2010).

In 2006, Boyacıoğlu is applied factor analysis to the surface water quality data sets obtained from Büyük Menderes River Basin by using SPSS program. In the study, water quality data sets are obtained from two hydrological periods that are low-flow and high-flow conditions. For high-flow conditions, 32 data is used while, 17 data is used for low-flow conditions. The study is pointed that water quality is controlled by agricultural pollutant sources under low-flow conditions while; pollutants are mainly originated from urban land use during high-flow period. The main reason for abundance of agricultural uses during low-flow period is the negative effect of runoff to surface water quality. As a result, both of the obtained pollution threats are defined as non-point sources therefore, measures should be taken to minimize these sources for improvement of water quality in the basin. Finally, it is revealed that factor analysis is a useful technique that can help to decision makers for determination of the extent pollution via practical pollution indicators.

Ouyang et al. used PCA and Principal Factor Analysis (PFA) to assess the seasonal changes in surface water quality of the main stream of St. Johns River in Florida. PCA is conducted to evaluate the seasonal correlations of water quality parameters while; PFA is used to extract the most important parameters for the assessment of seasonal variations of the river water quality. In the study, 22 monitoring stations and 16 physiochemical parameters obtained from each station are used from 1998 to 2001. The studied data indicates that seasonal variations should be considered when using DO as an indicator parameter for evaluation of surface water quality in St. Johns River. Another result obtained from the study is that the most important parameter in contributing to water quality variation for one season may not be important for another season except for dissolved organic carbon and electrical conductance, which were always the most important.

In 2009, Kazi et al. assessed the water quality of polluted Manchar Lake (Pakistan) using multivariate statistical methods that are cluster analysis (CA) and PCA. In the study, data of 36 parameters obtained from 5 monitoring sites during 2005-2006 is used. The aim of paper is to evaluate and interpret the complex data sets and to identify the pollution sources for getting better information about water quality and x a monitoring network design. In the study, CA uses PCA results to classify the samples. CA groups the five sampling sites in to three clusters of similar water quality characteristics. In this way, number of the sampling sites can be reduced in future. In addition, PCA explains the main pollution causes to the lake as discharge of industrial, agricultural wastes and municipal sewage water. Fishing and boating activities are also in the reasons for the deterioration of lake water quality.

Zhao et al. conduct a similar study with of Kazi et al. for Baiyangdian Lake (China) for assessment of water quality. PCA and hierarchical CA are used for evaluation. In this study, 21 physicochemical parameters obtained from 13 different sampling sites are analyzed during two years. As a result, 5 principal components are obtained from PCA that describe the 92% of both temporal and spatial changes. CA reduces the sampling site number into 5 from 13 based on similar water quality characteristics. According to the conducted PCA, discharge of industrial and agricultural wastes, domestic sewage from the upstream Fuhe River and pollution from local villages around the lake are the main causes for degradation of the lake. The study concluded that for an effective lake water management, multivariate statistical methods are useful for the analysis and interpretation of complex data sets, for identification of pollution sources and for determination of variations in water quality.

After some example studies, water quality parameters of Lake Köyceğiz and precipitation data are analyzed using factor analysis tool of Statistical Package for the Social Sciences (SPSS) program in this study. At the end of the analysis, five principal components are obtained for the lake. Because of the bilocation of some parameters in two principal components, varimax rotation process, which is the most common one, is applied to data set (Abdi, et al., 2010). The five components explain the 69.73% of the total variance before and after the rotation process. Explanations of total variances before and after rotation are given in Table 2.20. In addition, obtained components are summarized in Table 2.21.

Table 2.20: Explained variance with obtained components for the lake from PCA

Component	Total Variance Explained					
	Before Rotation			After Rotation		
	Eigenvalues	% of Variance	Cumulative %	Eigenvalues	% of Variance	Cumulative %
1	2.654	22.119	22.119	2.012	16.770	16.770
2	1.877	15.641	37.760	1.963	16.360	33.130
3	1.710	14.253	52.012	1.802	15.013	48.143
4	1.112	9.267	61.280	1.469	12.241	60.384
5	1.014	8.448	69.728	1.121	9.343	69.728

Table 2.21: Obtained components for the lake from PCA

Component_1		Component_2		Component_3	
Chl-a	0.813	TP	0.847	TC	0.849
NO_2-N	0.786	DO	-0.715	FC	0.777
T	-0.650	NH_4-N	0.710		
Component_4		**Component_5**			
TSS	0.886	TN	0.882		
EC	-0.542				
pH	0,538				

(T: Temperature; EC: Electrical Conductivity; DO: Dissolved Oxygen; TSS: Total Suspended Solids; TC: Total Coliform; FC: Fecal Coliform; TP: Total Phosphorus; TN: Total Nitrogen; NO_2-N)

Principal component (PC) 1 includes chl-a, and NO_2-N with high positive correlation and negative moderate correlation with T parameters. This component can be indicator of fertilizer use in the area, a sewage discharge or leakage from septic systems. Decrease in temperature can be explained that existence of high algae in water penetrates the absorption of sun light. Second component includes TP and NH_4-N parameters with high positive correlation while DO parameter with negative correlation. Existence of NH_4-N and TP in natural waters is regarded as indicative of domestic, industrial or agricultural pollution, primarily from fertilizers, organic matter or fecal matter. In natural systems, NH_4-N is rapidly oxidized to nitrite and nitrate however, this process requires presence of DO. As can be understood from the PC 2, due to absence of DO, NH_4-N concentration is high. TP parameter in water can be caused from fertilizer use in the area or domestic discharges including detergents. To conclude, PC 2 can represent the anthropogenic pollution sources to the lake. PC 3 consists of total and fecal coliforms parameters therefore; it can be interpreted as wastewater contribution to the lake. Considering PC 4, it can indicate the meromictic property of the lake. Existence of the salty water in the lake increases the TSS and EC parameters.

CHAPTER 3

PREDICTION OF WATER QUALITY PARAMETERS USING ARTIFICIAL NEURAL NETWORKS

This chapter starts with brief information about artificial neural networks. Then, preparation of data and application of artificial neural network to predict water quality parameters in Lake Köyceğiz are described.

3.1 Artificial Neural Networks (ANNs)

Initially in the late-1800s, neural network is defined as a concept to understand how the human brain works. Artificial neural networks (ANNs) are systems trying to imitate human brain and biological nervous system on computers. They can simulate humanistic features such as learning, storing of learned information and making generalization by use of a system composed of a lot of neurons connected with weights (Steyl, 2009).

3.1.1 General Characteristic of ANNs

In this subsection, benefits of ANNs are explained based on general characteristics of them. First of general characteristics is their non-linear structure. ANNs have the ability to converge the non-linear relationship between variables to a continuous function or its derivative without any predefinition of model. Therefore, they are defined as *Universal Function Approximators* (Beltratti, et al., 1996). Secondly, ANNs can learn the relationship between given input and output data and can store the learned information on their distributed memory with synaptic weights without any model predefinition (Çelik, 2008). In addition, ANNs can generalize for never before seen cases based on their learning property using known samples. They can produce solutions for noisy or missing data with a high performance on the contrary to many classical methods. Unlike statistical methods, ANNs can work with an unlimited number of variables and parameters without the need for any extra conversion. In this way, general solutions can be provided with an excellent forecasting accuracy. Also, fault tolerance of ANNs is extremely higher than traditional methods (Yurtoğlu, 2005). Moreover, an ANN trained to solve a specific problem can be retrained to account for the variations of problem (Saraç, 2004). Finally, ANNs used in different application areas can share similar learning algorithms and theories. This feature will provide an important convenience for solution of a problem using ANN (Subaşı, 2010).

3.1.2 Disadvantages of ANNs

ANNs have some disadvantages besides the advantages. Need of very fast parallel processors is the first of them. Especially in applications where response time is critical, the computer system should be very fast to perform an affective analysis. Secondly, ANNs may not solve all kinds of problems. Solution sets of some problems may never be solved by using ANN or created solution set can have no connection with real values (Çelik, 2008). For example, analysis of stochastic independent events such as chance games, coin flipping, ANNs can not produce better solutions than statistical methods (Navarro, et al., 2014). Other disadvantage is architectural problem of ANNs. There is no guarantee for that chosen ANN architecture will give the best solution for the problem because of subjective decision of ANN structure criteria such as neuron number, layer number, learning algorithm, etc. Another drawback of ANNs is interpretation problem. Although the test results obtained from conventional techniques can be interpreted with a set of tests and stability analysis can be performed, there is no such possibility for artificial neural networks. How to reach the output of the network is not known. Especially in applications where prediction is crucial, this feature of ANNs causes a major problem. For example, in a biological application, artificial neural networks used in cancer-related research, this obscurity leads to decrease in confidence using ANNs (Çelik, 2008). Finally, stopping of ANN training is other disadvantage. The training is stopped when reached to an acceptable error rate which is decided subjectively. In case of early completion of the training low predictive performance problem and in case of late completion over fitting (memorizing) problem arise. Overfitting or over-optimization is described as that any prediction system provides successful results for known sample however; for unknown data, the system is lack of ability to predict (Pardo, 1992).

3.1.3 Usage Area of ANNs

Artificial neural networks have gained a wide range usage area in applications for encountered problems in real life. Today, they can be used in many industries successfully. There no restrictions about usage area for ANNs however, they are mainly used in classification, forecasting and modeling areas. ANNs are being implemented for many serious problems and problem number gradually increases. Due to the ANNs' best description ability for the trends and data pattern, they are well suited for forecasting processes. Some examples for widespread applications of ANNs are quality control, financial forecasting, economic forecasting, credit rating, Laboratory Research, System Modeling, Fingerprint Recognition, Meteorological Interpretation, etc. (Yurtoğlu, 2005).

3.1.4 Basic Components of ANNs

In this part of the chapter, theoretical knowledge about artificial neural networks will be given and architectures of them will be explained.

ANNs are composed of artificial cells which are connected hierarchically and can work in parallel. Artificial cells are assumed same as neurons in biological nervous system also; they can be called as process elements. Each process element consists of five main components which are inputs, weights, summation function, activation/transfer function, and output of the neuron (Açıkalın, 2007).

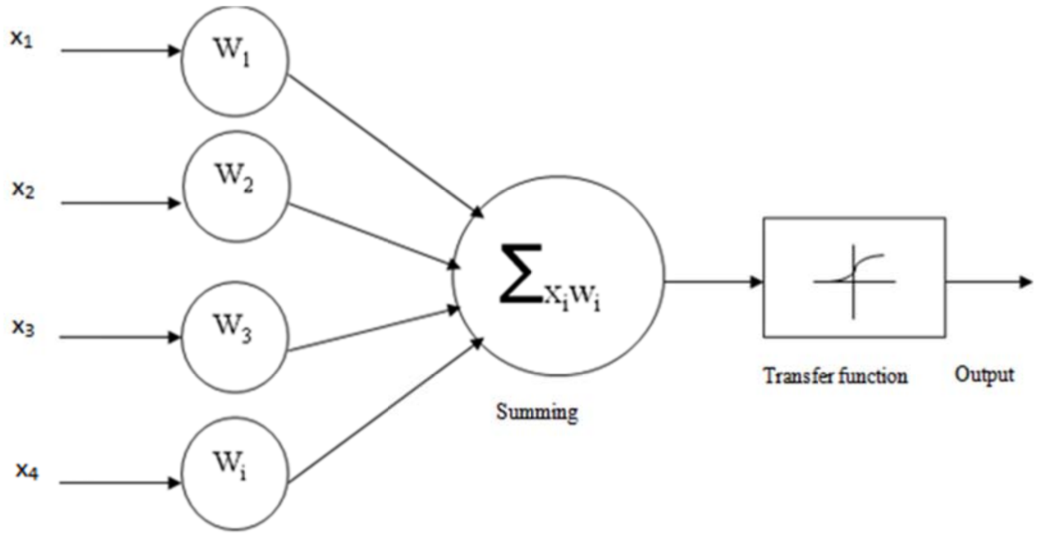


Figure 3.1 Schematic representation of an artificial neuron (process element) (Kashid, et al., 2012)

These shown neuron components of ANNs in
Figure 3.1 are explained clearly below.

a. Inputs

Inputs ($x_1, x_2, x_3, \ldots x_i$) are elements of artificial neural networks which take data from outer world. They do not have any other functions than to transform data to next step. In other words, inputs do not make any mathematical process on data and they work only as transmitter. Inputs are one of the two elements of ANNs which are related with outer world of network. A neuron can have unlimited number of inputs but, there must be only one output of every neuron (Çelik, 2008).

b. Weights

Weights (w_1, w_2, w_3,…w_i) are most important elements of mathematical neuron by itself and artificial neural network generally. Because learned data by the network is stored on the weights. ANNs can achieve their learning function through changing the weights. Accordingly, network's synaptic weight vector expressing as $W = [W_i]_{nx1}$ consists convertible values. Typically, initial values of weights are selected as a random value in (-1, 1) range. How to set up of a weight for learning relationship between the given variables is decided based on selected *learning rule*. Weights of ANNs can be thought as synapses in biological nervous system (Çelik, 2008).

c. Summation Function

Summation function is responsible for summation of all data coming from outer world and related weights. Summation function is shown in (3.1).

$$v_i = \sum_{i=1}^{n} x_i w_i - \theta \qquad (3.1)$$

Expression (3.1) is shown more clearly in (3.2).

$$v_i = x_1 w_1 + x_2 w_2 + x_3 w_3 + \cdots x_n w_n - \theta \qquad (3.2)$$

Summation function transfers the created *weighted input* (v_i) to activation function. (θ) value in (3.1) and (3.2) express the threshold value. Use of threshold value in summation function is not obligatory. Using of the threshold value is related with the request of network architecture designer.

Summation functions is the most used function in artificial neural network. However, there is no restriction for that the process must be summation function. Some other functions which can be used instead of summation function are given in Table 3.1.

Although there are many functions can be used, summation function is used in most of the research (Çelik, 2008).

Table 3.1 Some Functions can be used instead of summation function

Function	Weighted input (v_i)
Multiplication	$v_i = \prod_{i=1}^{n} x_i w_i$
Maximum	$v_i = \max(x_i w_i)$
Minimum	$v_i = \min(x_i w_i)$
Signum	$v_i = \mathrm{sgn}(x_i w_i)$

d. Activation (Transfer) Function

Activation (transfer) function is responsible for activation of coming weighted input (v_i) and determination of the final output value. As in summation function, functions which can be used as activation function vary based on type of problem. There is no a universal formula for which type of activation function should be used (Öztemel, 2006). Optimal activation function is determined as a result of attempts by the designer. The choice of activation function depends largely on available data and what the desired learning of network.

Coming weighted input to the cell is passed through the activation function and sent as final output to a new cell's input layer. Mathematical expression of activation function is shown in (3.3)

$$f(v_i) = y \tag{3.3}$$

Sigmoid and hyperbolic tangent functions are most commonly used in activation functions (Saraç, 2004).

e. Output Function

Output function is responsible for transfer of output value of activation function to outer world as network's final output value or to other connected neurons as their input values.

After brief information about artificial neuron components, artificial neural neurons come together to form an artificial neural network. In general, they come together in three layers and parallel in each layer to form ANN, not randomly. These layers are:

1. <u>Input layer</u>: neurons in this layer transfer the information coming from outer world to hidden layers. There is no data processing in this layer generally.
2. <u>Hidden layers</u>: Information coming from input layer is processed in these layers and sent to output layer. There can be more than one hidden layers for an ANN.
3. <u>Output layer</u>: Process elements in this layer, transferred information from hidden layers are handled and produce output for given data in input layer. Produced output is sent to outer world as result. (Açıkalın, 2007)

Layers are connected to each other with weights. Weight values are determined through learning step.

This procedure is summarized in Figure 3.2:

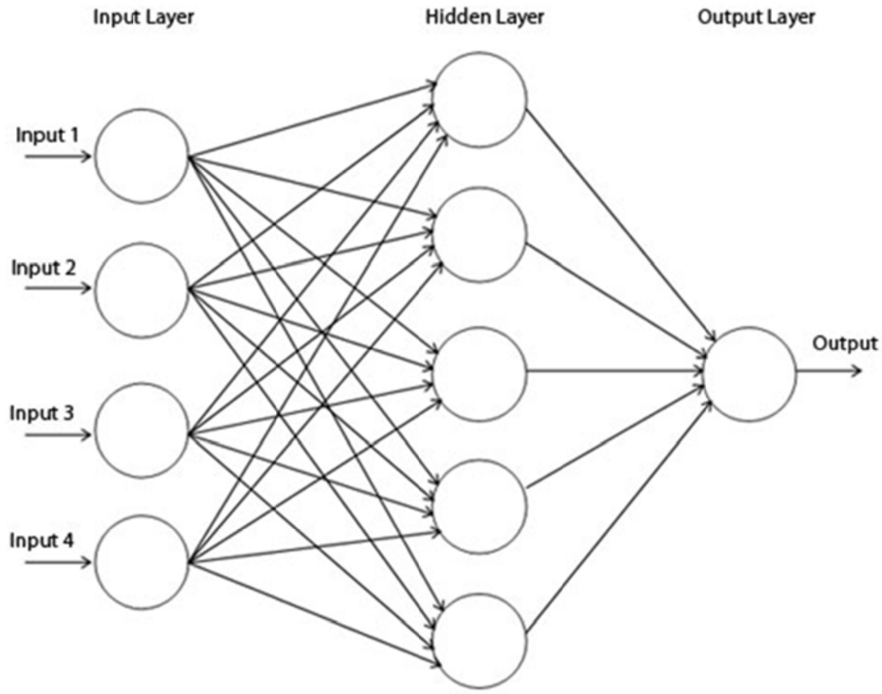


Figure 3.2 An artificial neural network representation with 3 layer (Sergiu, 2011)

General working principle of ANNs can be explained as conversion of an input data set to an output. To achieve this aim, ANN must be trained to get true output values for given input data set. The procedure for this conversion ability of ANNs can not be explained. Input data set is given then output value is taken therefore, ANNs is named as “black box” (Açıkalın, 2007).

3.1.5 ANN Types Based on Structure

As mentioned before, artificial neural networks are composed of neurons related with each other. This relation between neurons indicates the structure of ANNs. Connection structures and weight values are decided by learning algorithms. In general, artificial neural networks are divided in two groups as feed forward artificial neural networks and back forward artificial neural network.

f. Feed Forward Networks

In feed forward networks, neurons are located into layers generally. Inputs are sent from one layer to next one by one-way weights. Due to the one-way connection, a return back of the next layer's output to previous layer as input is impossible. Multi-Layer Perceptron (MLP) and Learning Vector Quantization (LVQ) networks are examples for feed forward networks.

In feed forward networks, input layer sends the information coming from outer world to the hidden layer without any process on data. Data is processed in hidden and output layers then output of the network is generated. A feed forward network with these structures performs a non-linear static function. Any continuous function can be converged via a feed forward artificial neural network with three layers can converge with desired accuracy. The most known back-propagation learning algorithm is used effectively in this type of artificial neural networks' training so; sometimes they are called as back-propagation networks (Saraç, 2004).

g. Feed-Back Networks

In a feed-back network, outputs of hidden and output layer are fed to input layer or previous hidden layers. By this way, inputs can be transferred in both forward and reverse directions. This type networks have dynamic memory and an output in a moment reflects inputs at that moment and previous inputs. Therefore, they are especially suitable for prediction applications. Feed-back networks were quite successful in the estimation of various types of time series. Hopfield, SOM (Self Organization Map), Elman and Jordan networks can be given as examples for these networks.

In a feedback artificial neural network, at least one output of a cell is given to itself or other cells as input and feedback is done by a delay element. Feedback can be between neurons in layers as well as between layers. ANN with this structure shows a non-linear dynamic behavior. In backward calculations, output value generated by the network is compared with expected results of the network. For decrease of total error to minimum, it must be distributed to the process elements causing this error. This means changing the weights of process elements (Saraç, 2004).

3.1.6 ANN Types Based on Learning Algorithm

Learning is defined as change in behavior of natural structure result of observation, training and movement. In that case, change of network weights should be supplied by some methods, rules and training. For this aim, generally three main learning methods and learning rules related these methods can be mentioned. These are supervised, unsupervised and reinforcement learnings.

In supervised learning, differences between output values generated by the network and target output values are considered as error and this error is attempted to be minimum. For this aim, connection weights are changed to obtain appropriate output from them. Due to this case, supervised learning needs a supervisor. (Subaşı, 2010).

In unsupervised learning type, only inputs are provided to the network. Target outputs are not given. The network develops the classification rules by itself based on obtained output from given sample to input layer. Then, the network sets the connection weights to form patterns exhibiting same properties. (Steyl, 2009).

Finally, reinforcement learning is similar with supervised learning method however, in reinforcement learning a score or grade indicating accuracy degree of network output instead of target output as in supervised learning. (Saraç, 2004).

3.1.7 Basic Learning Rules for ANN

There is a lot of learning algorithms can be used for development of an artificial neural network. The majority of these learning algorithms are mathematical-based and used for update of the network weights. There more than one hundred varieties of these algorithms which vary according to network architecture and encountered problem type. However, most of them have been developed by inspiration from the Hebb Rule, Delta Rule, Kohonen Rule, and Hopfield Rule (Bayır, 2006).

3.1.8 Design of an ANN

For creation of a neural network, neuron connection type (topology), summation and activation functions used by neurons, learning method, learning rule and algorithm should be determined. Models are designed based on available data and type of application. Success of the created model is directly related with the correctly selection of the network architecture. Therefore, ANN designer should make the following decisions related network structure and operation:

- Selection of network architecture, determination of its structural properties such as layer number, neuron number in layers,
- Determination of function characteristics of neurons,
- Identification of learning algorithm and parameters, and
- Preparation of training and test data sets.

If these decisions are not made correctly, system complexity will increase or stable results will not be obtained. Total reaction and training time of the network increase based on system complexity. Therefore, the network size should be as small as possible for minimum reaction time of it (Bayır, 2006).

h. Selection of Learning Algorithm and ANN Structure

At design process of an artificial neural network, network structure should be selected based on encountered problem. It is important that which network is suitable for which problem (Bayır, 2006).

The selection of appropriate ANN structure depends on considered learning algorithm. Many of the network structures can be used with only one learning algorithm therefore, used ANN structure will be selected with identification of the learning algorithm by necessity. For example, back propagation learning algorithm requires feed-forward network structure (Saraç, 2004).

Differences between network categories and which of the more common network structures are shown in Table 3.2.

Table 3.2 Network types and their intended use (Anderson, et al., 1992)

Indented Use of Network	**Networks**	**Use for Network**
Prediction	- Back-Propagation - Delta Bar Delta - Extended Delta Bar Delta - Directed Random Search - Higher Order Neural Networks - Self-Organizing Map (SOM) into Back-Propagation	Using of input values for prediction of some output
Classification	- Learning Vector Quantization (LVQ) - Counter-propagation - Probabilistic Neural Networks	Using of inputs for classification
Data Association	- Hopfield - Boltzmann Machine - Hamming Network - Bidirectional Associative Memory - Spatio-Temporal Pattern Recognition	Determination of incorrect values and completing of missing values in input data
Data Conceptualization	- Adaptive Resonance Network (ART) - Self-Organizing Map (SOM)	Analyze of for derivation of grouping relationships
Data Filtering	- Recirculation	Smooth of an input signal

This table should be used as a guide and is not contain all network types. Some of the networks can be used in different type of applications. For example, feed-forward back-propagation network is used to solve almost all types of problems and indeed is the most popular for the first four categories (Anderson, et al., 1992).

i. Determination of Layer and Neuron Numbers

Layer number and neuron number in layers identify the complexity of the network. Layer number and neuron number in layers are decided by trial and error method. Neuron number used in a layer should be small as possible. Small neuron number in neural network increases generalization ability of the network while, redundancy of neuron number causes the memorization of data. However, use of less neuron than necessity can cause problems such as failure of learning data pattern. To sum up, too many hidden neuron may lead to the problem of overfitting; too few neuron in the hidden layer may cause the problem of underfitting. The optimum neuron number is found by increasing or decreasing the initial neuron number until it reaches to desired performance. To sum up, as the number of layer and neuron in layers increases, learning and process ability of the network increases, convergence time also increases, generalization capability decreases and memorization problem occurs. For many problems, a two or three layered network is able to produce satisfactory results (Saraç, 2004). In network structure, more than four hidden layers do not have developmental impact on network performance according to a lot of studies (Çelik, 2008).

j. Selection of Functions

Identification of neurons' characteristic properties is very important step for an artificial neural network design. Selection of summation and activation functions is related with available data properties and desired learning type of the network. In activation functions, sigmoid and hyperbolic tangent activation functions are commonly used ones. For non-linear problems, non-linear functions are used (Bayır, 2006).

k. Normalization

Available data is subjected to the normalization process before being submitted to the network. For prevention of excessive fluctuations in the data and improvement of the system performance, this normalization process is applied. As logarithmic functions used for this purpose, the data is usually scaled in [0, 1] or [-1, +1] ranges. Meaning that compressing of data at current axis, data including excessive fluctuations can cause adverse effects on the neural network model of the problem. This negativity can make learning function unsuccessful (Saraç, 2004).

l. Identification of Performance Function

Performance functions calculate the cumulative values between the target outputs values and created outputs by the network. According to these calculated values, how the network close to the pattern of training set is observed and connection weights are changed by using these values. Therefore, performance functions are one of the important factors affects the learning performance. Mean Square Error (MSE) is commonly used performance function in feed forward networks.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}[e(t)]^2 \qquad (3.7)$$

Mean Error (ME), Root Mean Square (RMSE), Mean Absolute Error (MAE), Normalized Mean Absolute Error and Mean Square Error are some of the other performance functions can be used. These are expressed by the following equations, respectively:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n} e(t) \qquad (3.8)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[e(t)]^2} \qquad (3.9)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|e(t)| \tag{3.10}$$

$$e_t = \frac{y_t - t_t}{t_t} x100 \text{ then } \mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|e_t| \tag{3.11}$$

Where *e(t)* is forecast error at period *t*; *n* is number of periods; y_t is forecast and t_t is actual result at period *t* (Bayır, 2006).

3.2 Previous Studies on Using of ANNs for Water Quality

In literature, ANN has used in wide study areas in order to predict water quality (river, groundwater etc.). It is used in modeling parameters like TDS, EC, turbidity, pesticide concentration, salinity, T, DO, and chl-a, DO, BOD, algal bloom etc. There are some example studies given below which summarize the using areas of ANN:

Wei et al. (2001) have developed a model to see the interactions between abiotic factors and algal genera in Lake Kasumigaura in Japan by using ANN technology. According to the results obtained from study; timing and magnitude of algal blooms of *Microcystis, Phormidium* and *Synedra* in the lake could be predicted successfully. On the other hand, for the newly occurring dominant *Oscillatoria* results are not such successful. In the study external environmental factors namely COD, pH, total nitrogen and total phosphorus were used in modeling algal responses. It enabled to find desirable combinations of environmental factors that inhibit algal blooms and one example to the combination is given as by improving COD by 10% and lowering TN by 10% would decrease the cell density of *Microcytis* more than 80%.

In the study of Kuo et al. (2004) ANN is used in the forecast of groundwater quality in the blackfoot disease area in Taiwan. Model A, B and C is developed by BP ANN to evaluate models' learning performance. Model A included five concentration parameters as input variables to determine seawater intrusion and three to determine arsenic pollutant, respectively, whereas models B and C used only one concentration parameter for each. The results showed that the RMSE obtained by model C was lower than the other two models. It is understood that when the training test has the maximum and minimum data, the model will successfully forecast within the 90% confidence interval.

Sahoo et al. (2005) have used ANN model in order to predict pesticide in groundwater in North Carolina domestic wells. In this study, four neural network types (three and four layer feed forward BP, a radial basis function and an adaptive neural network-based fuzzy inference system (ANFIS) were tested. It was observed that in generalization of the process only from two or three input parameters was useless with ANN. For higher prediction performance of the neural network, the presence of all variables in the input data is helpful. It is also found by sensitivity analysis that the time of sample collection (month of the year), the depths of the wells and pesticide travel times are very important parameters for the prediction of the pesticide occurrences in rural domestic wells and this helps to make a generalization as the wells with shallow ground water table are more susceptible to pesticide occurrence.

In the simulation of regional seawater quality modeling, Palani et al. (2008) have also used ANN approach. The reason of selecting this model is mentioned as its ability to represent not only linear but also non-linear relationships and learning these relationships directly from the data used in modeling. Investigated parameters are salinity, T, DO, and chl-a in the study. For the training and overfitting test data, simulation accuracy (coefficient of efficiency-R^2) ranged between 0.8 and 0.9. This showed that ANN has great potential to simulate water quality variable.

In the study of Singh et al. (2009), training, validation and application of ANN models is carried out in order to figure dissolved oxygen (DO) and biochemical oxygen demand (BOD) levels in the Gomti river (India). Two different models were studied, validated and tested in river water by eleven input water quality variables which are measured over 10 years period. The coefficient of determination (R^2) and

RMSE values are calculated for DO and BOD models. For R^2 is between 0.70 and 0.76 for DO for training, validation and test, whereas it is between 0.77 and 0.85 for BOD for training, validation and test. On the other hand, RMSE is between 1.23 and 1.5 for DO and it is between 1.38 and 2.25 for BOD. This study shows that the neural networks are effective tool for estimation of river water quality.

In the study of Najah et al. (2009), a ANN model is constructed in order to make simulation and prediction of water quality parameters namely; total dissolved solids, electrical conductivity and turbidity at Johor River and its tributary monitoring stations in Malaysia. The main reason of modeling these parameters is defined as the statistical correlation analysis of the field data, the domain knowledge and the prediction accuracy of the water quality parameters. It is found that the absolute mean percentage error for the three parameters is 10% for different water bodies. Six Multilayer Perceptron (MLP) ANN architectures (one for each parameters at each location) were developed in this study and it is showed that this model is reliable and useful in prediction of total dissolved solids, electrical conductivity and turbidity parameters with different data input patterns for testing part. Given studies are summarized in Table 3.3

Table 3.3: Summary table for the previous studies

Author	Year	Aim	Study Area	Data No	Used ANN Architecture	Performance Function
Wei et al.	2001	Quantification of the interactions between abiotic factors and algal genera	Lake Kasumigaura, Japan	120	Two-layer of FFNN	Regression coefficient
Kuo et al.	2004	To forecast the variation of the quality of Groundwater in the blackfoot disease area	The Yun-Lin coastal area in Taiwan	672	BP - ANN	RMSE and goodness-of-fit
Sahoo et al.	2005	Prediction of pesticide occurrence in	North Carolina domestic	4426	Three and four layer feed-forward BP,	Correlation coefficient (R), RMSE

Author	Year	Aim	Study Area	Data No	Used ANN Architecture	Performance Function
		rural domestic wells from the available limited information	wells		a radial basis function (RBF) and an ANFIS	and detection efficiency (E_f)
Palani et al.	2008	Prediction and forecasting of salinity, T, DO and chl-a parameters	Singapore coastal waters	48	General regression neural networks (GRNNs) and The Wardnet (WN) ANN architecture consisted of BP	RMSE, the mean absolute error (MAE) and Nash–Sutcliffe coefficient of efficiency (R^2)
Singh et al.	2009	Computing the DO and BOD levels	Gomti River, India	960	Three-layer feed-forward neural networks (FFNN) with back propagation (BP) learning	R^2 value and the root mean square of error (RMSE)
Najah et al.	2009	Prediction of TSS, EC and turbidity parameters	Johor River, Malaysia	30	Multi-Layer Perceptron (MLP) Network	The absolute mean error

3.3 Application of Artificial Neural Networks in Lake Köyceğiz with MATLAB

In this part, water quality parameters of Lake Köyceğiz are predicted with parameters of Namnam and Yuvarlakçay creeks by using ANN tool of MATLAB. For this aim, the method applied is illustrated in Figure 3.3.

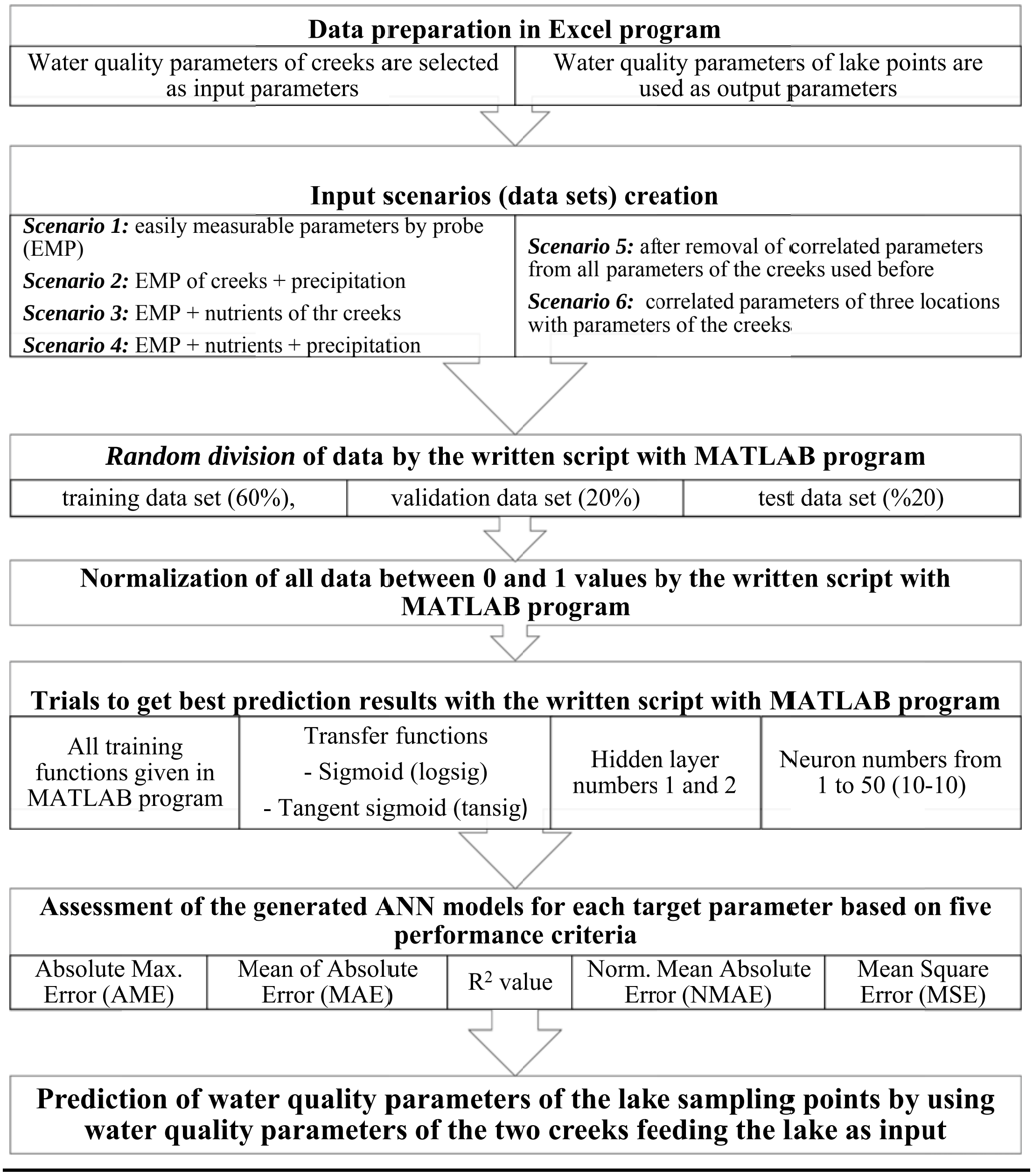


Figure 3.3: The applied method for ANN approach

In Figure 3.3, there are seven boxes to describe the applied methodology for prediction of water quality.

Box 1- Data preparation: Available data is prepared for Dalyan Entrance, Lake Center, Lake Beach, Yuvarlakçay and Namnam Creeks sampling points in Excel program. Water quality parameters of the creeks are chosen as input parameters while; water quality parameters of the lake points are chosen as output parameters. There are 63 data for each water quality parameters of the sampling points.

Box 2 - Creation of input scenarios: To predict water quality parameters of south, center and north points of the lake, parameters of Yuvarlakçay and Namnam creeks, which are two big creeks flowing to the lake, are used as input parameter. Total of six input scenarios (data sets) are created to obtain the best ANN structure. There are three sub-scenarios under first five scenarios based on sampling location that are Dalyan Entrance (south point of the lake), Lake Center, and Lake Beach (north point of lake). The first of the six scenarios consists of easily measurable parameters by a probe (T, pH, EC and DO) of the creeks as input for ANN. Following three scenarios are created by adding nutrients parameters (NH_4-N and TP) of the creeks and precipitation data to the first scenario. In addition, two more scenarios are created based on the obtained results from correlation analysis section. In this way, requirement and advantage of correlation analysis before ANN process can be assessed. In the fifth scenario, correlated parameters of the creeks between each other are removed from all parameters of them used in other scenarios. Input parameters of sixth scenario are correlated parameters of the creeks with parameters of the lake sampling locations. Scenarios are summarized in Table 3.4 and Table 3.5.

Scenario 1: easily measurable parameters by probe
Scenario 2: easily measurable parameters by probe of creeks + precipitation
Scenario 3: easily measurable parameters by probe + nutrients of creeks
Scenario 4: easily measurable parameters by probe + nutrients + precipitation
Scenario 5: after removal of correlated parameters from all parameters of the creeks used before
Scenario 6: correlated parameters of three locations with parameters of the creeks

Table 3.4: Scenarios for ANN

Scenarios	Sub-scenario	Input Parameters of Yuvarlakçay & Namnam Creek	OUTPUTS	
			Location	**Parameters**
1	south	T, pH, EC,DO	Dalyan Entrance	DO, NH_4-N, TP
	center	T, pH, EC,DO	Lake Center	DO, TN, TP, Chl-a
	north	T, pH, EC,DO	Lake Beach	DO, TN, TP, Chl-a, TC
2	south	T, pH, EC,DO + **Precipitation**	Dalyan Entrance	DO, NH_4-N, TP
	center	T, pH, EC,DO + **Precipitation**	Lake Center	DO, TN, TP, Chl-a
	north	T, pH, EC,DO + **Precipitation**	Lake Beach	DO, TN, TP, Chl-a, TC
3	south	T, pH, EC,DO + **NH_4-N** + **TP**	Dalyan Entrance	DO, NH_4_N, TP
	center	T, pH, EC,DO + **NH_4-N** + **TP**	Lake Center	DO, TN, TP, Chl-a
	north	T, pH, EC,DO + **NH_4-N** + **TP**	Lake Beach	DO, TN, TP, Chl-a, TC
4	south	T, pH, EC,DO + **NH_4-N** + **TP** + **Precipitation**	Dalyan Entrance	DO, NH_4-N, TP
	center	T, pH, EC,DO + **NH_4-N** + **TP** + **Precipitation**	Lake Center	DO, TN, TP, Chl-a
	north	T, pH, EC,DO + **NH_4-N** + **TP** + **Precipitation**	Lake Beach	DO, TN, TP, Chl-a, TC

Table 3.5: Scenarios 5 and 6 for ANN

	Sub-scenario	**Yuvarlakçay Creek**	**Namnam Creek**	**OUTPUTS**	
				Location	**Parameters**
5*	south	T, pH, EC, DO, NH_4_N, TP	pH, EC, TP	Dalyan Entrance	DO, NH_4_N, TP
	center			Lake Center	DO, TN, TP, Chl-a
	north			Lake Beach	DO, TN, TP, Chl-a, TC
	Sub-scenario	**Yuvarlakçay Creek**	**Namnam Creek**	**OUTPUTS**	
				Location	**Parameters**
6*	south	DO	DO	Dalyan Entrance	DO
	center	DO	DO	Lake Center	DO
	north	FC	FC, TC	Lake Beach	TC

(T: temperature; EC: electrical conductivity, DO: dissolved oxygen; NH_4-N: ammonium nitrogen; TP: total phosphorus; Chl-a: Chlorophyll-1; FC: fecal coliforms; TC: total coliforms)

*these scenarios are created based on correlation analysis conducted in the Section 2.5.

Box 3 - Random division of the data: After creation of input scenarios, 63 data is divided randomly with a written script using with MATLAB. The reason for random division of data is available data for chlorophyll-a and TP parameters that have detection limit measurements in sequence. This reason decreases the learning ability of generated ANN. As indicated in the most of literature, 60% of total data is selected as training set; 20% of total data as validation set and final 20% of total data is chosen as test data. ANN use training data set to realize the learning process. After that it chooses the best structure based on validation data set during training process. Finally, test data is used to evaluate the performance of the generated ANN structure. After random division, first 38 data (60% of all data) is selected as training data set; later 12 data (20% of all data) is chosen validation data set and last 13 data (20% of all data) is indicated as test data set. Figure 3.4 represents the data division as train, validation and test data sets, which is used in all best prediction result plots. The written script for random data division is given in APPENDIX C.

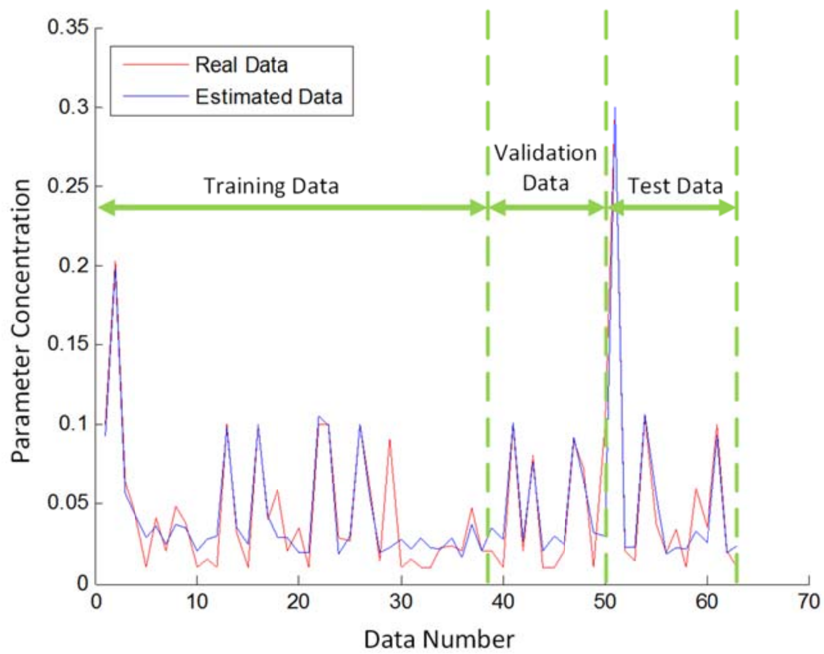


Figure 3.4: Representation of randomly divided data as train, validation and test

<u>Box 4 – Normalization of data:</u> After data division, all data sets are normalized between 0 and 1 values with a written script by using MATLAB program. Written scripts by using the MATLAB program for normalization and denormalization process of data is given in APPENDIX D.

Box 5 – Trials to get best prediction result: After creation of scenarios and data preparation, the main ANN script is written (APPENDIX E) by using the MATLAB program. In the written script, two transfer functions (log-sigmoid and tangent sigmoid), all training functions available in MATLAB program (Table 3.6) are tested. In addition, one and two hidden layers and 1 to 50 neurons (10-10) are tried to obtain the best simulation.

Table 3.6: Training functions in MATLAB for ANN

trainlm	Levenberg-Marquardt
trainbr	Bayesian Regularization
trainbfg	BFGS Quasi-Newton
trainrp	Resilient Backpropagation
trainscg	Scaled Conjugate Gradient
traincgb	Conjugate Gradient with Powell/Beale Restarts
traincgf	Fletcher-Powell Conjugate Gradient
traincgp	Polak-Ribiere Conjugate Gradient
trainoss	One Step Secant
traingdx	Variable Learning Rate Gradient Descent
traingdm	Gradient Descent with Momentum
traingd	Gradient Descent

Mean square error is used as stopping criteria to find the neural network structure with good performance.

Box 6 – Assessment of ANN performance: After finding the best ANN structure, its estimation performance examined by using five different performance criteria that are absolute maximum error (AME), mean of absolute error (MAE), R^2 value, normalized mean absolute error (NMAE) and mean square error (MSE).

Box 7 – Obtaining of the best prediction results: For selection of the best input scenario with the best performance criteria, followed steps are illustrated in Figure 3.5

.

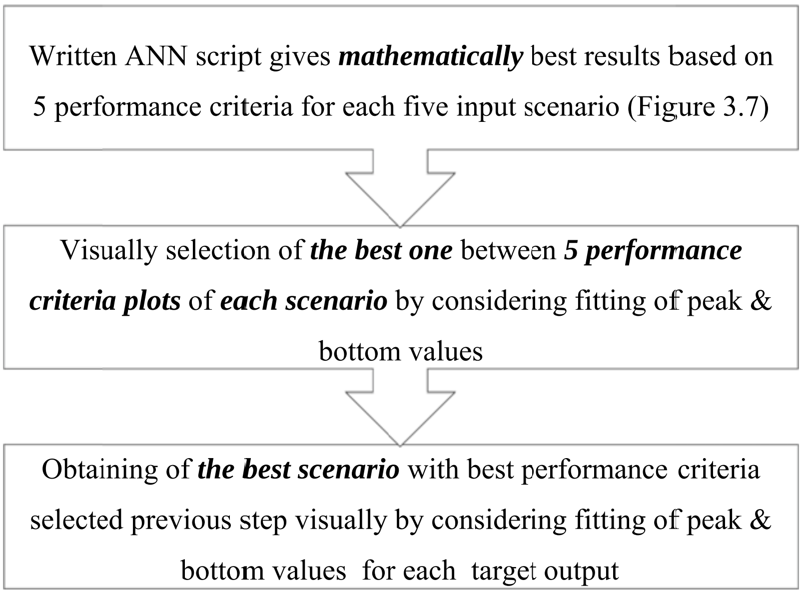


Figure 3.5: Selection of the best input scenario with the best performance criteria for output parameters

The written script gives the plots for each performance criteria based on minimum values of these criteria. For each scenario, the best plot is selected visually between these obtained five plots for each performance criteria. Then, the best scenario with best performance criteria selected in the previous step is chosen visually for each target output. Visual selection is conducted by considering fitting of peak & bottom values. This procedure is applied to each output parameter of three sampling points.

To clarify the mentioned process, selection of the best prediction result for DO concentration of Dalyan Entrance is given as example. In the example, third scenario, which in easily measurable parameters by a probe (T, pH, EC, DO) and nutrient parameters (NH_4-N and TP) of the creeks, is used as input data set. Firstly, Figure 3.6 shows the obtained five plots from the written script based on minimum values of five performance criteria.

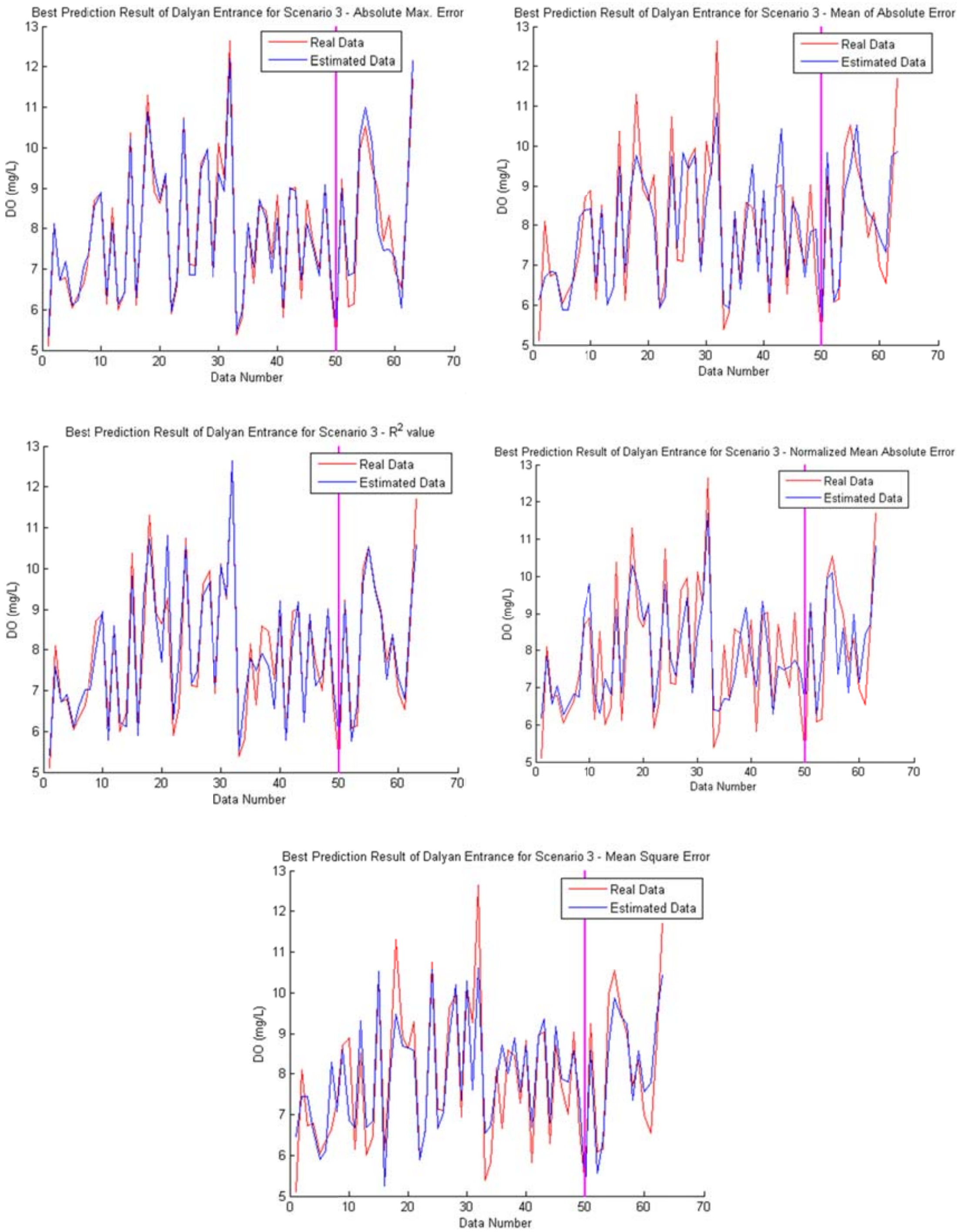


Figure 3.6: Obtained five plots based on minimum values of five performance criteria for DO parameter prediction of Dalyan Entrance with third scenario (left of the magenta line at 50^{th} data number shows test data set)

As can be seen from the above plots, the plot of performance criteria with R^2 value is the best prediction between them, visually. This visually performance criteria selection is made for each output parameter of each location for each scenario. Table 3.7 summarizes the selected performance criteria and obtained performance values for each output parameter of all locations for all five scenarios.

Table 3.7: Selected performance criteria and performance values for each output parameter of three locations for five scenarios

Location	Output Parameter	Scenario	Selected Performance Criteria	Performance Value Whole Data	Performance Value Test Data
Dalyan Entrance (South Point)	DO	1	MAE	0.5920	0.4505
		2	AME	1.1791	0.4206
		3	R^2	0.9134 (0,89*)	0.9520 (0,85*)
		4	MAE	0.7013	0.8495
		5**	MAE	0.6286	0.7507
		6**	MAE	0.7009	0.7239
	NH_4-N	1	R^2	0.8265 (0,80*)	0.9654 (0,92*)
		2	MAE	0.0109	0.0097
		3	MAE	0.0042	0.0038
		4	R^2	0.9108(0,89*)	0.9612(0,94*)
		5**	MAE	0.0129	0.0153
	TP	1	NMAE	340.479	432.113
		2	MSE	0.0071	0.0077
		3	MSE	0.0000	0.0000
		4	AME, MAE, R^2, MSE	0.0000, 0.0000, 1.0000(1.00*) 0.0000	0.0000, 0.0000, 1.0000(1.00*) 0.0000
		5**	AME, R^2	0.0000, 0.9998(1.00*)	0.0000, 0.9999(1.00*)

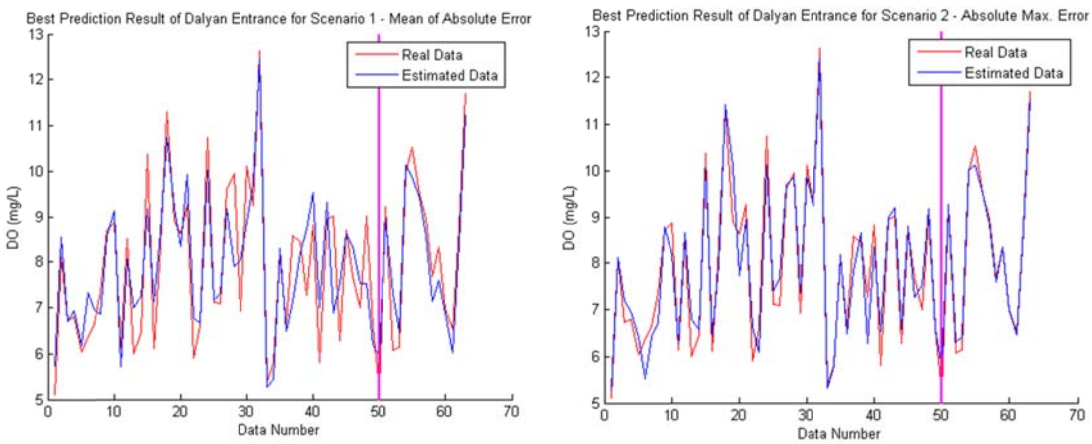
Location	Output Parameter	Scenario	Selected Performance Criteria	Performance Value Whole Data	Performance Value Test Data
Lake Center	DO	1	AME	1.6404	1.1259
		2	R^2	0.6696 (0.65*)	0.6857 (0.62*)
		3	MSE	0.1888	0.1138
		4	AME	3.0801	1.5101
		5**	AME	2.8243	2.5196
		6**	R^2	0.6106 (0.38*)	0.3158 (0.60*)
	TN	1	NMAE	18.744	17.697
		2	R^2	0.9416 (0.92*)	0.9888 (0.96*)
		3	NMAE	27.470	19.700
		4	MAE	0.1205	0.1003
		5**	R^2	0.9078(0.88*)	0.9535(0.96*)
	TP	1	MAE	0.0000	0.0000
		2	AME	0.0041	0.0041
		3	AME, MAE, R^2, MSE	0.0000, 0.0000, 1.0000(1.00*) 0.0000	0.0000, 0.0000, 1.0000(1.00*) 0.0000
		4	AME, R^2, MSE	0.0000, 1.0000(1.00*) 0.0000	0.00000, 1.00000(1.00*) 0.00000
		5**	MAE, R^2	0.0457, 1.0000(1.00*)	0.0257, 1.0000(1.00*)
	Chl-a	1	MAE	0.4490	0.4314
		2	MSE	0.0967	0.0351
		3	MAE	0.3915	0.4574
		4	MSE	0.2427	0.1960
		5**	AME	2.6893	2.6893

Location	Output Parameter	Scenario	Selected Performance Criteria	Performance Value Whole Data	Performance Value Test Data
Lake Beach (North Point)	DO	1	MSE	0.2019	0.1219
		2	NMAE	3.4208	2.3713
		3	MSE	0.0019	0.0070
		4	R^2	0.9291 (0.93*)	0.9421(0.97*)
		5**	MSE	0.4171	0.3397
	TN	1	MSE	0.0287	0.0381
		2	MAE	0.0462	0.0330
		3	AME	0.2373	0.0941
		4	AME	0.1431	0.1431
		5**	MAE	0.2277	0.2296
	TP	1	MAE	0,0047	0,0030
		2	MAE	0.0022	0.0019
		3	R^2	0.9778 (0.96*)	0.9929(0.95*)
		4	AME	0.0035	0.0007
		5**	R^2	0.8117(0.77*)	0.9021(0.87*)
	TC	1	R^2	0.9989 (1.00*)	0.9998(1.00*)
		2	AME	1303.1	344.00
		3	R^2	0.9914 (0.99*)	0.9992(1.00*)
		4	R^2	1.0000(0.99*)	1.0000(0.99*)
		5**	AME	77600	43500
		6**	MSE	1.0691	3.3412
	Chl-a	1	MAE	0.0000	0.0000
		2	MSE	0.0001	0.0000
		3	MAE	0.1737	0.1336
		4	AME, MAE, R^2, MSE	0.0000, 0.0016, 1.0000(1.00*) 0.0000	0.0000, 0.0014, 1.0000(1.00*) 0.0000
		5**	R^2	0.8656(0.85*)	0.9187(0.92*)

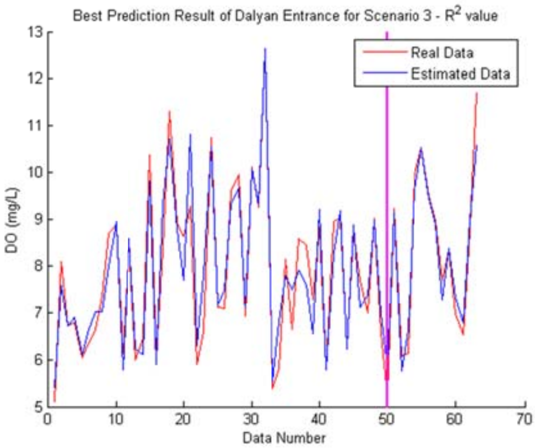
(MAE: Mean of Absolute Error; AME: Absolute Maximum Error; R^2: R^2 value; NMAE: Normalized Mean Absolute Error; MSE: Mean Square Error; T: temperature; EC: electrical conductivity, DO: dissolved oxygen; NH_4-N: ammonium nitrogen; TP: total phosphorus; Chl-a: Chlorophyll-1; TC: total coliforms)

*slope values are given in parenthesis
**these scenarios are created based on correlation analysis conducted in the Section 2.5.

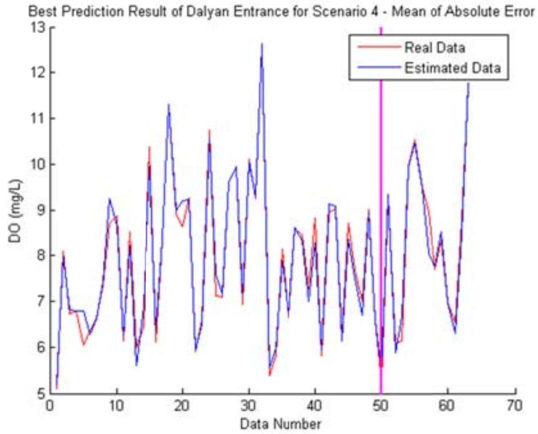
After performance criteria selection for each scenario, best scenario for prediction of each parameter of all three locations is selected visually. Again, DO parameter prediction of Dalyan Entrance is given as example to clarify this selection process. Figure 3.7 shows the obtained best performance criteria plots for five scenarios. For other output parameters, obtained best performance criteria plots are given in APPENDIX F.



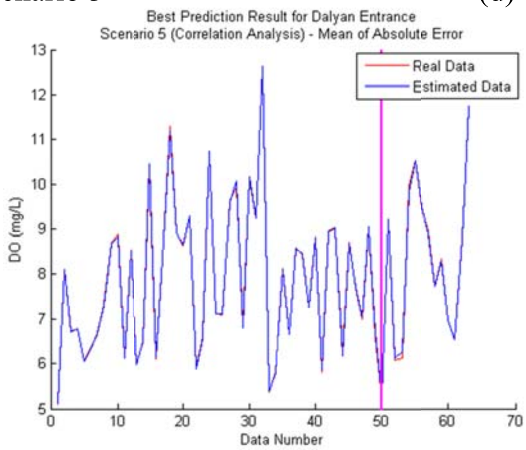
(a) Scenario 1 (b) Scenario 2



(c) Scenario 3



(d) Scenario 4

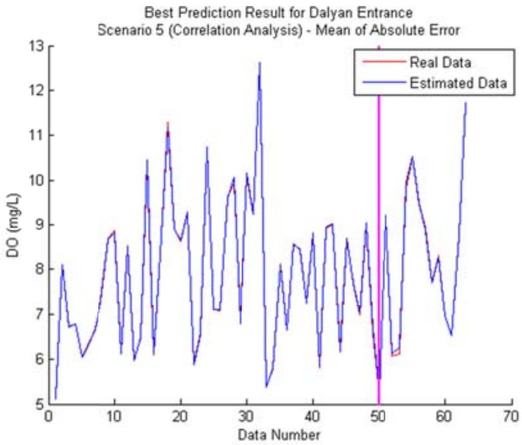


(e) Scenario 5

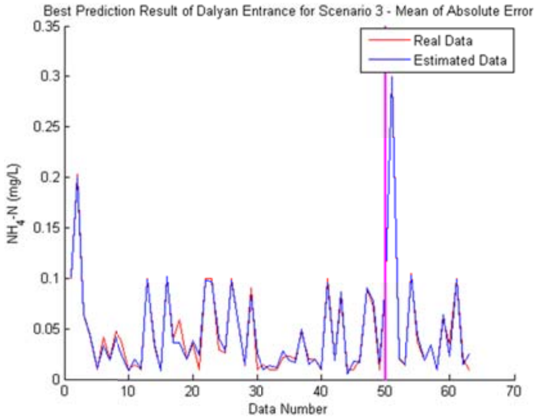
Figure 3.7: Selected best performance criteria for DO prediction of Dalyan Entrance for five scenarios (left of the magenta line at 50th data number shows test data set)

As can be seen from the best performance plots for each scenario, prediction of DO can be done with using the fifth input scenario give better result than other scenarios. Therefore, it can be said that fifth scenario is the best input data set to predict the DO parameter of Dalyan Entrance. In other words, DO of Dalyan Entrance can be predicted more precisely by MAE performance criteria and using input parameters of fifth scenario which are obtained from the correlation analysis conducted in Section 2.5. T, pH, EC, DO, NH_4-N and TP parameters of Yuvarlakçay Creek; pH, EC and

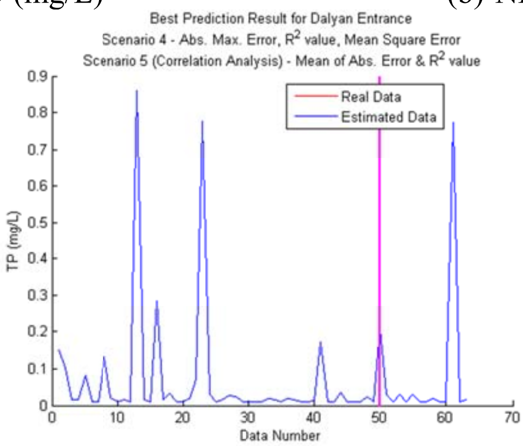
TP parameters of Namnam Creek are inputs parameters of the fifth scenario. Figure 3.8 gives the best prediction results of each output parameters for Dalyan Entrance.



(a) DO (mg/L)



(b) NH_4-N (mg/L)

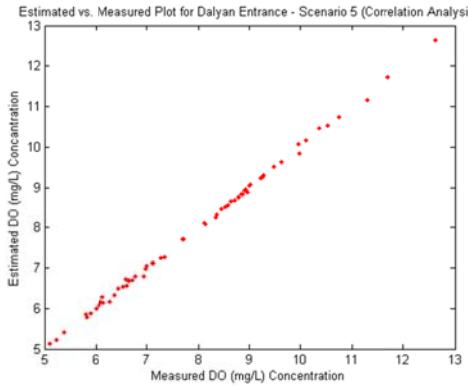


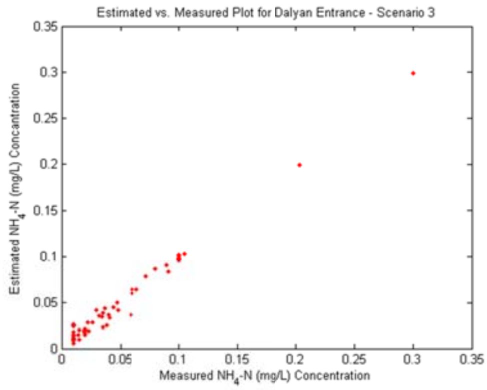
(c) TP (mg/L)

Figure 3.8: Best plots for (a) DO, (b) NH_4-N, (c) TP parameters of Dalyan Entrance (left of the magenta line at 50th data number shows test data set)

For Dalyan Entrance location, scenarios 5 and 3 are best to predict DO and NH4-N, respectively. Scenario 4 and 5 gives same results for TP parameter of Dalyan Entrance. For NH4-N parameter of Dalyan Entrance, easily measurable parameters by a probe (T, pH, EC and DO) and nutrients parameters (NH4-N and TP) of the creeks are sufficient as input. For DO parameter prediction, T, pH, EC, DO, NH4-N and TP parameters of Yuvarlakçay Creek; pH, EC and TP parameters of Namnam

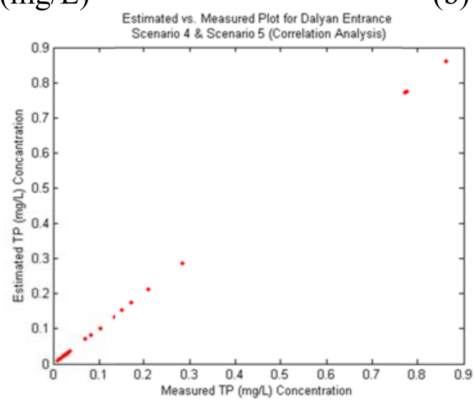
Creek are sufficient as input parameters. For TP prediction, there are two options. It can be predicted by using of easily measurable parameters by a probe (T, pH, EC and DO) and nutrients parameters (NH_4-N and TP) of the creeks and precipitation data as input or results obtained from the correlation analysis as in prediction of DO parameter of Dalyan Entrance. Also, MAE performance criteria gives the best plots for prediction of DO and NH_4-N parameters. For TP prediction, AME, MAE, R^2 value and MSE performance criteria give the same plots when scenario 4 is used, therefore four performance criteria can be used with fourth scenario for TP prediction. In addition, MAE and R^2 value give the same results with scenario 5. Measured vs. predicted plots of best prediction results and sampling date vs. concentration plots for each output for Dalyan Entrance location are given in
Figure 3.9 and Figure 3.10, respectively.





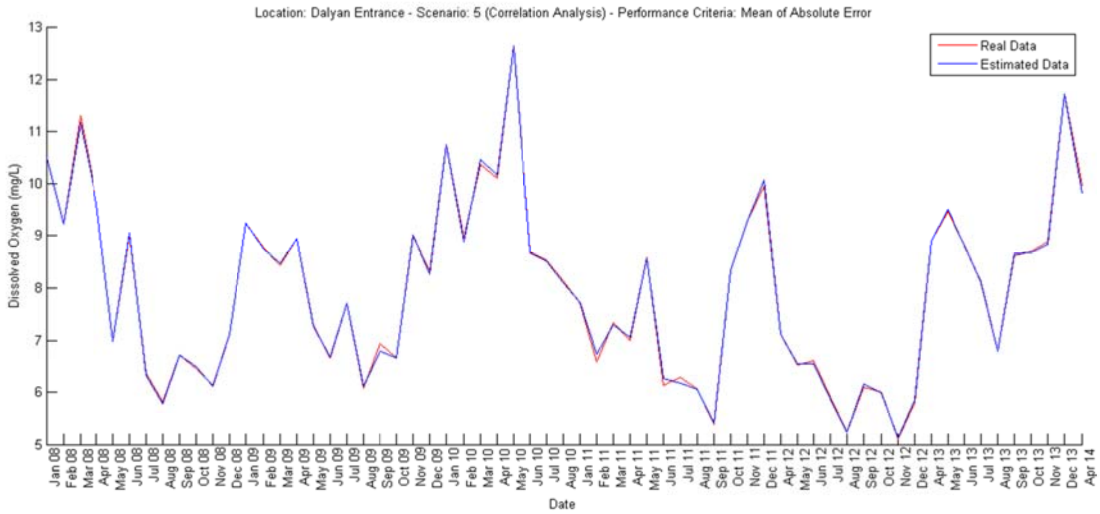
(a) DO (mg/L)

(b) NH_4-N (mg/L)

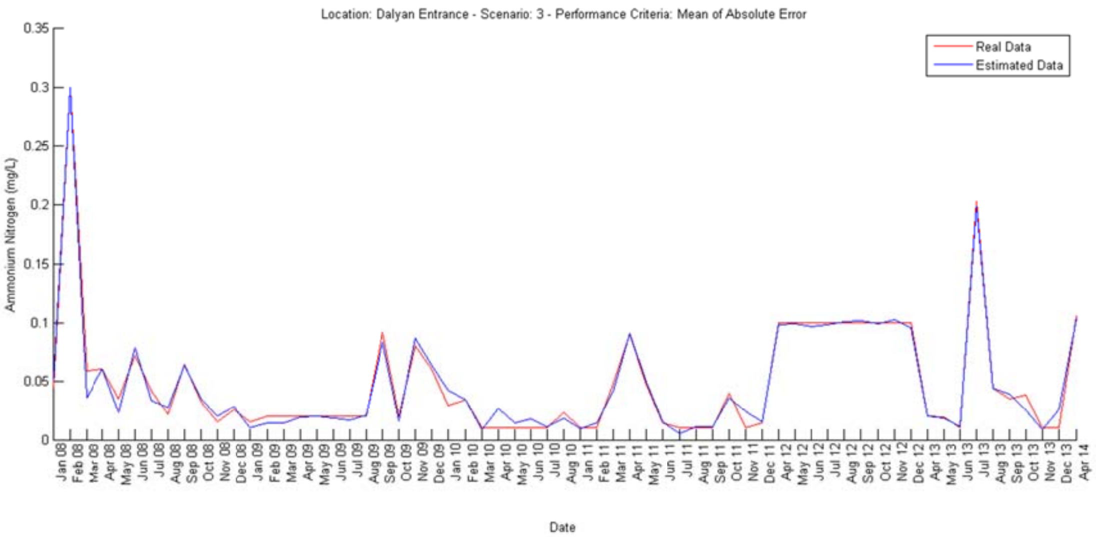


(c) TP (mg/L)

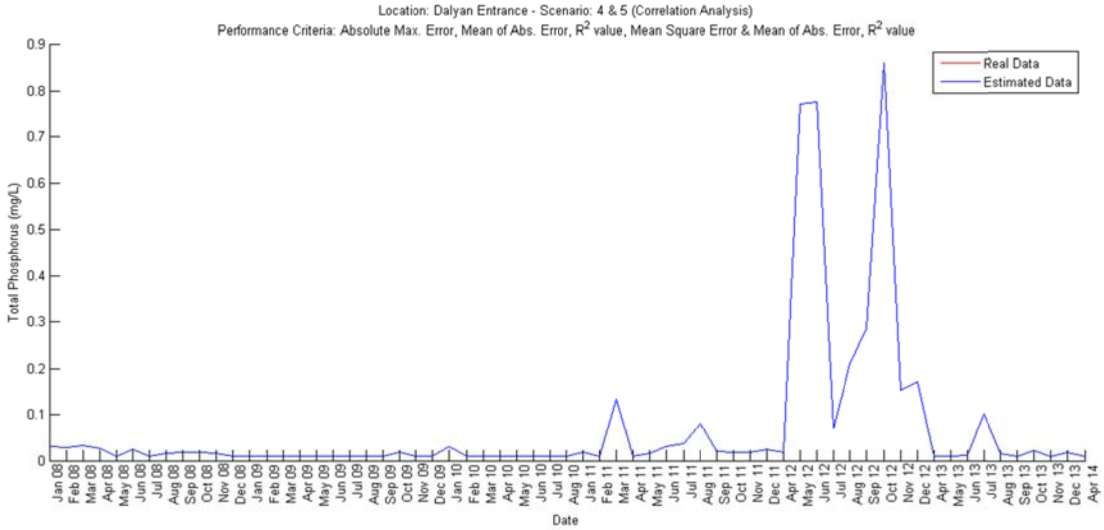
Figure 3.9: Measured vs. estimated plots for (a) DO, (b) NH_4-N, (c) TP parameters of Dalyan Entrance



(a) DO (mg/L)



(b) NH_4-N (mg/L)

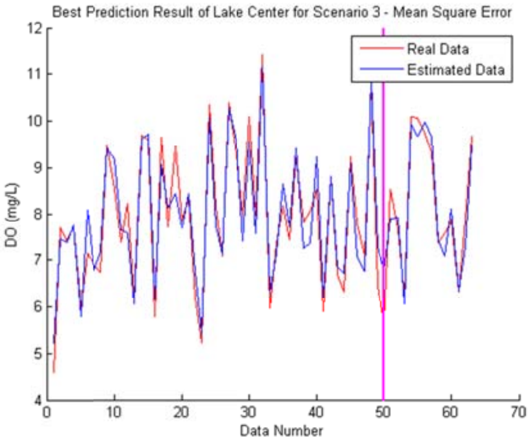


(c) TP (mg/L)

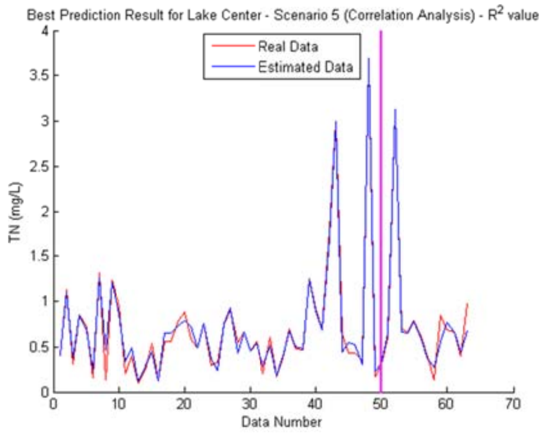
Figure 3.10: Sampling date vs. concentration plots for (a) DO, (b) NH_4-N, (c) TP parameters of Dalyan Entrance

To see changes of generated ANN performance according to the seasons, plots are given based on sampling data sequence. ANN performances for DO and NH4-N parameters of Dalyan Entrance have variations with respect to sampling date. In general, ANN performance is better for peak and bottom values than the rest of the values. ANN performance is very good for prediction of TP parameter. However, failures of the ANN occur in different seasons and years so, ANN performance can not be interpreted based on seasons.

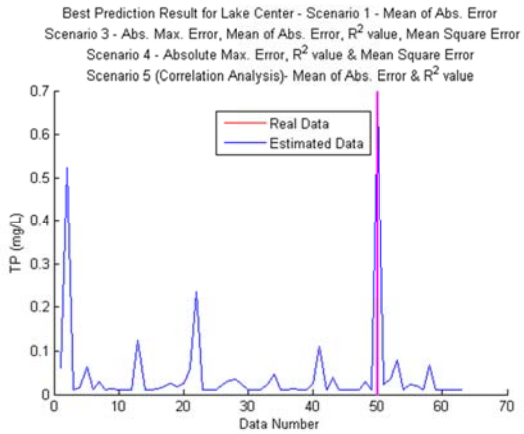
For second location that Lake Center, Figure 3.11 gives the best prediction results of output parameters (DO, TN, TP and Chl-a).



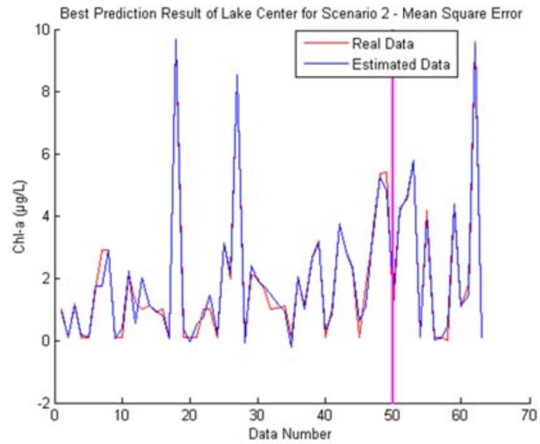
(a) DO (mg/L)



(b) TN (mg/L)



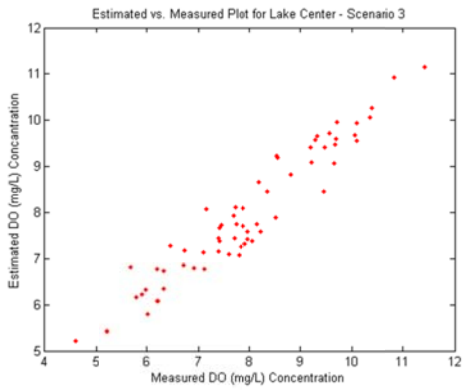
(c) TP (mg/L)



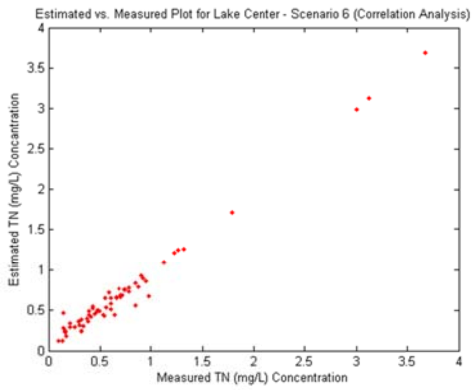
(d) Chl-a (µg/L)

Figure 3.11: Best plots for prediction of (a) DO, (b) TN, (c) TP and (d) Chl-a parameters for Lake Center (left of the magenta line at 50th data number shows test data set)

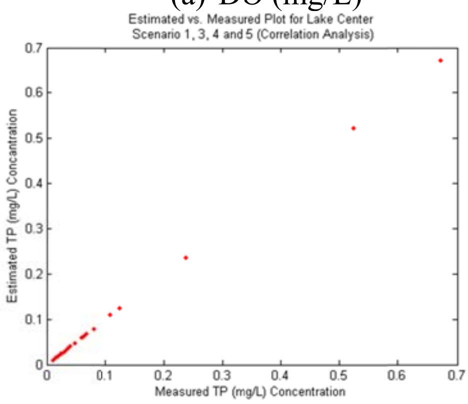
Scenarios 3, 5, and 2 are best to predict DO, TN and Chl-a parameters, respectively for Lake Center location. For Chlorophyll-a parameter, easily measurable parameters by a probe (T, pH, EC and DO) of the creeks and precipitation data are sufficient as input while; for DO parameter, addition of nutrients parameters (NH_4-N and TP) of the creeks is needed. For TN parameter prediction, T, pH, EC, DO, NH_4-N and TP parameters of Yuvarlakçay Creek; pH, EC and TP parameters of Namnam creek are sufficient as input parameters. For TP parameter, Scenarios 1, 3, 4 and 5 give same results visually. Therefore, only easily measurable parameters by a probe (T, pH, EC and DO) of the creeks can be enough as input for prediction of TP parameter. Based on performance criteria, MSE performance criteria gives the best plots for prediction of DO and Chl-a parameters. For TN prediction, R^2 value performance criteria gives best prediction plot. Finally, MAE criteria for first scenario; AME, MAE, R^2 value and MSE performance criteria for third scenario; AME, R^2 value and MSE performance criteria for fourth scenario; MAE and R^2 value for fifth scenario give the best plots for TP prediction. Measured vs. predicted plots of the best prediction results and sampling date vs. concentration plots for each output parameter of Lake Center location are given in Figure 3.12 and Figure 3.13, respectively.



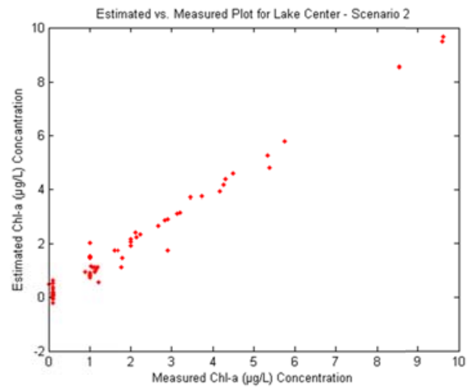
(a) DO (mg/L)



(b) TN (mg/L)

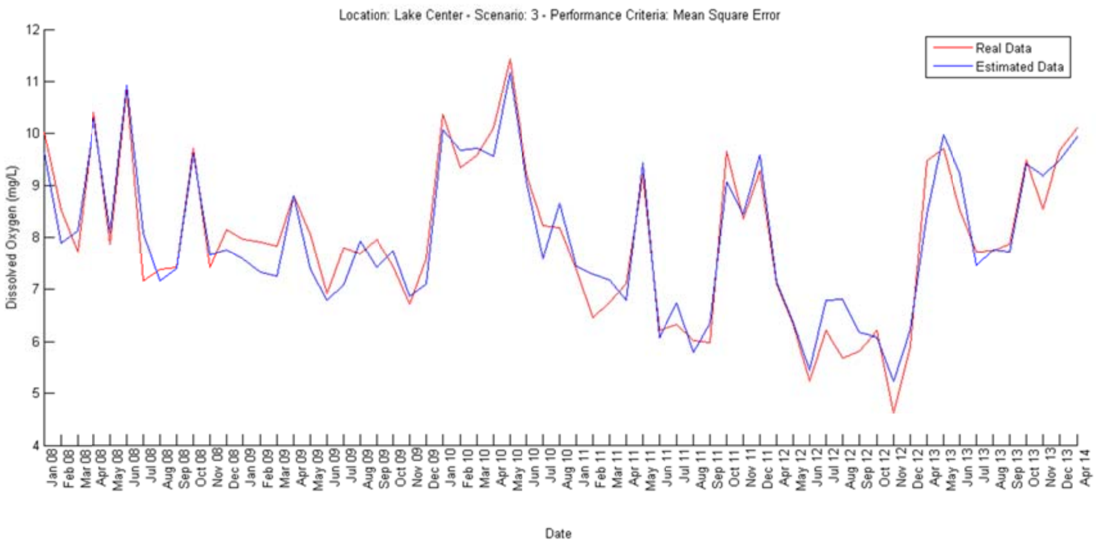


(c) TP (mg/L)

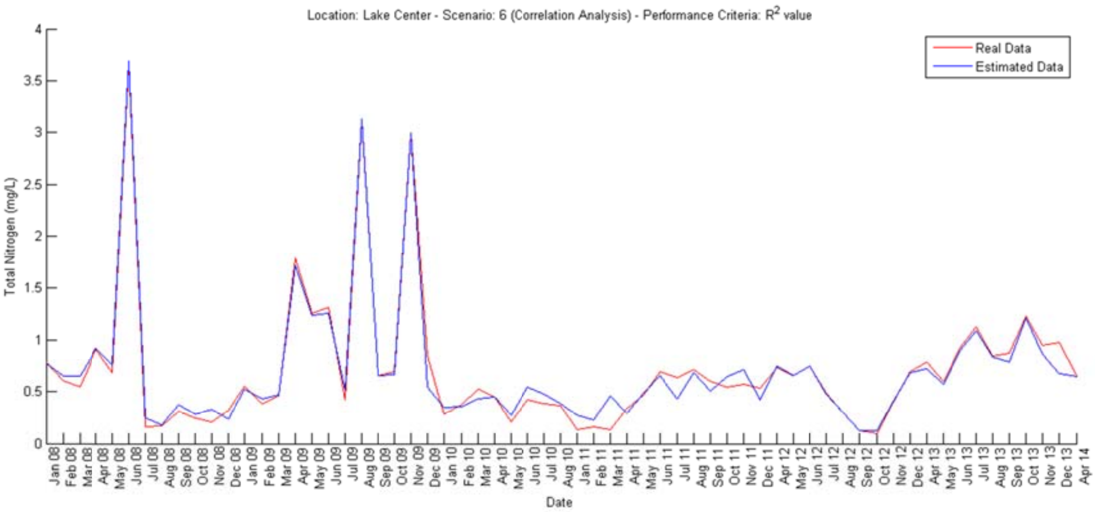


(d) Chl-a (µg/L)

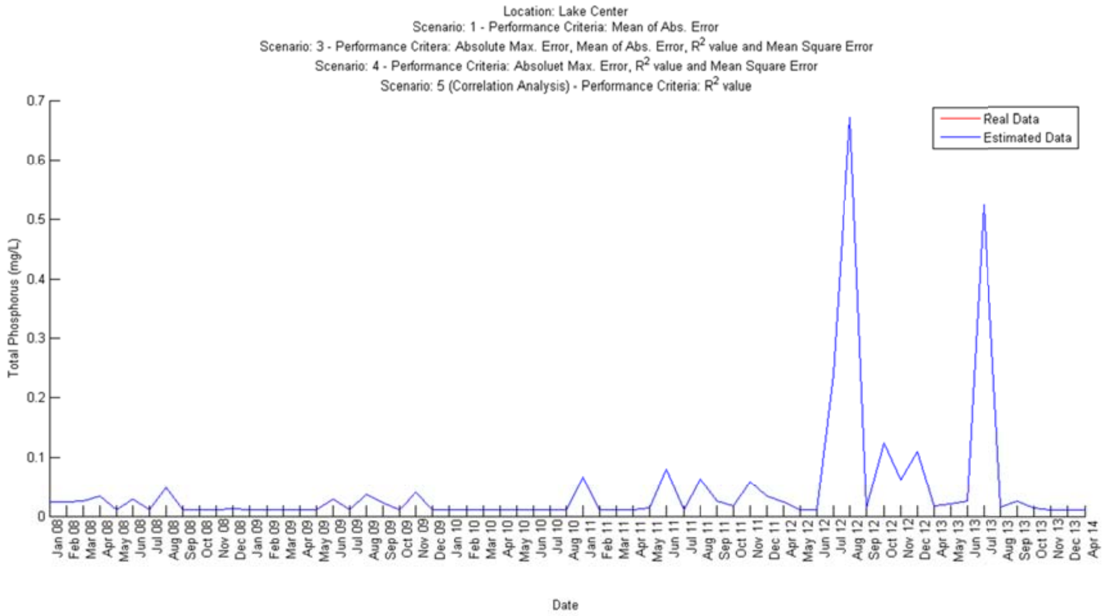
Figure 3.12: Measured vs. estimated plots for (a) DO, (b) TN, (c) TP, (d) Chl-a parameters of Lake Center



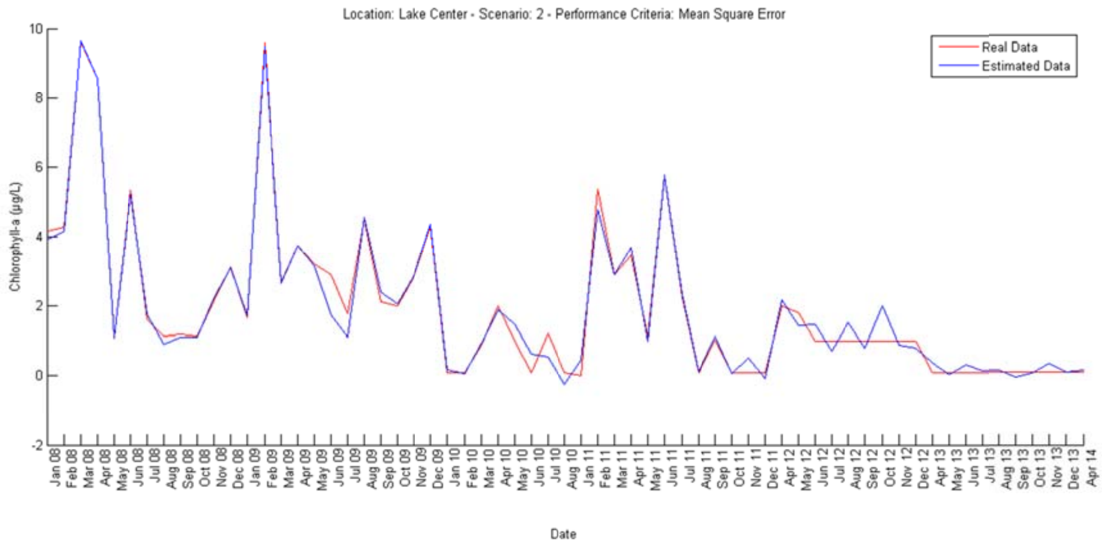
(a) DO (mg/L)



(b) TN (mg/L)



(c) TP (mg/L)

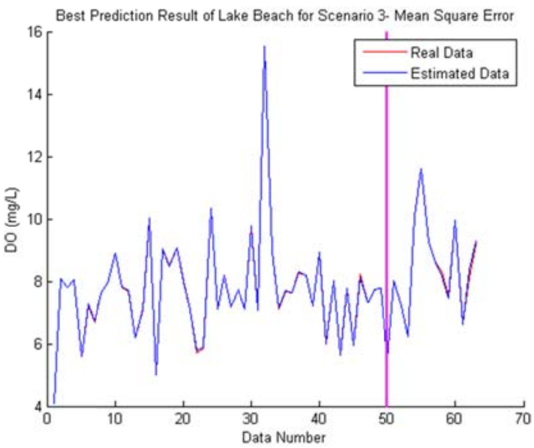


(d) Chl-a (µg/L)

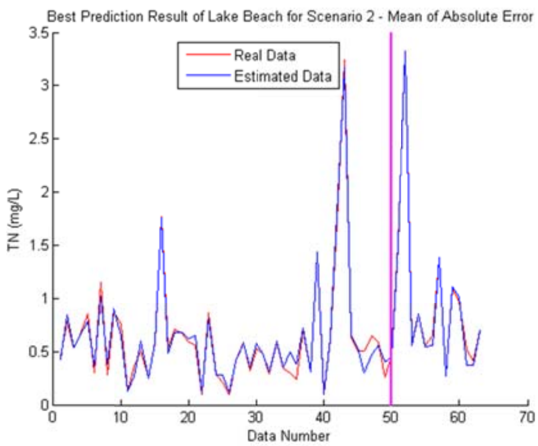
Figure 3.13: Sampling date vs. concentration plots for (a) DO, (**b**) NH_4-N, (c) TP, (d) Chl-a parameters of Lake Center

ANN performance for prediction of DO parameter gets worse than other parameters of Lake Beach. For DO and TN parameters, ANN follows the trend in general however; it can not make precise predictions. ANN performs excellent for TP prediction. In Chl-a plot, failures of ANN occurs in the year 2012 and 2013. The reason of the decrease in performance is existence of detection limit measurements in these years.

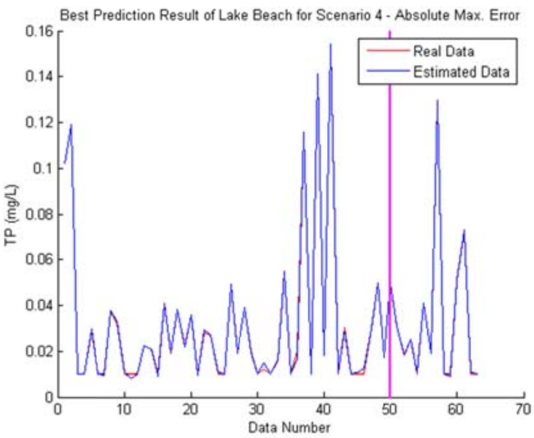
Figure 3.14 gives the best prediction results of five output parameters for the last location Lake Beach.



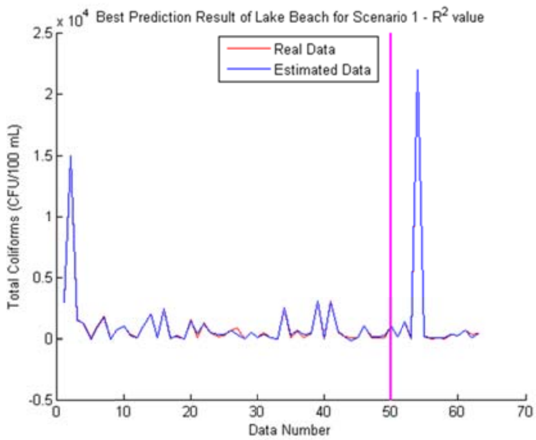
(a) DO (mg/L)



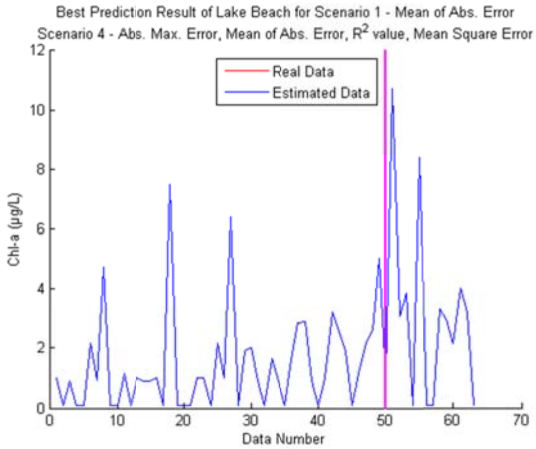
(b) TN (mg/L)



(c) TP (mg/L)



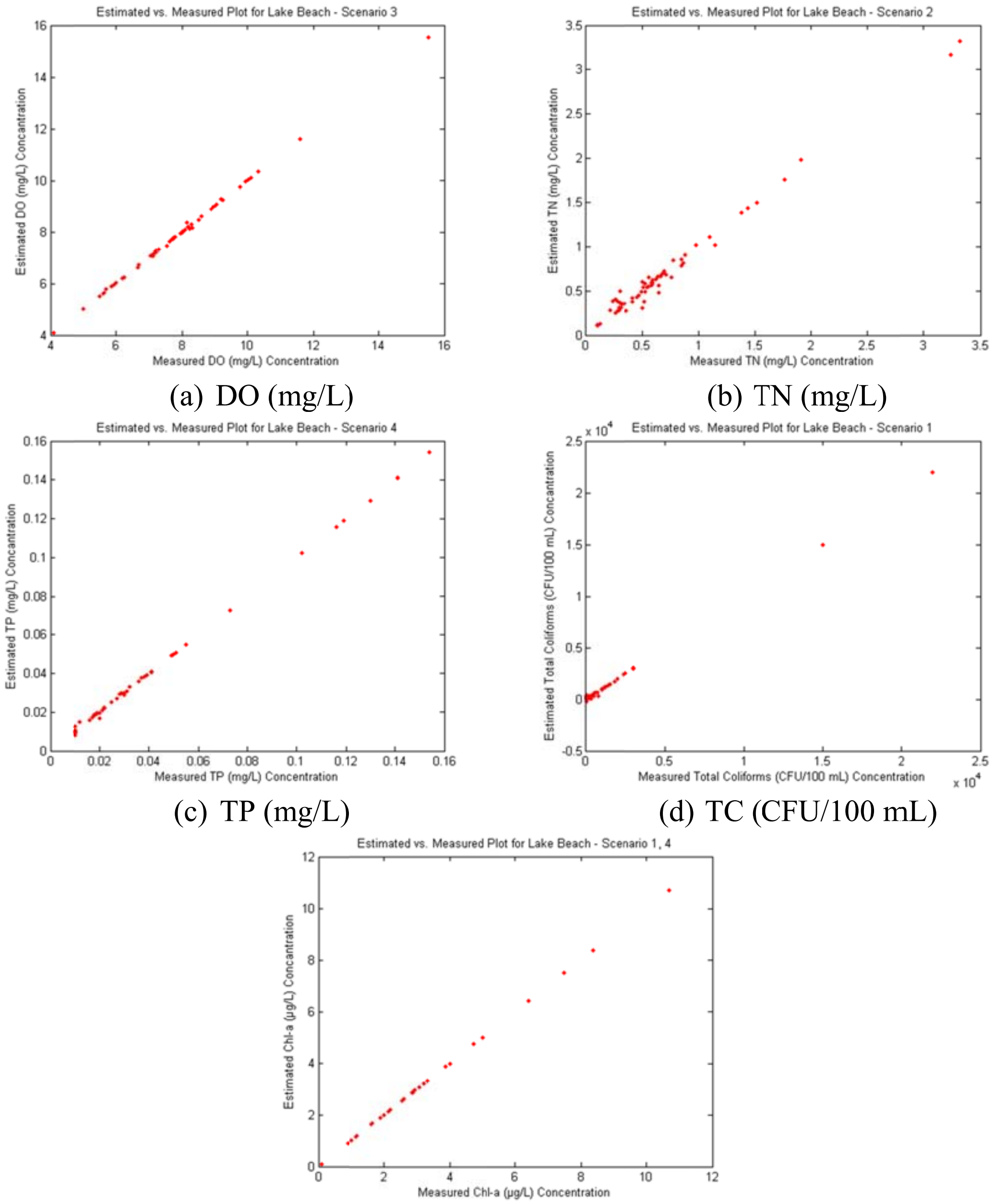
(d) TC (CFU/100 mL)



(e) Chl-a (µg/L)

Figure 3.14: Best plots for prediction of (a) DO, (b) TN, (c) TP, (d) TC and (e) Chl-a parameters of Lake Beach (left of the magenta line at 50th data number shows test data set)

For Lake Beach, scenarios 3, 2, 4, and 1 are best to predict DO, TN, TP and TC, respectively. For TC parameter of Lake Beach, only easily measurable parameters by a probe (T, pH, EC and DO) of the creeks are sufficient as input. Precipitation data should be added as inputs to find best prediction result of TN parameter. However, for DO prediction, addition of nutrient parameters (NH_4-N and TP) of the creeks is needed as input. Finally, precipitation data should be supplemented to easily measurable parameters by a probe (T, pH, EC and DO) and nutrient parameters (NH_4-N and TP) of the creeks as input for TP parameter. For prediction of Chl-a parameter, Scenarios 1 and 4 give same results visulally. Therefore, only easily measurable parameters by a probe (T, pH, EC and DO) of the creeks are sufficient as input for prediction of Chl-a parameter. Based on performance criteria, MSE, MAE, AME and R^2 value give the best plots for prediction of DO, TN, TP and TC parameters, respectively. Finally, for Chl-a prediction, MAE performance criteria for firts scenario; AME, MAE, R^2 value and MSE criteria for fourth scenario give same plots. Measured vs. predicted plots of the best prediction results and sapling date vs. concentration plots for each output of Lake Beach location are given in Figure 3.15 and Figure 3.16, respectively.



(a) DO (mg/L)

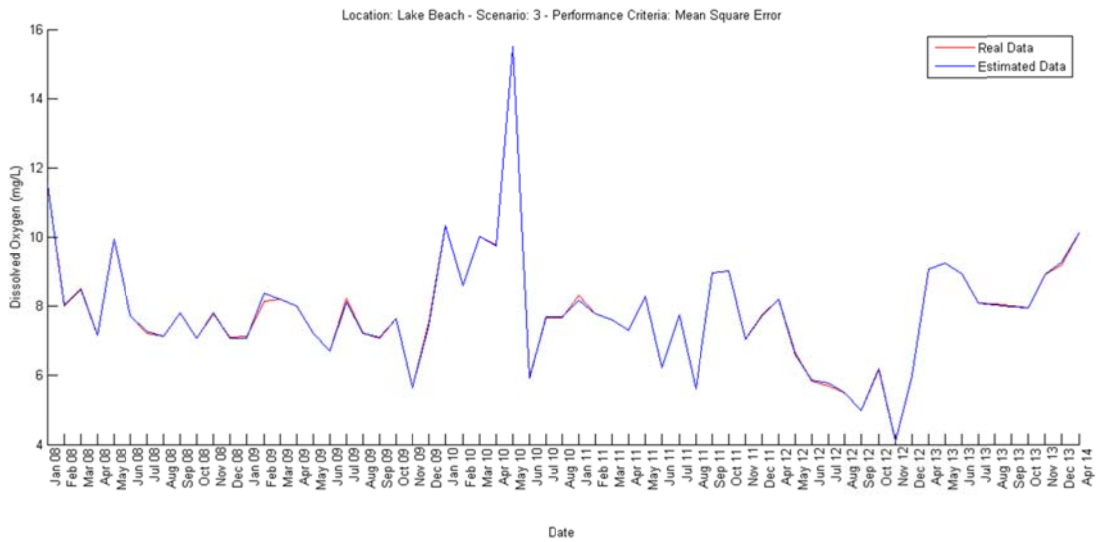
(b) TN (mg/L)

(c) TP (mg/L)

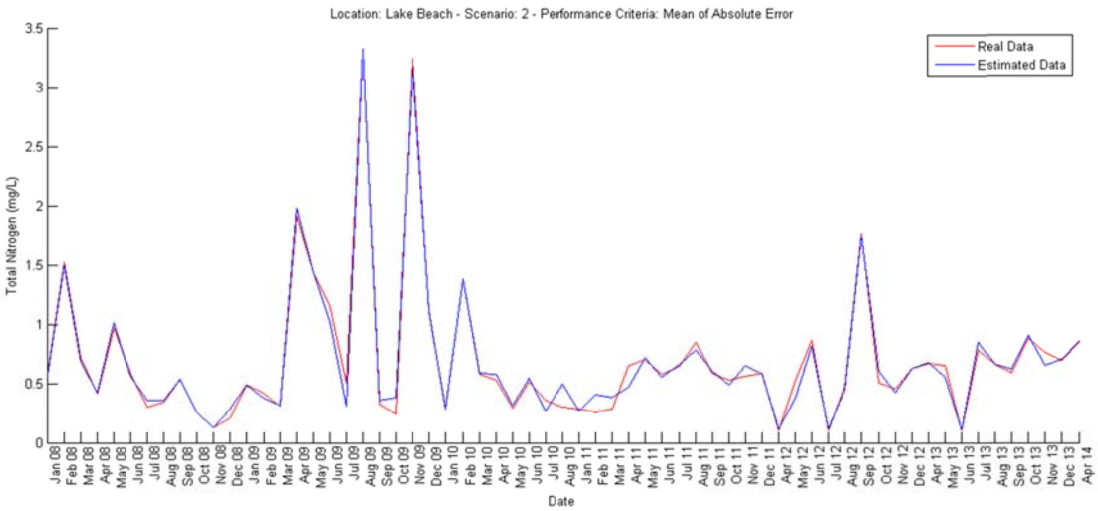
(d) TC (CFU/100 mL)

(e) Chl-a (μg/L)

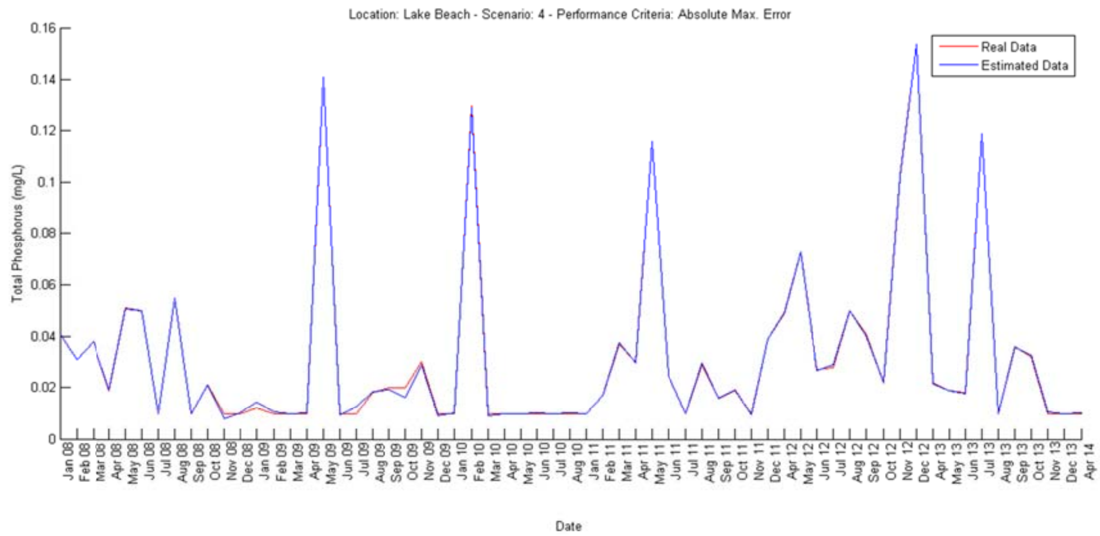
Figure 3.15: Measured vs. estimated plots for (a) DO, (b) TN, (c) TP, (d) TC and (e) Chl-a parameters of Lake Beach



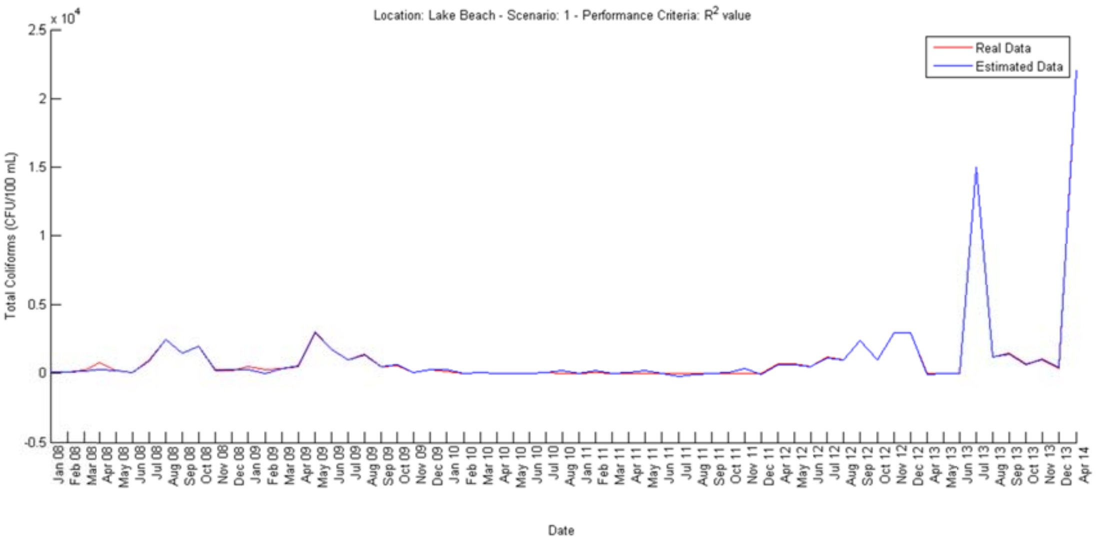
(a) DO (mg/L)



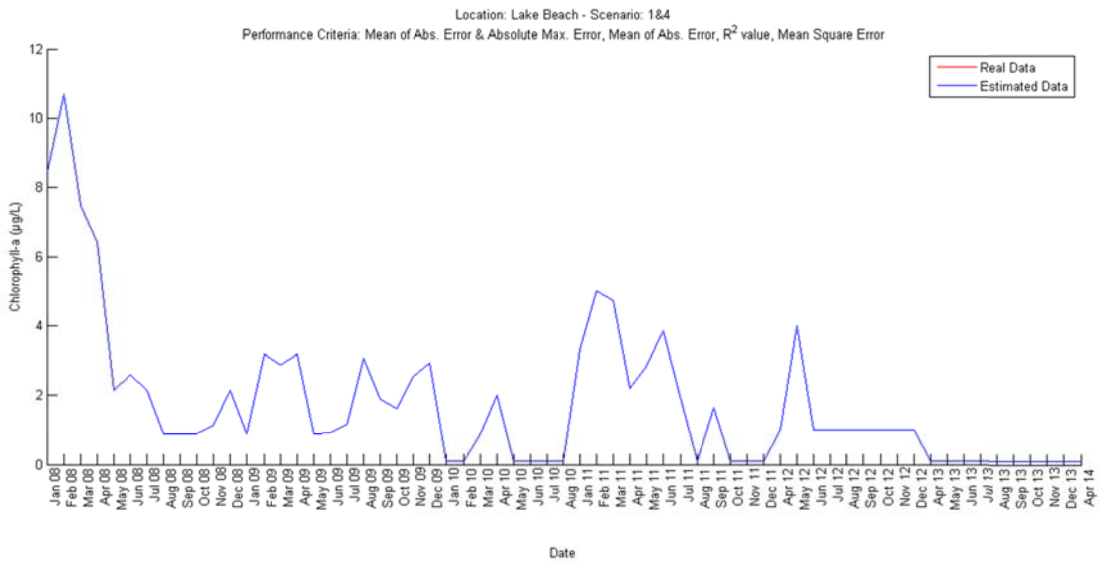
(b) TN (mg/L)



(c) TP (mg/L)



(d) TC (CFU/100 mL)



(e) Chl-a (µg/L)

Figure 3.16: Sampling data vs. concentration plots for (a) DO, (b) TN, (c) TP, (d) TC and (e) Chl-a parameters of Lake Beach

ANN performs well for DO, TP and TC predictions. However, deviations from the real data are observed during the year 2009 except a peak concentration. The reason of this fact can be wrong measurements for the year 2009. The performance for TN parameter is worse than the other parameters of the beach. Finally, ANN performance is perfect for Chl-a parameter prediction.

Visually selected scenarios based on performance criteria, their transfer function in hidden layer 1 and 2 and training functions, neuron numbers in hidden layer 1 and 2 are given in Table 3.8 for three locations' output parameters.

Table 3.8: Selected best ANN structures for Dalyan Entrance, Lake Center and Lake Beach

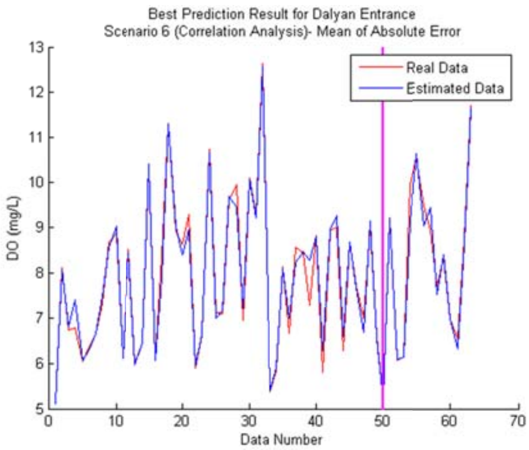
Location	Output Parameter	Scenario	Selected Performance Criteria	Transfer Function 1	Transfer Function 2	Training Function	Neuron Number 1	Neuron Number 2
Dalyan Entrance (south)	**DO**	5*	MAE	logsig	logsig	trainbr	1	41
	NH_4-N	3	MAE	logsig	tansig	trainbr	1	31
	TP	4	AME, MAE, R^2, MSE	tansig, logsig, logsig, logsig	tansig	trainbr	21, 11, 11, 11	21, 11, 11, 11
		5*	MAE, R^2	tansig	logsig	trainbr	21	21
Lake Center	**DO**	3	MSE	tansig	tansig	trainbr	1	31
	TN	5*	R^2	logsig	tansig	trainbr	1	41
	TP	1	MAE	tansig	logsig	trainbr	11	11
		3	AME, MAE, R^2, MSE	tansig	tansig	trainbr	11	11
		4	AME, R^2, MSE	logsig	tansig	trainbr	31	31
		5*	MAE, R^2	logsig	tansig	trainbr	11	31
	Chl-a	2	MSE	logsig	tansig	trainbr	1	41
Lake Beach (north)	**DO**	3	MSE	logsig	tansig	trainbr	1	41
	TN	2	MAE	logsig	tansig	trainbr	1	41
	TP	4	AME	tansig	tansig	trainbr	1	41
	TC	1	R^2	tansig	logsig	trainbr	1	41
	Chl-a	1	MAE	tansig	logsig	trainbr	41	11
		4	AME, MAE, R^2, MSE	tansig	tansig	trainbr	21	41

(MAE: Mean of Absolute Error; AME: Absolute Maximum Error; R^2: R^2 value; MSE: Mean Square Error)

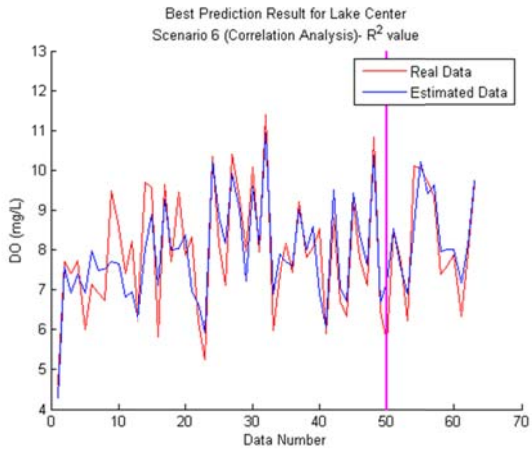
* the scenario is created based on correlation analysis conducted in the Section 2.5.

As can be seen from the table, trainbr (Bayesian Regularization) training function is best for all cases. Transfer function type show alterations between log-sigmoid and tangent-sigmoid transfer functions for hidden layer number 1 and 2 based on output parameter and performance criteria.

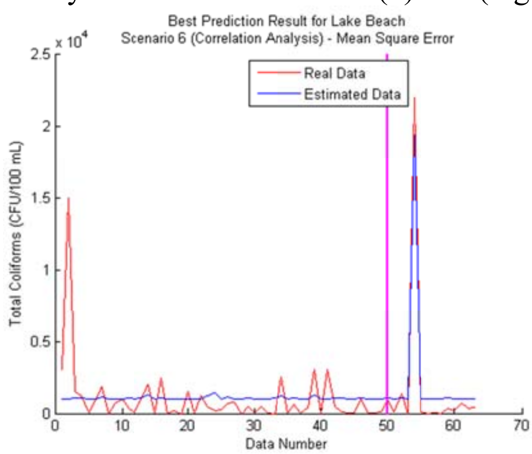
Results obtained with sixth scenario are given in Figure 3.17. They are best ones that chosen from five performance parameters. In this scenario, parameters of the creeks, which are correlated with output parameters of the three points of the lake, are used as input. At the end of the correlation analysis in Section 2.5, only output parameter, which is correlated with parameters of the creeks, is found as DO for Dalyan Entrance and Lake Center points. In addition, inputs are obtained as again DO parameter of the creeks. For Lake Beach, TC is reached as only output parameter correlated with parameters of Namnam and Yuvarlakçay Creeks. Besides, FC and TC parameters of Namnam Creek and FC parameter of Yuvarlakçay Creek are obtained from analysis as input.



(a) DO (mg/L) of Dalyan Entrance



(b) DO (mg/L) of Lake Center

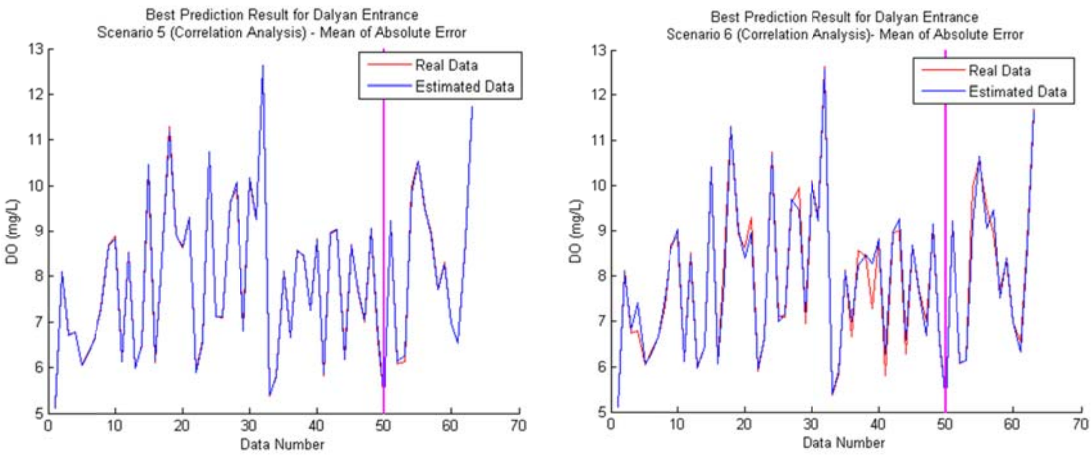


(c) TC (CFU/100 mL) of Lake Beach

Figure 3.17: Best prediction results obtained from Scenario 6 for three locations

As can be seen from the Figure 3.17, ANN generated based on correlation analysis result catch the general trend in measured DO parameter for Dalyan Entrance location.

Figure 3.18 compares the results of the fifth and sixth scenarios based on DO parameter prediction for Dalyan Entrance.

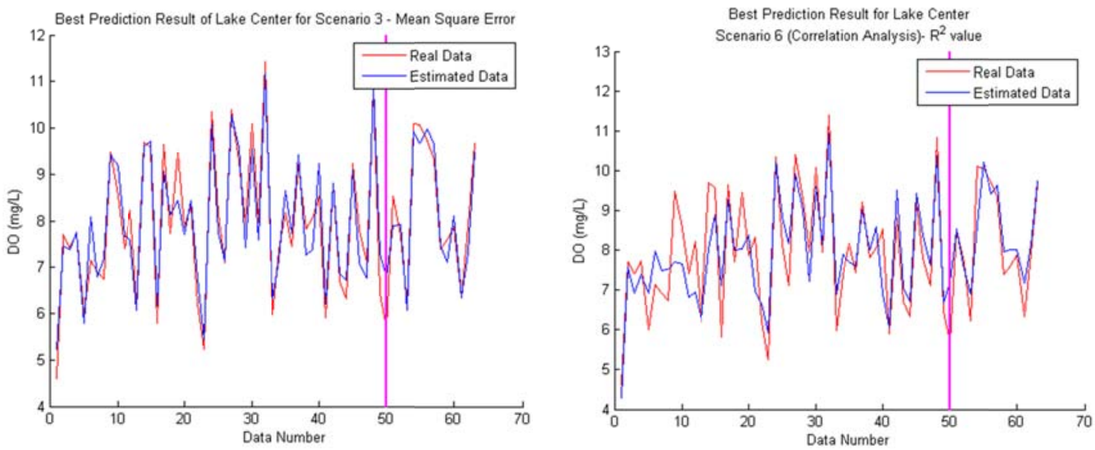


(a) Scenario 5 (b) Scenario 6

Figure 3.18: Comparison of Scenario 5 & 6 results for DO prediction of Dalyan Entrance

According to Table 3.7, performance value for MAE criteria is 0.6286 and 0.7507 for whole data and test data set, respectively in scenario 5. In scenario 6, performance value for MAE criteria is 0.7009 and 0.7239 for whole data and test data set, respectively. Therefore, scenario 5 is better than scenario 6 mathematically. According to Figure 3.18, the plot obtained by scenario 5 is better than scenario 6 visually. This means that, T, pH, EC, DO, NH_4-N and TP parameters of Yuvarlakçay Creek; pH, EC and TP parameters of Namnam Creek as input parameters give better prediction results for DO parameter of Dalyan Entrance.

According to Figure 3.17, generated ANN can catch the general trend at only a few points. Figure 3.19 compares the results of the third and sixth scenarios based on DO parameter prediction for Lake Center.

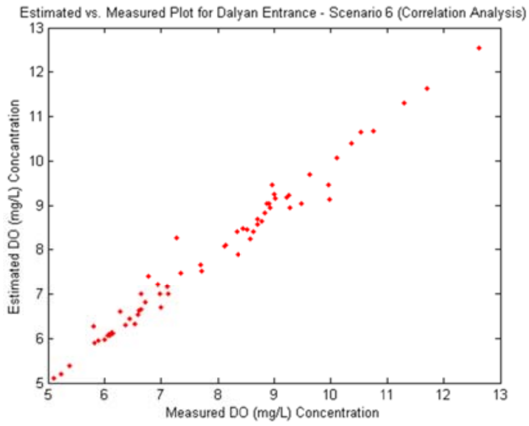


(a) Scenario 3 (b) Scenario 6

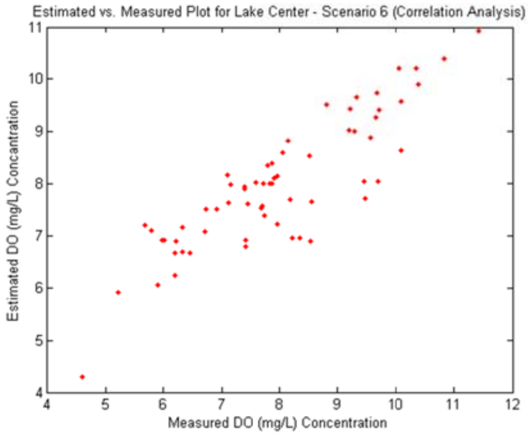
Figure 3.19: Comparison of Scenario 3 & 6 results for DO prediction of Lake Center

As can be seen from Figure 3.19, prediction of DO parameter for Lake Center is better with using input scenario 3.

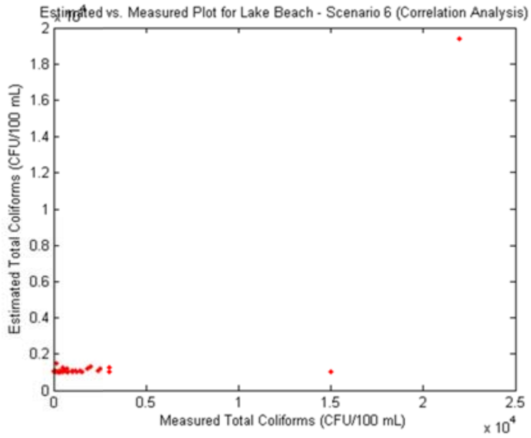
Finally, based on Figure 3.17, results obtained from correlation analysis are not sufficient for prediction of TC parameter of Lake Beach. According to Figure 3.14, the first scenario including only easily measurable parameters by a probe of the creeks gives excellent result for prediction of TC parameter for Lake Beach. Measured vs. predicted plots of the best prediction results for scenario 6 are given in Figure 3.20.



(a) DO (mg/L) of Dalyan Entrance



(b) DO (mg/L) of Lake Center



(c) TC (CFU/100 mL) of Lake Beach

Figure 3.20: Measured vs. estimated plots for obtained results from Scenario 6 for three lake locations

From the measured vs. estimated plots for scenario 6, performance of created ANNs can be verified. First graph shows a nearly linear line while; others can not catch the linearity.

Table 3.9 summarizes the required input parameters and performance criteria to obtain accurate prediction results for target output parameters of lake points.

Table 3.9: Summary of scenarios and performance criteria for accurate predictions

Predicted Parameter	Dalyan Entrance		
	Scenario No	**Input Parameters**	**Performance Criteria**
DO	5*	T, pH, EC, DO, NH_4-N, TP of Yuvarlakçay + pH, EC, TP of Namnam	MAE
NH_4-N	3	EMP + nutrients of the creeks	MAE
TP	5*	T, pH, EC, DO, NH_4-N, TP of Yuvarlakçay + pH, EC, TP of Namnam	MAE / R^2
Predicted Parameter	**Lake Center**		
	Scenario No	**Input Parameters**	**Performance Criteria**
DO	3	EMP + nutrients of the creeks	MSE
TN	5*	T, pH, EC, DO, NH_4-N, TP of Yuvarlakçay + pH, EC, TP of Namnam	R^2
TP	1	EMP	MAE
Chl-a	2	EMP + precipitation	MSE
Predicted Parameter	**Lake Beach**		
	Scenario No	**Input Parameters**	**Performance Criteria**
DO	3	EMP + nutrients of the creeks	MSE
TN	2	EMP + precipitation	MAE
TP	4	EMP + nutrients of the creeks + precipitation	AME
TC	1	EMP	R^2
Chl-a	1	EMP	MAE

EMP: Easily measurable parameters by a probe (T, pH, EC, DO)

According to the summary table, TP concentration of the Lake Center and TC amount and Chl-a concentration of Lake Beach can be predicted by using only T, pH, EC and DO parameters as input. Precipitation data is added to EMP as input for prediction of Chl-a concentration of Lake Center and TN concentration of Lake Beach. NH_4-N concentration of Dalyan Entrance, DO concentrations of Lake Center and Lake Beach can be predicted by addition of nutrients parameters to EMP as input. All parameters of the creeks should be used as input to predict TP concentration of Lake Beach. Finally, obtained input parameters from the correlation analysis should be used as input to predict DO and TP concentrations of Dalyan Entrance and TN concentration of Lake Center. By using the Table 3.9, accurate predictions can be obtained for water quality parameters of the lake sampling points by the decision makers.

CHAPTER 4

CONCLUSIONS AND DISCUSSION

With this thesis, a method was applied to Lake Köyceğiz to provide water quality assessment by using limited budget and data conditions.

In the first part of thesis, water quality assessment is conducted for Lake Köyceğiz and the two main creeks that flow to the lake. Firstly, available data was studied. EC of the lake sampling points are ten times higher than EC amount of the creeks due to the meromictic lake property of the lake. According to DO concentrations of the sampling points, higher concentrations than saturation value indicate the algal blooms in the region. Secondly, water quality parameters of the creeks are classified based on SWQMR. Classification results show that there should be agricultural pollution or wastewater discharge to the creeks. For Yuvarlakçay Creek, eight of twelve parameters are in class I, two of twelve parameters are in class II and remains are in class III. For Namnam Creek, five of eleven parameters are in class I while the rest are in class II. Yuvarlakçay is more contaminated than Namnam Creek based on NO_2-N and FC parameters. Finally, classification results show that Lake Köyceğiz in eutrophic state based on SWQMR and OECD criteria. The reasons of it can be transported pollution by the creeks to the Lake or existence of Köyceğiz waste water treatment plant at north side of the Lake or agricultural diffuse pollution to the Lake with surface runoff.

Hypothesis analysis was used to assess water quality in terms of spatial and temporal variations. It is observed that there is no spatial difference between TP, DO, TN, and Chl-a parameters in Lake Köyceğiz. However, there are seasonal variations for DO and Chl-a parameters. As a result of the Post Hoc Tests, Chl-a amounts are significantly different in fall months from spring and winter months while; DO concentrations obtained from samples in winter months are significantly difference from summer and fall months. Changes in DO concentrations should be resulted from temperature variations between seasons.

Correlation analysis demonstrated that the same water quality parameters of the Lake Köyceğiz and creeks of the lake are correlated with each other. This means that pollution carried with the creeks significantly impacts the water quality of the lake.

PCA was conducted to determine different water quality sets that can be used to explain the characteristics of water quality in Lake Köyceğiz. The components explain nearly 70% of total variance. According to the PCs, decision makers can use PC1 and PC2 to assess the anthropogenic pollution of the lake such as fertilizer use, sewage discharge, leakage from septic systems or industrial pollution; PC3 to specify waste water contribution to the lake and finally PC4 to understand impact of meromictic property on the lake.

At the second stage of this thesis, ANN approach was performed to predict water quality parameters of the lake using water quality parameters of two main creeks that feeding the lake as input. In this way, decisionmakers can predict easily lake parameters without sampling from lake points by using only parameters of Namnam and Yuvarlakçay Creeks with less time, lower budget and minor human resources. Obtained results indicate that ANN is powerful prediction tool for water quality due to the low error values and high correlation amounts between measured and predicted values. However, that should be careful about these points such as input parameter decision, hidden layer number selection, data division and performance criteria determination due to the subjectively selection of them. In general, ANN is good for prediction of peak and bottom values. In other words, ANN works better when variations are high between sampling dates however, errors are increasing for less variations.

Six scenarios were created and different input data sets were tried to find the best predcition result. Input scenario 5, which is created based on correlation analysis, is successful for DO and TP parameters prediction for Dalyan Entrance location and TN parameter prediction for Lake Center point. However, scenario 6 is not sufficient to predict DO parameter of Dalyan Entrance and Lake Center locations; TC parameter of Lake Beach point.

Acquired ANN results can be used by the decisionmakers based on two intended purposes. First of them is to get more accurate results by using the results obtained from the trials. To make precise predictions about lake sampling points, indicated input parameters and performance criteria given in Table 3.9 should be used. Second purpose is to get general ideas about the lake with less accurate results. In water quality assessment part it is concluded that there is no spatial difference between three points of the lake from the conducted hypothesis testing. By using this conclusion, the scenario required less input data of a location can be used to get a general idea about other points of the lake with using less parameter. This purpose will be less expensive than first one due to the less parameter using.

CHAPTER 5

RECOMMENDATIONS FOR FUTURE STUDIES

An important component of most water quality monitoring projects is flow rate measurement of surface water. The flow data is used for many purposes such as assessment of treatment needs, design of management measures, problem assessment, watershed project planning, targeting source areas and project evaluation (Meals, et al., 2008). Therefore, flow rate data should be measured regularly for the creeks feeding Lake Köyceğiz and ANN approach can be tried by using the flow rate parameter. In this way, better predictions can be obtained by less input parameter with flow rate data.

In the water quality parameters, real measured value should be used instead of 'less than detection limit value' term. By this way, ANN performance in learning and prediction processes can be increased.

After provision of required data, system can be made more complex with addition of other creeks and drainage channels that flow to Lake Köyceğiz. In this way, more factors affecting the water quality of the lake can be assessed.

Finally, in the future studies, same ANN procedure can be applied to another case study for comparison and the results and making generalizations about the obtained prediction results.

REFERENCES

Abdi, H., & Williams, L. (2010). Principal Component Analysis. *WIREs Comp Stat*, 2: 433–459.

Açıkalın, S. (2007). Estimation of Efficiency Urban Wastewater Treatment Plant by Artificial Neural Network. Sakarya.

Anderson, D., & McNeill, G. (1992). *Artificial Neural Networks Technology.* New York: Kaman Sciences Corporation.

Bayır, F. (2006). *An Application upon Artificial Neural Network Modeling and Forecasting*.

Beltratti, A., Margarita, S., & Terna, P. (1996). *Neural Networks for Economic and Financial Modelling.* London: International Thomson Computer Press.

Boyacıoğlu, H. (2006). Surface Water Quality Assessment using Factor Analysis. *Water SA*(32), 389-393.

Buhan, E. (1996). Development of Lagoon Management by the Investigation of Current Situation and Mullet Populations in Köyceğiz Lagoon System. İzmir.

Çelik, B. (2008). Time Series Analysis with Artificial Neural Networks Methodology: Theory and Application. İstanbul.

Dalkılıç, Y., & Harmancıoğlu, N. (2008). Application Opportunities of European Union Water Framework Directive in Turkey. *UCTEA 2. Water Policies Congress*, (pp. 415-424). Ankara.

Demuth, H., Beale, M., & Hagan , M. (R2015b). Neural Network Toolbox 6, User's Guide. The MathWorks.

Ekdal, A., Gürel, M., & Tanık, A. (2011). Application of WASP and SWAT Models for a Mediterranean Coastal Lagoon with Limited Seawater Exchange. *Journal of Coastal Research*(64), 1023-1027.

Elmas, Ç. (2003). *Artificial Networks (Theory, Architecture, Training, Application).* Ankara: Seçkin Publications.

Environmental Protection Agency for Special Areas. (2007). *Detection of Biological Diversity and Preperation of Management Plan Köyceğiz–Dalyan Specially Protected Environmental Area.* Final Report, Ministry of Environment and Forestry .

GDPNH. (2015). *Waste Water Management Services*. Retrieved April 13, 2015, from General Directorate for Preservation of Natural Heritage: http://www.csb.gov.tr/gm/tabiat/index.php?Sayfa=sayfahtml&Id=966

Google Earth 7.1.2.2041 (2013). *Lake Köyceğiz 36°53'04.66"N, 28°36'53.95"W, elevation 0m.* [Viewed 02 December 2015].

Gönenç, İ., Ertürk, A., Ekdal, A., Tümay, A., Tanık, A., Baykal, B., et al. (2002). *Modelling of Köyceğiz-Dalyan Lagoon and Watershe and Land Planing.* İstanbul.

Gümrah, F., Öz, B., Güler, B., & Evin, S. (2000). The Application of Artificial Neural Networks for the Prediction of Water Quality of Polluted Aquifer. *Water, Air, and Soil Pollution*(119), 275-294.

Gürel, M., Ertürk, A., Şeker, D. Z., Ekdal, A., Yüceil, K. B., Tanık, A., et al. (2002). *İGEM.* Retrieved April 2015, from İGEM Web Site: http://igemportal.org/

Hebb, D. O. (1949). *The Organization of Behaviour.* New York: Wiley.

Hicks, S. (2010). *Pollution.* Retrieved October 29, 2015, from Lake Scientist: http://www.lakescientist.com/lake-facts/water-quality/

Jordan, M. (1995). *Why the Logistic Function a Tutorial Discussion on Probabilities and Neural Networks.* MIT Computational Cognitive Science Report.

Kashid, S., & Kumar, S. (2012). *Applications of Artificial Neural Network to Sheet Metal Work - A Review.* Retrieved October 15, 2015, from Scientific and Academic Publishing: www.sapub.org

Kazancı, N., & Dügel, M. (2000). An Evaluation of the Water Quality of Yuvarlakçay Stream, in the Köycegiz-Dalyan Protected Area, SW Turkey. *Turkish Journal of Zoology*(24), 69-80.

Kazancı, N., İzbırak, A., Çağlar, S., & Gökçe, D. (1992). *Investigation of Köyceğiz-Dalyan Special Protected Area Aquatic Ecosystem Hydrobiologically.*

Hacettepe University Science Faculty Biology Department. Ankara: Özyurt Press.

Kazi, T., Arain, M., Jamali, M., Jalbani, N., Afridi, H., Sarfraz, R., et al. (2009). Assessment of Water Quality of Polluted Lake Using Multivariate Statistical Techniques: A Case Study. *Ecotoxicology and Environmental Safety*(72), 301-309.

Köyceğiz Municipality. (2015). Retrieved Agust 10, 2015, from http://www.koycegiz.bel.tr/web/index4.asp

Kuo, Y., Liu, C., & Lin, K. (2004). Evaluation of the Ability of an Artificial Neural Network Model to Asses the Variation of Groundwater Quality in an Area of Blackfoot Disease in Taiwan. *Water Research(38)*, 148-158. Taiwan.

Mack, J. (2011). *Eutrophication.* Retrieved October 29, 2015, from Lake Scientist: www.lakescientist.com/lake-facts/water-quality/

Meals, D. W., & Dressing, S. A. (2008). *Surface Water Flow Measurement for Water Quality Monitoring Projects.* Retrieved November 23, 2015, from Tech Notes 3: https://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/TechNotes/technote3_surface_flow.pdf

Ministry of Forestry and Water Affairs. (2012). Surface Water Quality Management Regulation. *(28483)*.

Najah, A., Elshafie, A., Karim, O., & Jaffar, O. (2009). Prediction of Johor River Water Quality Parameters Using Artificial Neural Networks. *Euro Journals Publishing(422-435)*. Malasia.

Navarro, H., & Bennun, L. (2014). Descriptive Examples of the Limitations of Artificial Neural Networks Applied to the Analysis of Independent Stochastic Data.

OECD. (1982). *Eutrophication of Waters. Monitoring, Assessment and Control.* Final Report, Paris.

OECD. (2008). *OECD Environmental Performance Reviews: Turkey 2008.* OECD Publishing.

Ouyang, Y., Nkedi-Kizza, P., Wu, Q., Shinde, D., & Huang, C. (2006). Assessment of Seasonal Variations in Surface Water Quality. *Water Research*(40), 3800-3810.

Özdemir, N., & De Sa Matos Paixao Çiçek, L. (2012). Water Quality Investigation of Namnam Creek that Feed Köyceğiz Lake. *Water Resources and Wetlands.* Tulcea (Romania).

Özdemir, N., Yılmaz, F., Barlas, M., & Yorulmaz, B. (2003). Fish Fauna and Ecological Propeties of Namnam Creek (Köyceğiz). *XII. Ulusal Su Ürünleri Sempozyumu.* Elazığ.

Öztemel, E. (2006). *Artificial Neural Networks.* İstanbul: Papatya Publishing.

Palani, S., Liong, S., & Tkalich, P. (2008). An ANN application for Water Quality Forecasting. *Marine Pollution Bulletin*, 56. Singapore.

Papoulis, A. (1991). *Probability, Random Variables, and Stochiastic Processes* (3 ed.). McGraw-Hill.

Pardo, R. (1992). *Design, Testing and Optimization of Trading Systems.* USA: John Wiley & Sons Inc.

Sahoo, G., Ray, C., & Wde, H. (2005). Pesticide Prediction in Groundwater in North Carolina Domestic Wells Using Artificial Neural Networks. *Ecological Modelling(183)*, 29-46. USA.

Saraç, T. (2004). *Artificial Neural Networks.* Seminar Project, Gazi University, Industrial Engineering Department, Ankara.

Sergiu, C. (2011). *Financial Predictor via Neural Network.* Retrieved October 16, 2015, from Code Project: http://www.codeproject.com/

Singh, K., Basant, A., Malik, A., & Jain, G. (2009). Artificial Neural Network Modeling of the River Water Quality-A Case Study. *Ecological Modelling(220)*, 888-895. India.

Steyl, G. (2009). *Application of Artificial Neural Networks in the Field of Geohydrology.* Master Thesis, University of the Free State.

Subaşı, H. (2010). *Modelling of Wastewater Treatment Performance using Artficial Neural Network.* Master Thesis, Çukurova University, Institute of Science, Adana.

Tekinay, A., Güroy, D., & Çevik, N. (2009). The Environmental Effect of a Land-Based Trout Farm on Yuvarlakçay, Turkey. *Ekoloji*(73), 65-70.

The Ministry of Forestry and Water Affairs. (2015). General Directorate of Water Management. *Projects for the years 2014 and 2015*. Ankara.

Türedi, M. (2006). *Lake Köyceğiz (Limnological Survey)*, 122. İstanbul: Marmara University.

Webb, S., Croft, T., Mustoe, L., & Ward, J. (2005). *Engineering Mathematics Open Learning Project Workbook.* Loughborough, UK.

Wei, B., Sugiura, N., & Maekawa, T. (2001). Use of Artificial Neural Network in the Prediction of Algal Blooms. *Water Resources(35)*, 2022-2028. Japan.

WHO, & EC. (2002). Eutrophication and Health Booklet. Luxemburg: Office for Official Publications of the European Communities.

Wilson, H., & Recknagel, F. (2001). Towards a Generic Artificial Neural Network Model for Dynamic Predcitions of Algal Abundance in Freshwater Lakes. *Ecological Modeling(146)*, 69-84. Australia.

Yurtoğlu, H. (2005). *Artificial Neural Network Modeling and Forecasting: The Case of Turkey for some macroeconomic variables.* Ankara: SPO Publication.

Zhang, G., Patuwo, B., & Hu, M. (1998). Forecasting with Artificial Neural Networks:The State of the Art. *International Journal of Forecasting*, 35–62.

Zhao, Y., Xia, X., Yang, Z., & Wang, F. (2012). Assessment of Water Quality in Baiyangdian Lake Using Multivariate Statistical Techniques. *Procedia Environmental Sciences*(8), 1213-1226.

APPENDIX A

DESCRIPTIVE STATISTICS

A.1. Descriptive statistics for the water quality data of Namnam and Yuvarlakçay Creeks

Parameter	Unit	Number of Data	Yuvarlakçay Creek					Namnam Creek				
			Min	Max	Mean	Median	St. Deviation	Min	Max	Mean	Median	St. Deviation
Temperature	°C	63	11.30	25.90	18.84	19.00	3.61	9.00	31.10	20.03	19.60	5.55
pH	-	63	6.70	9.01	8.25	8.28	0.40	6.62	9.63	8.44	8.54	0.53
Electrical Conductivity	µS/cm	63	228	5470	560	489	637	302	8970	790	541	1170
Disslved Oxygen	mg/L	63	5.32	12.88	8.51	8.11	1.63	4.52	13.99	8.20	8.27	1.87
Ammonium Nitrogen	mg/L	63	0.01	0.22	0.05	0.03	0.05	0.01	0.19	0.04	0.02	0.04
Total Phosphorus	mg/L	63	0.01	0.42	0.04	0.03	0.06	0.01	0.97	0.05	0.01	0.14

A.2. Descriptive statistics for the water quality data of Dalyan Entrance (south point of the lake)

Parameter	Unit	Number of Data	Dalyan Entrance				
			Min	Max	Mean	Median	St. Deviation
Temperature	°C	63	8.50	31.50	21.77	21.60	6.32
pH	-	63	7.24	9.21	8.50	8.47	0.37
Electrical Conductivity	µS/cm	63	537	1057	5459	5330	2041
Dissolved Oxygen	mg/L	63	5.10	12.63	7.94	8.11	1.71
Ammonium Nitrogen	mg/L	63	0.01	0.30	0.05	0.03	0.05
Total Phosphorus	mg/L	63	0.01	0.86	0.07	0.02	0.17

A.3. Descriptive statistics for the water quality data of Lake Center and Lake Beach (north point of the lake)

Parameter	Unit	Number of Data	Lake Center					Lake Beach				
			Min	Max	Mean	Median	St. Deviation	Min	Max	Mean	Median	St. Deviation
Temperature	°C	63	9.80	31.45	21.42	20.70	6.19	11.00	32.70	22.22	22.20	6.11
pH	-	63	7.12	9.21	8.52	8.54	0.37	7.00	9.16	8.48	8.50	0.39
Electrical Conductivity	µS/cm	63	768	9870	5080	5150	2062	-	-	-	-	-
Dissolved Oxygen	mg/L	63	4.60	11.42	7.97	7.87	1.50	4.10	15.52	7.84	7.75	1.69
Total Nitrogen	mg/L	63	0.10	3.67	0.72	0.58	0.66	0.10	3.32	0.69	0.56	0.60
Total Phosphorus	mg/L	63	0.01	0.67	0.05	0.01	0.11	0.01	0.15	0.03	0.02	0.03
Chlorophyll-a	µg/L	63	0.00	9.61	1.97	1.12	2.24	0.10	10.70	1.80	1.00	2.15
Total Coliforms	CFU/ 100 mL	63	-	-	-	-	-	0.00	22000	1196	300	3321

APPENDIX B

WATER QUALITY CLASSIFICATION

B.1. Calculated statistical values and SWQMR classes for Yuvarlakçay Creek

Parameters	Normal Distribution						Log-normal Distribution						Mean after removal of > %c95 value	SWQMR Class
	Mean	Median	St. Dev.	Variance	Skewness	X_{95} value	Mean	Median	St. Dev.	Variance	Skewness	X_{95} value		
Temperature (°C)	**19.21**	**19.15**	**3.66**	**13.42**	**0.06**	**22.26**	1.28	1.28	0.09	0.01	-0.29	25.26	18.87	I
pH	**8.26**	**8.28**	**0.36**	**0.13**	**-1.14**	**8.84**	0.92	0.92	0.02	0.00	-1.46	8.84	8.23	I
Electrical Conductivity (µS/cm)	560	489	637	405358	7.63	1611	**2.69**	**2.69**	**0.17**	**0.03**	**3.64**	**927**	481	II
Dissolved Oxygen (mg O_2/L)	8.49	8.01	1.58	2.50	0.34	11.10	**0.92**	**0.90**	**0.08**	**0.01**	**-0.07**	**11.36**	8.36	I
Oxygen Saturation (%)	95.21	92.65	18.91	357.51	0.92	126.41	**1.97**	**1.97**	**0.08**	**0.01**	**0.40**	**127.9**	91.21	I
Ammonium Nitrogen (mg NH4+-N/L)	**0.05**	**0.03**	**0.05**	**0.00**	**1.39**	**0.13**	-1.44	-1.50	0.39	0.15	0.17	0.16	0.045	I
Nitrite Nitrogen (mg NO2--N/L)	**0.02**	**0.02**	**0.04**	**0.00**	**6.12**	**0.08**	-1.99	-1.83	0.54	0.29	-0.27	0.08	0.016	III
Nitrate Nitrogen (mg NO3--N/L)	0.29	1.09	0.99	0.98	1.39	2.92	**-0.03**	**0.03**	**0.38**	**0.14**	**-0.38**	**3.91**	1.163	I
Total Phosphorus (mg P/L)	0.06	0.03	0.15	0.02	6.11	2.00	**-1.58**	**-1.60**	**0.42**	**0.18**	**1.59**	**0.13**	0.027	I
Chemical Oxygen Demand (COD) (mg/L)	10.53	8.00	8.42	70.97	2.80	24.43	**0.93**	**0.90**	**0.26**	**0.07**	**0.91**	**22.96**	9.88	I
Fecal Coliform (CFU/100 mL)	316	119	432	186604	2.24	1029	**2.25**	**2.33**	**0.61**	**0.37**	**-0.37**	**1794**	294	III
Total Coliform (CFU/100 mL)	1574	1100	2132	4544649	3.50	5092	**2.89**	**3.15**	**0.68**	**0.46**	**-0.80**	**10100**	1301	II

B.2. Calculated statistical values and SWQMR classes for Namnam Creek

Parameters	Normal Distribution						Log-normal Distribution						Mean after removal of >%c95 value	SWQMR Class
	Mean	Median	St. Dev.	Variance	Skewness	X_{95} value	Mean	Median	St. Dev.	Variance	Skewness	X_{95} value		
Temperature (°C)	**20.25**	**20.00**	**5.31**	**28.19**	**0.24**	**29.01**	1.29	1.30	0.12	0.01	-0.27	30.55	19.772	I
pH	**8.46**	**8.54**	**0.49**	**0.24**	**-1.28**	**9.26**	0.926	0.931	0.03	0.001	-1.28	9.32	8.427	I
Electrical Conductivity (µS/cm)	790	541	1170	1369369	5.96	2721	**2.78**	**2.73**	**0.25**	**0.06**	**2.83**	**1513**	532.03	II
Dissolved Oxygen (mg O_2/L)	8.06	7.91	1.78	3.15	0.44	10.99	**0.90**	**0.90**	**0.10**	**0.01**	**-0.32**	**11.40**	7.881	II
Oxygen Saturation (%)	90.97	90.10	15.67	245	0.29	117	**1.95**	**1.96**	**0.08**	**0.01**	**-0.47**	**120**	89.87	II
Ammonium Nitrogen (mg NH_4^+-N/L)	**0.04**	**0.02**	**0.04**	**0.00**	**2.00**	**0.106**	-1.59	-1.70	0.37	0.14	0.71	0.10	0.032	I
Nitrite Nitrogen (mg NO_2^--N/L)	**0.012**	**0.005**	**0.027**	**0.001**	**5.952**	**0.056**	-2.273	-2.301	0.513	0.263	0.520	0.037	0.008	II
Nitrate Nitrogen (mg NO_3^--N/L)	0.300	0.230	0.323	0.105	3.266	0.83	**-0.677**	**-0.638**	**0.371**	**0.137**	**-0.400**	**0.860**	0.231	I
Total Phosphorus (mg P/L)	**0.047**	**0.010**	**0.131**	**0.017**	**5.511**	**0.262**	-1.751	-2.000	0.437	0.191	2.231	0.093	0.021	I
Fecal Coliform (CFU/100 mL)	363	23.50	2427	5889211	8.97	4367	**1.740**	**1.653**	**0.658**	**0.434**	**1.151**	**670.62**	70.04	II
Total Coliform (CFU/100 mL)	1313	600	4404	19394267	8.51	8580	**2.648**	**2.874**	**0.721**	**0.520**	**-0.474**	**6893**	836.64	II

APPENDIX C

CORRELATION COEFFICIENTS

C.1. Detailed Table for Namnam and Yuvarlakçay Creeks

	T_YC	pH_YC	EC_YC	DO_YC	NH_4_YC	NO_2_YC	TSS_YC	TP_YC	FC_YC	TC_YC	T_NC	pH_NC	EC_NC	DO_NC	NH_4_NC	NO_2_NC	TSS_NC	TP_NC	FC_NC	TC_NC
T_YC	1 63	,216 ,089 63	,169 ,185 63	-,130 ,310 63	,048 ,708 63	-,030 ,814 63	-,238 ,061 63	-,061 ,635 63	,160 ,209 63	,160 ,209 63	,847 ,000 63	,076 ,552 63	,226 ,074 63	-,505 ,000 63	,018 ,891 63	,094 ,465 63	,151 ,236 63	,057 ,660 63	,013 ,917 63	,047 ,714 63
pH_YC	,216 ,089 63	1 63	-,102 ,428 63	,059 ,643 63	-,234 ,065 63	-,025 ,845 63	-,020 ,878 63	,181 ,155 63	-,143 ,262 63	,136 ,286 63	,186 ,144 63	,482 ,000 63	-,005 ,971 63	-,080 ,532 63	,024 ,851 63	-,098 ,447 63	,201 ,114 63	,198 ,120 63	,137 ,283 63	,132 ,302 63
EC_YC	,169 ,185 63	-,102 ,428 63	1 63	,086 ,500 63	-,051 ,689 63	,130 ,311 63	-,057 ,659 63	-,091 ,477 63	-,027 ,836 63	-,054 ,677 63	,250 ,048 63	,076 ,553 63	-,003 ,978 63	,049 ,702 63	-,113 ,378 63	,846 ,000 63	,034 ,789 63	-,038 ,767 63	-,024 ,853 63	-,009 ,943 63
DO_YC	-,130 ,310 63	,059 ,643 63	,086 ,500 63	1 63	-,160 ,211 63	,121 ,344 63	,120 ,350 63	-,168 ,189 63	-,330 ,008 63	-,114 ,374 63	-,104 ,418 63	,078 ,543 63	,161 ,206 63	,555 ,000 63	-,438 ,000 63	-,078 ,546 63	,052 ,684 63	-,160 ,209 63	-,069 ,592 63	-,128 ,316 63
NH_4_YC	,048 ,708 63	-,234 ,065 63	-,051 ,689 63	-,160 ,211 63	1 63	-,038 ,767 63	,001 ,996 63	,210 ,098 63	,198 ,119 63	,084 ,512 63	,042 ,743 63	,002 ,988 63	,251 ,047 63	-,059 ,646 63	,577 ,000 63	,018 ,891 63	-,187 ,143 63	,216 ,089 63	,113 ,379 63	,144 ,261 63
NO_2_YC	-,030 ,814 63	-,025 ,845 63	,130 ,311 63	,121 ,344 63	-,038 ,767 63	1 63	-,036 ,777 63	-,041 ,751 63	-,130 ,308 63	-,021 ,868 63	,120 ,348 63	,110 ,391 63	-,016 ,901 63	,142 ,265 63	-,108 ,400 63	,133 ,298 63	-,095 ,460 63	-,100 ,436 63	-,088 ,494 63	-,085 ,507 63
TSS_YC	-,238 ,061 63	-,020 ,878 63	-,057 ,659 63	,120 ,350 63	,001 ,996 63	-,036 ,777 63	1 63	,040 ,756 63	-,089 ,489 63	-,070 ,585 63	-,194 ,128 63	-,038 ,766 63	-,048 ,711 63	,150 ,241 63	-,087 ,496 63	-,034 ,792 63	-,144 ,259 63	,264 ,037 63	-,035 ,784 63	-,042 ,745 63
TP_YC	-,061 ,635 63	,181 ,155 63	-,091 ,477 63	-,168 ,189 63	,210 ,098 63	-,041 ,751 63	,040 ,756 63	1 63	-,054 ,676 63	,107 ,404 63	-,140 ,272 63	,016 ,899 63	-,083 ,516 63	,057 ,657 63	,278 ,027 63	-,080 ,531 63	,095 ,460 63	,365 ,003 63	-,071 ,582 63	-,026 ,840 63
FC_YC	,160 ,209 63	-,143 ,262 63	-,027 ,836 63	-,330 ,008 63	,198 ,119 63	-,130 ,308 63	-,089 ,489 63	-,054 ,676 63	1 63	,238 ,061 63	,052 ,688 63	-,348 ,005 63	-,037 ,771 63	-,084 ,514 63	,191 ,135 63	,015 ,909 63	-,095 ,457 63	-,087 ,497 63	,664 ,000 63	,690 ,000 63
TC_YC	,160 ,209 63	,136 ,286 63	-,054 ,677 63	-,114 ,374 63	,084 ,512 63	-,021 ,868 63	-,070 ,585 63	,107 ,404 63	,238 ,061 63	1 63	,047 ,712 63	-,108 ,399 63	-,076 ,555 63	-,100 ,436 63	,069 ,592 63	-,036 ,781 63	,247 ,051 63	,075 ,559 63	,190 ,135 63	,208 ,102 63
T_NC	,847 ,000 63	,186 ,144 63	,250 ,048 63	-,104 ,418 63	,042 ,743 63	,120 ,348 63	-,194 ,128 63	-,140 ,272 63	,052 ,688 63	,047 ,712 63	1 63	,210 ,099 63	,367 ,003 63	-,575 ,000 63	-,005 ,967 63	,179 ,160 63	,138 ,282 63	,029 ,820 63	-,088 ,491 63	-,063 ,622 63

	T_YC	pH_YC	EC_YC	DO_YC	NH_4_YC	NO_2_YC	TSS_YC	TP_YC	FC_YC	TC_YC	T_NC	pH_NC	EC_NC	DO_NC	NH_4_NC	NO_2_NC	TSS_NC	TP_NC	FC_NC	TC_NC
pH_NC	,076 ,552 63	,482 ,000 63	,076 ,553 63	,078 ,543 63	,002 ,988 63	,110 ,391 63	-,038 ,766 63	,016 ,899 63	-,348 ,005 63	-,108 ,399 63	,210 ,099 63	1 63	,002 ,986 63	-,098 ,443 63	,155 ,226 63	,084 ,515 63	,205 ,106 63	,228 ,073 63	-,097 ,448 63	-,128 ,316 63
EC_NC	,226 ,074 63	-,005 ,971 63	-,003 ,978 63	,161 ,206 63	,251 ,047 63	-,016 ,901 63	-,048 ,711 63	-,083 ,516 63	-,037 ,771 63	-,076 ,555 63	,367 ,003 63	,002 ,986 63	1 63	-,167 ,192 63	-,016 ,902 63	-,046 ,719 63	-,089 ,486 63	-,057 ,658 63	-,028 ,828 63	-,021 ,870 63
DO_NC	-,505 ,000 63	-,080 ,532 63	,049 ,702 63	,555 ,000 63	-,059 ,646 63	,142 ,265 63	,150 ,241 63	,057 ,657 63	-,084 ,514 63	-,100 ,436 63	-,575 ,000 63	-,098 ,443 63	-,167 ,192 63	1 63	-,278 ,028 63	-,024 ,850 63	-,345 ,006 63	-,128 ,318 63	,076 ,553 63	,047 ,714 63
NH_4_NC	,018 ,891 63	,024 ,851 63	-,113 ,378 63	-,438 ,000 63	,577 ,000 63	-,108 ,400 63	-,087 ,496 63	,278 ,027 63	,191 ,135 63	,069 ,592 63	-,005 ,967 63	,155 ,226 63	-,016 ,902 63	-,278 ,028 63	1 63	-,004 ,976 63	-,001 ,992 63	,233 ,066 63	,196 ,123 63	,224 ,077 63
NO_2_NC	,094 ,465 63	-,098 ,447 63	,846 ,000 63	-,078 ,546 63	,018 ,891 63	,133 ,298 63	-,034 ,792 63	-,080 ,531 63	,015 ,909 63	-,036 ,781 63	,179 ,160 63	,084 ,515 63	-,046 ,719 63	-,024 ,850 63	-,004 ,976 63	1 63	,052 ,683 63	-,009 ,947 63	,017 ,896 63	,051 ,694 63
TSS_NC	,151 ,236 63	,201 ,114 63	,034 ,789 63	,052 ,684 63	-,187 ,143 63	-,095 ,460 63	-,144 ,259 63	,095 ,460 63	-,095 ,457 63	,247 ,051 63	,138 ,282 63	,205 ,106 63	-,089 ,486 63	-,345 ,006 63	-,001 ,992 63	,052 ,683 63	1 63	,038 ,770 63	-,092 ,472 63	-,103 ,421 63
TP_NC	,057 ,660 63	,198 ,120 63	-,038 ,767 63	-,160 ,209 63	,216 ,089 63	-,100 ,436 63	,264 ,037 63	,365 ,003 63	-,087 ,497 63	,075 ,559 63	,029 ,820 63	,228 ,073 63	-,057 ,658 63	-,128 ,318 63	,233 ,066 63	-,009 ,947 63	,038 ,770 63	1 63	-,041 ,750 63	,009 ,942 63
FC_NC	,013 ,917 63	,137 ,283 63	-,024 ,853 63	-,069 ,592 63	,113 ,379 63	-,088 ,494 63	-,035 ,784 63	-,071 ,582 63	,664** ,000 63	,190 ,135 63	-,088 ,491 63	-,097 ,448 63	-,028 ,828 63	,076 ,553 63	,196 ,123 63	,017 ,896 63	-,092 ,472 63	-,041 ,750 63	1 63	,990** ,000 63
TC_NC	,047 ,714 63	,132 ,302 63	-,009 ,943 63	-,128 ,316 63	,144 ,261 63	-,085 ,507 63	-,042 ,745 63	-,026 ,840 63	,690** ,000 63	,208 ,102 63	-,063 ,622 63	-,128 ,316 63	-,021 ,870 63	,047 ,714 63	,224 ,077 63	,051 ,694 63	-,103 ,421 63	,009 ,942 63	,990** ,000 63	1 63
T_DE	,875** ,000 63	,198 ,120 63	,212 ,095 63	-,187 ,142 63	,060 ,640 63	,025 ,847 63	-,191 ,133 63	-,137 ,285 63	,173 ,174 63	,180 ,158 63	,933** ,000 63	,132 ,304 63	,261* ,039 63	-,627** ,000 63	,044 ,732 63	,132 ,304 63	,214 ,093 63	,098 ,446 63	-,023 ,857 63	,010 ,939 63
pH_DE	,220 ,084 63	,441** ,000 63	,011 ,933 63	,171 ,181 63	-,216 ,089 63	,098 ,444 63	-,128 ,318 63	-,166 ,195 63	-,247 ,051 63	,070 ,586 63	,315* ,012 63	,399** ,001 63	,034 ,792 63	-,240 ,059 63	-,066 ,605 63	,037 ,776 63	,280* ,026 63	-,240 ,059 63	,077 ,547 63	,048 ,709 63

	T_YC	pH_YC	EC_YC	DO_YC	NH_4_YC	NO_2_YC	TSS_YC	TP_YC	FC_YC	TC_YC	T_NC	pH_NC	EC_NC	DO_NC	NH_4_NC	NO_2_NC	TSS_NC	TP_NC	FC_NC	TC_NC
EC_DE	-,109 ,395 63	-,319 ,011 63	-,010 ,936 63	-,057 ,660 63	,125 ,328 63	,080 ,535 63	-,012 ,928 63	-,151 ,239 63	,002 ,990 63	-,235 ,064 63	-,175 ,170 63	-,443 ,000 63	,224 ,077 63	,121 ,346 63	,039 ,760 63	-,095 ,460 63	-,340 ,006 63	-,249 ,049 63	-,112 ,384 63	-,088 ,493 63
DO_DE	-,495 ,000 63	-,001 ,995 63	-,091 ,477 63	,616 ,000 63	-,102 ,428 63	,065 ,614 63	,199 ,119 63	-,084 ,513 63	-,130 ,310 63	-,051 ,690 63	-,460 ,000 63	,112 ,384 63	-,159 ,213 63	,757 ,000 63	-,265 ,036 63	-,124 ,335 63	-,205 ,107 63	-,182 ,153 63	,140 ,273 63	,078 ,544 63
NH_4_DE	-,046 ,720 63	,059 ,644 63	-,013 ,918 63	-,259 ,040 63	,486 ,000 63	-,028 ,826 63	-,062 ,627 63	,206 ,104 63	,092 ,475 63	,014 ,913 63	-,035 ,785 63	,242 ,056 63	,029 ,821 63	-,090 ,484 63	,397 ,001 63	,038 ,766 63	-,078 ,542 63	,227 ,073 63	,138 ,282 63	,158 ,215 63
NO_2_DE	-,197 ,122 63	-,147 ,249 63	-,036 ,782 63	-,055 ,669 63	,189 ,138 63	,190 ,136 63	-,075 ,558 63	,029 ,821 63	,219 ,085 63	,003 ,982 63	-,281 ,026 63	-,106 ,409 63	-,027 ,831 63	,213 ,094 63	,039 ,762 63	-,013 ,917 63	-,161 ,207 63	-,090 ,485 63	,215 ,090 63	,208 ,101 63
TSS_DE	,099 ,439 63	,030 ,817 63	,004 ,978 63	-,160 ,210 63	,010 ,937 63	-,028 ,825 63	-,157 ,221 63	,055 ,670 63	-,152 ,234 63	,151 ,236 63	,173 ,174 63	,175 ,169 63	,007 ,956 63	-,345 ,006 63	,132 ,301 63	-,001 ,994 63	,332 ,008 63	-,028 ,828 63	-,122 ,342 63	-,142 ,267 63
TP_DE	,134 ,294 63	,290 ,021 63	-,061 ,635 63	-,398 ,001 63	,300 ,017 63	-,056 ,663 63	-,030 ,816 63	,261 ,039 63	,017 ,898 63	,075 ,558 63	,204 ,108 63	,241 ,057 63	-,021 ,867 63	-,329 ,009 63	,376 ,002 63	,164 ,198 63	,115 ,368 63	,128 ,318 63	-,049 ,704 63	-,003 ,982 63
FC_DE	,079 ,538 63	,026 ,838 63	,206 ,105 63	-,134 ,295 63	,076 ,555 63	-,114 ,375 63	-,051 ,693 63	-,144 ,260 63	,605 ,000 63	,094 ,466 63	,036 ,780 63	-,103 ,423 63	,054 ,672 63	,032 ,806 63	,144 ,260 63	,305 ,015 63	-,065 ,612 63	-,013 ,922 63	,778 ,000 63	,781 ,000 63
TC_DE	,079 ,539 63	,153 ,230 63	,112 ,380 63	,123 ,336 63	-,038 ,768 63	,051 ,693 63	-,037 ,773 63	,007 ,958 63	-,040 ,757 63	,712 ,000 63	,021 ,871 63	,037 ,771 63	-,069 ,590 63	,046 ,718 63	-,077 ,548 63	,112 ,384 63	,204 ,109 63	-,027 ,833 63	,027 ,835 63	,023 ,860 63
T_LC	,874 ,000 63	,219 ,084 63	,218 ,087 63	-,184 ,149 63	,039 ,759 63	,029 ,823 63	-,257 ,042 63	-,101 ,433 63	,150 ,239 63	,164 ,200 63	,922 ,000 63	,114 ,373 63	,239 ,059 63	-,623 ,000 63	,051 ,692 63	,111 ,387 63	,227 ,074 63	,108 ,401 63	-,013 ,918 63	,017 ,892 63
pH_LC	,235 ,063 63	,520 ,000 63	,029 ,823 63	,127 ,322 63	-,297 ,018 63	,130 ,309 63	-,074 ,566 63	-,212 ,096 63	-,187 ,143 63	,194 ,127 63	,284 ,024 63	,439 ,000 63	-,125 ,328 63	-,201 ,115 63	-,100 ,434 63	-,043 ,738 63	,253 ,045 63	-,191 ,135 63	,104 ,417 63	,070 ,587 63
EC_LC	-,182 ,153 63	-,209 ,100 63	-,021 ,873 63	,002 ,989 63	-,015 ,906 63	,043 ,737 63	-,024 ,850 63	-,164 ,198 63	-,045 ,725 63	-,255 ,044 63	-,274 ,030 63	-,382 ,002 63	,224 ,078 63	,168 ,189 63	-,200 ,115 63	-,143 ,263 63	-,370 ,003 63	-,280 ,026 63	-,099 ,442 63	-,085 ,509 63

	T_YC	pH_YC	EC_YC	DO_YC	NH_4_YC	NO_2_YC	TSS_YC	TP_YC	FC_YC	TC_YC	T_NC	pH_NC	EC_NC	DO_NC	NH_4_NC	NO_2_NC	TSS_NC	TP_NC	FC_NC	TC_NC
DO_LC	-,343 ,006 63	-,150 ,241 63	-,014 ,914 63	,552 ,000 63	-,098 ,445 63	,095 ,460 63	,185 ,146 63	-,133 ,299 63	-,015 ,907 63	,043 ,739 63	-,350 ,005 63	-,119 ,353 63	-,132 ,301 63	,647 ,000 63	-,228 ,073 63	-,123 ,337 63	-,305 ,015 63	-,197 ,122 63	,173 ,174 63	,135 ,293 63
TN_LC	,003 ,982 63	-,125 ,331 63	,051 ,689 63	,100 ,437 63	,051 ,693 63	-,141 ,269 63	-,084 ,515 63	-,061 ,637 63	,193 ,129 63	,032 ,801 63	-,015 ,909 63	-,076 ,552 63	-,089 ,487 63	,044 ,732 63	-,006 ,964 63	-,005 ,970 63	,093 ,467 63	-,105 ,411 63	,021 ,873 63	,008 ,951 63
TSS_LC	,182 ,154 63	,136 ,287 63	-,007 ,954 63	-,092 ,475 63	-,077 ,548 63	-,144 ,262 63	,033 ,797 63	,107 ,403 63	,048 ,707 63	,051 ,693 63	,260 ,039 63	,204 ,109 63	-,059 ,648 63	-,393 ,001 63	,023 ,855 63	-,041 ,751 63	,406 ,001 63	,064 ,618 63	-,091 ,479 63	-,109 ,394 63
TP_LC	,173 ,175 63	,145 ,255 63	-,053 ,678 63	-,175 ,169 63	,330 ,008 63	-,091 ,479 63	-,042 ,746 63	-,002 ,987 63	,010 ,937 63	,059 ,644 63	,205 ,107 63	,270 ,033 63	-,068 ,598 63	-,270 ,032 63	,159 ,214 63	-,055 ,669 63	,110 ,389 63	,663 ,000 63	-,050 ,700 63	-,037 ,772 63
TC_LC	-,006 ,965 63	,203 ,110 63	-,078 ,541 63	-,222 ,081 63	,169 ,186 63	-,148 ,246 63	-,032 ,800 63	,010 ,940 63	,594 ,000 63	,200 ,116 63	-,063 ,623 63	,034 ,791 63	-,018 ,886 63	-,046 ,718 63	,295 ,019 63	-,030 ,818 63	-,034 ,790 63	,366 ,003 63	,834 ,000 63	,844 ,000 63
Chla_LC	-,195 ,126 63	-,132 ,301 63	-,076 ,554 63	-,053 ,681 63	,095 ,461 63	,029 ,819 63	-,044 ,733 63	,100 ,436 63	,011 ,932 63	-,184 ,148 63	-,303 ,016 63	-,127 ,323 63	-,108 ,401 63	,216 ,090 63	,137 ,283 63	-,122 ,340 63	-,197 ,123 63	-,111 ,388 63	-,098 ,447 63	-,106 ,407 63
T_LB	,879 ,000 63	,214 ,092 63	,239 ,059 63	-,211 ,097 63	,014 ,914 63	-,005 ,970 63	-,200 ,116 63	-,096 ,456 63	,139 ,277 63	,143 ,263 63	,911 ,000 63	,133 ,299 63	,187 ,141 63	-,639 ,000 63	,074 ,567 63	,131 ,305 63	,240 ,058 63	,114 ,373 63	-,031 ,810 63	,002 ,989 63
pH_LB	,189 ,138 63	,534 ,000 63	,144 ,260 63	,126 ,326 63	-,231 ,068 63	-,106 ,409 63	-,034 ,790 63	,070 ,587 63	-,247 ,051 63	,104 ,417 63	,123 ,338 63	,384 ,002 63	-,061 ,636 63	-,042 ,744 63	-,043 ,738 63	,103 ,421 63	,220 ,084 63	,024 ,855 63	,054 ,672 63	,027 ,836 63
DO_LB	-,233 ,067 63	-,114 ,375 63	,028 ,828 63	,483 ,000 63	-,124 ,334 63	,037 ,773 63	,174 ,173 63	-,047 ,712 63	-,070 ,583 63	-,015 ,910 63	-,222 ,081 63	-,129 ,315 63	-,028 ,825 63	,562 ,000 63	-,301 ,017 63	-,039 ,764 63	-,207 ,103 63	-,127 ,320 63	,158 ,215 63	,122 ,340 63
TN_LB	-,118 ,356 63	,015 ,904 63	,008 ,948 63	-,046 ,719 63	,001 ,994 63	-,168 ,189 63	-,086 ,504 63	-,148 ,248 63	,151 ,236 63	-,002 ,990 63	-,024 ,854 63	-,046 ,721 63	-,013 ,917 63	,013 ,921 63	,060 ,643 63	-,014 ,911 63	,024 ,852 63	-,113 ,380 63	,070 ,583 63	,067 ,599 63
TSS_LB	-,071 ,581 63	-,041 ,748 63	-,032 ,805 63	-,150 ,240 63	-,071 ,583 63	-,094 ,465 63	,313 ,012 63	,069 ,594 63	-,103 ,420 63	-,045 ,727 63	,020 ,876 63	,108 ,401 63	,003 ,980 63	-,272 ,031 63	,010 ,938 63	-,039 ,763 63	,253 ,045 63	,115 ,369 63	-,118 ,357 63	-,140 ,275 63

	T_YC	pH_YC	EC_YC	DO_YC	NH_4_YC	NO_2_YC	TSS_YC	TP_YC	FC_YC	TC_YC	T_NC	pH_NC	EC_NC	DO_NC	NH_4_NC	NO_2_NC	TSS_NC	TP_NC	FC_NC	TC_NC
TP_LB	-,064 ,617 63	-,176 ,167 63	-,092 ,471 63	-,230 ,070 63	,232 ,067 63	,000 ,997 63	-,077 ,550 63	,026 ,839 63	,056 ,664 63	,037 ,776 63	-,049 ,704 63	,042 ,742 63	-,093 ,470 63	-,146 ,253 63	,140 ,274 63	-,084 ,511 63	,060 ,641 63	,081 ,527 63	-,072 ,575 63	-,064 ,618 63
TC_LB	,114 ,374 63	,041 ,749 63	-,002 ,990 63	-,129 ,312 63	,308 ,014 63	-,076 ,553 63	-,052 ,684 63	-,075 ,557 63	,565 ,000 63	,149 ,245 63	,068 ,596 63	-,077 ,550 63	,011 ,935 63	-,033 ,797 63	,174 ,171 63	,004 ,973 63	-,052 ,683 63	,018 ,888 63	,804 ,000 63	,804 ,000 63
Chla_LB	-,327 ,009 63	-,201 ,114 63	-,070 ,588 63	-,029 ,821 63	,242 ,056 63	,103 ,423 63	-,097 ,449 63	,047 ,713 63	-,061 ,634 63	-,209 ,101 63	-,394 ,001 63	-,031 ,808 63	-,098 ,443 63	,248 ,050 63	,094 ,463 63	-,050 ,697 63	-,252 ,046 63	-,120 ,349 63	-,106 ,406 63	-,118 ,357 63
Precipitation_15 days	-,551 ,000 63	-,246 ,052 63	-,124 ,334 63	-,011 ,932 63	-,094 ,463 63	-,009 ,943 63	,082 ,522 63	-,040 ,756 63	-,043 ,740 63	-,009 ,944 63	-,517 ,000 63	-,122 ,341 63	-,177 ,166 63	,290 ,021 63	,128 ,316 63	-,114 ,375 63	-,046 ,722 63	-,154 ,229 63	,041 ,749 63	,025 ,848 63

C.2. Detailed Table for Dalyan Entrance Lake Center and Lake Beach

	T_DE	pH_DE	EC_DE	DO_DE	NH$_4$_DE	NO$_2$_DE	TSS_DE	TP_DE	FC_DE	TC_DE	T_LC	pH_LC	EC_LC	DO_LC
T_YC	,875 ,000 63	,220 ,084 63	-,109 ,395 63	-,495 ,000 63	-,046 ,720 63	-,197 ,122 63	,099 ,439 63	,134 ,294 63	,079 ,538 63	,079 ,539 63	,874 ,000 63	,235 ,063 63	-,182 ,153 63	-,343 ,006 63
pH_YC	,198 ,120 63	,441 ,000 63	-,319 ,011 63	-,001 ,995 63	,059 ,644 63	-,147 ,249 63	,030 ,817 63	,290 ,021 63	,026 ,838 63	,153 ,230 63	,219 ,084 63	,520 ,000 63	-,209 ,100 63	-,150 ,241 63
EC_YC	,212 ,095 63	,011 ,933 63	-,010 ,936 63	-,091 ,477 63	-,013 ,918 63	-,036 ,782 63	,004 ,978 63	-,061 ,635 63	,206 ,105 63	,112 ,380 63	,218 ,087 63	,029 ,823 63	-,021 ,873 63	-,014 ,914 63
DO_YC	-,187 ,142 63	,171 ,181 63	-,057 ,660 63	,616 ,000 63	-,259 ,040 63	-,055 ,669 63	-,160 ,210 63	-,398 ,001 63	-,134 ,295 63	,123 ,336 63	-,184 ,149 63	,127 ,322 63	,002 ,989 63	,552 ,000 63
NH$_4$_YC	,060 ,640 63	-,216 ,089 63	,125 ,328 63	-,102 ,428 63	,486 ,000 63	,189 ,138 63	,010 ,937 63	,300 ,017 63	,076 ,555 63	-,038 ,768 63	,039 ,759 63	-,297 ,018 63	-,015 ,906 63	-,098 ,445 63
NO$_2$_YC	,025 ,847 63	,098 ,444 63	,080 ,535 63	,065 ,614 63	-,028 ,826 63	,190 ,136 63	-,028 ,825 63	-,056 ,663 63	-,114 ,375 63	,051 ,693 63	,029 ,823 63	,130 ,309 63	,043 ,737 63	,095 ,460 63
TSS_YC	-,191 ,133 63	-,128 ,318 63	-,012 ,928 63	,199 ,119 63	-,062 ,627 63	-,075 ,558 63	-,157 ,221 63	-,030 ,816 63	-,051 ,693 63	-,037 ,773 63	-,257 ,042 63	-,074 ,566 63	-,024 ,850 63	,185 ,146 63
TP_YC	-,137 ,285 63	-,166 ,195 63	-,151 ,239 63	-,084 ,513 63	,206 ,104 63	,029 ,821 63	,055 ,670 63	,261 ,039 63	-,144 ,260 63	,007 ,958 63	-,101 ,433 63	-,212 ,096 63	-,164 ,198 63	-,133 ,299 63

	TN_LC	TSS_LC	TP_LC	TC_LC	Cha_LC	T_LB	pH_LB	DO_LB	TN_LB	TSS_LB	TP_LB	TC_LB	Cha_LB	Prec
T_YC	,003 ,982 63	,182 ,154 63	,173 ,175 63	-,006 ,965 63	-,195 ,126 63	,879 ,000 63	,189 ,138 63	-,233 ,067 63	-,118 ,356 63	-,071 ,581 63	-,064 ,617 63	,114 ,374 63	-,327 ,009 63	-,551 ,000 63
pH_YC	-,125 ,331 63	,136 ,287 63	,145 ,255 63	,203 ,110 63	-,132 ,301 63	,214 ,092 63	,534 ,000 63	-,114 ,375 63	,015 ,904 63	-,041 ,748 63	-,176 ,167 63	,041 ,749 63	-,201 ,114 63	-,246 ,052 63
EC_YC	,051 ,689 63	-,007 ,954 63	-,053 ,678 63	-,078 ,541 63	-,076 ,554 63	,239 ,059 63	,144 ,260 63	,028 ,828 63	,008 ,948 63	-,032 ,805 63	-,092 ,471 63	-,002 ,990 63	-,070 ,588 63	-,124 ,334 63
DO_YC	,100 ,437 63	-,092 ,475 63	-,175 ,169 63	-,222 ,081 63	-,053 ,681 63	-,211 ,097 63	,126 ,326 63	,483 ,000 63	-,046 ,719 63	-,150 ,240 63	-,230 ,070 63	-,129 ,312 63	-,029 ,821 63	-,011 ,932 63
NH$_4$_YC	,051 ,693 63	-,077 ,548 63	,330 ,008 63	,169 ,186 63	,095 ,461 63	,014 ,914 63	-,231 ,068 63	-,124 ,334 63	,001 ,994 63	-,071 ,583 63	,232 ,067 63	,308 ,014 63	,242 ,056 63	-,094 ,463 63
NO$_2$_YC	-,141 ,269 63	-,144 ,262 63	-,091 ,479 63	-,148 ,246 63	,029 ,819 63	-,005 ,970 63	-,106 ,409 63	,037 ,773 63	-,168 ,189 63	-,094 ,465 63	,000 ,997 63	-,076 ,553 63	,103 ,423 63	-,009 ,943 63
TSS_YC	-,084 ,515 63	,033 ,797 63	-,042 ,746 63	-,032 ,800 63	-,044 ,733 63	-,200 ,116 63	-,034 ,790 63	,174 ,173 63	-,086 ,504 63	,313 ,012 63	-,077 ,550 63	-,052 ,684 63	-,097 ,449 63	,082 ,522 63
TP_YC	-,061 ,637 63	,107 ,403 63	-,002 ,987 63	,010 ,940 63	,100 ,436 63	-,096 ,456 63	,070 ,587 63	-,047 ,712 63	-,148 ,248 63	,069 ,594 63	,026 ,839 63	-,075 ,557 63	,047 ,713 63	-,040 ,756 63

	T_DE	pH_DE	EC_DE	DO_DE	NH_4_DE	NO_2_DE	TSS_DE	TP_DE	FC_DE	TC_DE	T_LC	pH_LC	EC_LC	DO_LC	TN_LC	TSS_LC	TP_LC	TC_LC	Cha_LC	T_LB	pH_LB	DO_LB	TN_LB	TSS_LB	TP_LB	TC_LB	Cha_LB	Prec
FC_YC	,173 ,174 63	-,247 ,051 63	,002 ,990 63	-,130 ,310 63	,092 ,475 63	,219 ,085 63	-,152 ,234 63	,017 ,898 63	,605 ,000 63	-,040 ,757 63	,150 ,239 63	-,187 ,143 63	-,045 ,725 63	-,015 ,907 63	,193 ,129 63	,048 ,707 63	,010 ,937 63	,594 ,000 63	,011 ,932 63	,139 ,277 63	-,247 ,051 63	-,070 ,583 63	,151 ,236 63	-,103 ,420 63	,056 ,664 63	,565 ,000 63	-,061 ,634 63	-,043 ,740 63
TC_YC	,180 ,158 63	,070 ,586 63	-,235 ,064 63	-,051 ,690 63	,014 ,913 63	,003 ,982 63	,151 ,236 63	,075 ,558 63	,094 ,466 63	,712 ,000 63	,164 ,200 63	,194 ,127 63	-,255 ,044 63	,043 ,739 63	,032 ,801 63	,051 ,693 63	,059 ,644 63	,200 ,116 63	-,184 ,148 63	,143 ,263 63	,104 ,417 63	-,015 ,910 63	-,002 ,990 63	-,045 ,727 63	,037 ,776 63	,149 ,245 63	-,209 ,101 63	-,009 ,944 63
T_NC	,933 ,000 63	,315 ,012 63	-,175 ,170 63	-,460 ,000 63	-,035 ,785 63	-,281 ,026 63	,173 ,174 63	,204 ,108 63	,036 ,780 63	,021 ,871 63	,922 ,000 63	,284 ,024 63	-,274 ,030 63	-,350 ,005 63	-,015 ,909 63	,260 ,039 63	,205 ,107 63	-,063 ,623 63	-,303 ,016 63	,911 ,000 63	,123 ,338 63	-,222 ,081 63	-,024 ,854 63	,020 ,876 63	-,049 ,704 63	,068 ,596 63	-,394 ,001 63	-,517 ,000 63
pH_NC	,132 ,304 63	,399 ,001 63	-,443 ,000 63	,112 ,384 63	,242 ,056 63	-,106 ,409 63	,175 ,169 63	,241 ,057 63	-,103 ,423 63	,037 ,771 63	,114 ,373 63	,439 ,000 63	-,382 ,002 63	-,119 ,353 63	-,076 ,552 63	,204 ,109 63	,270 ,033 63	,034 ,791 63	-,127 ,323 63	,133 ,299 63	,384 ,002 63	-,129 ,315 63	-,046 ,721 63	,108 ,401 63	,042 ,742 63	-,077 ,550 63	-,031 ,808 63	-,122 ,341 63
EC_NC	,261 ,039 63	,034 ,792 63	,224 ,077 63	-,159 ,213 63	,029 ,821 63	-,027 ,831 63	,007 ,956 63	-,021 ,867 63	,054 ,672 63	-,069 ,590 63	,239 ,059 63	-,125 ,328 63	,224 ,078 63	-,132 ,301 63	-,089 ,487 63	-,059 ,648 63	-,068 ,598 63	-,018 ,886 63	-,108 ,401 63	,187 ,141 63	-,061 ,636 63	-,028 ,825 63	-,013 ,917 63	,003 ,980 63	-,093 ,470 63	,011 ,935 63	-,098 ,443 63	-,177 ,166 63
DO_NC	-,627 ,000 63	-,240 ,059 63	,121 ,346 63	,757 ,000 63	-,090 ,484 63	,213 ,094 63	-,345 ,006 63	-,329 ,009 63	,032 ,806 63	,046 ,718 63	-,623 ,000 63	-,201 ,115 63	,168 ,189 63	,647 ,000 63	,044 ,732 63	-,393 ,001 63	-,270 ,032 63	-,046 ,718 63	,216 ,090 63	-,639 ,000 63	-,042 ,744 63	,562 ,000 63	,013 ,921 63	-,272 ,031 63	-,146 ,253 63	-,033 ,797 63	,248 ,050 63	,290 ,021 63
NH_4_NC	,044 ,732 63	-,066 ,605 63	,039 ,760 63	-,265 ,036 63	,397 ,001 63	,039 ,762 63	,132 ,301 63	,376 ,002 63	,144 ,260 63	-,077 ,548 63	,051 ,692 63	-,100 ,434 63	-,200 ,115 63	-,228 ,073 63	-,006 ,964 63	,023 ,855 63	,159 ,214 63	,295 ,019 63	,137 ,283 63	,074 ,567 63	-,043 ,738 63	-,301 ,017 63	,060 ,643 63	,010 ,938 63	,140 ,274 63	,174 ,171 63	,094 ,463 63	,128 ,316 63
NO_2_NC	,132 ,304 63	,037 ,776 63	-,095 ,460 63	-,124 ,335 63	,038 ,766 63	-,013 ,917 63	-,001 ,994 63	,164 ,198 63	,305 ,015 63	,112 ,384 63	,111 ,387 63	-,043 ,738 63	-,143 ,263 63	-,123 ,337 63	-,005 ,970 63	-,041 ,751 63	-,055 ,669 63	-,030 ,818 63	-,122 ,340 63	,131 ,305 63	,103 ,421 63	-,039 ,764 63	-,014 ,911 63	-,039 ,763 63	-,084 ,511 63	,004 ,973 63	-,050 ,697 63	-,114 ,375 63
TSS_NC	,214 ,093 63	,280 ,026 63	-,340 ,006 63	-,205 ,107 63	-,078 ,542 63	-,161 ,207 63	,332 ,008 63	,115 ,368 63	-,065 ,612 63	,204 ,109 63	,227 ,074 63	,253 ,045 63	-,370 ,003 63	-,305 ,015 63	,093 ,467 63	,406 ,001 63	,110 ,389 63	-,034 ,790 63	-,197 ,123 63	,240 ,058 63	,220 ,084 63	-,207 ,103 63	,024 ,852 63	,253 ,045 63	,060 ,641 63	-,052 ,683 63	-,252 ,046 63	-,046 ,722 63

	T_DE	pH_DE	EC_DE	DO_DE	NH_4_DE	NO_2_DE	TSS_DE	TP_DE	FC_DE	TC_DE	T_LC	pH_LC	EC_LC	DO_LC
TP_NC	,098 ,446 63	-,240 ,059 63	-,249 ,049 63	-,182 ,153 63	,227 ,073 63	-,090 ,485 63	-,028 ,828 63	,128 ,318 63	-,013 ,922 63	-,027 ,833 63	,108 ,401 63	-,191 ,135 63	-,280 ,026 63	-,197 ,122 63
FC_NC	-,023 ,857 63	,077 ,547 63	-,112 ,384 63	,140 ,273 63	,138 ,282 63	,215 ,090 63	-,122 ,342 63	-,049 ,704 63	,778 ,000 63	,027 ,835 63	-,013 ,918 63	,104 ,417 63	-,099 ,442 63	,173 ,174 63
TC_NC	,010 ,939 63	,048 ,709 63	-,088 ,493 63	,078 ,544 63	,158 ,215 63	,208 ,101 63	-,142 ,267 63	-,003 ,982 63	,781 ,000 63	,023 ,860 63	,017 ,892 63	,070 ,587 63	-,085 ,509 63	,135 ,293 63
Prec_15_days	-,549 ,000 63	-,163 ,202 63	,316 ,012 63	,213 ,093 63	-,089 ,487 63	,016 ,902 63	-,117 ,360 63	-,146 ,252 63	-,048 ,711 63	,022 ,862 63	-,541 ,000 63	-,149 ,244 63	,121 ,345 63	,231 ,069 63

	TN_LC	TSS_LC	TP_LC	TC_LC	Cha_LC	T_LB	pH_LB	DO_LB	TN_LB	TSS_LB	TP_LB	TC_LB	Cha_LB	Prec
TP_NC	-,105 ,411 63	,064 ,618 63	,663 ,000 63	,366 ,003 63	-,111 ,388 63	,114 ,373 63	,024 ,855 63	-,127 ,320 63	-,113 ,380 63	,115 ,369 63	,081 ,527 63	,018 ,888 63	-,120 ,349 63	-,154 ,229 63
FC_NC	,021 ,873 63	-,091 ,479 63	-,050 ,700 63	,834 ,000 63	-,098 ,447 63	-,031 ,810 63	,054 ,672 63	,158 ,215 63	,070 ,583 63	-,118 ,357 63	-,072 ,575 63	,804 ,000 63	-,106 ,406 63	,041 ,749 63
TC_NC	,008 ,951 63	-,109 ,394 63	-,037 ,772 63	,844 ,000 63	-,106 ,407 63	,002 ,989 63	,027 ,836 63	,122 ,340 63	,067 ,599 63	-,140 ,275 63	-,064 ,618 63	,804 ,000 63	-,118 ,357 63	,025 ,848 63
Prec_15_days	,029 ,820 63	-,112 ,384 63	-,152 ,234 63	,035 ,784 63	,082 ,521 63	-,501 ,000 63	-,098 ,444 63	,000 ,998 63	,132 ,303 63	-,049 ,701 63	,183 ,152 63	-,013 ,916 63	-,066 ,610 63	1 63

APPENDIX D

RANDOM DIVISION SCRIPT

randomnpoints(A,B,N) function returns a vector of N random integers in the interval [A,B].

```
function  randomvector = randomnpoints(A, B, N)
    %select N random points uniformly distributed between A and B
    PointCount = 0;
    while PointCount < N
        flag = 0;
        %select one random point
        point = round(unifrnd(A, B, [1 1]));
        if (PointCount >= 1)
            if max(randomvector == point) > 0
                flag = 1;
            end
        end
        if flag == 0
            %insert point
            PointCount = PointCount + 1;
            randomvector(PointCount)= point;
        end
    end
end %function()
```

APPENDIX E

NORMALIZATION AND DENORMALIZATION SCRIPT

The scripts for normalizing a data matrix in the interval [0,1] and denormalizing data matrix back are provided below:

```
function [NormalizedData, A, B] = normalize_0_1(InputData)

    [NVariable, NDay] = size(InputData);

    [MIN,I] = min(InputData,[],2);
    [MAX,I] = max(InputData,[],2);

    a = MIN;
    b = MAX - MIN;

    A = repmat(a,1,NDay);
    B = repmat(b,1,NDay);

    NormalizedData = (InputData - A) ./ B;
end


function [DeNormalizedData] = unnormalize_0_1(NormalizedData, A, B)

 DeNormalizedData = NormalizedData.B + A;

end
```

APPENDIX F

ANN SCRIPT

The main function of the thesis is given below:

```
close all;
clear all;
clc;
tic;
tStart = tic;

[AllData_Numeric,AllData_TXT,AllData_RAW]=xlsread('input.xlsx');

%define input-output
InputIndices = [5 11] - 1;
TargetIndices = [23] - 1;

InputData = AllData_Numeric(:, InputIndices)';
TargetData = AllData_Numeric(:, TargetIndices)';

%Normalize input & Target
[InputNormalized, A_Input, B_Input] = normalize_0_1(InputData);
[TargetNormalized, A_Target, B_Target] = normalize_0_1(TargetData);

%define train test sets
[NVariables, NDays] = size(InputData);

load randomData;

TrainIndex = randomData(1:round(NDays0.6));
ValidationIndex = randomData(round(NDays0.6)+1:round(NDays0.8));
TestIndex = randomData(round(NDays0.8)+1:NDays);


TargetNormalizedTrain = TargetNormalized(:,TrainIndex);
TargetNormalizedValidation = TargetNormalized(:, ValidationIndex);
TargetNormalizedTest = TargetNormalized(:, TestIndex);

InputNormalizedTrain = InputNormalized(:,TrainIndex);
InputNormalizedValidation = InputNormalized(:,ValidationIndex);
InputNormalizedTest = InputNormalized(:,TestIndex);

%parameters
numberOfParemeters = 7;
transferFunction1 = {'logsig', 'tansig'};
```

```
transferFunction2 = {'logsig', 'tansig'};
trainingFunction = {'trainlm', 'trainbfg', 'trainrp', 'traingd',
'trainscg', 'traincgb', 'trainbr', 'traincgp', 'trainoss',
'traingdx', 'traingdm'};
learningFunction = {'learngdm'};
performanceFunction = {'mse'};
numberOfHiddenLayers = 2;
numberOfNeurons1 = 1:10:50;
numberOfNeurons2 = 1:10:50;
numberOfPerformanceParameters = 5; %

counter = 0;
Parameters = [];
Performances =
zeros(size(transferFunction1,2)size(trainingFunction,2)size(learning
Function,2)size(performanceFunction,2)...

size(numberOfNeurons1,2)size(numberOfNeurons2,2)size(transferFunctio
n2,2),size(TargetData,1),numberOfPerformanceParameters);

%loop over parameters to find best set
for transferFunctionIndex1 = 1:size(transferFunction1,2)
    for trainingFunctionIndex = 1:size(trainingFunction,2)
        for learningFunctionIndex = 1:size(learningFunction,2)
            for performanceFunctionIndex =
1:size(performanceFunction,2)
                 for numberOfNeuronsIndex1 =
1:size(numberOfNeurons1,2)  %number of neurons in hidden layer
                     for numberOfNeuronsIndex2 =
1:size(numberOfNeurons2,2)  %number of neurons in hidden l.
                         for transferFunctionIndex2 =
1:size(transferFunction2,2)

                        %update counter
                        counter = counter + 1

                        %get parameter
                        DummyParameters =
cell(numberOfParemeters,1);
                        DummyParameters{1} =
transferFunction1{1,transferFunctionIndex1};
                        DummyParameters{2} =
trainingFunction{1,trainingFunctionIndex};
                        DummyParameters{3} =
learningFunction{1,learningFunctionIndex};
                        DummyParameters{4} =
performanceFunction{1,performanceFunctionIndex};
                        DummyParameters{5}
=numberOfNeurons1(1,numberOfNeuronsIndex1);
                        DummyParameters{6} =
numberOfNeurons2(1,numberOfNeuronsIndex2);
                        DummyParameters{7} =
transferFunction2(1,transferFunctionIndex2);

                        Parameters = [Parameters DummyParameters];
```

```
                        %loop over target values
                        dummyPerformance = 0;
                        for targetIndex =  1: size(TargetData,1)

                            %create neural network
                            trf = cell(1,2);
                            trf{1} = DummyParameters{1};
                            trf{2} = DummyParameters{7};

                            net = newff(InputNormalized,
TargetNormalized(targetIndex,:), ...
                                [DummyParameters{5}
DummyParameters{6}],...
                                [trf{1} trf{2}], ...
                                DummyParameters{2}, ...
                                DummyParameters{3}, ...
                                DummyParameters{4}, {}, {} );

                            net.trainParam.showWindow = false;
                            net.trainParam.showCommandLine = false;
                            net.trainParam.epochs = 10000;
                            net.trainParam.goal = 0.00;

                            %train neural network by default toolbox
                            %divides input into 3 as training,
validation
                            %and test data. So all input and target
data
                            %can be feeded.
                            %define validation and test structures
for train function
                            validation.P =
InputNormalizedValidation;
                            validation.T =
TargetNormalizedValidation(targetIndex,:);

                            test.P = InputNormalizedTest;
                            test.T =
TargetNormalizedTest(targetIndex,:);

                            net = train(net, ...

InputNormalized(:,TrainIndex),TargetNormalized(targetIndex,TrainInde
x),[],[],validation,test);

                            %simulate neural network
                            TargetEstimatedNormalized =
sim(net,InputNormalized(:,TestIndex));

                            %Performance 1: Absolute Max Error
                            Performances(counter, targetIndex, 1) =
...

max(abs(TargetNormalized(targetIndex,TestIndex) -
TargetEstimatedNormalized));
```

```
                            %Performance 2: Mean of Absolute Error
                            Performances(counter, targetIndex, 2) =
...
mean(abs(TargetNormalized(targetIndex,TestIndex) -
TargetEstimatedNormalized));

                            %Performance 3:  R2 value
                            [M, B, R] = ...
                                 postreg( TargetEstimatedNormalized,
TargetNormalized(targetIndex,TestIndex), 'hide');
                             %negated so that minimum value is aimed
                             %instead of max for R2
                             Performances(counter, targetIndex, 3) =
-R^2;

                            %Performance 4: Normalized Mean Absoulte
Error
                            Performances(counter, targetIndex, 4) =
...
                                mean(abs( (
TargetEstimatedNormalized - TargetNormalized(targetIndex,TestIndex)
) ./ ...

TargetNormalized(targetIndex,TestIndex)  100 ));

                            %Performance 5: Mean Square Error
                             Performances(counter, targetIndex, 5) =
...
                                mean(
(TargetNormalized(targetIndex,TestIndex) -
TargetEstimatedNormalized).^2 );

                            %denormalize target
                            %TargetEstimated =
unnormalize_0_1(TargetEstimatedNormalized,
A_Target(targetIndex,TestIndex), B_Target(targetIndex,TestIndex));
                            %Perf = perform(net,
TargetData(targetIndex,TestIndex), TargetEstimated);

                            %Perf = perform(net,
TargetNormalized(targetIndex,TestIndex), TargetEstimatedNormalized);
                            %Perf = mean(
(TargetNormalized(targetIndex,TestIndex) -
TargetEstimatedNormalized).^2 );

                            %dummyPerformance = dummyPerformance +
Perf;

                        end


                    end
                   %pause;
```

```
                    end
                end
            end
        end
    end
end
%get best parameter
for targetIndex =  1: size(TargetData,1)

    %get performances of target
    targetPerformance = reshape(Performances(:,targetIndex,:),
size(Performances,1),numberOfPerformanceParameters );

    %get minimum values for errors
    [Y,I] = min(targetPerformance);

    %get best parameter
    BestParameter{targetIndex} = Parameters(:,I);
end
 plotCounter = 0;
%perform estimation with best parameters
for targetIndex =  1: size(TargetData,1)

    parameter = BestParameter{targetIndex};

    %loop over performance parameters
    for perfIndex = 1:numberOfPerformanceParameters

        perParameter = parameter(:,perfIndex);

        trf = cell(1,2);
        trf{1} = perParameter{1};
        trf{2} = perParameter{7};

    net = newff(InputNormalized, TargetNormalized(targetIndex,:),
...
        [perParameter{5} perParameter{6}],...
        [trf{1} trf{2}], ...
        perParameter{2}, ...
        perParameter{3}, ...
        perParameter{4}, {}, {} );
    net.trainParam.showWindow = false;
    net.trainParam.showCommandLine = false;
    net.trainParam.epochs = 10000;
    net.trainParam.goal = 0.00;
    %train neural network by default toolbox
    %divides input into 3 as training, validation
    %and test data. So all input and target data
    %can be feeded.
    validation.P = InputNormalizedValidation;
    validation.T = TargetNormalizedValidation(targetIndex,:);

    test.P = InputNormalizedTest;
    test.T = TargetNormalizedTest(targetIndex,:);
    net = train(net, ...
```

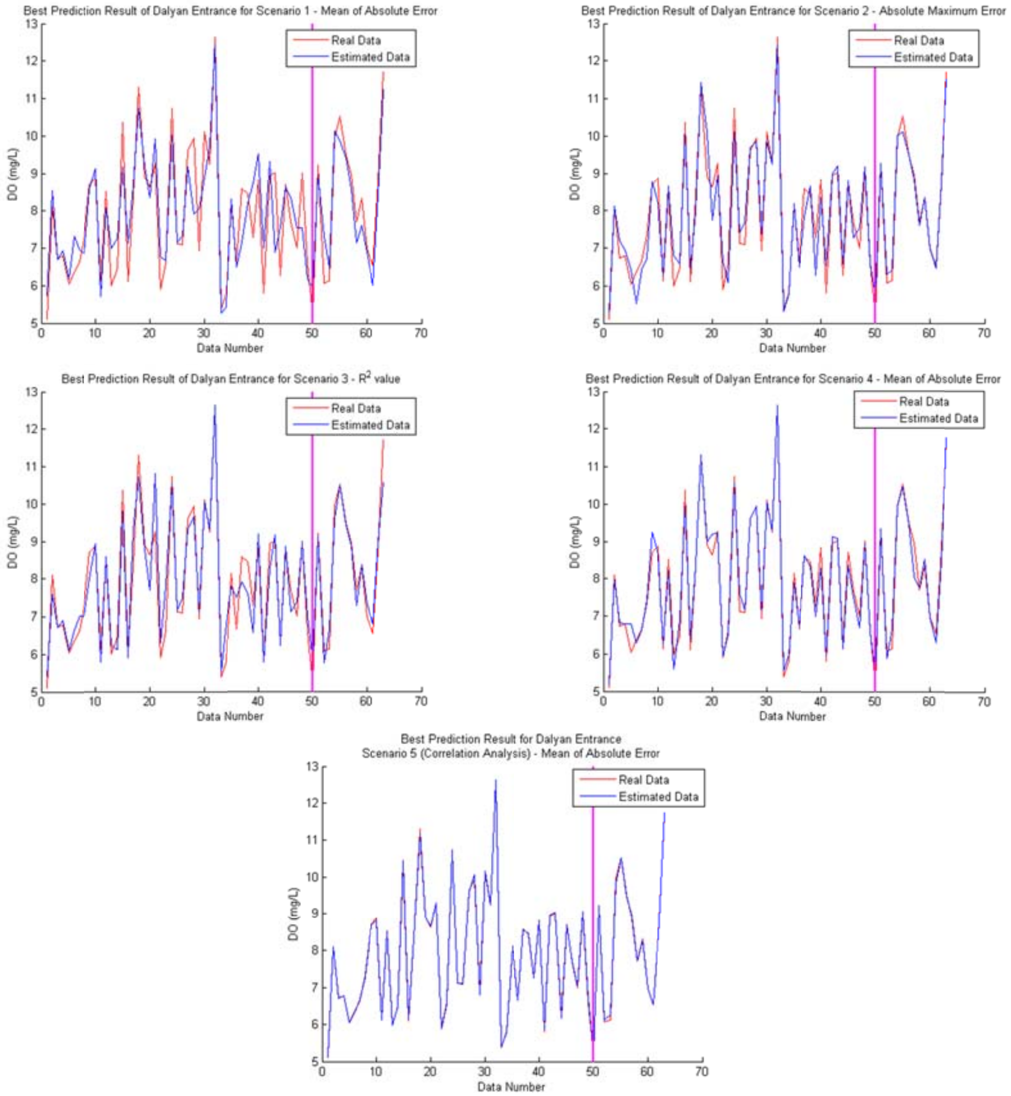
```
InputNormalized(:,TrainIndex),TargetNormalized(targetIndex,TrainInde
x), [],[],validation,test);
    %simulate neural network
    %TargetEstimatedNormalized =
sim(net,InputNormalized(:,TestIndex));
    TargetEstimatedNormalized = sim(net,InputNormalized(:,:));

    %denormalize target
    TargetEstimated = unnormalize_0_1(TargetEstimatedNormalized,
A_Target(targetIndex,:), B_Target(targetIndex,:));

    %plot
    figure; hold on;
    plot(TargetData(targetIndex,:), 'r');
    plot(TargetEstimated, 'b');
    title(['Target Index = ' num2str(targetIndex) ' Performance
Parameter = ' num2str(perfIndex)]);
    plotCounter = plotCounter + 1;
    saveas(gcf, ['Result_5B_Tarih_' num2str(plotCounter)], 'fig');

    figure; plot(TargetData(targetIndex,:), TargetEstimated(1,:),
'r.');
    xlabel('Target Data');
    ylabel('Target Estimated');
    title(['Target Index = ' num2str(targetIndex) ' Performance
Parameter = ' num2str(perfIndex)]);
    plotCounter = plotCounter + 1;
    saveas(gcf, ['Result_5B_Tarih' num2str(plotCounter)], 'fig');

    %random plot
    figure; hold on;
    plot(TargetData(targetIndex,randomData), 'r');
    plot(TargetEstimated(1,randomData), 'b');
    title(['Target Index = ' num2str(targetIndex) ' Performance
Parameter = ' num2str(perfIndex)]);
    plotCounter = plotCounter + 1;
    saveas(gcf, ['Result_5B_Random_' num2str(plotCounter)], 'fig');

    figure;
    plot(TargetData(targetIndex,randomData),
TargetEstimated(1,randomData), 'r.');
    xlabel('Target Data');
    ylabel('Target Estimated');
    title(['Target Index = ' num2str(targetIndex) ' Performance
Parameter = ' num2str(perfIndex)]);
    plotCounter = plotCounter + 1;
    saveas(gcf, ['Result_5B_Random_' num2str(plotCounter)], 'fig');

    end
end

timeElapsed = toc(tStart);

save('result_senaryo_5_B.mat');
```

APPENDIX G

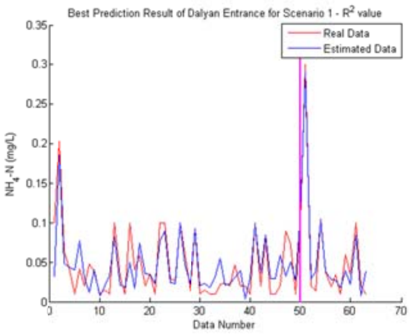
ANN PLOTS FOR SCENARIOS

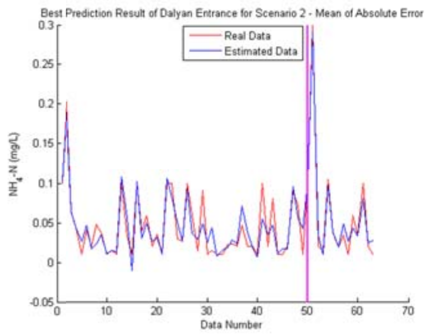
G.1. Selected Best Performance Criteria Plots of five scenarios for each output of Dalyan Entrance, Lake Center and Lake Beach

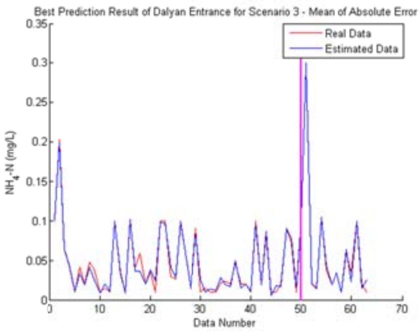
Dalyan Entrance (south point of the lake) - Dissolved Oxygen (DO)

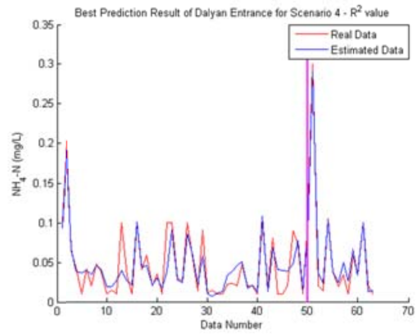


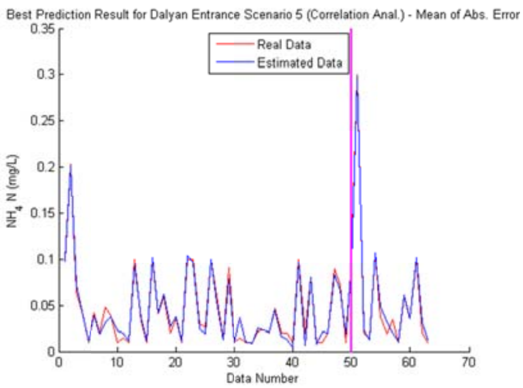
Dalyan Entrance (south point of the lake) – Ammonium Nitrogen (NH_4-N)



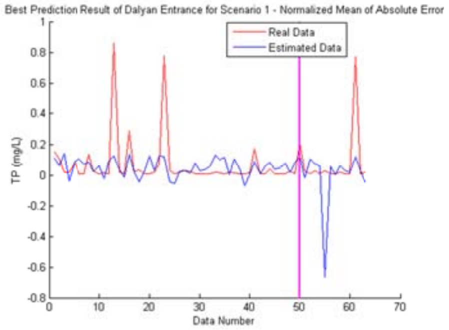


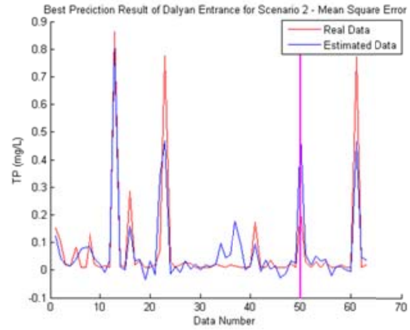


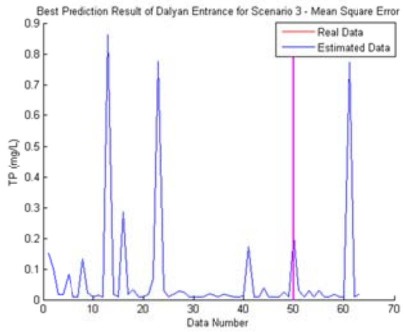


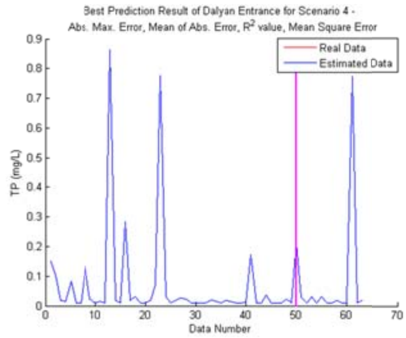


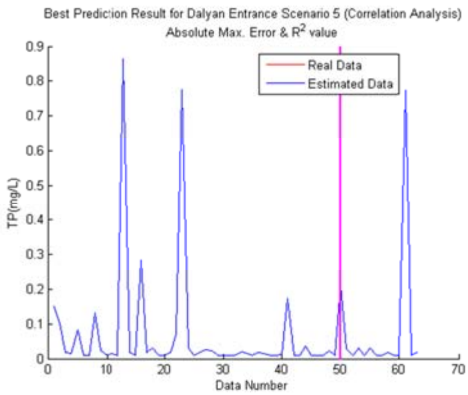
Dalyan Entrance (south point of the lake) – Total Phosphorus (TP)



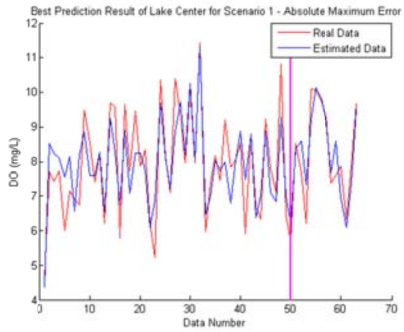


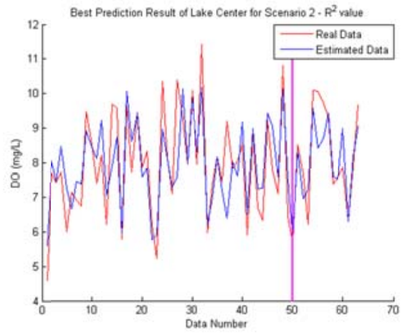


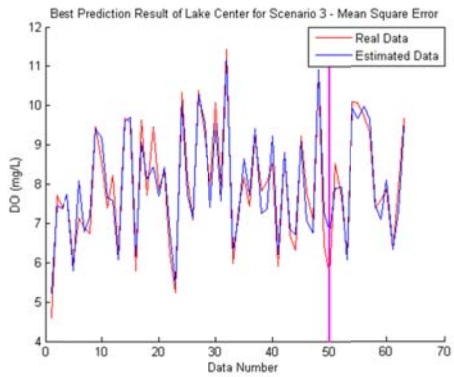


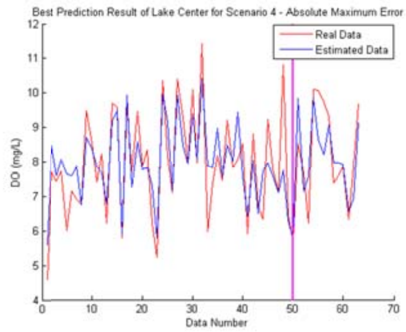


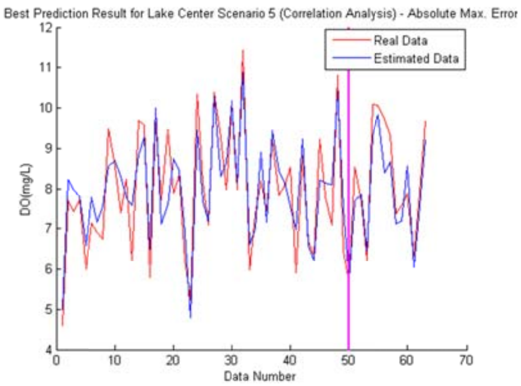
Lake Center – Dissolved Oxygen (DO)



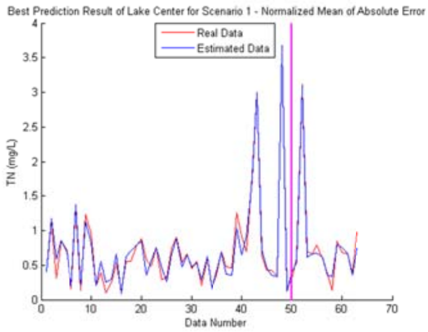


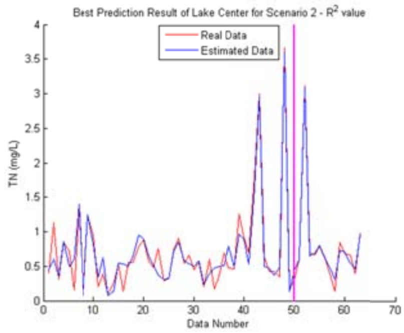


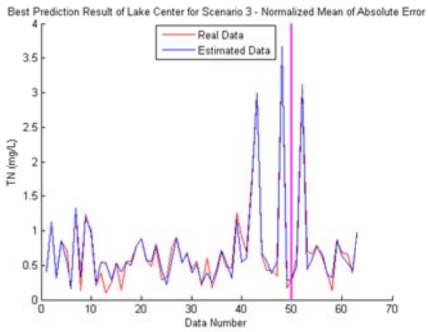


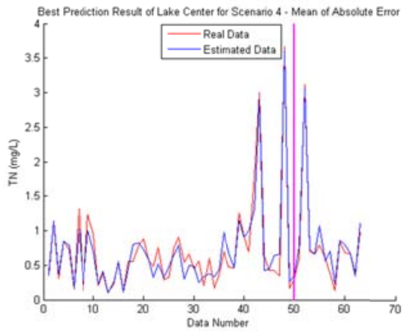


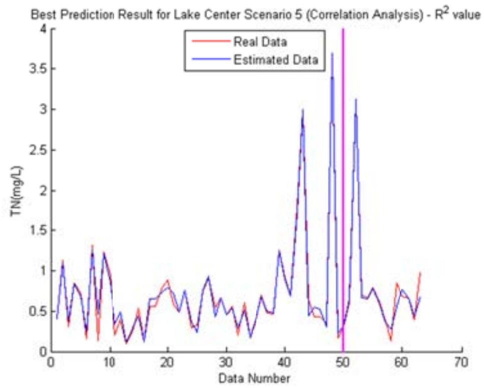
Lake Center – Total Nitrogen (TN)



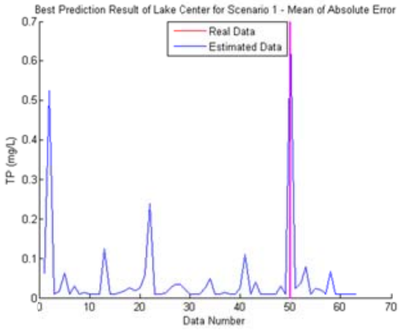


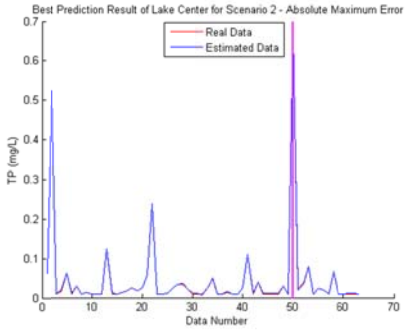


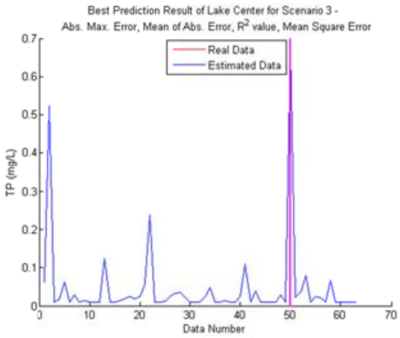


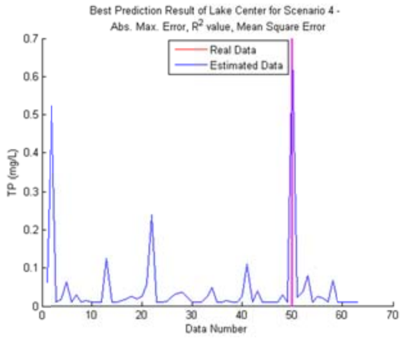


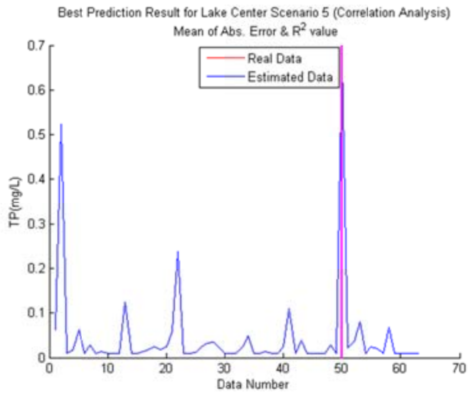
Lake Center – Total Phosphorus (TP)



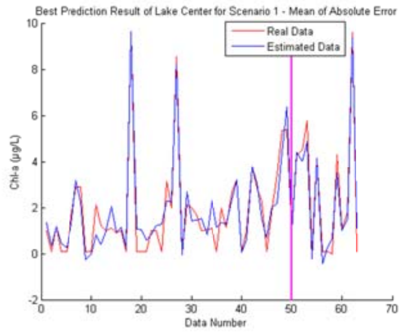


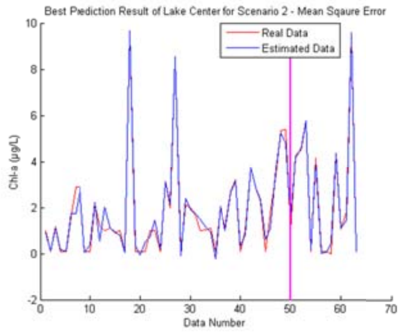


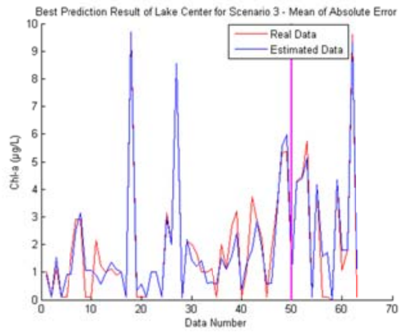


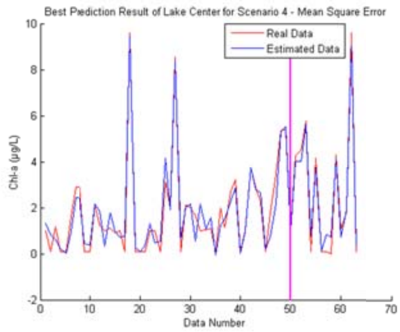


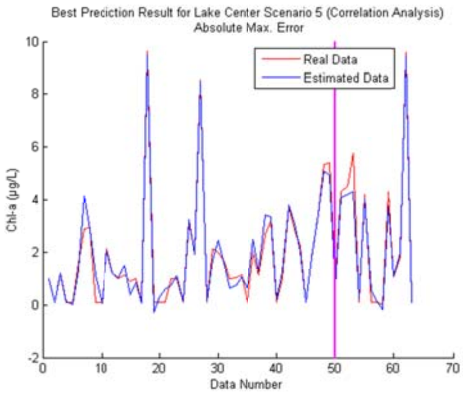
Lake Center – Chlorophyll-a (Chl-a)



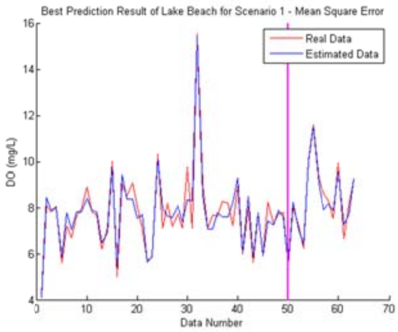


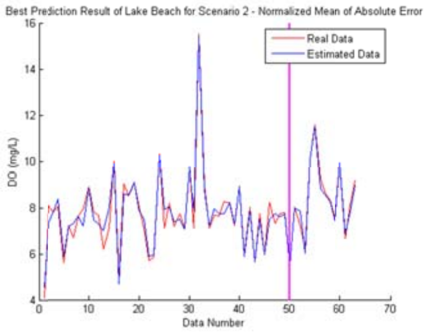


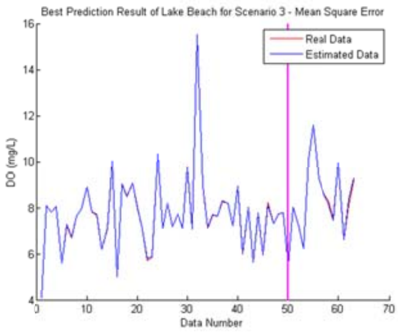


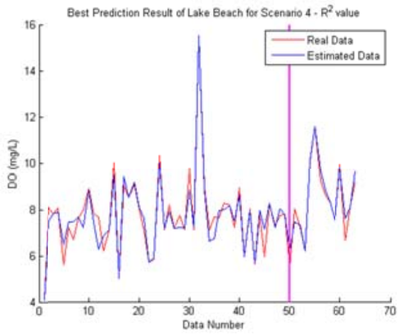


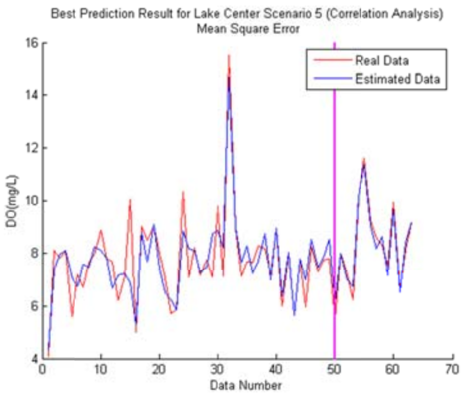
Lake Beach – Dissolved Oxygen (DO)



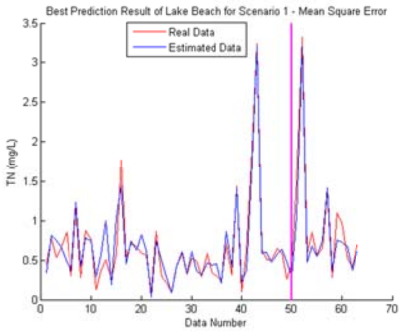


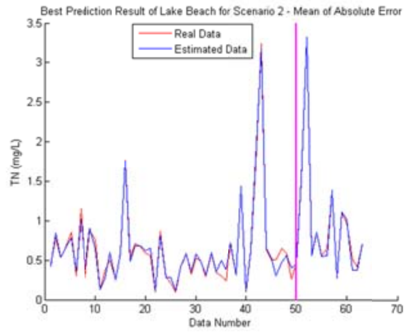


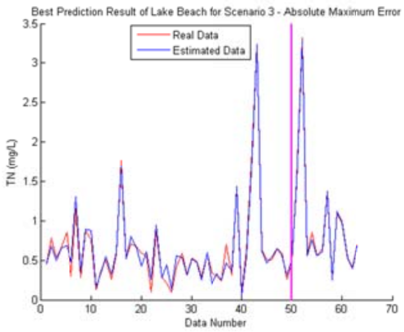


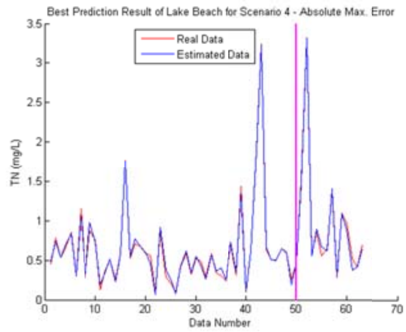


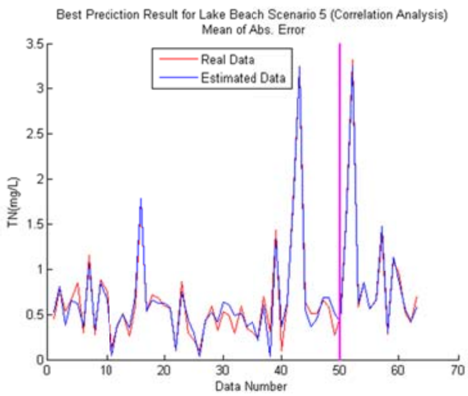
Lake Beach – Total Nitrogen (TN)



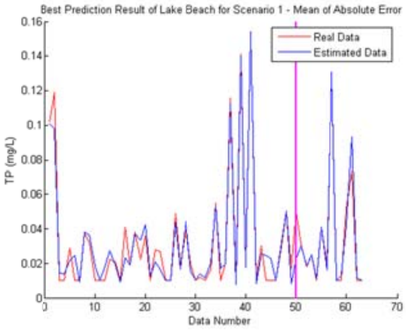


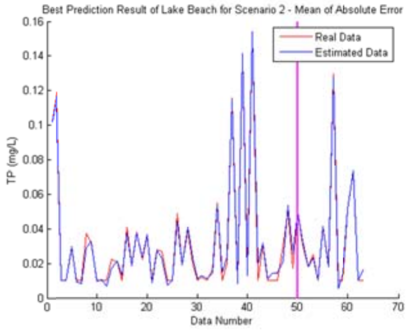


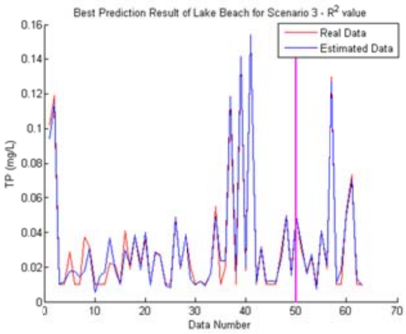


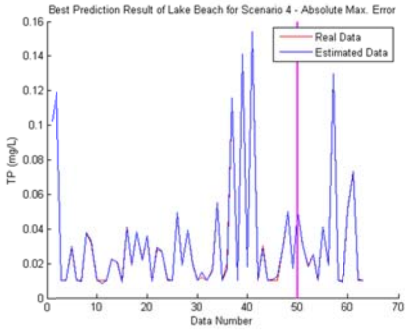


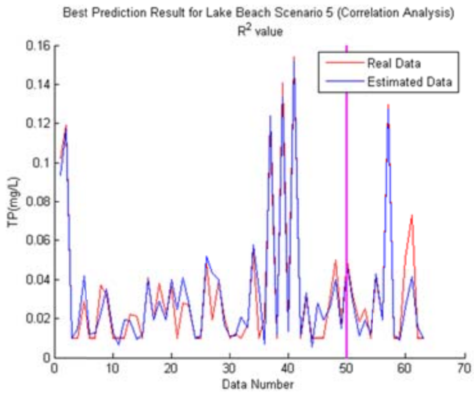
Lake Beach – Total Phosphorus (TP)



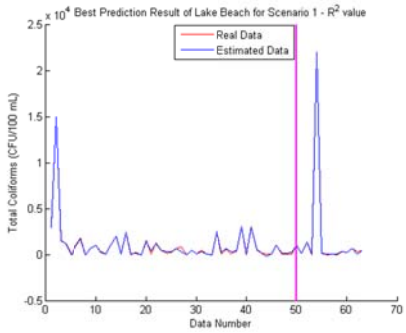


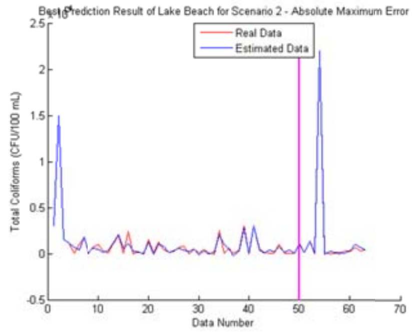


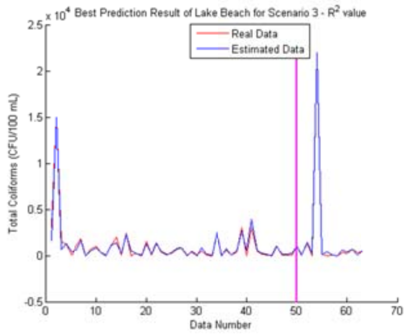


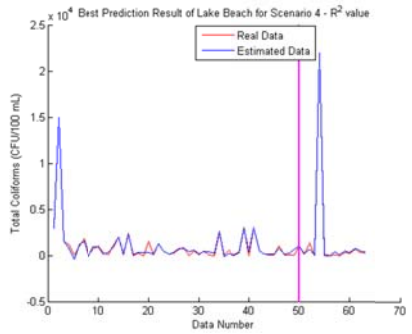


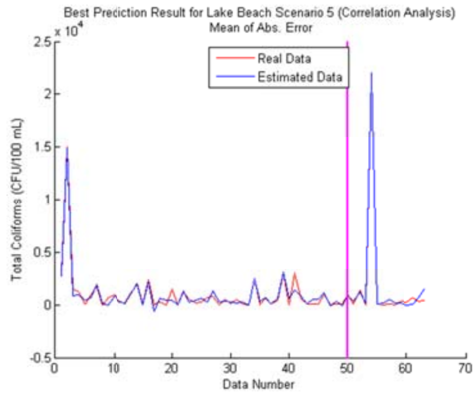
Lake Beach – Total Coliforms (TC)











Lake Beach – Chlorophyll-a (Chl-a)