WASTEWATER TREATMENT PLANT PERFORMANCE ANALYSIS USING ARTIFICIAL INTELLIGENCE – AN ENSEMBLE APPROACH

**A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES
OF
NEAR EAST UNIVERSITY**

**By
SANI ISAH ABBA**

**In Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Science
in
Civil Engineering**

NICOSIA, 2019

SANI ISAH ABBA

WASTEWATER TREATMENT PLANT PERFORMANCE ANALYSIS USING ARTIFICIAL INTELLIGENCE – AN ENSEMBLE APPROACH

NEU
2019

WASTEWATER TREATMENT PLANT PERFORMANCE ANALYSIS USING ARTIFICIAL INTELLIGENCE – AN ENSEMBLE APPROACH

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By
SANI ISAH ABBA

In Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Science in

Civil Engineering

NICOSIA, 2019

Sani Isah ABBA: WASTEWATER TREATMENT PLANT PERFORMANCE ANALYSIS USING ARTIFICIAL INTELLIGENCE – AN ENSEMBLE APPROACH

Approval of Director of Graduate School of Applied Sciences

Prof. Dr. Nadire CAVUS

We certify this thesis is satisfactory for the award of the degree of Doctor of Philosophy in Science in Civil Engineering

Examining Committee in Charge:

Prof. Dr. Hüseyin Gökçekuş	Committee Chairman, Dean of Civil and Environmental Engineering, NEU
Assoc. Prof. Dr. Gozen Elkiran	Supervisor, Department of Civil Engineering, NEU
Prof. Dr. Vahid Nourani	Co-Supervisor, University of Tabriz, Iran
Prof. Dr. Derin Orhon	NEU- Gene. Mish. Dept. of Civil Engineering, NEU
Prof. Bülent Bilgehan	NEU- For E.E, Dept. of Elec. and Electronic, Eng.
Prof. Dr. Umut Türker	EMU- Civil Eng. Dept. Faculty of Engineering
Assoc. Prof. Dr. Mustafa Ergil	EMU- Civil Eng. Dept. Faculty of Engineering

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 to this work.

Name, Last Name: Sani Isah Abba

Signature:

Date:

Dedicated to My Parents…

AKNOWLEDGEMENTS

First and foremost, Glory, and thanks to Allah, the Almighty, for His showers of blessings throughout my studies and to complete my Ph.D. thesis work successfully. I would like to thank my supervisors, Assoc. Prof. Dr. Gozen Elkiran and Prof. Vahid Nourani, for their patience, motivation, enthusiasm, and immense contribution, toward achieving this success.

Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Dr. Hüseyin Gökçekuş (Committee Chairman and Dean of Civil and Environmental Engineering, NEU), Prof. Dr. Vahid Nourani (Co-Supervisor, University of Tabriz, Iran) and Prof. Bülent Bilgehan (Committee Member, Department of Electrical and Electronic, NEU), for their encouragement and knowledgeable comments.

Above all, my endless thanks and heartfelt love would be dedicated to my family, friends, and colleagues for their loyalty and their great confidence in me. I'm greatly indebted to my Late. Father and indeed my lovely Mother, who is indeed my inspiration, who led me to the treasures of knowledge. I really appreciate them for giving me support, encouragement and constant prayers that sustained me throughout my life.

I must express my gratitude to Kano State Government and Near East University for giving me the opportunity to obtain this valuable diploma, Near East University in general and Nicosia wastewater treatment plant for proving the necessary data and documents to accomplish this task.

ABSTRACT

In the present study, three different artificial intelligence-based non-linear models, i.e. FFNN, ANFI), SVM approaches and MLR method were applied for predicting the performance of Nicosia wastewater treatment plant (NWWTP), in terms of effluent biological oxygen demand (BODeff), chemical oxygen demand (CODeff) and total nitrogen (TNeff). The results showed that in prediction of BODeff, the ensemble models of simple averaging ensemble (SAE), weighted averaging ensemble (WAE) and neural network ensemble (NNE), increased the performance efficiency of artificial intelligence (AI) modeling up to 14%, 20% and 24% at verification phase, respectively, and less than or equal to 5% for both CODeff and TNeff in calibration phase. This shows that NNE model is more robust and reliable ensemble method for predicting the NWWTP performance due to its non-linear averaging kernel. Secondly, the ELM integrated with PCA and compare with the MLP neural network and MLR models. The comparison results demonstrated that ELM and MLP revealed higher prediction accuracy than the MLR model, and the ELM model comparably outperformed MLP model. Overall results indicated that both the PCs-ELM and two scenarios could be alternatives and reliable tools for modeling the performance of Nicosia MWWTP. The study also proposed two types of non-linear system identification (NSI) models i.e., HW and NARX model with the classical method known as AR model to estimate effluents characteristic of total suspended solids (TSS_{eff}) and pH_{eff}. For comparison with the traditional AR, the results indicated that both HW and NARX are more accurate than the AR model. Hence, the outcomes determined that the NSI model (HW and NARX) are reliable modeling tools that could be adopted for the simulation of pH_{eff} and TSS_{eff}, respectively.

***Keywords*: A**rtificial intelligence; Black box model; Ensemble learning; Nicosia wastewater treatment plant; Wastewater

ÖZET

Bu çalışmada üç farklı yapay zeka temelli doğrusal olmayan modeller olan (FFNN), ANFIS, SVM, yöntemleri ile klasik MLR modeli kullanılarak Lefkoşa atık su arıtma tesisinin (NWWTP) performansı, biyolojik oksijen ihtiyacı (BODeff), kimyasal oksijen ihtiyacı (CODeff) ve toplam azot (TNeff) karakteristikleri dikkate alınarak tahmin edilmeye çalışılmıştır. Sonuçlar, BODeff'in tahmininde, basit ortalama (SAE), ağırlıklı ortalama (WAE) ve sinir ağı (NNE) topluluk modellerinin yapay zeka (AI) modellemesinin performans verimliliğini doğrulama safhasında, sırasıyla % 14, % 20 ve % 24 arttırdığını, kalibrasyon aşamasında ise hem CODeff hem de TNeff için% 5'den küçük veya ona eşit olduğunu göstermiştir. Bu, NNE topluluk modelinin NWWTP performansını öngörmede daha sağlam ve güvenilir bir yöntem olduğunu göstermektedir. Ayrıca, yeni kullanılmaya başlanan bir kara kutu modeli olan aşırı öğrenme makinesi (ELM), birleşik temel bileşen analizi (PCA) ile kurulmuş ve geleneksel çok katmanlı algılayıcı (MLP) sinir ağı ve çoklu doğrusal regresyon (MLR) modelleri ile karşılaştırılmıştır. Karşılaştırma sonuçları, ELM ve MLP'nin MLR modelinden daha güvenilir sonuçlar verdiğini ve ELM modelinin MLP modeline kıyasla daha iyi performans gösterdiği sonucunu vermiştir. Genel sonuçlar her iki modellemenin de Lefkoşa atık su arıtma tesisinin performansını ölçmek için alternatif ve güvenilir birer araç olabileceklerini ortaya çıkarmıştır. göstermiştir. Çalışma aynı zamanda, atık su karakteristikleri olan toplam askıda katı madde (TSSeff) ve pHeff'i tahmin etmek için oto regressif (AR) model olarak bilinen klasik yöntemle, doğrusal olmayan sistem tanımlama (NSI) sinir ağı yöntemini (Hammerstein-Weiner Modeli (HW) ve NARX) kullanmıştır. Geleneksel AR ile karşılaştırıldığında, sonuçlar hem HW hem de NARX'ın AR modelinden daha doğru sonuçlar verdiği sonucunu göstermiştir. Dolayısıyla, NSI modelinin (HW ve NARX), pHeff ve TSSeff'in simülasyonu için benimsenebilecek güvenilir modelleme araçları olarak kullanılabilirler.

Anahtar Kelimeler: Yapay zeka, Kara kutu modeli; Topluluk öğrenme modeli; Lefkoşa atık su arıtma tesisi; Atık su

TABLE OF CONTENTS

AKNOWLEDGEMENTS .... **ii**

ABSTRACT .... **iii**

ÖZET .... **iv**

TABLE OF CONTENTS .... **v**

LIST OF TABLES .... **ix**

LIST OF FIGURES .... **x**

LIST OF SYMBOLS .... **xiii**

CHAPTER 1 : INTRODUCTION **1**

1.1 Overview 1

1.2 Problem Statement 4

1.3 Aim of the Study 5

1.4 Objectives of the Study 6

1.5 Hypothesis 6

1.6 Significance of the Study 7

CHAPTER 2 : LITERATURE REVIEW **9**

2.1 Previous reviews for new Nicosia WWTP 9

2.2 Application of ANN in Water and wastewater treatment plant 10

2.3 Application of ANFIS in Water and wastewater treatment plant 15

2.4 Application of SVM in the wastewater treatment plant 20

2.5 Application of data-driven models in the wastewater treatment plant 21

CHAPTER 3 : MATERIAL AND METHODS..25

3.1 Study area and location..25

3.2 Plant and Operational line of Treatment..26

3.3 Data used and Pre-processing..28

3.4 Data Processing and Statistical analysis..29

3.5 Parameters Description..30

3.5.1 Influent and Effluents concentration of pH at Nicosia WWTP..30

3.5.2 Influent and Effluents concentration of COD at Nicosia WWTP..31

3.5.3 Influent and Effluents concentration of BOD at Nicosia WWTP..33

3.5.4 Influent and Effluents concentration of NH4-N at Nicosia WWTP..34

3.5.5 Influent and Effluents concentration of TSS at Nicosia WWTP..36

3.5.6 Influent and Effluents concentration of TN at Nicosia WWTP..37

3.5.7 Influent and Effluents concentration of TP at Nicosia WWTP..39

3.5.8 Influent and Effluents concentration of Conductivity at Nicosia WWTP.......40

3.5.9 Influent and Effluents concentration of SS at Nicosia WWTP..42

3.6 Proposed Methodology..43

3.7 Reasons for combining linear and non-linear models..47

3.8 Input variable and models selection..47

3.9 Methodology for the scenario I models (FFNN, SVM, ANFIS, MLR, and Ensemble) 52

3.9.1 Feed Forward Neural Network (FFNN)..52

3.9.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)..54

3.9.3 Support Vector Machine (SVM)..57

3.9.4 Multilinear regression analysis (MLR)..59

3.9.5 Ensemble Learning Approach......60

3.10 Methodology for scenario II models (ELM, MLP, and PCA)......63

3.10.1 Extreme Learning Machine (ELM)......63

3.10.2 Multi-layer perceptron (MLP) neural network......66

3.10.3 Principal Component Analysis (PCA)......67

3.11 Methodology for scenario III models (AR, HW, and NARX)......68

3.11.1 Autoregressive (AR) model......68

3.11.2 Hammerstein- Weiner Model (HW)......69

3.11.3 Nonlinear autoregressive with exogenous (NARX) neural network......69

CHAPTER 4 : RESULTS AND DISCUSSION......71

4.1 Results for the scenario I models (FFNN, SVM, ANFIS, MLR, and Ensemble) 71

4.1.1 Results of single model predictions of BOD_{eff}, $COD_{eff,}$ and TN_{eff}......71

4.1.2 Results of ensemble predictions for BOD_{eff}, $COD_{eff,}$ and TN_{eff}......79

4.2 Results for scenario II models (ELM, MLP and PCA)......83

4.2.1 Implementation of scenario I......83

4.2.2 Implementation of scenario II......95

4.3 Results for scenario II models (AR, HW, and NARX)......100

CHAPTER 5: CONCLUSION AND RECOMMENDATION......110

5.1 Conclusion for scenario I models (FFNN, SVM, ANFIS, MLR and Ensemble) 110

5.2 Conclusion for scenario II models (ELM, MLP and PCA)......111

5.3 Conclusion for scenario II models (AR, HW, and NARX)......112

REFERENCES 113

LIST OF TABLES

Table 3.1: Parameter used (Influents and Effluents) 28

Table 3.2: Descriptive Statistic of the Parameters 30

Table 3.3: Pearson Correlation matrix between the effluent and influent quality 48

Table 3.4: Selected input variables in FFNN, ANFIS, SVR and MLR 49

Table 3.5: Pearson and Spearman correlation analysis of the parameters 51

Table 3.6: The development of input variables models 51

Table 3.7: Classification of KMO coefficients 67

Table 4.1: Performance efficiency of single FFNN, ANFIS, SVM, and MLR models 74

Table 4.2: Results of the proposed ensemble techniques 80

Table 4.3: Eigenvalue and percentage of data explained by each factor. 84

Table 4.4: Results of ELM, and PCs-ELM for BOD_{eff}, COD_{eff}, $TN_{eff,}$ and TP_{eff} 86

Table 4.5: Results of MLP and MLR models for BODeff, CODeff, TNeff, and TPeff 96

Table 4.6: Results of NSI models for pH_{eff} and TSS_{eff} 102

LIST OF FIGURES

Figure 3.1: Study location showing the map of Cyprus 26
Figure 3.2: Schematic of the Nicosia NWWTP line of treatments... 27
Figure 3.3: Time-series and Box-plot of pH concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 31
Figure 3.4: Time-series and Box-plot of COD concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 32
Figure 3.5: Time-series and Box-plot of BOD concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 34
Figure 3.6: Time-series and Box-plot of NH4-N concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 35
Figure 3.7: Time-series and Box-plot of TSS concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 37
Figure 3.8: Time-series and Box-plot of TN concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 38
Figure 3.9: Time-series and Box-plot of TP concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 40
Figure 3.10: Time-series and Box-plot of Conductivity concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).. 41
Figure 3.11: Time-series and Box-plot of SS concentrations of the plant at the (a) inlet (influent).. 42
Figure 3.12: Flow chart of the proposed scenario I 44
Figure 3.13: Flow chart of the proposed scenario II.. 45
Figure 3.14: Flow chart of the proposed scenario III 46
Figure 3.15: Schematic diagram of ANN-FFNN showing Input, Hidden and Output Layers 53
Figure 3.16: Structure of Adaptive Neuro-Fuzzy Inference 55
Figure 3.17: Conceptual Network Architecture of SVR Algorithms 59
Figure 3.18: Schematic of Simple Averaging Ensemble 62

Figure 3.19: Schematic Structure of Weighted Average Ensemble 62
Figure 3.20: Schematic Structure of Neural Ensemble Method..63
Figure 3.21: The structure of the extreme learning machine network (Yaseen et al., 2018)...66
Figure 3.22: Multilayer perceptron neural network (MLPNN) architecture...........................67
Figure 3.23: Schematic of Hammerstein-Wiener model 69
Figure 3.24: Architectures of the NARX neural network (Men et al., 2014) 70
Figure 4.1: Observed vs predicted time series obtained by best single models for (a) $BOD_{eff,}$ (b) COD_{eff} and (c) TN_{eff}..76
Figure 4.2: Observed vs predicted Scatter plot for the best single model obtained by best single models for (a) $BOD_{eff,}$ (b) COD_{eff} and (c) TN_{eff} 78
Figure 4.3: RMSE for the best model in the verification phase 78
Figure 4.4: Scatter plot and time series plots of results obtained by NNE techniques for a) BOD_{eff}, (b) COD_{eff} and (c) TN_{eff} 81
Figure 4.5: Radar chart of DC for SAE, WAE, and NNE for both calibration and verification phases (BODeff, CODeff, and TNeff) 83
Figure 4.6: (a) shows the percentage variance vs the number of factors and eigenvalue vs the number of factors (b) orthonormal loadings biplot of the first two components of the PCA model...85
Figure 4.7: The comparison box-plot of the observed and all the predicted models 87
Figure 4.8: Taylor diagram showing the degree of prediction in terms of R and SD for (a) BODeff (b) CODeff (c) TNeff and (d) TPeff..90
Figure 4.9: Scatter plots of observed and computed values for the best model of BODeff, CODeff, TNeff and TPeff 92
Figure 4.10: Error bar chart plots for the simulated BODeff, CODeff, TNeff and TPeff in both calibration and verification phase. 94
Figure 4.11: Scatter plots of observed and computed values for the best model of BODeff, CODeff, TNeff and TPeff 99
Figure 4.12: The autocorrelation and partial autocorrelation functions of pH_{eff} and TSS_{eff}.. 101

Figure 4.13: The time series plots for the best observed versus computed models for HW, NARX and AR models.. 104
Figure 4.14: Taylor diagrams for evaluating the performances of best models for (a) pHeff (HW-M1 and M2) and (b) TSSeff (NARX-M1 and M2) 105
Figure 4.15: Radar chart for CC in both calibration and verification phase for (a) pHeff and (b) TSSeff 107
Figure 4.16: The box pots the observed and computed value of effluent pH_{eff} and TSS_{eff} 108
Figure 4.17: Error bar chart plots for the simulated pHeff and TSSeff in both calibration and verification phase. 109

LIST OF ABBREVIATION

%:	Percentage
^{0}C:	Degree centigrade
AI:	Artificial intelligence
ANFIS:	Adaptive neuro fuzzy inference system
ANN:	Artificial neural network
ASCE:	American Society of Civil Engineer
Autoregressive:	AR
ARIMA:	Auto Regressive Integrated Moving Average
BODinf:	Influent Biological oxygen demand
BODeff:	Effluents Biological oxygen demand
BP:	Back propagation
Condinf:	Influent conductivity
Condeff:	Effluent conductivity
CODinf:	Influent Chemical oxygen demand
CODeff:	Effluents Chemical oxygen demand
DC:	Determination coefficient
ELM:	extreme learning machine
FFNN:	Feed forward neural network
FL:	fuzzy logic
GA:	Genetic algorithm
H-W:	Hammerstein-Wiene
KMO:	Kaiser–Meyer–Olkin
mg/L:	Milligram per litre
MLFF:	Multi-layer feed forward
MBR:	Membrane Bioreactor
MWWTP:	Municipal wastewater Treatment Plant
MLR:	Multi-linear regression
MLP:	Multi-layer Perceptron

MSE:	Mean square error
MAPE:	Mean absolute percentage error
NWWTP:	Nicosia wastewater treatment plant
NNE:	Neural network ensemble
NARX:	Nonlinear autoregressive with exogenous
NNAR:	Neural Network Auto-regression
NSI:	Non-linear system identification
pHinf:	Influent pH
pHeff:	Effluent pH
PCA:	Principal Component Analysis
R:	Correlation coefficient
RMSE:	Root Mean Square Error
R^2:	Determination coefficient
SVR:	Support Vector Regression
SVM:	Support vector machine
SS:	Suspended Solids
SAE:	Simple averaging ensemble
SLFNs:	Single Layer Feed-Forward Networks
TNinf:	Influent Total Nitrogen
TNeff:	Effluents Total Nitrogen
TSS:	Total Suspended Solid
TN:	Total Nitrogen
TP:	Total Phosphates
UNESCO:	United Nation Educational, Scientific and Cultural Organization
UNDP:	United Nation Development Programs
WAE:	Weighted Averaging Ensemble
WWTPs:	Wastewater Treatment Plants
WNN:	Wavelet neural network

CHAPTER 1
INTRODUCTION

1.1 Overview

Water is essential to sustain life; therefore, affordable and adequate supply of water must be available (WHO and UNICEF 2012). Wastewater Treatment Plant (WWTP) is the process that removes the contaminants from the untreated domestic wastewater with the goal of safeguarding the public health and natural environment (Gómez et al., 2017; Mesdaghinia et al., 2015; Nourani et al., 2018a). wastewater management is important to protect our environment from deteriorating as well as improving the water scarcity which exists in a place where the water is insufficient to meet satisfy requirements demands (Gozen and Turkman (2008). WWTP is extremely complex and dynamic process due to its intricacy of the treatment method., Appropriate action, maintenance and control of WWTPs is very vital for monitoring the environmental and ecological health (Gaya et al., 2014a).

A satisfactory treatment plant is quite vital in order to avoid the discharge of high pollutants and meet the required standards by law. The parameters combination from physical, chemical and biological characteristics are often the major factors affecting the operation and control of WWTPs (Mikosz, 2016). Due to various composition and characteristic of wastewater treatment plant (WWTP) variables, it is performance can be assessed by considering certain sensitive variables such as Total Nitrogen (TN), Biological Oxygen Demand (BOD), Total Suspended Solids (TSS) and Chemical Oxygen Demand (COD). Yet, the available literature and published studies for predicting the WWTP used these parameters (Tumer and Edebali, 2015; Nourani et al., 2018). The quality of untreated and treated sewage has a great effect on the operation and efficiency of any WWTP. However, WWTP comprises of large numbers of parameters and operations which are complex in terms of measurement and evaluation (Abba and Elkiran, 2017). Hence modelling this system is considered difficult due to the nature of the process and most of the available traditional models are based on the assumptions, estimation and requiring too much time and money, as such a reliable and appropriate tool are indispensable in predicting the performance of MWWPTP.

Due to the importance of wastewater management, planning, and control, the modelling field in this remains dynamic and active of study. Basically, the models applied in hydro-environmental studies can be categorized into two, namely, physical-based and data-driven models. Physical models are based on the concept of distributed (white-box) models addressed the physical process and interaction for simulating the hydro-environmental system. In contrast, data-driven models are based lumped (black-box) models that acquire the optimal links between inputs and outputs but neglect the physical process (Hadi and Tombul, 2018). Various efforts have been presented to improve the accuracy and reliability of the effluent variables in the field of hydro-environmental studies, but no individual method has been proved applicable in modelling environmental process (Danandeh et al., 2018). With regards to this perspective, it could be pronounced that there are no acceptable single models which can perform better than the other in the different hydro-environmental system, due to the dynamic and complex nature of the data. This has necessitated the development of reliable and efficient models using the available data (Yaseen et al., 2015; Govindaraju, 2000). In addition, the process of WWTP have both deterministic and stochastic system, stochastic time series model such as Multilinear regression analyses (MLR), Autoregressive (AR) models have been used in modelling and prediction of hydrological process especially time-series process (Hadi and Tombul, 2018). The AR is widely known by it is moderation and simplicity among the linear models and is employed in several modelling studies (Kişi, 2008). Owing to its linear nature, AR may not reliably and properly model the possibly intricate processes taking place in WWTP.

Based on the established WWTP, linear and conventional regression tools have been widely used, but they have been generally associated with low accuracy levels, giving room to the development of the AI methods which are considered as accurate and non-linear hydrologic tools (Nourani 2018a). Meanwhile, several researchers have established different types of intelligence techniques such artificial intelligence (AI) which have been gradually applied for modelling and estimation in various discipline of hydrology and environmental engineering in order to rescue the existing traditional models (Lermontov et al., 2009; Abba et al., 2017; Nourani et al., 2018a; Nourani et al., 2018b; Elkiran et al., 2018; Nazir et al., 2019; Zhu et al., 2019). On the other side, the artificial intelligence (AI) models play a vital role and created great variations for forecasting several environmental and hydrological phenomena (Solgi et al., 2017; Elkiran et al., 2018; Nourani et al., 2018b; Jeihouni et al., 2019; Tiantian et al., 2019). Meanwhile, recent researches showed that the black-box models like Support Vector Machine (SVM), Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) could be proper alternatives for any WWTPs performance analysis.

For example, Maleki et al. (2018) predicted the influent parameters in WTP using Auto-Regressive Integrated Moving Average (ARIMA) and Neural Network Auto-regression (NNAR) models, despite an acceptable performance of ARIMA model, the results observed better prediction performance for NNAR with regard to ARIMA. Chen et al. (2001) developed ANN, genetic algorithm (GA) and fuzzy logic (FL) as a new method for modelling the industrial WWTPs at Taiwan. The proposed new method served as the control strategies in the successful management of the WWTP. Verma et al. (2013) demonstrated the ability of five different data mining approaches includes, MLP, K-nearest neighbour, SVM, random forest and multivariate adaptive regression spline to estimates the total suspended solids (TSS) in a WWTP using different input parameters. The obtained results depicted that MLP outperformed all the models.

Memon et al., (2012) developed an artificial neural network (ANN) with multi-layer perceptron (MLP) model to forecast the treated and untreated pH using the 17 measured input parameters in water treatment plant (WTP). The outcomes proved the suitability of MLP in modelling the drinking WTP parameters. Granata et al., (2017) made an attempt using several types of algorithms (i.e., support vector regression (SVR) and regression tree (RT)) to simulate wastewater quality indicators such as BOD, COD, TDS, and TSS. From the outcomes, it was observed that both models showed the robustness and reliability in the prediction; however, a significant performance of SVR was observed with regards to RT in modelling the effluent TSS, TSS, and COD.

Similarly, Gaya et al. (2017) developed the first implementation of ANN and Hammerstein-Wiene (H-W) models for forecasting the influent turbidity in Tamburawa WTP using different input parameters. The results indicated that ANN could outperform the H-W model and could serve as an acceptable tool for modelling the turbidity of WTP. Guo et al. (2015) used the influents of pH, Temperature, COD, and SS to predict the concentration of TN effluent from the WWTP in Ulsan, Korea, by employing ANN and SVM models and concluded that, AI models could be reliable methods for prediction of the effluent conditions of the WWTPs. Civelekoglu et al. (2009) applied ANN and ANFIS methods to model the COD removal in biological WWTP, and the overall results indicated that ANFIS is a suitable model for prediction of the WWTP performance. Hamed et al. (2004) used the BOD and TSS values recorded at various positions as input parameters and outlet BOD and TSS as target variables to predict the performance of WWTP using ANN model. The results proved the ability of the ANN model for predicting WWTP performance.

As the literature review shows, there is no unique model to be superior to others in all cases, and the performances of different models may be different according to the condition of each WWTP. Therefore, it is tested and verified that the combination of outputs (from different models) through an ensemble method may lead to more accurate results. The idea of such an ensemble model has been already used in different fields of engineering, environmental and water quality modelling (Cloke and Pappenberger, 2009; Sharghi et al., 2018). However, since the pronouncement of ensemble methods in engineering, to the best of the authors' knowledge, there is no published study in the technical literature indicating the application of AI-based ensemble approach in WWTP modelling.

1.2 Problem Statement

Cyprus suffers from water scarcity, and it is, therefore, no coincidence that a substantial share of the EU's aid programmed for Turkish Cypriots has been allocated to the water sector, including for wastewater collection and treatment. At normal condition, when water is scared reusing, it is beneficial for everyone. In many countries, water and wastewater management has proved to be a powerful incentive to overcome political and cultural tensions and build trust and peace between the different communities. Likewise, in Cyprus, wastewater management has been at the core of bi-communal cooperation between the two communities of Nicosia since the 1960s. In recent years, due to the growing urban development in both the side of Nicosia, the existing WWTP has begun to increasingly experience capacity overload and could not meet the European union effluent quality requirements. This led to heavy environmental burdens for the neighbouring areas, and unpleasant odour had become a serious nuisance to Nicosia residents. As such a state-of-art was urgently installed to control the situation.

I. Problems of water scarcity from all sectors
II. bi-communal cooperation
III. Growing urban developments
IV. Operation and control of WWTPs is difficult and time-consuming
V. The general WWTP system is Complex
VI. The traditional linear model is based on rough estimation, linear approximation, and assumption.
VII. There is a need for a reliable and convenient modelling tool

1.3 Aim of the Study

The thesis aimed to develop and compare the potential of some AI-based models (feed-forward neural network (FFNN), support vector machine (SVM), and adaptive neuro-fuzzy inference system (ANFIS) conventional multi-linear model (MLR) for prediction of the Nicosia WWTP performance considering four different combinations of input parameters. To establish and apply three ensemble techniques using the outputs of the aforementioned single models in order to improve the overall efficiency of the prediction performance. In this way, simple linear averaging, weighted linear averaging and non-linear neural ensemble techniques are proposed to combine the outputs of the methods. In addition, the other selected data-driven- approach were also such as extreme learning machine (ELM), multilayer perceptron (MLP) neural network, non-linear system identification (NSI) models (Hammerstein- Weiner Model (HW) and Nonlinear autoregressive with exogenous (NARX) neural network model) and a classical method known as autoregressive (AR) were also employed to achieved the same aim.

1.4 Objectives of the Study

- To perform the sensitivity analysis or employ PCA techniques to determine the most dominant parameters.
- To develop an independent model for BOD_{eff}, COD_{eff}, TN_{eff}, TP_{eff}, TSS_{eff}, pH_{eff},
- To determine the performance of WWTP using BOD_{eff}, $COD_{eff,}$ and TN_{eff}
- To determine the performance of WWTP using BOD_{eff}, COD_{eff}, $TN_{eff,}$ and TP_{eff}
- To determine the performance of WWTP using TSS_{eff} and pHeff
- To develop and compare different AI-based models in modelling the performance of WWTPs
- To establish and apply three ensemble techniques in order to improve the overall efficiency of the prediction performance.

1.5 Hypothesis

- ❖ Conventional models are capable of modelling WWTPs performance
- ❖ AI-based model is capable of modelling WWTPs
- ❖ Ensemble techniques can improve the performance of both AI and conventional models

1.6 Significance of the Study

According to the United Nations Educational, Scientific and Cultural Organization (UNESCO) 2015, WWTPs is paramount important for sustainable development and critical for human health ecosystems. This study will overcome the problems of water scarcity in Cyprus, particularly Northern Cyprus. The will serves as another important benchmark that will create and bring the two communities of Nicosia together. In another important significance of this study is the agricultural production, the treated effluents will substantially be used for irrigation, other farming activities and recharging the aquifers. Finally, the study may serve as the background for researchers carrying out further studies in new Nicosia WWTP.

CHAPTER 2
LITERATURE REVIEW

2.1 Previous reviews for new Nicosia WWTP

Abba and Elkiran (2017) implemented a study on the prediction of effluent chemical oxygen demand from the new Nicosia WWTP using different input combination models. The input parameters of ANNs are influents COD, BOD, pH, Conductivity, Total Nitrogen (T-N), Total Phosphates (T-P), Total suspended solids (TSS), Suspended solid (SS). The ANN performance has been evaluated using statistical techniques (DC and RMSE), the result of ANNs model was compared with the MLR, and the efficiency revealed that ANNs model showed the prominent accuracy and better performance in predicting the effluent COD over the MLR model.

Nourani et al. (2018) proposed different types of data-driven algorithms, including Feed Forward Neural Network (FFNN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Support Vector Machine (SVM) and conventional Multi-Linear Regression (MLR) for modeling and forecasting the performance of Nicosia wastewater treatment plant (NWWTP), in terms of effluent Biological Oxygen Demand (BOD_{eff}), Chemical Oxygen Demand (COD_{eff}) and Total Nitrogen (TN_{eff}). The daily data were used to develop single and ensemble models to improve the prediction ability of the methods. The obtained results of single models proved that the ANFIS model provides effective outcomes in comparison with single models. In the ensemble modelling, simple averaging ensemble, weighted averaging ensemble, and neural network ensemble techniques were proposed subsequently to improve the performance of the single models. The results showed that in prediction of BOD_{eff}, the ensemble models of simple averaging ensemble (SAE), weighted averaging ensemble (WAE) and neural network ensemble (NNE), increased the performance efficiency of Artificial Intelligence (AI) modeling up to 14%, 20% and 24% at verification phase, respectively and less than or equal to 5% for both COD_{eff} and TN_{eff} in calibration phase. This shows that NNE model is more robust and reliable ensemble method for predicting the NWWTP performance due to its non-linear averaging kernel.

2.2 Application of ANN in Water and wastewater treatment plant

Dogan et al. (2008) employed the ANN model to predict the measure daily influent BOD in WWTPs using the daily instances set of 354 records for the year 2005. Several input combinations were developed for this purpose which includes COD, discharge (Q), SS, TN and TP from a local WWTP. The outcomes are compared in terms of mean square error (MSE), DC and average absolute relative error with a classical MLR model. The obtained shows that, ANN model produced better results than MLR and could be applied successfully in predicting daily BOD in the WWTP.

Hamed et al. (2004) applied the ANN model to modelling the performance of a conventional WWTP located in Cairo, Egypt, based on the historical records. The daily measures record data for 10 months were obtained from the WWTP includes BOD and SS at various places with the treatment plant. A different exploratory analysis was carried out to identified the links between data and determination coefficient as indicators of performance accuracy were also used. The obtained results depicted that, ANN model is capable of predicting the performance of WWTP and therefore marked as a reliable tool for modelling the conventional WWTP in the Greater Cairo district, Egypt.

Moreno-Alfonso and Redondo (2001) developed a concept based on intelligent WWT, which are supported by two different ANNs with the intention of managing and controlling the treatment plant. For this purpose, historical data of various parameters were obtained. The first ANN was applied for the steady control of plant, and the second was used as natural (automatic) and was devoted to monitoring the sensitive parameters. The target regulates whether the employed decision is necessary. The results also proved that the two ANN system prove useful in the management of WWTPs.

Zhang and Stanley (1999) implemented a study on the ability of ANNs to determine the performance and control of WWTP processes. The employed model was found to be the major element that controls the number of processes such as coagulation, sedimentation, and flocculation. The obtained results also showed that the model was found quite reliable for optimum alum prediction.

Güçlü and Dursun, (2010) studied ANNs models trained with BP algorithm to estimates COD, SS and aeration tank-mixed liquor suspended solids (MLSS) concentrations at Ankara central WWTPs. The study involves the determination of the desired architecture through various calibration and verification of the models. The results obtained indicated that ANN models produced suitable prediction in terms of RMSE, MAE, and MAPE. According to the above performance indices, the obtained results also justified that ANNs models are satisfactory, precise and reliable approaches to control the processes and concentrations of Ankara WWTP.

Mjalli et al. (2007) studied the black- box viz: ANN model to obtain the real WWTP knowledge base and implement it as a process model. The obtained data from local WWTP called Doha WWTP were implemented for training and validation of the model, which comprises of effluents BOD, COD, TSS as the inputs of the model and effluent stream as the model output. The results indicated that using the above combinations, and the prediction became better. In addition, the prediction demonstrated the ability of ANN in capturing the operational process and the characteristics with good performance accuracy.

Pisa et al. (2018) proposed ANN model with Recurrent Neural Networks (RNNs) to estimate and monitor the outflow concentrations of WWTP based on total nitrogen and ammonia as the predicted target of the plant. The online data measured by a sensor positioned in WWTP was used for this analysis using different influent parameters (temperature, ammonia, overflow rate, and internal cycle flow rate). The used MAPE index as the measure of performance efficiency. The proposed approach was found to provides satisfactory prediction results in term of error accuracy.

Hamoda et al. (1999) studied the complex processes taking place in WTP with the emphasis on a data-driven algorithm of ANN. Among different types of ANN, a BP model was employed to simulate the MWWTP at Ardiya Kuwait City in term of TSS and BOD. The daily recorded data (228 instances) for the period of 16 months were used in the analysis. For the purpose of this research, correlation coefficient (R) was applied as the performance efficiency of the model. The obtained results justified the superiority of ANN-BP in modelling and simulation of Ardiya WWTP in comparison with alternative conventional methodology.

Hanbay et al. (2008) developed a model based on intelligence approaches to simulate the performance of WWTP at Malatya WWTP, Turkey. The wavelet neural network (WNN) package, NN-MLP and wavelet entropy (used of data input characterization) were employed to achieve the desired aim. The daily data records used in this paper were obtained from a laboratory of WWTP in Malatya, Turkey. Different input parameters were involved in predicting the performance in term of TSS as the target output. The obtained result depicted that, intelligence approaches are suitable for and robust tool to estimate the performance of WWTP.

Hassen and Asmare, (2019) proposed the application of two different ANN learning algorithms, including feed-forward neural network (FFNN) and BP, for modelling the effluent quality of WWTP located in Ethiopia. Both the influent and effluent parameters were recorded for approximately 11 months between 2016-2017 as the historical data from Habesha brewery WWTP. The input-outputs parameters include the combination of pH, COD, and TN, while the R and MSE were employed to estimate the performance of the models. The results demonstrated high-performance accuracy by achieving the value of R more than 0.9 in both training and calibration. The obtained results also concluded that ANN has the ability to predict the quality of the effluent of Habesha WWTP; therefore, served as the predictive tool to determine the performance of the plant.

Jami et al. (2012) studied ANN models to determine the performance of WWTP in Malaysia using influent and effluent variables. The data were obtained from the Bandar Tun Razak Sewage Treatment Plant of Indah Water Konsortium Sdn Bhd (IWK) which covered two-year span with BOD, SS, COD as the input of the model while output contained the combined parameters. RMSE was used to analyses the performance of the model, and the obtained results show satisfactory prediction accuracy.

Pakrou *et al (2014)* suggested that improper functionality of sewage treatment might create serious problems for the environment and general health, they simulated the performance of Mashhad Industrial Town's wastewater treatment plant using multilayer perceptron neural networks (MLP-NN), which are among the most popular ANN used in environmental problems. Findings of the study showed that the provided neural networks model has an acceptable ability in predicting the performance of industrial treatment plants.

Nasr et al., (2012) considered Alexandria Wastewater Treatment Plant with the purpose of minimizing utility cost and evaluating environmental balance stability in treatment plants and increasing its performance, concentrating on making use of the artificial neural networks (ANN). Such parameters as chemical oxygen demand (COD), biochemical oxygen demand (BOD), and total suspended solids (TSS), and data collected in studies over a one-year period were considered. This study implies that ANN can increase plant performance with a correlation coefficient of R between observed and predicted output variables to 0.90.

Raduly et al., (2007) obtained treatment plants to evaluate performance and reliability via simulation in artificial neural networks, considering such factors as season temperature, rainfall amount, rainfall severity and duration, and effects of holidays on treatment plants, and they obtained similar results. Results show that artificial neural networks simulation offers acceptable results for wastewater system of an urban context, with an error of less than 10%.

Pakrou et al. (2015) employed ANN for estimating the efficiency performance of Tabriz WWTP using several input combinations viz: wastewater temperature, pH, turbidity, alkalinity, SS, COD, BOD, flow rate and small flow from sewage and wastewater particle. The 6-month available data were obtained from the WWTP engineering department for all the analysis, and the performance accuracy of the models was assessed based on RMSE, MAPE, R, and MAE. The obtained results demonstrated that ANN provided a satisfactory prediction with high correlation values.

Zidan et al. (2015) used ANNs model to determine the appropriates model for BOD, COD, and TSS in three different underflow wastewaters situated in Samaha village, Egypt and the input parameters for the models are influent concentration (Ci), loading rate (q), media surface area (As), and actual velocity (v). Afterwards, three approaches were adopted for comparison and feasibility (gravel, plastic, and tires) as different treatment media for the wetland. The data obtained from the experiment contained 300 instances, of which 240 and 60 were employed for training and validation, respectively. MSE and percentage error were used to determine the performance efficiency of the models. The results depicted that, ANN model demonstrated the good accuracy and reliable for the simulation of effluents concentration. The outcomes also showed that plastic media is superior to gravel and rubber with regards to performance accuracy.

Güçlü and Dursun, (2008) employed two different methods, including ANN and activated sludge model to evaluates the effluent COD concentration in Ankara central wastewater treatment plant (ACWTP). The historical data obtained from the plants and experimental laboratory contained several inputs variables. The study employed R, RMSE, MAE, and MAPE to determine the prediction performance efficiency of the models. The obtained results confirmed that the hybrid model (ASM-ANN) provides better performance accuracy and explain the operational process better than ASM.

Oliveira-Esquerre et al. (2002) acquired acceptable predictions in Brazil of the BOD for the output stream of paper production and a local biological wastewater treatment plant for the blend. The data in the backpropagated neural network (BPNN), was preprocessed by the use of principal component analysis.

Hong et al. (2003) applied the Kohonen Self-Organizing Feature Maps (KSOFM) neural network to detect the inter-relationship of the method variables in an actual simulated sludge procedure and to study the multidimensional procedure data. To quote knowledge from the multidimensional procedure of a large-scale WWTP, the authors came to the conclusion that the KSOFM method gives an effective study and an investigative means to know the behavior of the system.

Zhu et al. (2009) obtained the effectiveness of a biological treatment procedure for BOD amounts, a time-delay neural network (TDNN) was recommended. The one-line training of the neural network model brings about an upgrade in the prediction accuracy, as shown from the outcomes of the authors using actual procedure data.

2.3 Application of ANFIS in Water and wastewater treatment plant

Pai et al. (2011) studied three types of adaptive neuro-fuzzy inference system (ANFIS) to predict effluent suspended solids (SS_{eff}), chemical oxygen demand(COD_{eff}), and pH_{eff} from a wastewater treatment plant in an industrial park. For comparison, the artificial neural network (ANN) was also used. The data were obtained WWTP in the industrial park locating in the middle part of Taiwan. Among the total number of data, the numbers for training and testing (predicting) were 130 and 30, respectively. The input parameters included influent pH (pH_{inf}), influent temperature ($Temp_{inf}$), influent SS (SS_{inf}), and influent COD (COD_{inf}). The output parameters included effluent SS (SS_{eff}), effluent COD (COD_{eff}) and effluent pH(pH_{eff}). The performance criteria used in this study were R, RMSE, and MAPE. The results indicated that ANFIS statistically outperformed ANN in terms of effluent prediction. The minimum mean absolute percentage errors of 2.67%, 2.80%, and 0.42% for SS_{eff}, COD_{eff}, and pH_{eff} could be achieved using ANFIS. The maximum values of the correlation coefficient for SS_{eff}, COD_{eff}, and pH_{eff} were 0.96, 0.93, and 0.95, respectively. The minimum mean square errors of 0.19, 2.25, and 0.00, and the minimum root mean square errors of 0.43, 1.48, and 0.04 for SS_{eff}, COD_{eff}, and pH_{eff} could also be achieved. ANFIS's architecture can overcome the limitations of traditional neural network. It also revealed that the influent indices could be applied to the prediction of effluent quality.

Erdirencelebi and Yalpir (2011) developed and evaluated three adaptive network fuzzy inference system (ANFIS) models for a laboratory-scale anaerobic digestion system outputs with using different input selection approaches. The aimed was the investigation of the feasibility of the approach-based-control system for the prediction of effluent quality from a sequential up-flow anaerobic sludge bed reactor (UASBR) system that produced a strong nonlinear ship between its inputs and outputs. As ANFIS demonstrated its ability to construct any nonlinear function with multiple inputs and outputs in many applications, its estimating performance was investigated for a complex wastewater treatment process at increasing organic loading rates from 1.1 to 5.5 g COD/L d. Approximation of the ANFIS models was validated using correlation coefficient, MAPE, and RMSE. ANFIS was successful to unsteady model data for pH and acceptable for COD within anaerobic digestion limits with multiple input structure. The prediction performance showed high feasibility of the model-based-control system on the anaerobic digester system to produce an effluent amenable for a consecutive aerobic treatment unit.

Wan et al. (2011) employed advanced neuro-fuzzy modelling, namely an adaptive network-based fuzzy inference system (ANFIS), was employed to develop models for the prediction of suspended solids (SS) and chemical oxygen demand (COD) removal of a full-scale wastewater treatment plant treating process wastewaters from a paper mill. In order to improve the network performance, fuzzy subtractive clustering was used to identify model's architecture and optimize fuzzy rule. Meanwhile, principal component analysis(PCA) was applied to reduce the input variable dimensionality. Input variables were reduced from six to four for COD and SS models, by considering PCA results and linear correlation matrices among input and output variables. The results indicate that reasonable forecasting and control performances have been achieved through the developed system. The minimum mean absolute percentage errors of 1.003% and 0.5161% for COD_{eff} and SS_{eff} could be achieved using ANFIS. The maximum correlation coefficient values for COD_{eff} and SS_{eff} were 0.9912 and 0.9882, respectively. The minimum mean square errors of 1.2883 and 0.0342 and the minimum RMSEs of 1.135 and 0.1849 for COD_{eff} and SS_{eff} could also be achieved.

Civelekoglu et al. (2007) studied advanced neuro-fuzzy modelling, namely an adaptive network-based fuzzy inference system (ANFIS), to develop models for the prediction of carbon and nitrogen removal in the aerobic biological treatment stage of a full-scale wastewater treatment plant treating process wastewaters from the sugar production industry.

A total of six independent ANFIS models were developed with or without principal component analysis (PCA) using the correlations among the influent and effluent data from the plant. Input variables were reduced from eight to four and from eleven to nine for chemical oxygen demand (COD) and NH_4^+–N–TN (total nitrogen) models, respectively, by considering PCA results and linear correlation matrices among input and output variables. Correlation coefficients (R) were not in good agreement with root mean square error (RMSE) and average percentage error (APE) values without PCA. For the COD model after PCA; RMSE, APE and R values were 9.4 mg/L, 8.37 and 0.978%, respectively. Such values for the TN model were 4.3 mg/L, 23.65 and 0.992%. The results overall indicated that the simulated effluent COD, NH_4^+–N, and TN concentrations well fit measured concentrations. The ANFIS modelling approach may have application potential for performance prediction and control of treatment processes in treatment plants.

Pai et al. (2009) explored three types of adaptive network-based fuzzy inference system (ANFIS) in which the online monitoring parameters served as the input variable was employed to predict suspended solids (SS_{eff}), chemical oxygen demand (COD_{eff}), and pH_{eff} in the effluent from a biological wastewater treatment plant in industrial park. Artificial neural network (ANN) was also used for comparison. The results indicated that ANFIS statistically outperforms ANN in terms of effluent prediction. When predicting, the minimum mean absolute percentage errors of 2.90, 2.54 and 0.36% for SS_{eff}, COD_{eff} and pH_{eff} could be achieved using ANFIS. The maximum values of the correlation coefficient for SS_{eff}, COD_{eff}, and pH_{eff} were 0.97, 0.95, and 0.98, respectively. The minimum mean square errors of 0.21, 1.41 and 0.00, and the minimum root mean square errors of 0.46, 1.19 and 0.04 for SS_{eff}, COD_{eff}, and pH_{eff} could also be achieved.

Yel and Yalpir (2011) studied a fuzzy-logic-based diagnosis system to determine the primary treatment effluent quality in a municipal wastewater treatment plant (MWTP). The measured data of variables were implemented into the Fuzzy Inference System (FIS) with Mamdani’s method. The fuzzy control rule base was shaped to define essential quality parameters monitored as pH, COD, BOD, and SS outputs. The output approximations to real data remained in an acceptable range for an MWTP performance (89–96%). The averages and standard deviations of the model were also approximated closely as 93–98% and 89–97%, respectively. The resulting configuration proved a good modelling approach for MWTP effluent quality prediction.

Perendeci et al. (2007) proposed a conceptual neural-fuzzy model based on adaptive-network-based fuzzy inference system (ANFIS) to estimate effluent chemical oxygen demand (COD) of a full-scale anaerobic wastewater treatment plant for a sugar factory operating at unsteady state. The fitness of simulated results was improved by adding two new input variables into the model; phase vectors of the operational period and effluent COD values of the last five days (history). In modelling studies, the individual contribution of each input variable to the resulting model was evaluated. The addition of phase vectors and history of five days into the input variable matrix in ANFIS modelling for anaerobic wastewater treatment was applied for the first time in the literature to increase the prediction power of the model. By this way, the correlation coefficient between estimated and measured values of the output variable (COD) could be increased to the value of 0.8940, which is considered a good fit.

Rahimzadeh et al. (2016) used ANFIS as a powerful tool for modelling complex and nonlinear systems to predict the permeate volume of oil/water membrane separation process. The data used for modelling the flux behaviour consisted of three inputs (TMP, oil concentration, filtration time) and experimental permeation values as the output. First type Gaussian membership function was used for fuzzification of input variables, and the hybrid algorithm was chosen for the learning method of input-output data. Very well agreements were observed between experimental and simulation results. From the results, the ANFIS can be used as a reliable tool for the prediction of microfiltration systems' behaviour. The coefficient of determination (R^2) between the experimental and predicted values were greater than 0.99, and the mean percentage error was less than 2%, showing the great efficiency and reliability of the developed model.

Huang et al. (2012) employed a predictive control system based on an adaptive network-based fuzzy inference system (ANFIS) to develop models for predicting and controlling the performance of a paper-making wastewater treatment process. The system includes an ANFIS predictive model and an ANFIS controller. In order to improve the network performance, fuzzy subtractive clustering, Euclidean distance clustering, and principal component analysis (PCA) were used to identify model architecture and extract and optimize the fuzzy rule of the model. For the developed predictive model, when predicting, mean absolute percentage error (MAPE) lay 6.06% adopting ANFIS, root means square normalized error (RMSE) was 24.4485 and R was 0.9731. The control model, taking into account the difference between the predicted value of chemical oxygen demand (COD) and the setpoint, can effectively change the additive dosages. In order to verify the developed predictive control model, a paper-making wastewater treatment process was picked up to support operational performance. When the influent COD value or inflow flow rate was changed, the dosage could be accurately adjusted to make the effluent COD remain at the setpoint, and its MAPE was only 5.19%. The results indicated that reasonable forecasting and controlling performances had been achieved through the developed system.

Tay and Zhang (2000) combined fuzzy systems, and neural networks were in modelling the difficult method of anaerobic biological treatment of wastewater. The power of the method in two case studies of up-flow anaerobic sludge blanket an anaerobic fluidized bed reactor was illustrated. By looking at the previous information, the fuzzy neural model simulated the performance of the system well and delivered satisfactory outcomes of the prediction, even though, a high requirement on the excellence of the training data is the disadvantage of the model.

Mushiri et al. (2014) proposed an automated control for industrial wastewater treatment using fuzzy logic control is presented. Fuzzy control concepts have been adopted, and control algorithms based on pH and temperature variations provided accurate and reliable treatment of wastewater. Two inputs, pH and temperature, and four outputs, hot and cold-water valve and acid and base valve were given to the system. The behaviour of the system was observed whenever an input came in. the pH was to be kept at a neutral status of between 6.5 and 7.5, and the temperature had to be kept within a range of between 250C and 350C. The fuzzy rules managed to keep the system stable and therefore recommended satisfactory for the analysis.

2.4 Application of SVM in the wastewater treatment plant

Ribeiro et al. (2013) proposed a studied that was implemented to predict the performance of a WWTP located in northern Portugal, serving a population of about 45,000 inhabitants. The data we used were recorded based on the daily averaged values of the measured parameters during the period of one year. The predictive models were developed supported by two implementations of Support Vector Machines (SVM) methods for regression, due to the presence of two lines of treatment in the selected case of study, using two of the most relevant output parameters of a WWTP, the biochemical oxygen demand (BOD) and the total suspended solids (TSS). The showed that SVM model is capable of estimating the performance of WWTP in term of BOD and TSS.

Bao-lei et al. (2011) proposed the Least Square Support Vector Machine (LS-SVM) predict model of sewage outflow COD. The input parameters are the amount of water, PH, temperature, COD, sulfide, and mixed liquor suspended solids, and the output is the outflow COD. Compared with BP neural network, the experimental results verify that LS-SVM method has effectively improved performance in predicting sewage outflow COD. Some researches on the empirical application have been done with the monitoring data in a wastewater treatment plant to verify the effectiveness and feasibility of the model.

Li-juan and Chao-bo, (2008) adopted regression support vector machine (SVM) to set up a prediction model of a sewage treatment plant with a popular process Cyclic Activated Sludge System (CASS). Kernel function of the prediction model is radial basis function, and parameters of the kernel function are optimally determined by cross-validation. Then the prediction model is used to predict the effluent quality of the sewage treatment plant with CASS process. The test result of the case study shows that the prediction model works well, and the regression SVM is powerful in predicting effluent quality of CASS process sewage treatment plant with small sample learning ability and good generalization.

Huang et al. (2009) explored an improved least squares support vector machines for regression (LS-SVR) is proposed. Benchmark Simulation Model No.1 (BSM1) is used to generate input-output data, then effluent parameters, COD (chemical oxygen demand), BOD (biochemical oxygen demand), TN (total nitrogen), SNH (ammonium nitrogen) and TSS (total suspended solids) forecast model is built. The parameters of LS-SVR are optimized by particle swarm optimization (PSO) in order to obtain a more accurate model. The study shows that the improved LS-SVR modelling approach is capable of predicting the wastewater treatment plant effluent parameters with a good degree of accuracy and is adapted to the changes in the weather.

Hong et al. (2008) concluded that Least squares support vector machine (LS-SVM) combined with genetic algorithm (GA) are presented in this paper, and this new algorithm can be a classifier in the wastewater treatment process. The LS-SVM can overcome some shortcoming in the multilayer perception; meantime, the GA can be used to tune the parameters of LS-SVM automatically and can escape from the blindness of man-made choice of the parameters of LS-SVM. The numerical experiments for classifying the operational state of the wastewater treatment process show that the proposed algorithm is effective and has less prediction error.

2.5 Application of data-driven models in the wastewater treatment plant

Liu et al. (2013) demonstrated the application of hybrid learning method combining a genetic algorithm with an adaptive neuro-fuzzy inference system (GA-ANFIS) to estimate effluent nutrient concentrations in a full-scale biological wastewater treatment plant. The real data collected from Korean Daewoo nutrient removal wastewater treatment plant were used to demonstrate the prediction efficiency of the proposed soft sensor with the aid of three performance indices of root mean square error, mean absolute percentage error, and squared correlation coefficient. The results indicate that the hybrid GA-ANFIS soft sensors outperform ANFIS-based soft sensors in terms of effluent prediction accuracy.

Civelekoglu et al. (2009) evaluated artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) modelling methods to estimate organic carbon removal using the correlation among the past information of influent and effluent parameters in a full-scale aerobic biological wastewater treatment plant. Model development focused on providing an adaptive, useful, practical and alternative methodology for modelling of organic carbon removal. For both models, measured and predicted effluent COD concentrations were strongly correlated with determination coefficients over 0.96. The results overall indicated that the ANFIS modelling approach may be suitable to describe the relationship between wastewater quality parameters and may have application potential for performance prediction and control of aerobic biological processes in wastewater treatment plants.

Simsek, (2016) employed artificial intelligence techniques, such as adaptive neuro-fuzzy inference systems, multilayer perceptron, radial basis neural networks (RBNN), and generalized regression neural networks to estimates Biodegradable dissolved organic nitrogen (DON) in WWTPs. The input parameters include nitrite, nitrate, ammonium, TDN, and DON data were used as input neurons. Wastewater samples were collected from four different locations in the plant. Model performances were evaluated using root mean square error, mean absolute error, mean bias error, and coefficient of determination statistics. Modelling results showed that the R^2 values were higher than 0.85 in all four models for all wastewater samples, except the only R^2 in the final effluent sample for RBNN modelling was low (0.52). Overall, it was found that all four computing techniques could be employed successfully to predict BDON.

Oliveira et al. (2002) proposed ANN and PCA to predict the biochemical oxygen demand (BOD) of the output stream of the biological wastewater treatment plant in Brazil. The two years data from the plant were reduced to 71 instances due to many missing data and R, and MSE were used to predict the performance of the model. The best prediction performance is achieved when the data are preprocessed using principal components analysis (PCA) before they are fed to a backpropagated neural network. The influence of input variables is analyzed, and satisfactory prediction results are obtained for an optimized situation.

Areerachakul, (2012) compared between the predictive ability of the Adaptive Neuro-Fuzzy Inference System (ANFIS) model and Artificial Neural Network (ANN) model to estimate the Biochemical Oxygen Demand (BOD) on data from 11 sampling sites of Saen Saep canal in Bangkok, Thailand. The data is obtained from the Department of Drainage and Sewerage, Bangkok Metropolitan Administration, during 2004-2011. The five parameters of water quality namely Dissolved Oxygen (DO), Chemical Oxygen Demand (COD), Ammonia Nitrogen (NH3N), Nitrate Nitrogen (NO3N), and Total Coliform bacteria (T-coliform) are used as the input of the models. These water quality indices affect the biochemical oxygen demand. The experimental results indicated that the ANN model provides a higher correlation coefficient (R=0.73) and a lower root mean square error (RMSE=4.53) than the corresponding ANFIS model.

Chen et al. (2001) developed three data-driven models that include fuzzy logic (FL), ANN and genetic algorithms (GA) for modelling the industrial WWTPs. The industrial WWTP in Taiwan were considered as the case study to determine the applicability of the proposed models. the obtained results from this research proved that GA based controller could produce a better performance than the other two models in terms of economic and environmental point of views. The research also highlighted that this method could be applied to any WWTP processes by applying only little modifications.

Soltaninezhad et al. (2012) studied experimental data on daily measurements using the artificial intelligence of the wastewater treatment plant in Minorca, Spain, recorded over 1990-1991. The WWTP in an activated sludge system of this city was modelled using a data-mining process and results were compared. Target problem was to find a suitable algorithmic model for predicting output wastewater quality considering input wastewater quality of the same day.

Boyd et al. (2019) explored the ability of Autoregressive Integrated Moving Average (ARIMA) model for the prediction of WWTP performance in five different sample stations (Woodward, Niagara, North Davis, and two confidential plants) across North America. The daily influent-effluent data were used from each plant, and the performance efficiency was determined using the commonly used indices (R^2, RMSE, and MAPE). The predictive results demonstrated that ARIMA models can satisfactorily meet the required expectation for the prediction of daily influents flow.

CHAPTER 3
MATERIAL AND METHODS

3.1 Study area and location

Cyprus is an island in the Mediterranean Sea, the third biggest in the Mediterranean Sea, and is near the Middle East (and is sometimes included in the region geographically), south of Turkey. The latitude of Cyprus is 34° 33’ - 35° 34’ north and its longitude 32° 16’ - 34° 37’ east. The physical setting for life on the island is dominated by the mountain masses and the central plain they encompass, the Mesaoria (Hoşkara et al., 2007). The Troodos Mountains cover most of the southern and western portions of the island and account for roughly half its area. The narrow Kyrenia Range, extending along the northern coastline, occupies substantially less area, and elevations are lower. The two mountain systems run generally parallel to the Taurus Mountains on the Turkish mainland; whose silhouette is visible from Northern Cyprus (Hoşkara et al., 2007).

It was reported that management, control, and planning of wastewater exist to be the highest tool for bi-communal collaboration among the two peoples of Nicosia since the 1960s. The New Nicosia (NWWTP) is a bi-communal project serving two different communities between Turkish Cypriot and Greek Cypriot. The project was jointly founded by the sewage board of Nicosia and the European Union (EU) and implemented by United Nations Development Programs (UNDP). For sustainable development and recycling purposes, more than 300 tons/yr. will be generated. A total of about 10 million m^3 of quality effluent can be reused for different agricultural purpose (Nourani et al., 2018a; UNDP, 2014). Figure. 3.1 shows the map of Cyprus and the location of the study area. The operation of the plant will, therefore, be partly powered by renewable energy (10-20% on average), reducing its Carbon Dioxide (CO_2) emissions.

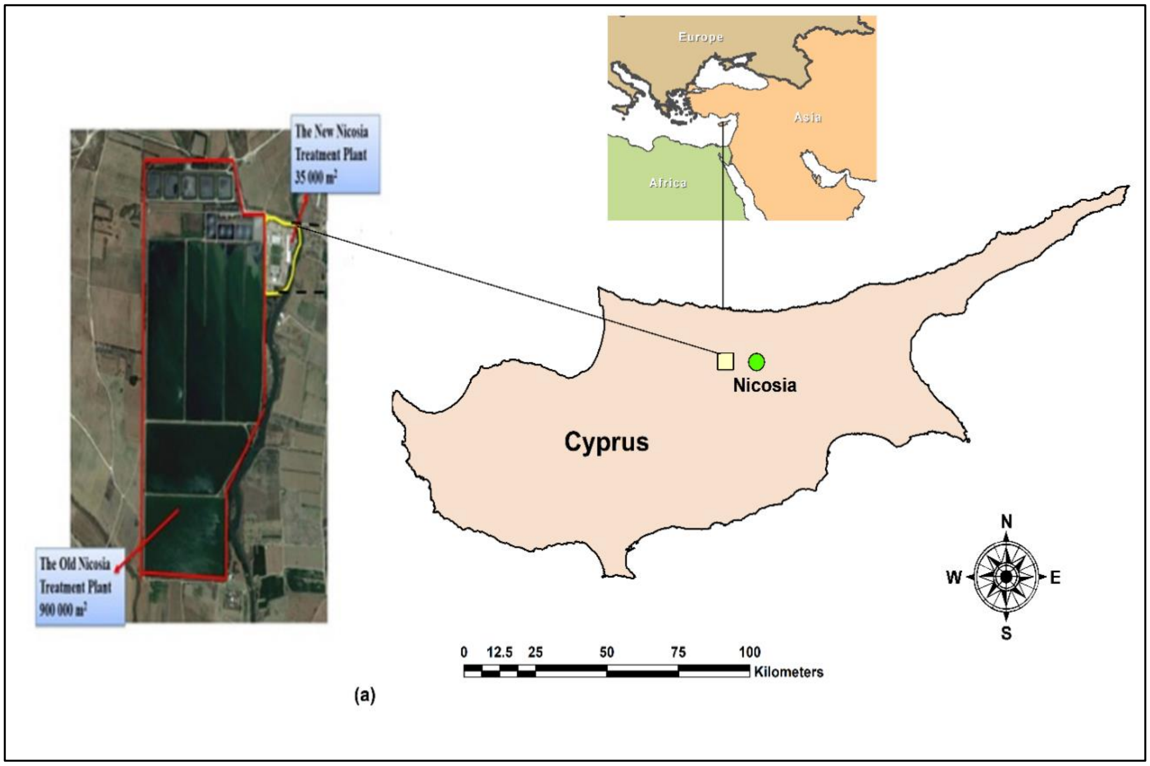


Figure 3.1: Study location showing the map of Cyprus

3.2 Plant and Operational line of Treatment

In new Nicosia MWWTP, the line of treatment comprises of 11 stages from the raw to treated sewage effluents as shown in Figure. 3.2 First of all, the sewage is separated into liquid and solid waste and goes through the first chamber called the *screening chamber* (1) where the solids larger than 6mm will be removed. The inflow then flows down slowly so that the heavy solids (grit, sand) can fall to the bottom and oil and grease float to the surface at *grit and grease chamber* (2). The pump station (3) was used to pump up the water to the next unit called *fine sieve* (4), this unit removes solids larger than 2mm. The next step is the *biological treatment* of waste which is the stage that creates the condition to encourage bacteria to consume the waste comprising of three units (5, 6 and 7). Stage (8) is where the separation and treatment of the by-products of the hall process into clean water, fertilizer, and biogas is taking place named *membrane treatment*. After that, the water is disinfected at a *chlorine contact tank* (10). Tank (11) is the treated sewage effluent before it gets discharged into the river (UNDP, 2014). Figure. 3.2 shows the schematic of the Nicosia NWWTP line of treatments.

The summary of the line of treatment are as follows:

- Raw effluent
- Screening chamber
- Grit and grease chamber
- Fine sieve
- Biological treatment
- Membrane treatment
- Chlorine contact tank
- Treated effluent

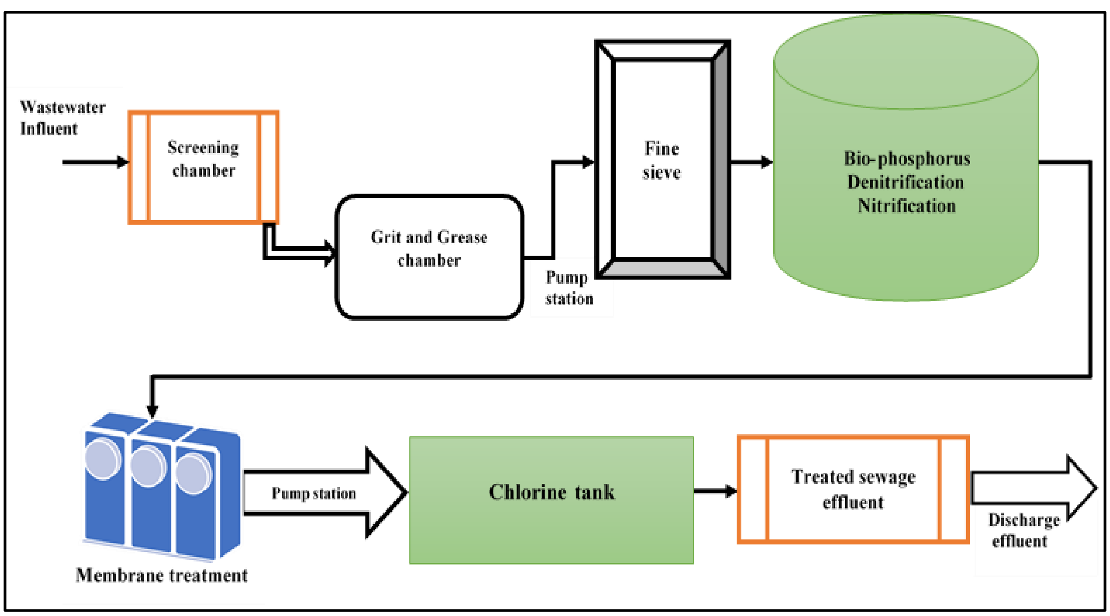


Figure 3.2: Schematic of the Nicosia NWWTP line of treatments

3.3 Data used and Pre-processing

For this thesis, the available daily data were obtained from the new NWWTP (United Nation Development Programs (UNDP), 2014). The measurement of the selected parameters covers all the seasonal variations and consists of various sets of inputs and outputs parameters. The daily measured data obtained from new Nicosia MWWTP which includes (pH_{inf}, $Conductivity_{inf}$ (Condinf), Biological Oxygen Demand (BOD_{inf}), Chemical Oxygen Demand (COD_{inf}), Total Nitrogen ($T\text{-}N_{inf}$), Total Phosphorite ($T\text{-}P_{inf}$), Ammonia nitrogen ($NH4\text{-}N_{inf}$), Suspended Solid (SS_{inf}) and Total Suspended Solid (TSS_{inf}) as the input variables and (BOD_{eff}, COD_{eff}, $T\text{-}N_{eff}$, $T\text{-}P_{eff}$, pH_{eff}, TSS_{eff}) as the corresponding output respectively. Table 3.1 shows the input and output parameters used in this thesis. Table 3.1 shows the general parameters used for the study. Note the selection of the parameters are quite line with the previous research done in the field of WWTP performance analysis. However, others studied engaged some of these parameters in the control and management of the WWTP from the angle of AI-based models.

Table 3.1: Parameter used (Influents and Effluents)

Parameters	Influents Parameter	Effluents Parameters
PHinf	pHinf	pHeff
Conductivityinf	Condinf	Condeff
Biological Oxygen Demand	BODinf	BODeff
Chemical Oxygen Demand	CODinf	CODeff
Total Nitrogen	T-Ninf	T-Neff
Ammonia nitrogen	NH4-Ninf	T-Peff
Total Phosphorite	T-Pinf	TSSeff
Suspended Solid	SSinf	
Total Suspended Solid	TSSinf	

3.4 Data Processing and Statistical analysis

Data processing is the process of turning the raw data into appropriate and meaningful information, prior to the model training, the data must be scale between the internal of 0 and 1 this process is called normalization (see, equation 3.1). The process was applied in order to deduce the data redundancy and increased data integrity (Abba and Elkiran, 2017). The normalized data were divided into 75% and 25% for both calibration and verification, respectively, over a period of 2015-2016 (contained 362 instances). The validation methods are implemented using a different approach. This study employed a holdout approach which is known as leave-group-out. In this approach, the data randomly assigned to two sets generally named calibration and verification and can be regarded as another version of k- fold cross-validation (Sargent, 2009; Tsioptsias et al., 2016, Nourani et al. 2018a).

$$\boldsymbol{y} = 0.5 + \left(0.5X\left(\frac{x-\bar{x}}{x_{max}-x_{min}}\right)\right) \tag{3.1}$$

Where y is the normalized data. x is the measured data, $\bar{x}$ is the mean of the measured data, x_{max} is the maximum value of the measured data and xmin is the minimum value.

Statistical analysis used to explain the data trend series are smoothing and normalization, the former was carried out by fitting the data into regression function to eliminate the noise from the data and latter was to ensure the uniformity of the input-output value (scaling to fall within a small, specified range). The descriptive statistic of the selected parameter can be presented in Table 3.2. Every data analysis concerning AI base models relies normally on historical data. Therefore, the data and statistical analysis of the input-output is essentials because it identifies the type and strength of the relations between inputs and outputs. In order to efficiently train AI base model, these data need to be clean and filtered properly, because the raw data often comprised of missing records, outliers, noise, discrepancies of codes and names or was infected by all kind of error including human and instrumental.

Table 3.2: Descriptive Statistic of the Parameters

Parameters	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
pHinf	7.5835	7.6000	8.2000	5.6000	0.2297	-1.9568	17.1633
CODinf	910.2262	960.0000	1463.0000	100.0000	273.1302	-1.5426	5.7670
BODinf	669.4396	411.0000	11685.0000	156.0000	1597.7780	6.0310	37.7502
NH4-Ninf	57.3728	58.0000	83.0000	20.0000	7.6068	-0.6818	5.9514
TSSinf	287.3830	280.0000	720.0000	70.0000	118.7209	0.8008	3.5792
T-Ninf	85.5270	85.0000	121.0000	50.0000	10.2677	-0.0325	3.7428
T-Pinf	11.1183	11.0000	19.0000	7.0000	1.4019	1.5312	9.0087
SSinf	11.0393	10.0000	30.0000	5.0000	4.1955	1.4789	5.5211
Condinf	3.2546	3.4000	4.0000	1.4000	0.4419	-1.4737	4.6842
pHeff	8.1091	8.1000	8.5000	6.1000	0.2652	-3.5165	26.9442
CODeff	21.9460	21.0000	55.0000	2.0000	3.9337	2.1321	18.8251
BODeff	1.2634	0.0000	27.0000	0.0000	2.7442	5.4596	45.2100
T-Neff	8.1978	8.0000	75.0000	0.0000	3.9007	12.9723	222.5533
T-SSeff	1.2482	1.0000	9.6000	0.1000	1.5692	2.4431	10.3038
T-Peff	0.7039	0.6300	2.2400	0.0700	0.4604	0.8535	3.3991
NH4-Neff	0.6874	0.3300	8.7100	0.0600	1.0726	4.4407	26.8939
Condeff	2.8918	3.0000	3.7000	0.4000	0.5649	-2.5871	10.4000

3.5 Parameters Description

3.5.1 Influent and Effluents concentration of pH at Nicosia WWTP

pH is a measure of the acidic and alkaline condition of a water body that affects its productivity. It is considered to be of great practical importance as it influences most of the chemical and biochemical reactions. The test for the pH value of wastewater (WW) is carried out to determine whether it is acidic or alkaline in nature. Fresh WW is generally alkaline in nature (its pH value between 7.3 to 7.5). However, as time passes, pH value tends to fall due to the production of acids by bacterial action and the WW tends to become acidic. After oxidation, the pH will relatively become alkaline again. Properly oxidize effluent should have a pH value of about 7.3

or so. The determination of the pH value of WW is important since certain treatment methods depend on the proper pH value of WW for their efficient workings. High or low pH values in a WW have been reported to affect its biota, impede recreational uses of water and alter the toxicity of other pollutants in one form or the other.

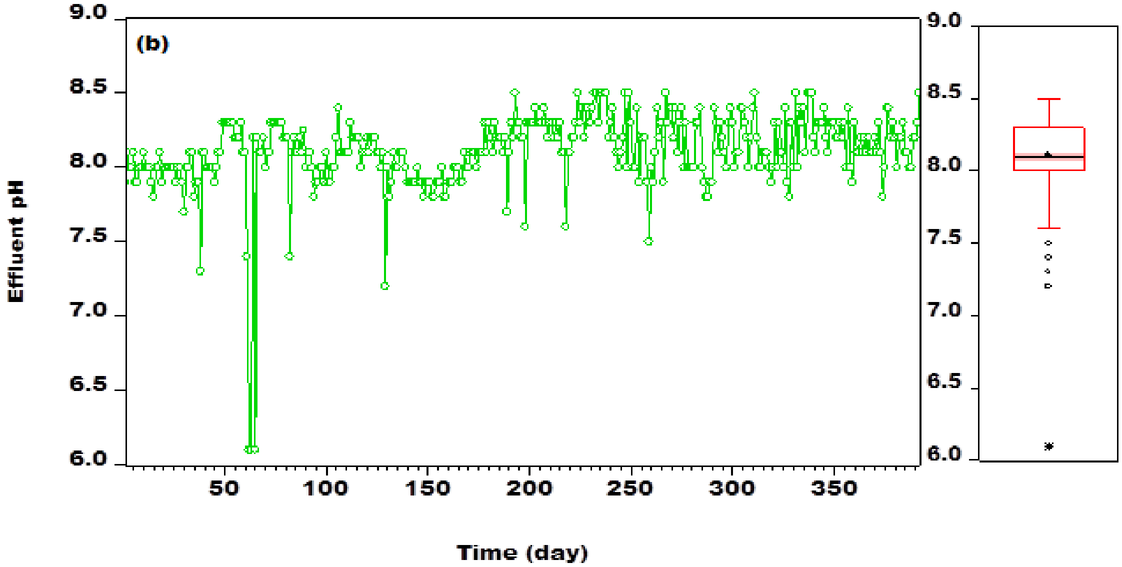


Figure 3.3: Time-series and Box-plot of pH concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.2 Influent and Effluents concentration of COD at Nicosia WWTP

Chemical oxygen demand (COD) is one of the most important parameters of water quality assessment employed for estimating the organic pollution of water. The COD is widely used as a measure of the susceptibility to oxidation of the organic and inorganic materials present in the water bodies. COD has been utilized to determine the content of organic matter of bathwater, wastewater and natural water, due to the time consuming of biological oxygen demand (BOD) test, COD became an alternative in controlling the treatment process. For the oxidation of both organic and inorganic matter, COD may be expressed as one of the demand parameters.

However, COD results are generally higher than BOD values since the test will oxidize material such as fats and lignin, which are only slowly biodegradable. No clear correlation exists between

BOD and COD in general, but at specific treatment plants, a correlation is possible. When once correlation has been established, the COD measurement, which can be concluded within 3 hours, can be used to good advantage for the control and operation of those treatment plants. For the typical untreated domestic WW, the ratio of COD/BOD is found to vary from 1.25 to 2.5. A higher value of the ratio indicates that WW is difficult to biodegrade. For non-biodegradable WW, the ratio exceeds 10. The limiting value of COD of WW generally specified by the authorities is 250mg/L.

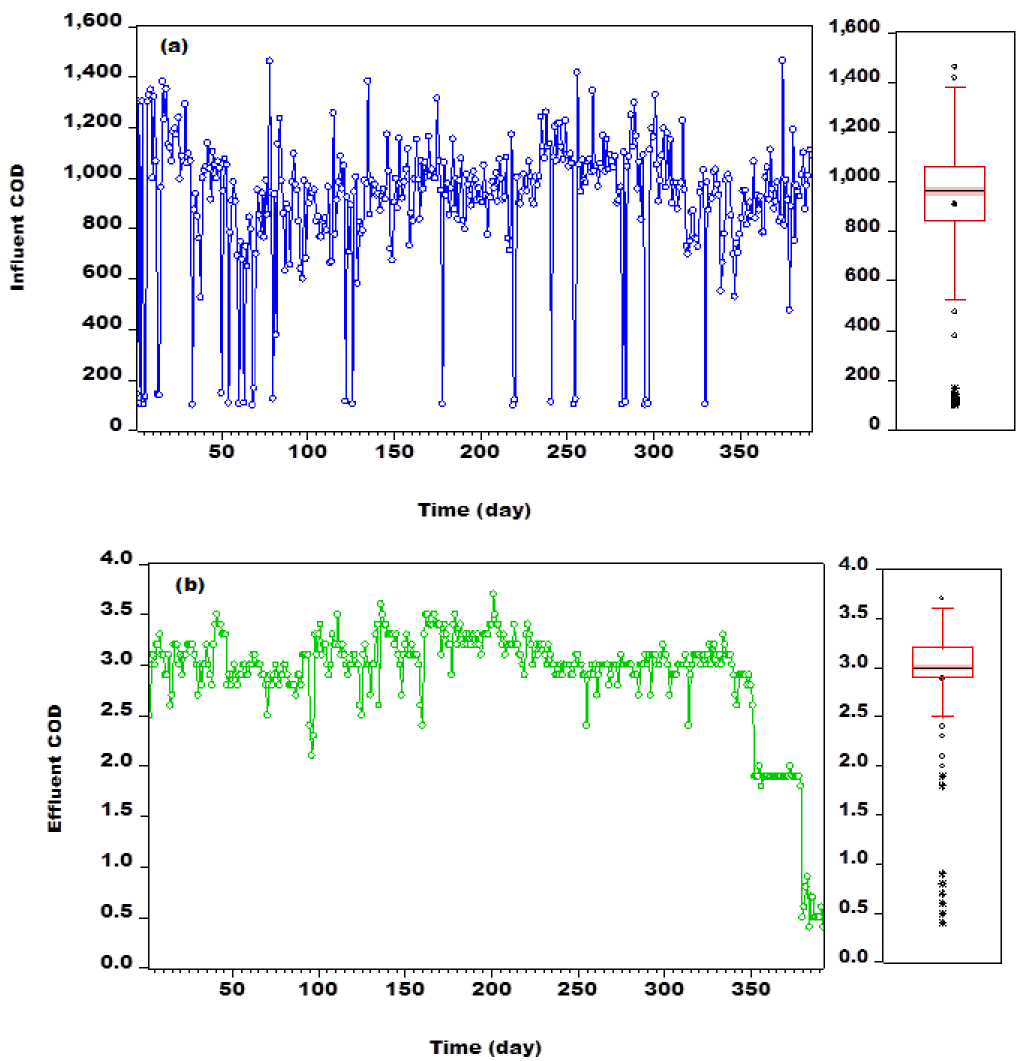


Figure 3.4: Time-series and Box-plot of COD concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.3 Influent and Effluents concentration of BOD at Nicosia WWTP

The biochemical oxygen demand (BOD) is an approximate measure of the amount of oxygen required by the aerobic micro-organisms to stabilize the biochemically degradable organic matter to a stable inorganic form present in any water sample, WW or treated effluents, therefore, it is taken as an approximate measure of the amount of biochemically degradable organic matter present in the aquatic systems, which adversely affects the water quality and biodiversity, the greater the decomposable organic matter present, the greater the oxygen demand and greater the BOD.

In another word, BOD is a measure of the oxygen demand required to oxidize the organic matter present in a sample through the action of micro-organism contained in a sample of WW. It is the widest parameter of pollution applied to the WW as well as surface water. The BOD test results are used for the following purposes

I. Determination of the approximate quantity of oxygen required for the biological stabilization of organic matter present in WW
II. Determination of the size of WW treatment facilities
III. Measurement of efficiency of some treatment process
IV. Determination of strength of sewage
V. Determination of the amount of clear water required for the efficient disposal of WW by dilution

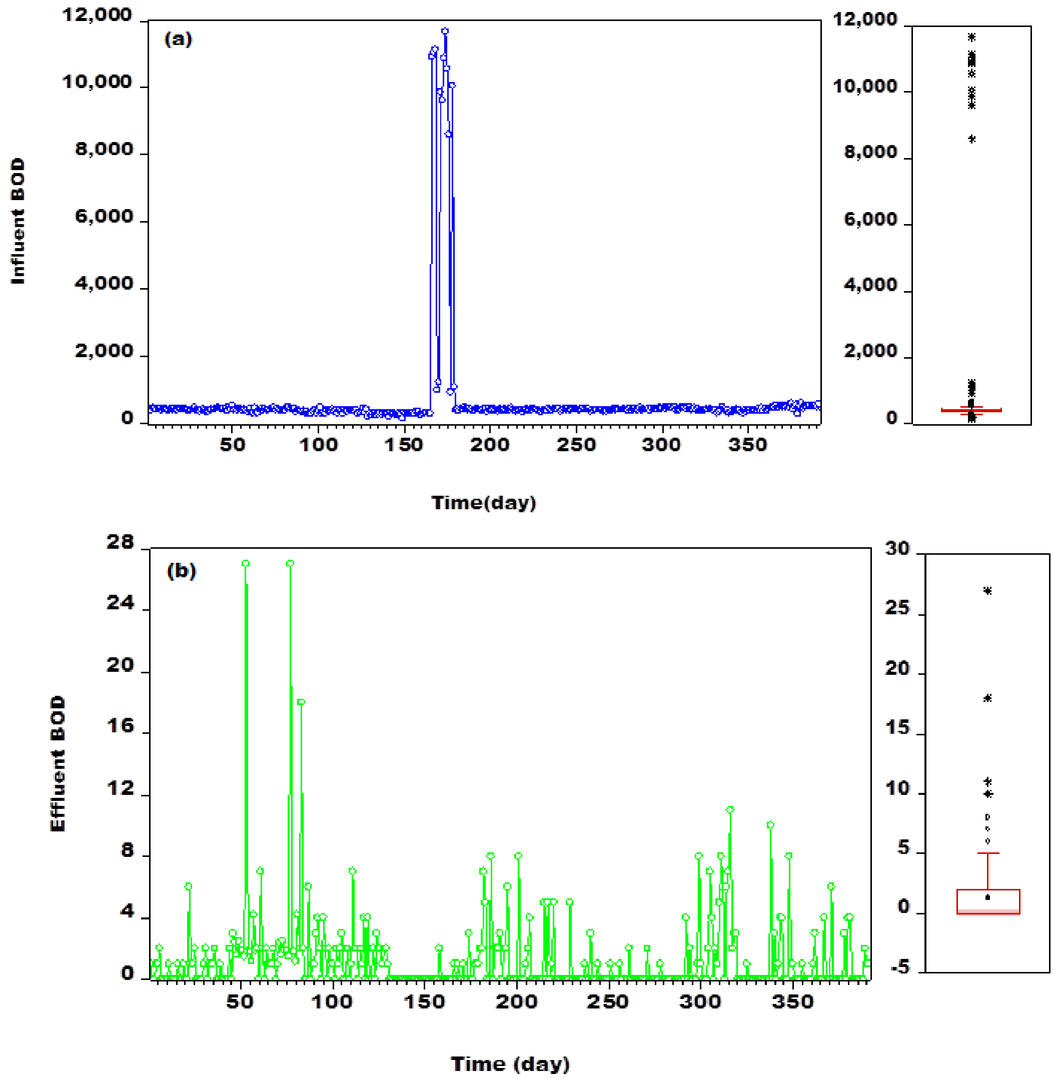


Figure 3.5: Time-series and Box-plot of BOD concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.4 Influent and Effluents concentration of NH4-N at Nicosia WWTP

Ammonia nitrogen (NH4-N) is an important parameter for water quality assessment; generally, the presence of nitrogen in WW indicates the presence of organic matter in it. Nitrogen is essential to the growth of Protista and plants, and such is known as nutrient or biostimulant. Since nitrogen is an essential building element in the synthesis of protein, nitrogen data is required to evaluate the treatability of WW by biological process. Nitrogen appears in the following five different forms in WW.

I. Ammonia nitrogen or free ammonia
II. Organic nitrogen
III. Albuminoid nitrogen
IV. Nitrites nitrogen and
V. Nitrates nitrogen

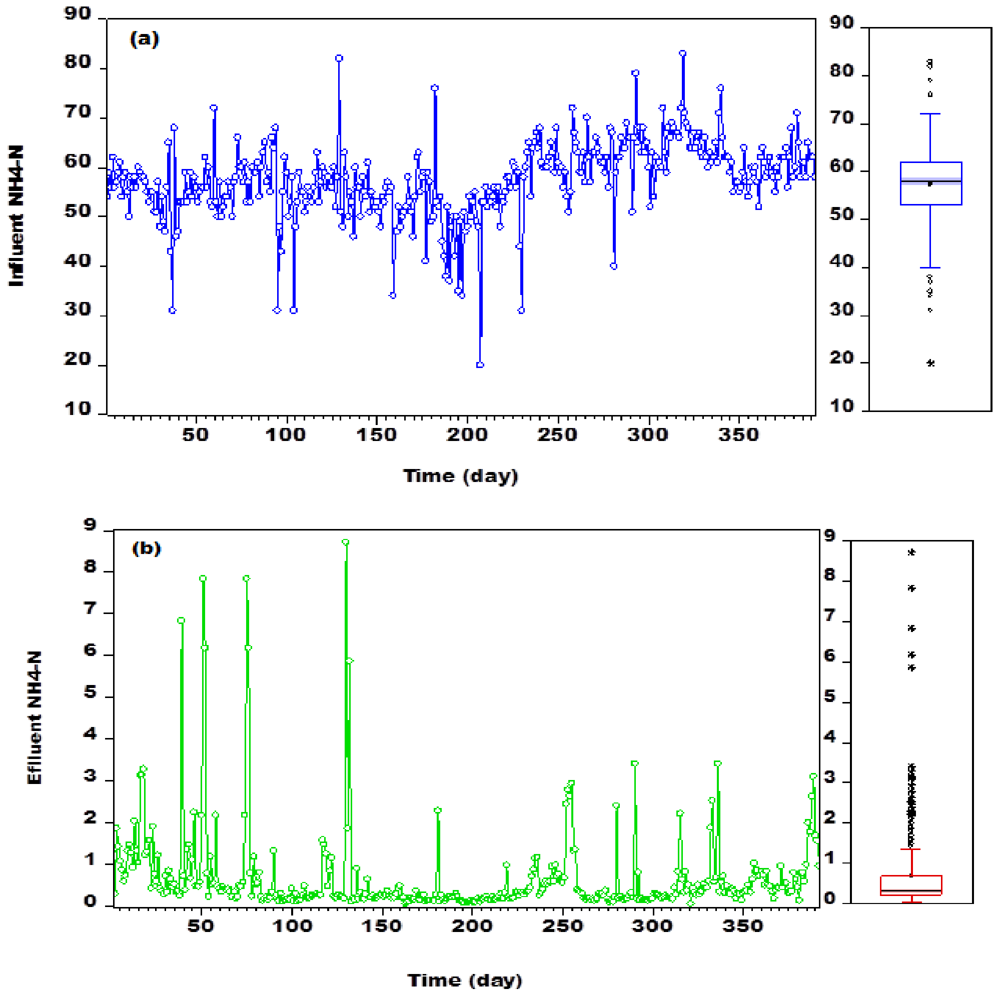
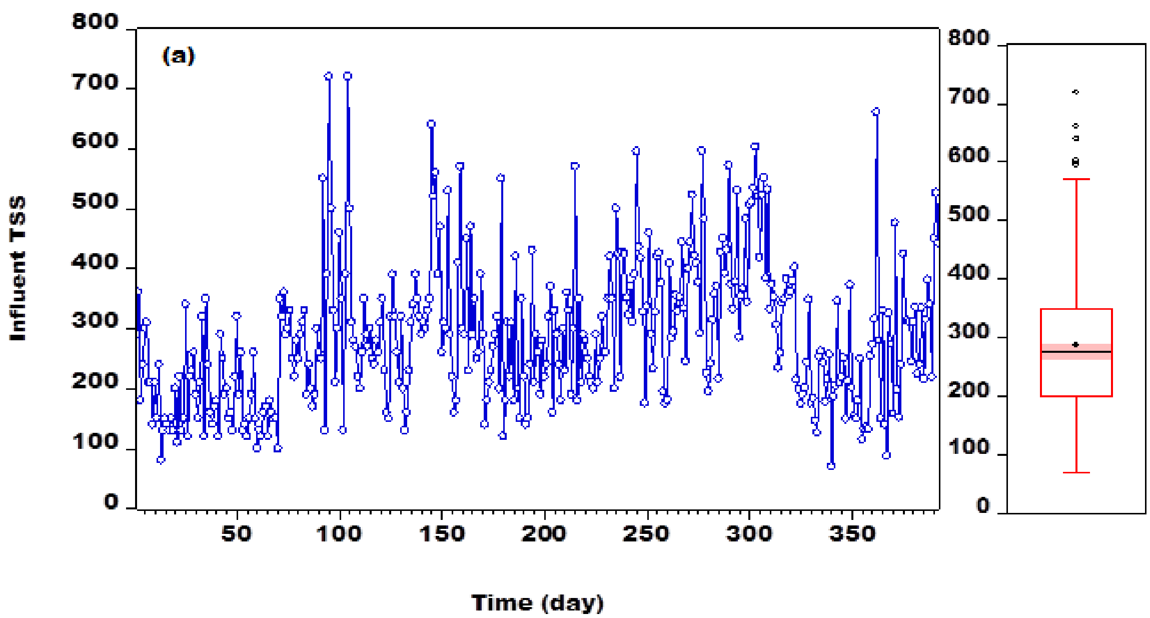


Figure 3.6: Time-series and Box-plot of NH4-N concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.5 Influent and Effluents concentration of TSS at Nicosia WWTP

Total suspended solids (TSS) are particles that are larger than 2 microns found in the water column. Total suspended solids (TSS) are considered to be one of the major pollutants that contribute to the deterioration of water quality, contributing to higher costs for water treatment, decreases in fish resources, and the general aesthetics of the water. The activities associated with wastewater treatment include control of water quality, protection of the shoreline, and identification of economic life of protective structures. Predicting suspended sediments is important in controlling the quality of wastewater. TSS is an important parameter because excess TSS depletes the dissolved oxygen (DO) in the effluent water. Thus, it is imperative to know the values of influent TSS at future time horizons in order to maintain the desired characteristics of the effluent.



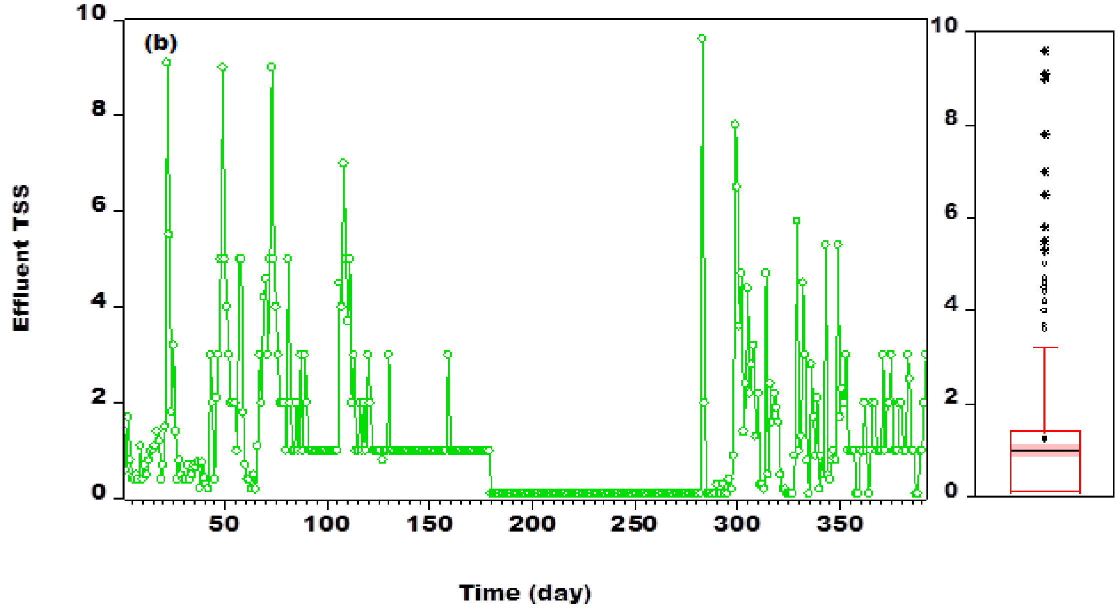


Figure 3.7: Time-series and Box-plot of TSS concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.6 Influent and Effluents concentration of TN at Nicosia WWTP

Total Nitrogen (TN) is an essential nutrient for plants and animals. However, an excess amount of nitrogen in a waterway may lead to low levels of dissolved oxygen and negatively alter various plant life and organisms. Sources of nitrogen include wastewater treatment plants, runoff from fertilized lawns and croplands, failing septic systems, runoff from animal manure and storage areas, and industrial discharges that contain corrosion inhibitors. Total Nitrogen (TN) is the sum of nitrate-nitrogen (NO3-N), nitrite-nitrogen (NO2-N), ammonia-nitrogen (NH3-N) and organically bonded nitrogen. Total Nitrogen (TN) should not be confused with TKN (Total Kjeldahl Nitrogen) which is the sum of ammonia-nitrogen plus organically bound nitrogen but does not include nitrate-nitrogen or nitrite-nitrogen**.**

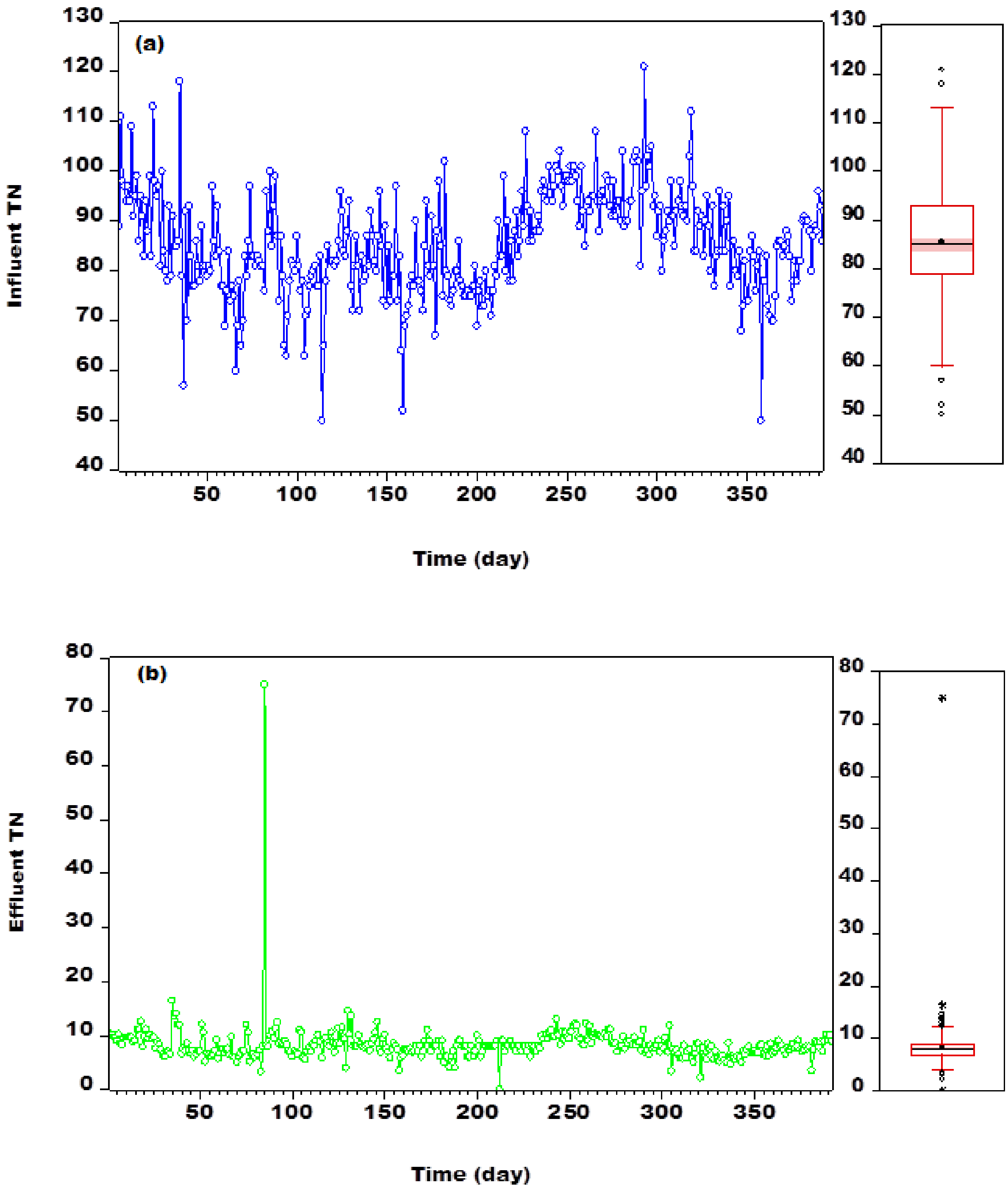
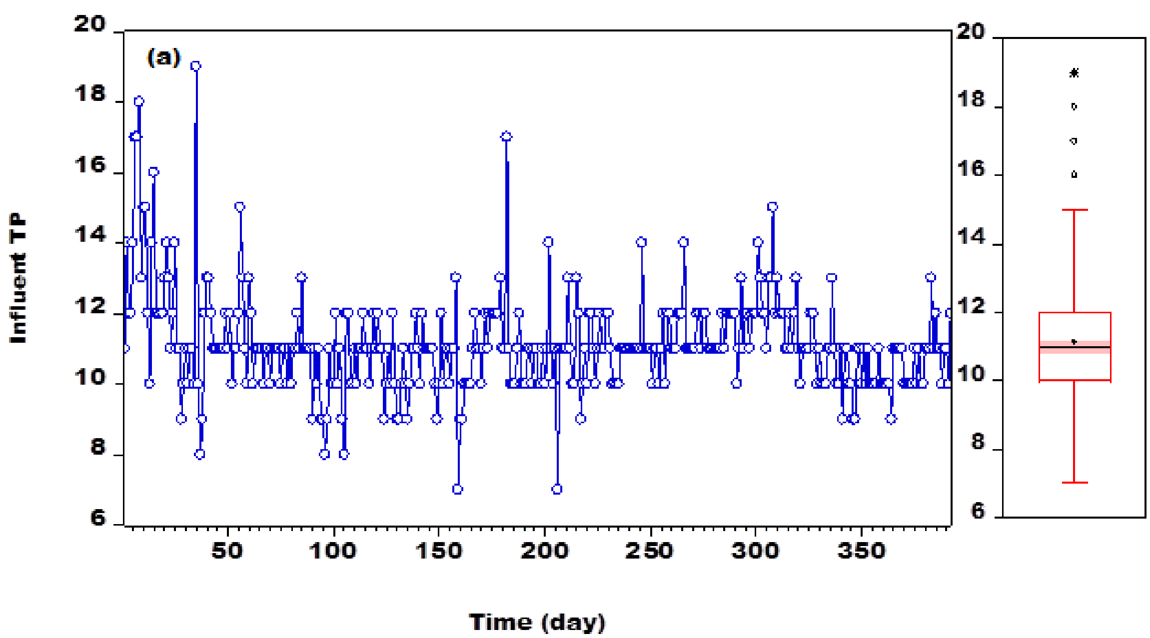


Figure 3.8: Time-series and Box-plot of TN concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.7 Influent and Effluents concentration of TP at Nicosia WWTP

Total Phosphorus is an essential nutrient for plants and animals. It is naturally limited in most freshwater systems because it is not as abundant as carbon and nitrogen; introducing a small amount of additional phosphorus into a waterway can have adverse effects. Sources of phosphorus include soil and rocks, wastewater treatment plants, runoff from fertilized lawns and cropland, runoff from animal manure storage areas, disturbed land areas, drained wetlands, water treatment, decomposition of organic matter, and commercial cleaning preparations.

Wastewater is relatively rich in phosphorus compounds. Phosphorus is a nutrient used by organisms for growth. It occurs in natural water and wastewater bound to oxygen to form phosphates. Phosphates come from a variety of sources including agricultural fertilizers, domestic wastewater, detergents, industrial process wastes, and geological formations. The discharge of wastewater containing phosphorus may cause algae growth in quantities sufficient to cause taste and odour problems in drinking water supplies. Dead and decaying algae can cause oxygen depletion problems which in turn can kill fish and other aquatic organisms in streams. For this reason, phosphorus removal is an essential role in wastewater treatment plants and testing for phosphorus in the plant effluent is critical.



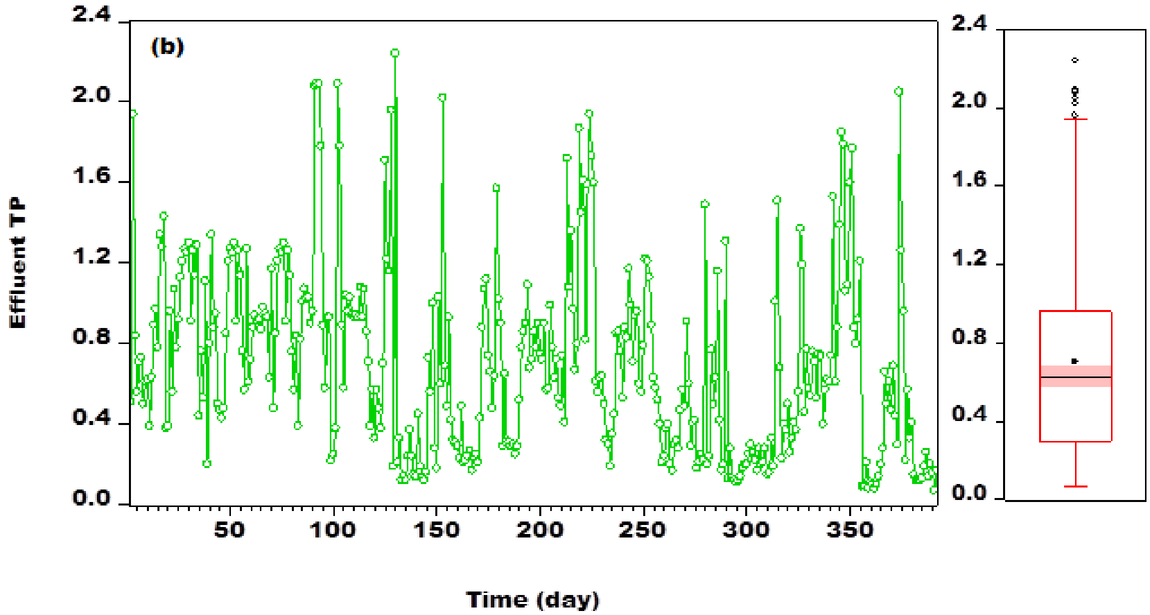


Figure 3.9: Time-series and Box-plot of TP concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

3.5.8 Influent and Effluents concentration of Conductivity at Nicosia WWTP

Electrical Conductivity is also known as specific conductance. It is defined as a measure of the ability of a water sample to convey an electric current. The electrical conductivity of industrial wastewaters, treatment plant effluents, and polluted water is due to the presence of ionic solutes. Electrical conductivity is a rapid and reasonably precise determination and values are always expressed at a standard temperature of 25^0C. The unit of Electrical Conductivity is µS/cm. Electrical conductivity measurements are often used to determine the salinity of natural and wastewaters.

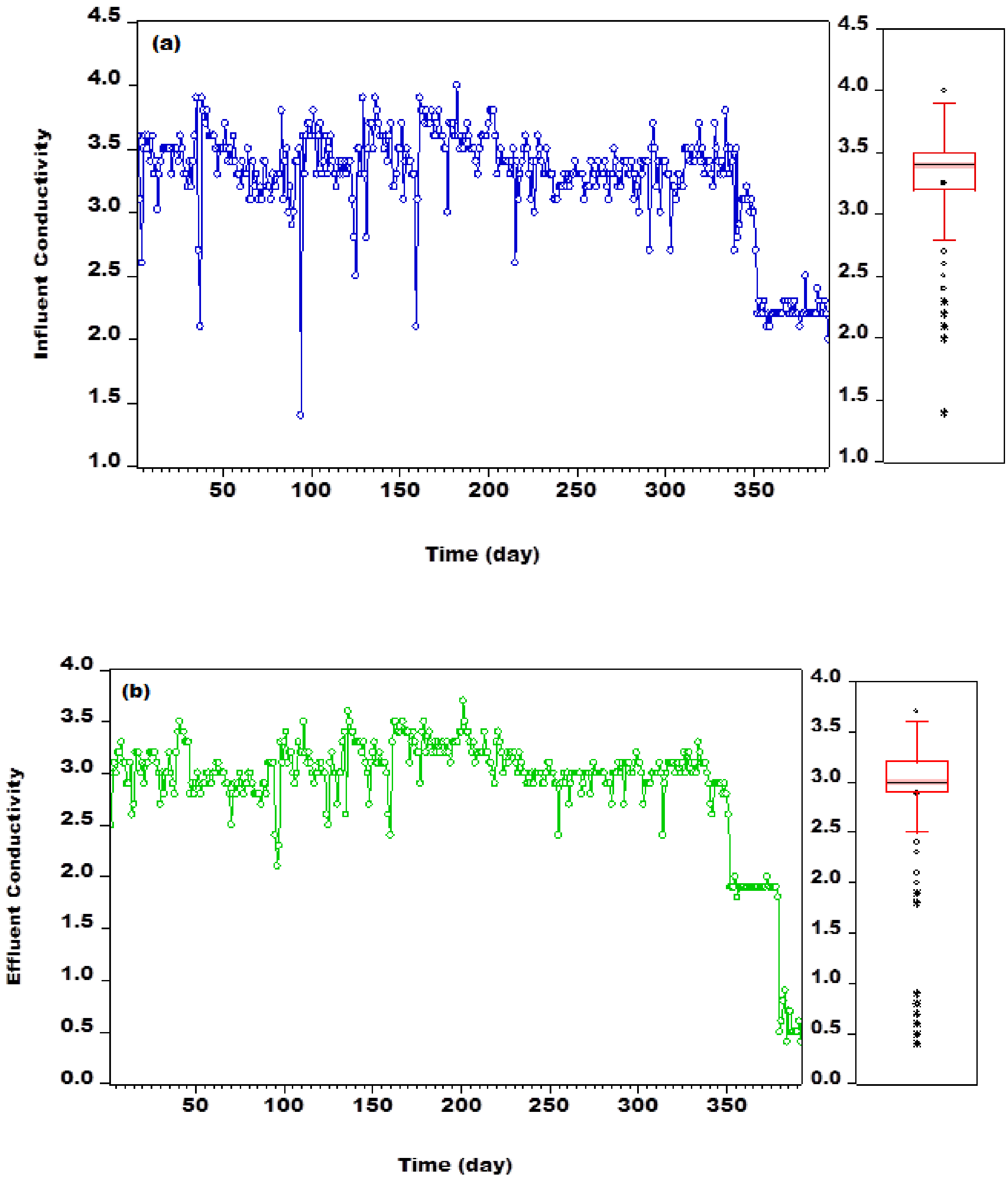


Figure 3.10: Time-series and Box-plot of Conductivity concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent)

3.5.9 Influent and Effluents concentration of SS at Nicosia WWTP

Sewage normally contained 99.9% of water and 0.1 % of solids. Measuring suspended solids (SS) in water is used for control of various treatment processes and for the examination of wastewater quality. The level of suspended solids (or total suspended solids) in water and wastewater affect the quality of the water and how it can be used. Total solids in WW exist in three different forms:

I. **Suspended Solids**
II. **Colloidal Solids**
III. **Dissolved Solids**

SS may be further subdivided into (a) Settleable solid and (b) Non-settleable solids.

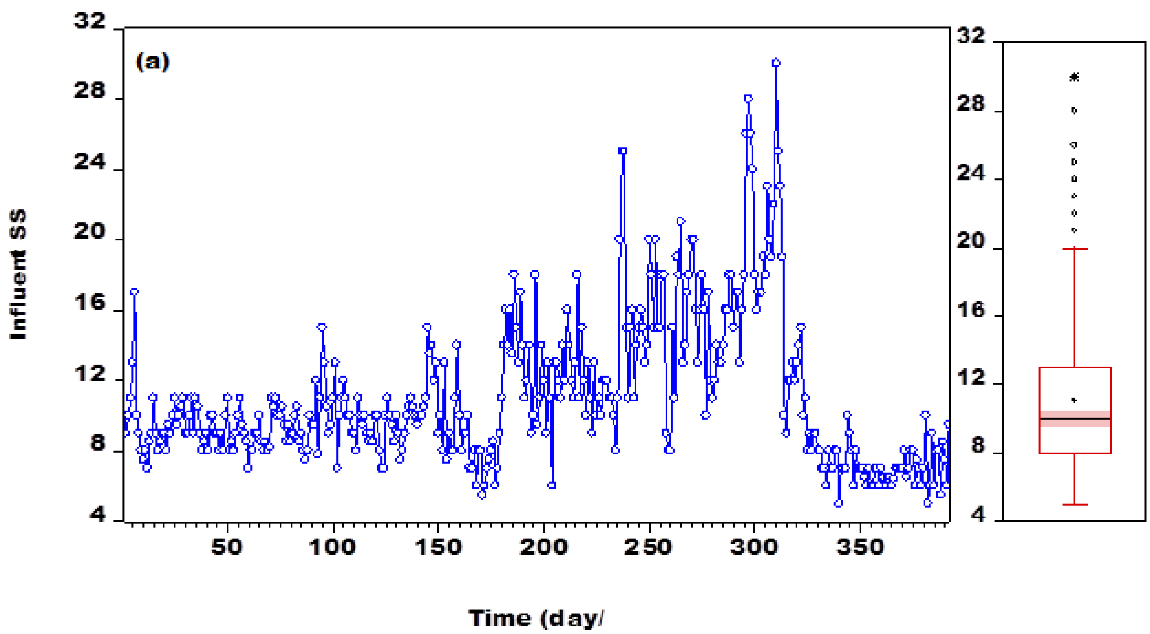


Figure 3.11: Time-series and Box-plot of SS concentrations of the plant at the (a) inlet (influent)

3.6 Proposed Methodology

In this thesis, three different scenarios were proposed separately for modelling and prediction the performance of new Nicosia WWTP. The first scenario I, explored the application of data-driven algorithms (i) to develop and compare the potential of some AI-based models (FFNN, SVM, and ANFIS) and conventional multi-linear model (MLR) for prediction of the Nicosia WWTP performance considering four different combinations of input parameters. Other feasible alternatives models may also be used, but they were adopted here due to their outstanding performances in various literature in hydro-environmental studies. Some alternatives are Genetic programing, ARIMA models (Olyaie et al., 2017). (ii) to establish and apply three ensemble techniques using the outputs of the aforementioned single models in order to improve the overall efficiency of the prediction performance. In this way, simple linear averaging, weighted linear averaging and non-linear neural ensemble techniques are proposed to combine the outputs of the methods. The flowchart of the scenario I can show in Figureure 3.11. The second scenario II, employed two different approach (i) The scenario I was aimed to develop the potential of ELM with PCA to predict the performance of new Nicosia MWWTP based on multi-parametric effluents modelling of BOD_{eff}, COD_{eff}, Total Nitrogen (TN_{eff}) and Total Phosphorite (TP_{eff}), the advantage of introducing PCA is for choosing the proper inputs of the models and to understand whether it is feasible to enhance the prediction accuracy of the ELM model. (ii) In approach II, the traditional multilayer perceptron (MLP) neural network and multiple linear regression (MLR) models were established for comparison using the same input combinations of scenario I. For the development of the model for the scenario II, Figure. 3.12, show the flowchart of the used model. And lastly, the third scenario III implemented the ability of data-intelligence approaches such as AR and NSI models (i.e. (Hammerstein- Weiner Model (HW) and Nonlinear autoregressive with exogenous (NARX) neural network model) to estimates the performance in terms of TSS and pH effluents in New Nicosia WWTP. Figureure 3.13 presented the flowchart of scenario III.

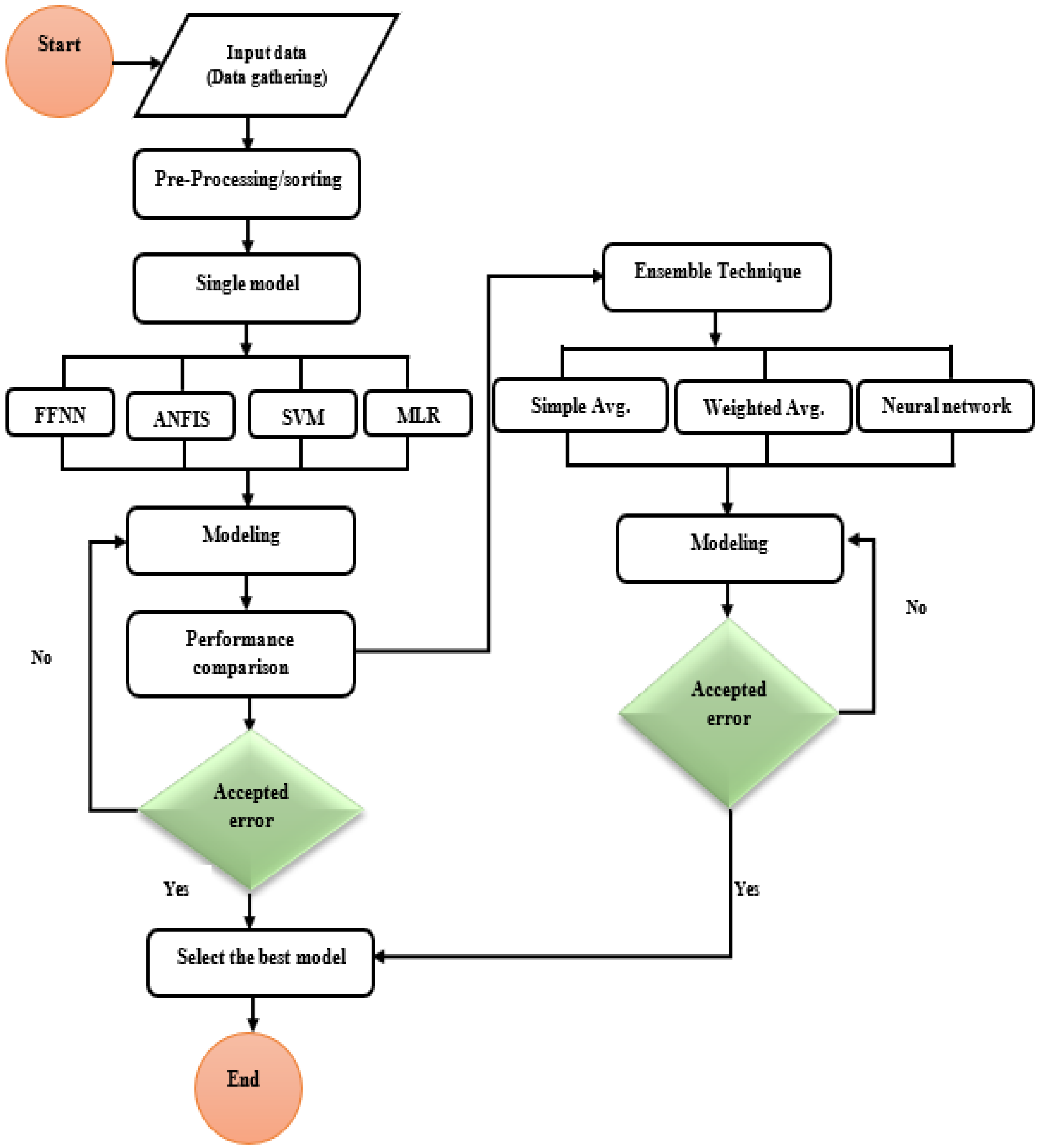


Figure 3.12: Flow chart of the proposed scenario I

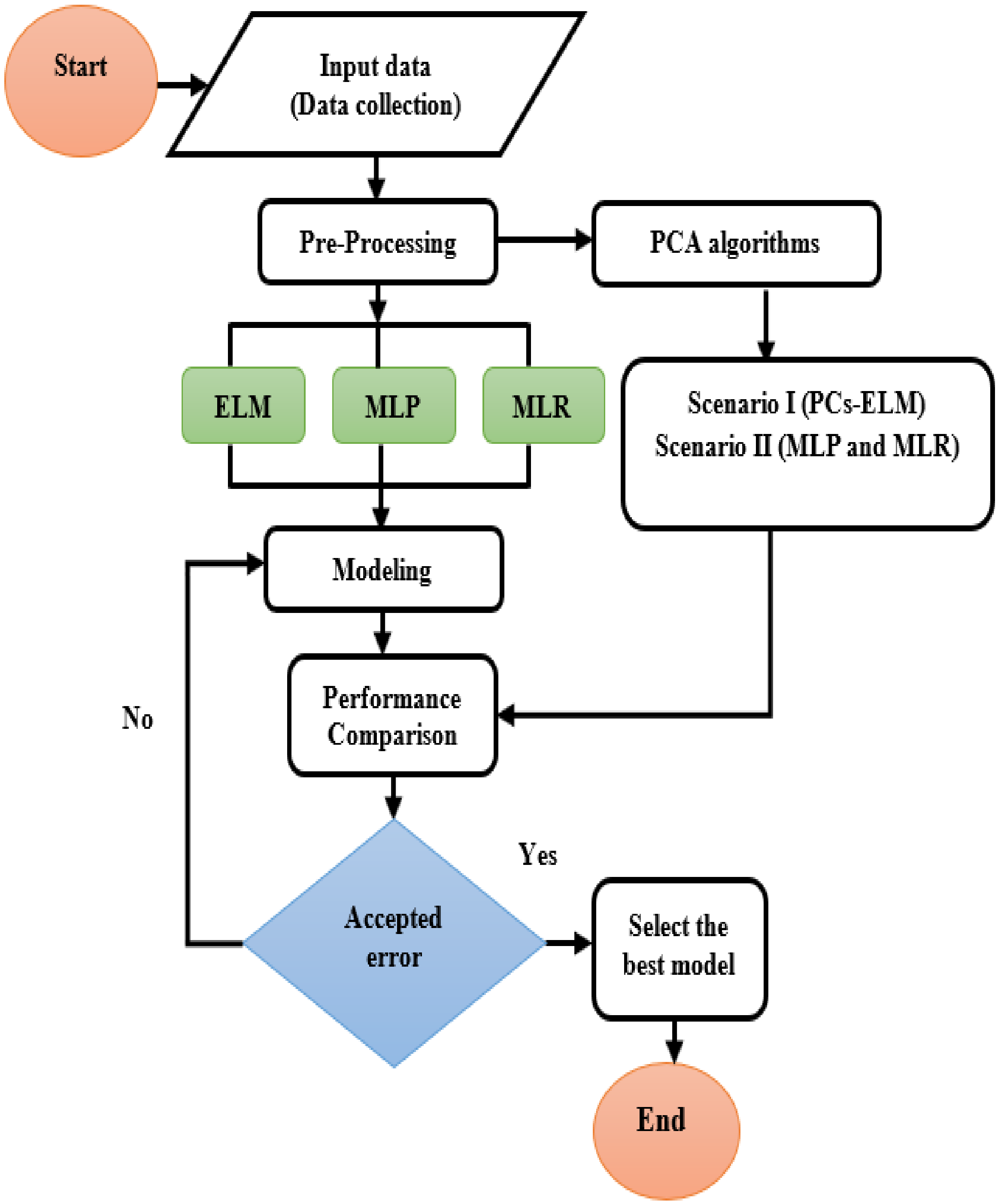


Figure 3.13: Flow chart of the proposed scenario II

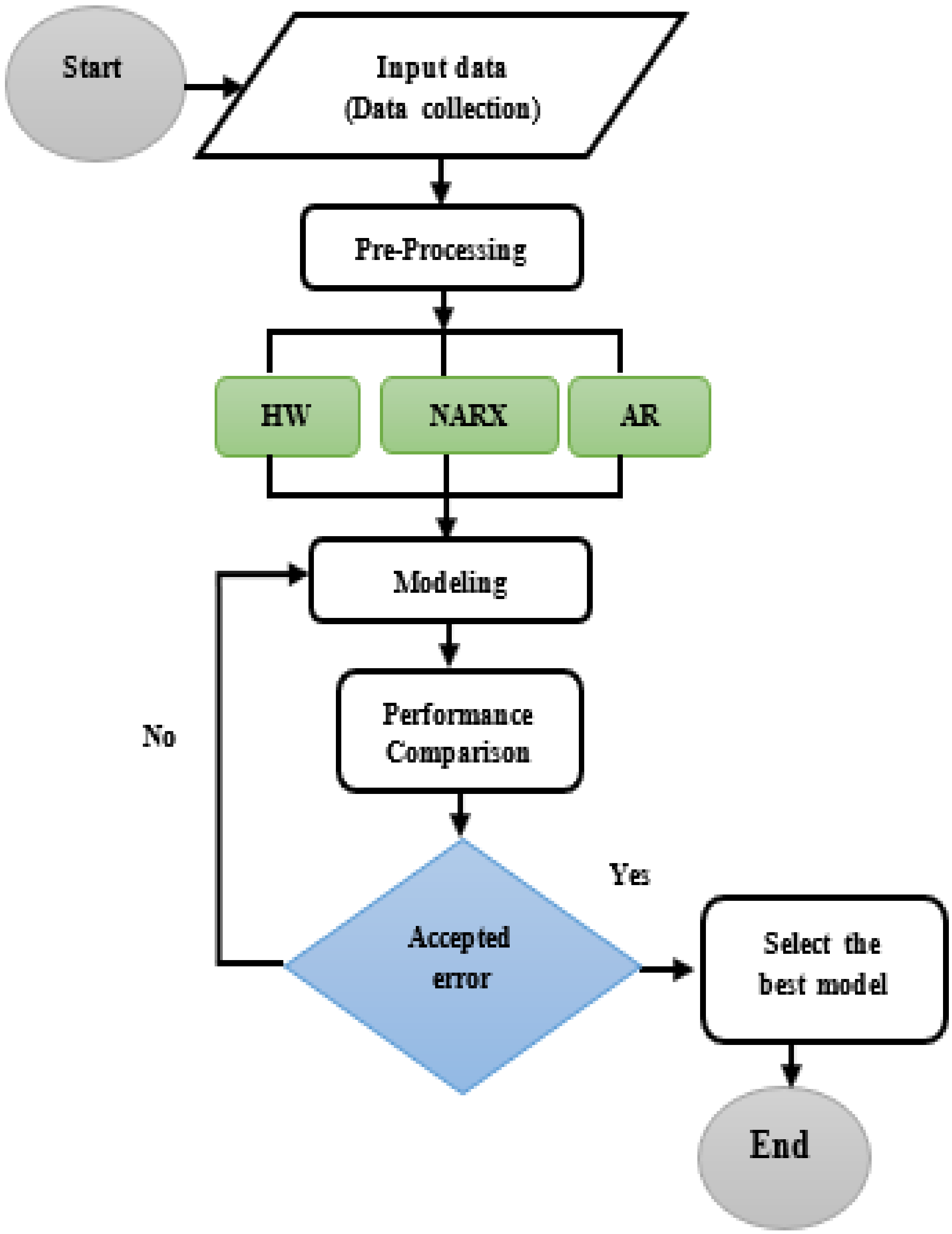


Figure 3.14: Flow chart of the proposed scenario III

3.7 Reasons for combining linear and non-linear models

It is difficult to determine in practice whether one particular model is better than others. Thus, selecting the proper method for a particular case is a difficult task for the predictors. The complexity of selecting the appropriate models can be resolved by choosing to assemble various models. The traditional linear models are still used despite the inability to provide accurate outputs due to their various limitations and inconsistencies to handle non-stationary and non-linearity data. Such linear models are still applicable because, a) traditional linear models are economical, uncomplicated, and the natural phenomenon can be employed in a functional linear system, b) non-linear models magnify the noise for additional time steps while the linear models increase the noise included in the data linearly. Therefore, applying the traditional linear model for linear portions of the process is recommended. The natural and real-world processes may contain both the linear and non-linear characteristic. As such, ARIMA, AR and MLR models are not capable of handling non-linear system solely. On the other hand, an AI model may expand the noise of the linear pattern, and therefore both of the models cannot adequately estimate the time series of the process individually. Hence, by combining results of the traditional model and AI models, the magnify non-linear behaviors of the noise, and complex architecture can be addressed in a simple approach.

3.8 Input variable and models selection

For this thesis, the input variables and model selection can be categorized based on the three different scenarios.

For the scenario I, the performance of NWWTP based on the daily measured data set comprising of influent of pH_{inf}, Conductivity ($Cond_{inf}$), BOD_{inf}, COD_{inf}, Total Nitrogen (TN_{inf}) as the inputs of models, and three effluent values like BOD_{eff}, COD_{eff}, and TN_{eff} as targets. In the descriptive statistical analysis and correlation coefficient (R) as the most commonly used techniques were calculated to measure the strength and degree of the linear relationship between two variables, which can be served as the preliminary indication of probable linear correlation between a set of variables. However, the weakness of the computed R values depicts that, the application of conventional linear techniques to handle such non-linear complex interactions cannot be recommended and there is a great need to introduce more robust nonlinear tools in input variable selection. Therefore, instead of some other studies which used linear correlation coefficient between input and output parameters to select the dominant inputs of non-linear, different combinations of input parameters are examined through the used methods (FFNN, ANFIS, SVM, and MLR) in this scenario.

As presented in Table 3.3, the R-value with bold marked is stationary and significant variables with probability less than 0.05 ($P<0.05$) that indicates high strength of linear correlations. Also, the negatives R-values indicate an inverse relationship between two variables. Hence, the weakness of R-value depicted that, the application of conventional techniques in modelling such complex interactions is insignificant, and there is a great need for introducing more robustness tools. The input selection and combination of FFNN, ANFIS, SVR, and MLR are presented in Table 3.4 base on sensitivity analysis.

Table 3.3: Pearson Correlation matrix between the effluent and influent quality

Parameters	pH_{inf}	$Cond_{inf}$	BOD_{inf}	COD_{inf}	$T\text{-}N_{inf}$
BOD_{eff}	**0.0831**	**0.0673**	**0.0872**	**0.0563**	0.0128
COD_{eff}	-0.1391	**0.2398**	-0.0125	-0.1129	**0.1652**
$T\text{-}N_{eff}$	-0.0051	0.0120	-0.0109	-0.0230	**0.1626**

Table 3.4: Selected input variables in FFNN, ANFIS, SVR and MLR

Model Input Variables	Model Structure	Model Output Variable
pH_{inf}, $Cond_{inf}$	FFNN, ANFIS, SVR, MLR (2-1)	
pH_{inf}, $Cond_{inf}$, BOD_{inf}	FFNN, ANFIS, SVR, MLR (3-1)	
pH_{inf}, $Cond_{inf}$, BOD_{inf}, COD_{inf}	FFNN, ANFIS, SVR, MLR (4-1)	BOD_{eff}
pH_{inf}, $Cond_{inf}$, BOD_{inf}, COD_{inf}, $T\text{-}N_{inf}$	FFNN, ANFIS, SVR, MLR (5-1)	
pH_{inf}, $Cond_{inf}$	FFNN, ANFIS, SVR, MLR (2-1)	
pH_{inf}, $Cond_{inf}$, BOD_{inf}	FFNN, ANFIS, SVR, MLR (3-1)	
pH_{inf}, $Cond_{inf}$, BOD_{inf}, COD_{inf}	FFNN, ANFIS, SVR, MLR (4-1)	COD_{eff}
pH_{inf}, $Cond_{inf}$, BOD_{inf}, COD_{inf}, $T\text{-}N_{inf}$	FFNN, ANFIS, SVR, MLR (5-1)	
pH_{inf}, $Cond_{inf}$	FFNN, ANFIS, SVR, MLR (2-1)	
pH_{inf}, $Cond_{inf}$, BOD_{inf}	FFNN, ANFIS, SVR, MLR (3-1)	
pH_{inf}, $Cond_{inf}$, BOD_{inf}, COD_{inf}	FFNN, ANFIS, SVR, MLR (4-1)	$T\text{-}N_{eff}$
pH_{inf}, $Cond_{inf}$, BOD_{inf}, COD_{inf}, $T\text{-}N_{inf}$	FFNN, ANFIS, SVR, MLR (5-1)	

For scenario II, the daily measured data obtained for analyzing the performance of Nicosia MWWTP which includes (pH_{inf}, $Conductivity_{inf}$, BOD_{inf}, COD_{inf}, $Total\text{-}N_{inf}$, $Total\text{-}P_{inf}$, $NH4\text{-}N_{inf}$, SS_{inf}, and TSS_{inf}) as the input variables and (BOD_{eff}, COD_{eff}, $Total\text{-}N_{eff}$, $Total\text{-}P_{eff}$) as the corresponding output respectively. For the purpose of scenario II, two different approaches were employed for modelling the performance of Nicosia MWWTP, as mentioned above. In approaches I, data is directly imposed into the ELM model, and modelling are carried out using all the inputs variables. If the attained error is acceptable based on performance criteria, then the best models are selected if they are not acceptable, then the modelling is repeated by adjusting the model parameters. In the second stage of the flowchart, PCA algorithms were employed for proper selection of inputs variable and to improve the ELM model by using the new principal components (PCs) variables as the new input variables of ELM models (see, equation 3.2).

Finally, the procedure is repeated for the selected PCs as that of ELM models. For approaches II, two most commonly used linear and non-linear models (i.e. MLR and MLP) were also introduced for comparison with the novel ELM using the same input combination and PCs variables.

$$\left.\begin{aligned}\boldsymbol{4PCs-ELM} &= pH_{inf}, BOD_{inf}, COD_{inf}, TN_{inf}\\ \boldsymbol{6PCs-ELM} &= pH_{inf}, BOD_{inf}, COD_{inf}, TN_{inf}, TP_{inf}, TSS_{inf}\\ \boldsymbol{All(9\ inputs)} &= pH_{inf}, BOD_{inf}, COD_{inf}, TN_{inf}, TP_{inf}, TSS_{inf}, SS_{inf}, NH4N_{inf}, Cond_{inf}\end{aligned}\right\}$$

(3.2)

For scenario III, Different methods were reported for input selection, such as (i) Pearson and Spearman correlation analysis to determine the strength and relations between inputs and outputs (see, Table 3.5) (ii) auto-correlation function (ACF) and the partial autocorrelation function (PACF) (Yaseen et al., 2016). Subsequently, a set of two different models were derived on the basis of significant input variables as in Table 3.6 and the value of PACF.

For any time-series modelling identifying the proper time lags is an essential part of selecting the appropriate model inputs combinations, as such autocorrelation function (ACF) and partial ACF (PACF) are used (see, Figure. 4). In a time-series, autocorrelation is considered as the correlation between the time series, previous and forthcoming data points. For instance, for a time X, the correlation (R) of the first lag (lag 1) is considered as the R between X_t and X_{t-1}; for the second lag (lag 2) is considered as R between X_t and X_{t-2}. On the other hand, the partial correlation is R-value of a parameter with its own lag that is yet to described by the R of the lower legs. (Hadi and Tombul, 2018).

Spearman Pearson correlation describes how well the relationship between the variables can be described using the linear and monotonic function. The strength of the correlation is not dependent on the direction or sign. A positive coefficient indicates that increase in the first parameter would correspond to an increase in the second parameter while the negative correlation indicates an inverse relationship whereas one parameter increases and the second parameter decreases (Eisinga et al., 2013). It can be seen from Table 3.5 that, after performing correlation analysis (R) for selecting the initial input variables, the significant R was observed between the variables.

Table 3.5: Pearson and Spearman correlation analysis of the parameters

Parameters	**pHinf**	**CODinf**	**TNinf**	**NH4-Ninf**	**SSinf**	**TSSinf**	**pHeff**	**TSSeff**
pHinf	1							
CODinf	-0.1599	1						
TNinf	0.0527	0.13532	1					
NH4-Ninf	0.2512	0.10088	0.50372	1				
SSinf	-0.0051	0.10151	0.33462	0.09191	1			
TSSinf	0.0098	0.17259	0.15469	0.04088	0.48520	1		
pHeff	0.0252	0.35638	0.13865	0.14749	0.40214	0.0766	1	
TSSeff	0.1426	-0.58359	-0.590883	0.06649	-0.674290	0.08003	0.05563	1

Table 3.6: The development of input variables models

Model output	**Model**	**Input variables**
	M1 (4)	SS+TN+COD+NH4-N
Effluent pH (pH_{eff})	M2 (6)	SS+TN+COD+NH4-N+ pH+ TSS
	M1 (4)	SS+TN+COD+NH4-N
Effluent TSS (TSS_{eff})	M2 (6)	SS+TN+COD+NH4-N+ pH+ TSS

3.9 Methodology for the scenario I models (FFNN, SVM, ANFIS, MLR, and Ensemble)

3.9.1 Feed Forward Neural Network (FFNN)

An artificial neural network as AI-based model is a mathematical model aims to handle the non-linear relationship of an input-output dataset. Historically, ANN is information processing tools derived from analogy with the biological nervous system of the brain, with the basic component called neuron (node) (Sirhan and Koch, 2015). ANN has proved to be effective with regards to complex function in various field, including prediction, pattern recognition, classification, forecasting, control system and simulation (Govindaraju, 2000, Solgi, et al.,2017). Among the different classifications of ANN algorithms, Feed-Forward Neural Network (FFNN) with Backpropagation (BP) is widely applied and the most common classes.

In FFNN-BP the network is trained with the training input data which process through the system and process passed to the output layer, an error is generated which is propagated back to the network until the desired output is archived (Abdullahi and Elkiran, 2017). The main concept of FFNN-BP is to reduce the error so that the network learns the training data and can predict the correct output 2012, ASCE Task Committee, 2000). FFNN consists of three layers; input, hidden and output layers as in Figure 3.14. In the FFNN process, the initial weights are multiplied by the inputs and the subsequent value moves to the next layer, till it gets to the output layer, shown by following equation 3.3 (Gaya et al.,2014).

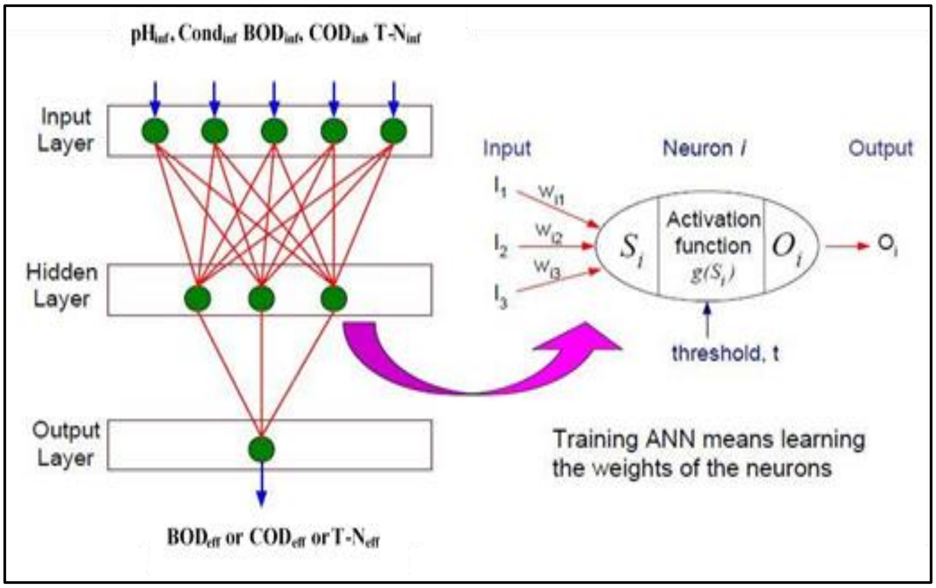


Figure 3.15: Schematic diagram of ANN-FFNN showing Input, Hidden and Output Layers

$$\boldsymbol{z_i = \sum_{j=1}^{m} w_{ij}x_{ij}} \tag{3.3}$$

where $\boldsymbol{w_{ij}}$ represents the weight moved from j^{th} input to the i^{th} node, $\boldsymbol{x_{ij}}$ depicts the input while $\mathbf{z_i}$ denotes the resultant summation of outputs of the i^{th} node. Therefore, the error value is determined through the back-propagation process by calculating the difference between predicted values and target value. It starts back from the output layer to the input layer. The difference is represented by the symbol *δ(l)j*, showing the error of node *j* in layer *l*. The error term for a training set (*xj*, *yj*) is mathematical express in the equation:

$$\boldsymbol{e_p = y_d - y_a} \tag{3.4}$$

where y_d is the desired output of the neuron p and y_a is the actual output of the calibration

However, too many neurons in the hidden layer may affect the generalization ability of the neural network and increase the computational burden, whereas too low neurons may not produce the required prediction accuracy (Gaya et al., 2014).

The continuous process by which the connection weights and biases are adjusted until you get a required output is called the learning the process. The learning process may be either supervised or unsupervised. Supervised learning was used because of its ability to minimize the error between desire and computed values (Hamed et al., 2004). Learning rate define the intelligence of the network, which plays an essential role in the convergence of the network and overcomes the problem of a local minimum. Both the architecture (the number of neurons, the number of layers, transfer function) and learning rate were determined by using the trial-and-error method. The sigmoid activation function is used for input and hidden layers while linear activation function is applied for the output layer, the activation function was introduced in each neuron in order to convert the linear function to none- linear function which is a mathematical function (Yetilmezsoy et al., 2011).

3.9.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

The combination of Artificial neural network with the fuzzy system creates a robust hybrid system that is able to solve the complex nature of the relationship (Akrami et al., 2014). Adaptive Neuro-Fuzzy Inference System (ANFIS) is a Multi-Layer Feed-Forward (MLFF) neural network using the integration of neural network and fuzzy logic algorithms in order to map inputs with outputs (Solgi et al., 2017). ANFIS uses Takagi–Sugeno type Fuzzy Inference System (FIS), where the output of each fuzzy rule can be a linear combination of input variables plus a constant term. Two types of learning algorithms are generally employed in ANFIS as AI-based model, i.e. the backpropagation and hybrid learning. The back-propagation learning is used similar to that of backpropagation in ANN (Yetilmezsoy et al., 2011). The hybrid learning consists of a combination of backpropagation and least squares method (Parmar and Bhardwaj, 2015). ANFIS hybrid learning algorithm was used because it is much faster to converge than the conventional backpropagation method. However, the most utilized membership function in

ANFIS model are triangular (Trimf), trapezoidal (tramf), gaussian (gaussm), bell-shaped (gbellmf) and sigmoidal (sigmf) membership function, for more information on this membership function see Solgi et al.,(2017) (Solgi et al., 2017). The main advantage of ANFIS rule system is basically classified into Mamdani and Takagi-Sugeno-Kang, which are normally expressed into a linguistic variable and mathematical function respectively. The de-fuzzification process is needed in Mamdani rule while there is no need for de-fuzzification in Sugeno process (Takagi and Sugeno, 1993; Mandani, 1974). The general structure of ANFIS is shown in Figure 3.15, despite the fact that the basic concept of an AI-based model is ANN and fuzzy logic, the ANFIS model-derived the merit of both the two methods. Since ANFIS combine the topology of ANN and Fuzzy logic, it covers their methodology and limitations. Therefore, the ANFIS model supplies the optimum desired outcomes quickly with less error and without any uncertainty and vagueness. In addition, in term of learning duration ANFIS model is very short in comparing with the ANN model (Yetilmezsoy et al., 2011).

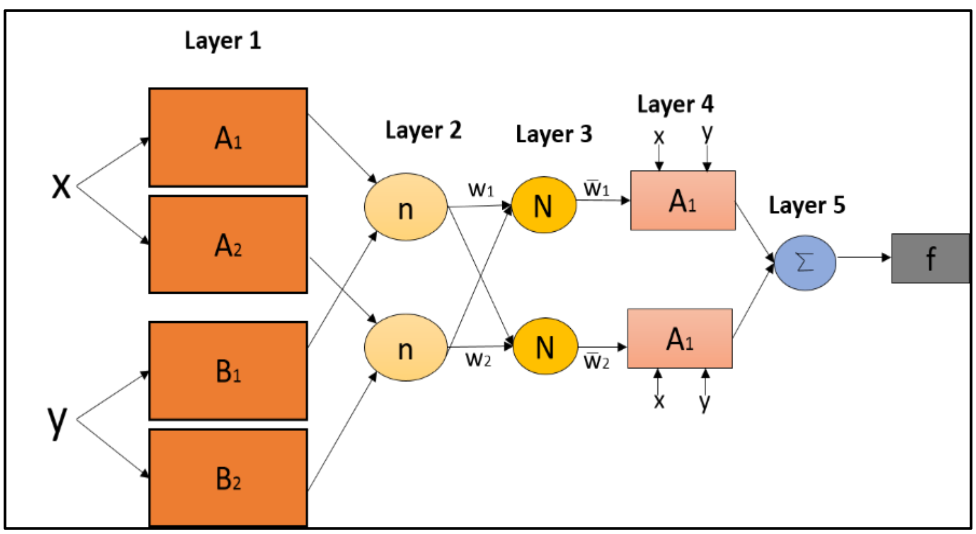


Figure 3.16: Structure of Adaptive Neuro-Fuzzy Inference

Assume the FIS contains two inputs ‘*x*’ and ‘*y*’ and one output ‘*f* ‘a first-order Sugeno fuzzy has following rules.

Rule (1): if $\mu(x)$ is A_1 and $\mu(y)$ is B_1; then $f_1 = p_1x + q_1y + r_1$ (3.5)

Rule (2): if $\mu(x)$ is A_2 and $\mu(y)$ is B_2; then $f_2 = p_2x + q_2y + r_2$ (3.6)

A_1, B_1, A_2, B_2 parameters are membership functions for x and y inputs

$p_1, q_1, r_{1,} p_2, q_2, r_{2,}$ are outlet functions' parameters, the structure and formulation of ANFIS followed a five-layer neural network arrangement.

Layer 1: In this layer, every node i is an adaptive node having a node function for

$$Q_i^1 = \mu_{Ai}(x) \text{ for } i = 1,2 \text{ or } Q_i^1 = \mu_{Bi}(x) \text{ for } i = 3,4 \tag{3.7}$$

Where $Q^1{}_i$ is the membership grade for input x or y. The membership function chosen was Gaussian because it has the lowest prediction error.

Layer 2: In this layer, every rule between inputs are connected by T-Norm operator that performs as 'AND' operator

$$Q^2{}_i = w_i = \mu_{Ai}(x). \mu_{Bi}(y) \text{ for } i = 1,2 \tag{3.8}$$

Layer 3: In this layer, every neuron is labelled Norm, and the output is called 'Normalized firing strength.''

$$Q_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2} \quad i\text{=1, 2} \tag{3.9}$$

Layer 4: Every node i in this layer is an adaptive node and performs the consequent of the rules.

$$Q_i^4 = \overline{w}_i(p_i x + q_i y + r_i) = \overline{w}_i f_i \tag{3.10}$$

$p_1, q_1, r_{1,}$ are irregular parameters referred to as consequent parameters

Layer 5: In this layer, the overall output is computed as the summation of all incoming signals

$$Q_i^5 = \overline{w}_i(p_i x + q_i y + r_i) = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \qquad (3.11)$$

3.9.3 Support Vector Machine (SVM)

Learning in the context of Support Vector Machine (SVM) was proposed and introduced by Vapnik (1995), which provides a satisfactory approach to the problems of prediction, classification, regression and pattern recognition. SVM is based on the concept of machine learning comprises of a data-driven model (Cortes and Vapnik, 1995). The structural risk minimization and statistical learning theory are two useful functions of support vector machine. This makes it differ from ANN because of its ability to reduce the error, complexity and increases the generalization performance of the network (Hong et al.,2008). In SVMs kernel function have been widely applied in the field of rainfall forecasting and various engineering area. Generally, SVMs could be categorized into linear support vector regression (L-SVR) and non-linear support vector regression (N-SVR) (Granata et al., 2017). Therefore, support vector regression (SVR) is a foam of SVM base on the two basic structural layers; the first layer is kernel function weighting on the input variable while the second function is a weighted sum of kernel outputs(Cortes and Vapnik, 1995). In SVR, first a linear regression is fitted on the data, and then the outputs go through a non-linear kernel to catch the non-linear pattern of the data. Given a set of training data $\{(x_i, d_i)\}_i^N$ (x_i is the input vector, d_i is the actual value, and N is the total number of data patterns), the general SVR function is given in equation 3.12.

$$y = f(x) = w\varphi(x_i) + b \qquad (3.12)$$

where $\varphi(x_i)$ indicates feature spaces, non-linearly mapped from input vector x (Cortes and Vapnik, 1995). Regression parameters of b and w may be determined by assigning positive values for the slack parameters of ξ and ξ^* and minimization of the objective function as in equation 3.13 (Wang et al., 2015).

$$\text{Minimize: } \frac{1}{2}\|w\|^2 + C\left(\sum_i^N (\xi_i + \xi_i^*)\right) \tag{3.13}$$

$$\text{Subject to: } \begin{cases} w_i\varphi(x_i) + b_i - d_i \leq \varepsilon + \xi_i^* \\ d_i - w_i\varphi(x_i) + b_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \end{cases} \qquad i=1,2,\ldots,N$$

Where $\frac{1}{2}\|w\|^2$ are the weights vector norm and C is referred to the regularized constant determining in SVR the main aim is to find a function $f(x)$ that is characterized by a maximum ε deviation from the actually obtained target values y_i for all the training data. Moreover, $f(x)$ has to be as flat as possible. The errors smaller than ε are tolerated, while the errors greater than ε are generally unacceptable. The mentioned optimization problem can be changed to the dual quadratic optimization problem by defining Lagrange multipliers α_i and α_i^*. Vector w in Eq. (3.14) can be computed after solving the quadratic optimization problem as (Wang et al., 2015). The general conceptual model structure of SVR can be illustrated in Figure 3.16.

$$w^* = \sum_{i=1}^N (\alpha_i - \alpha_i^*)\varphi(x_i) \tag{3.14}$$

So, the final form of SVR can be expressed as (Wang et al. 2013):

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^N (\alpha_i - \alpha_i^*)K(x, x_i) + b \tag{3.15}$$

Where α_i^+ and α_j^- are Lagrange multipliers, $k(x_i, x_j)$ is the kernel function performing the non-linear mapping into feature space, and b is bias term. One commonly used kernel function is the Gaussian Radial Basis Function (RBF) kernel as (Haghiabi et al.,2017) :

$$k(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2) \tag{3.16}$$

where γ is the kernel parameter.

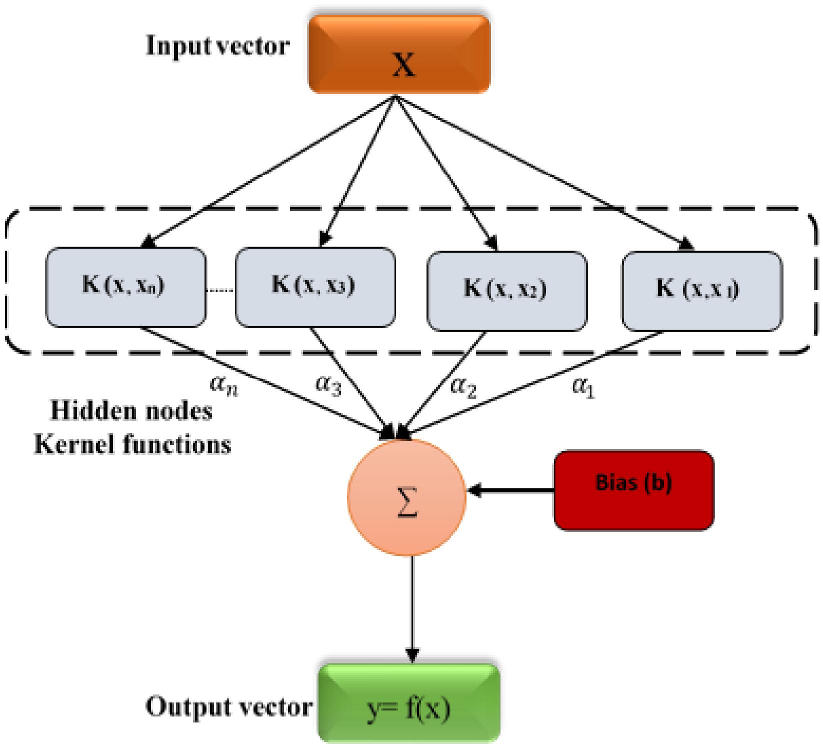


Figure 3.17: Conceptual Network Architecture of SVR Algorithms

3.9.4 Multilinear regression analysis (MLR)

MLR analysis is a model that is applied on a linear relationship between the dependent variable and independent variable, this model is based on the concept of least squares, which is the value of the predicted parameter is expressed as a linear function (Parmar and Bhardwaj, 2015; Chen and Liu, 2015). The general form of MLR can be represented in equation 3.17. However, linear regression can be categorized into simple and multiple linear regression. the model is assumed as the simple linear regression (SLR) If the aim is to estimate the linear correlation between one predictor and one criterion variable. While the model is called multiple linear regression if the goal is to predict the linear correlation between two or more predictors and still one criterion variable. It is essential to know that, MLR is the most common form of linear regression analysis and every value of the independent variable is associated with a value of a dependent variable (Khademi and Behfarnia, 2016).

$$\hat{Y} = a_o + \sum_{j=1}^{m} a_j X_j \tag{3.17}$$

where $\hat{Y}$ is the model 's output, X_j are the independent input variables to the model, and $a_o, a_1, \ldots a_m$ are partial regression coefficients (Khademi and Behfarnia, 2016).

3.9.5 Ensemble Learning Approach

It is clear that single models (e.g., AI-based models) produce different performances for same inputs based on the robustness or limitations. Hence ensemble modelling could effectively improve the general performance of the time series prediction. Ensemble methods have been already applied in some fields of science such as web ranking algorithm, classification, and clustering of time series and regression problems (Sharghi et al., 2018). Ensemble learning is a machine learning to combining the process of multiple predictors in order to enhance the final performance (Baba et al., 2015). Ensemble method has been proved to produce more accurate results than when a single model is used to solve the same problem. The branch of machine learning dealing with multiple homogenous or heterogeneous models is collectively termed as ensemble learning (Baba et al., 2015). The basic component of ensemble learning is a base learner who is created with a base learning algorithm (Kazienko et al., 2013).

This scenario employed three methods in order to improve the predicting performance of the model as (a) Simple Averaging Ensemble (SAE) for combining the FFNN, ANFIS, SVR and MLR predictors (b) Weighted Averaging Ensembling (WAE) and (c) The Non – Linear Neural Network Ensemble (NNE).

Technique 1*: Simple Averaging Ensemble (SAE)*

In the proposed SAE technique, first the FFNN, ANFIS, SVM, and MLR models are trained and tested separately, then the average of FFNN, ANFIS, SVM, and MLR outputs is compared and tested against the test observed values (see Figure. 3.17). The general formula for SAE is given as:

$$P_{(t)} = \frac{1}{N}\sum_{i=1}^{N} p_i\,(t) \tag{3.18}$$

Where N is the number of learners (here N=4) and p_i denotes to the output of a single model (i.e. FFNN, ANFIS, SVM, and MLR) at time *t*.

Technique 2: *Weighted Average Ensemble (WAE)*

Weighted averaging is predicted by assigning different weights to the individual outputs based on the relative significance of the outputs (see Figure. 3.18). The weight is assigned to each output based on relative importance which is not in the case of simple. The weighted averaging model is expressed as:

$$P_{(t)} = \sum_{i=1}^{N} w_i\, p(t) \tag{3.19}$$

Where w_i is the applied weight on the output of the i^{th} model, which can be determined based on the model performance as:

$$w_i = \frac{DC_i}{\sum_{i=1}^{N} DC_i} \tag{3.20}$$

DC_i is the performance efficiency of the i^{th} single model.

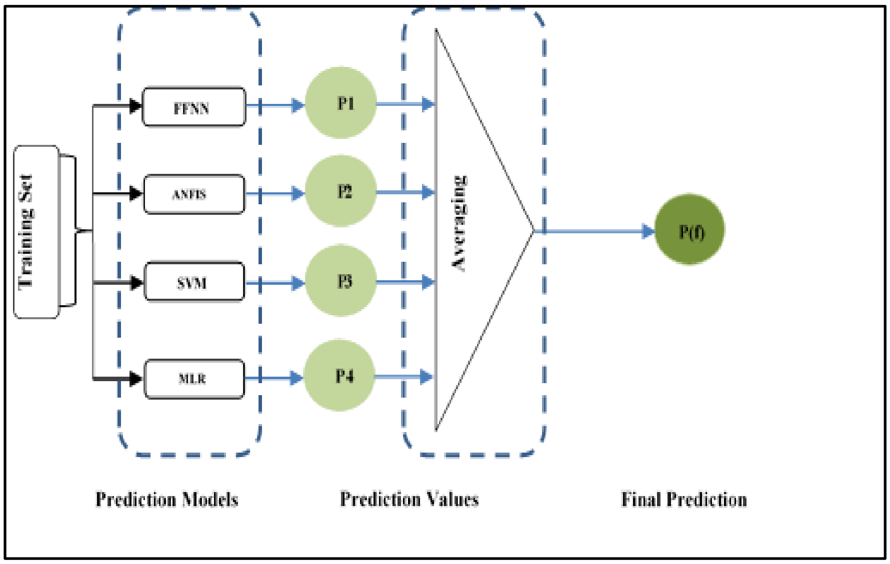


Figure 3.18: Schematic of Simple Averaging Ensemble

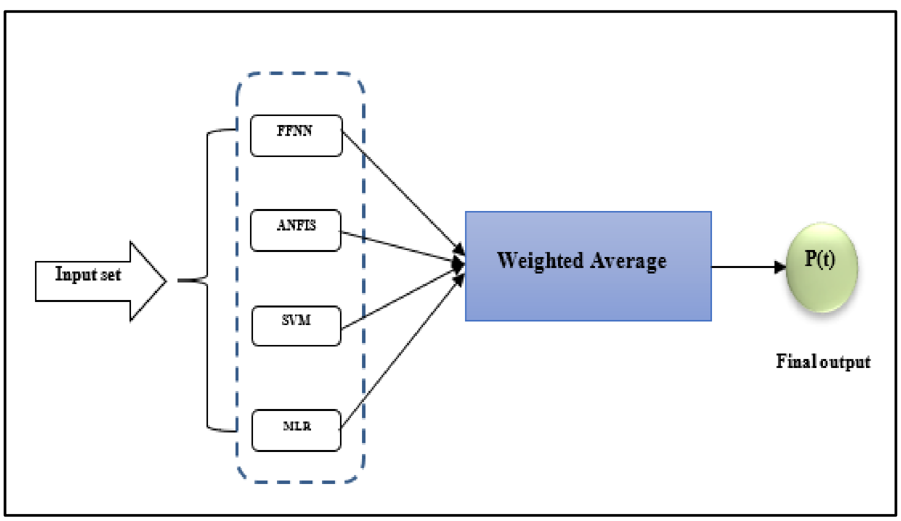


Figure 3.19: Schematic Structure of Weighted Average Ensemble

Approach 3: *Non –linear neural network ensemble (NNE)*

In the neural ensemble method, non-linear averaging is performed by training another neural network. The input layer of the neural ensemble model is fed by the outputs of considered models, each of which is assigned to one neuron in the input layer. A schematic of the proposed neural ensemble method is shown in Figure 3.19. In neural ensemble model like single FFNN, considering tangent sigmoid as activation functions of hidden and output layers, the network is trained using of BP algorithm and the best structure and epoch number of the ensemble network can be determined through the trial-error procedure.

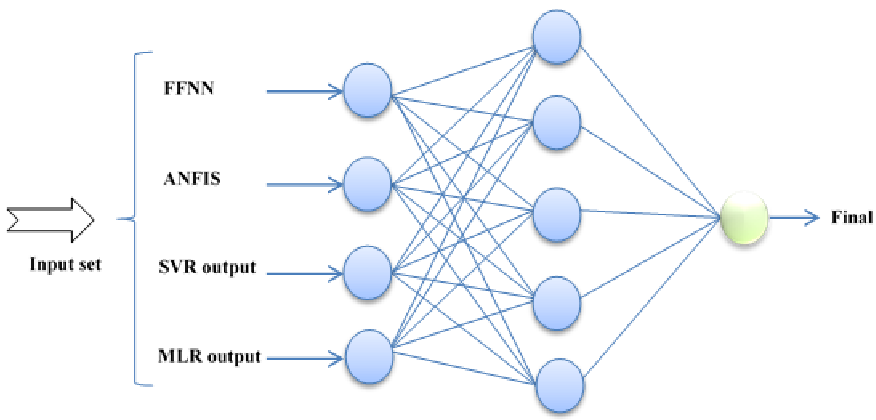


Figure 3.20: Schematic Structure of Neural Ensemble Method

3.10 Methodology for scenario II models (ELM, MLP, and PCA)

3.10.1 Extreme Learning Machine (ELM)

As a newly emerging black-box data-driven algorithms, the ELM was first proposed by (Huang et al. 2006) comprises of single hidden layer feedforward networks (SLFNs). The ELM is quite different from the traditional feed-forward neural network (FFNN) as it overcomes the problems of slow learning speed, local minima, and overfitting (Huang et al. 2015; Yaseen et al. 2018; Zhu et al., 2019). It is worth notable that the promising of ELM could be attributed to its generalization ability and fast learning speed (Nourani et al., 2017; Hou et al., 2018). Due to it is promising performance ability, ELM has been applied in various field of hydro-environmental studies (Yaseen et al. 2019). The structure of the ELM network used in this study is presented in Figure 3.20.

In this study, an ELM model was developed using calibration and validation data set, as mentioned above. For the set of N training samples (i.e. $t = 1, 2, \ldots, N$) in which $x_t \in \mathbb{R}^d$ and $y_t \in \mathbb{R}$, an SLFN with H hidden nodes, is mathematically expressed as (Huang et al. 2006):

$$\sum_{i=1}^{H} B_i \, g_i(\alpha_i . x_t + \beta_i) = z_t, \tag{3.21}$$

where $B \in \mathbb{R}^H$, $Z(z_t \in \mathbb{R})$ and $G(\alpha, \beta, x)$ represents the predicted weights in the output layer, model output and activation function of the hidden layer, respectively. while α_i, β_i, i and d indicate the weights of the randomized layers, biases of these randomized layers, the index of the specific node in the hidden layer and the number of inputs, respectively.

As mentioned above, the study employed activation function as:

$$G(x) = \frac{1}{1 + \exp(-x)} \tag{3.22}$$

,

In an ELM model a proper number of hidden neurons, randomized input layer weights (α), and randomized hidden layer biases (β) can lead to a zero error which, therefore produced the weights of the output layer can be obtained analytically for any training (Huang et al. 2006):

$$\sum_{t=1}^{N} \| z_t - y_t \| = 0, \tag{3.23}$$

The system of the linear equation can be used to obtain the value of B for any input-output training

samples:

$$Y = GB \tag{3.24}$$

In which

$$G(\alpha, \beta, x) = \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_N) \end{bmatrix} = \begin{bmatrix} g_1(\alpha_1 . x_1 + \beta_1) & \cdots & g_L(w_H . x_1 + \beta_H) \\ \vdots & \cdots & \vdots \\ g_1(\alpha_N . x_N + \beta_1) & \cdots & g_L(w_H . x_N + \beta_F) \end{bmatrix}_{N \times H} \tag{3.25}$$

and

$$B = \begin{bmatrix} B_1^T \\ \vdots \\ B_H^T \end{bmatrix}_{H \times 1} \tag{3.26}$$

and

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times 1} \tag{3.27}$$

where G here known as the hidden layer output, T is the transpose of the matrix. The output weights $\hat{B}$ can be estimated by inverting the matrix of the hidden layer using Moore-Penrose generalized inverse function (+):

$$\hat{B} = G^+ Y \tag{3.28}$$

Eventually, the estimated values $\hat{y}$ (i.e. represents a predicted value of BODeff, CODeff, TPeff, and TNeff) can be determined by:

$$\hat{y} = \sum_{i=1}^{H} \hat{B}_i \, g_i(\alpha_i . x_t + \beta_i) \tag{3.29}$$

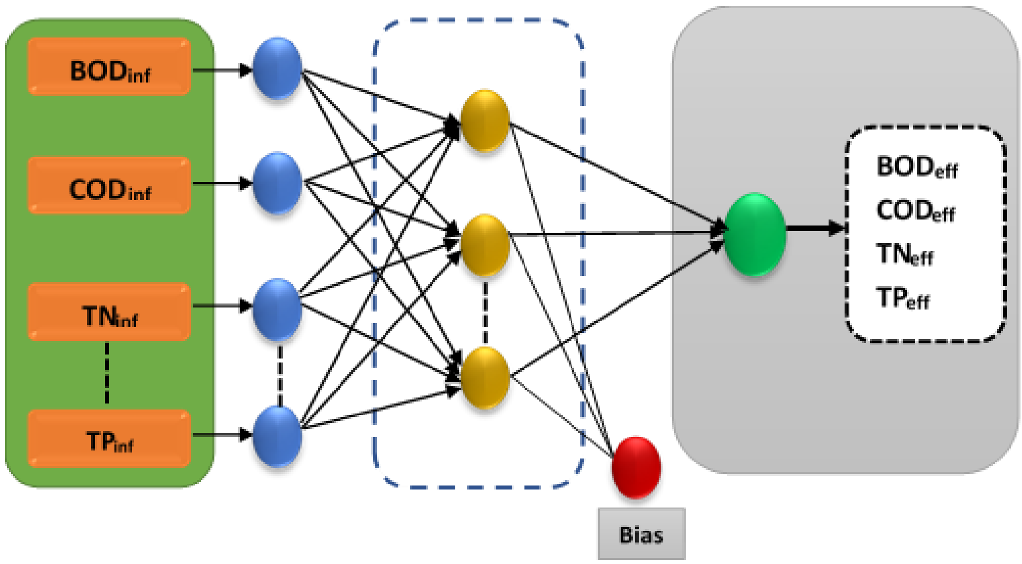


Figure 3.21: The structure of the extreme learning machine network (Yaseen et al., 2018).

3.10.2 Multi-layer perceptron (MLP) neural network

Multi-layer perceptron (MLP) neural network, as one of the most common kind ANN, has the capability to handle non-linear system and describe by numerous literatures as a universal approximator among the different categories of ANNs (Hornik 1991). As like the other traditional ANN, MLP consists of input, one or more hidden and output layers in its architecture (see, Figure. 3.13) (Kim and Singh 2014). The nodes of the input layer are connected to that of the hidden layer and subsequently, the output layer. The information and signals are processed and transmitted from the input to the output layer by the help of weight and biases through the sequential mathematical operation. The lavenberg-Marquardt algorithm is used as learning algorithms to optimize the error between the measured and computed values (ASCE, 2000). The training algorithms are iteratively repeated until the desired outcomes are achieved.

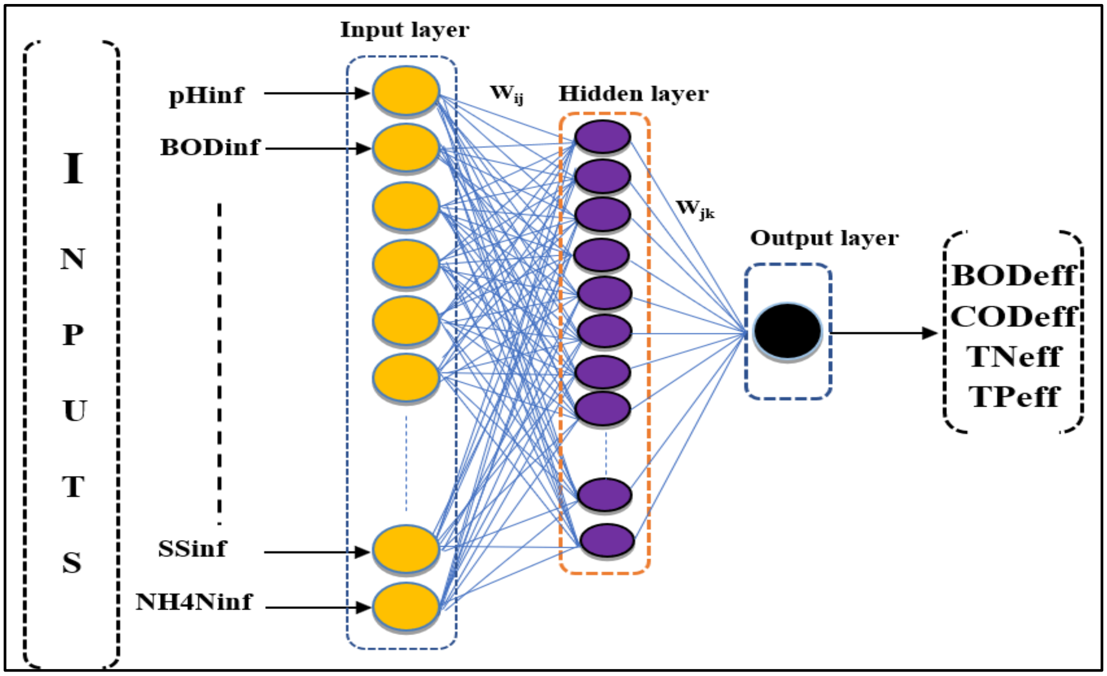


Figure 3.22: Multilayer perceptron neural network (MLPNN) architecture

3.10.3 Principal Component Analysis (PCA)

PCA as one of the multi-variate common statistical techniques for reducing the dimension of the high volume of data. The dimensionality reduction is normally achieved by randomly identifying the linear correlation between the variables (Hasanlou et al., 2015). By applying this method, input variables are changed and are used as independent PCs variables (Noori et al., 2009). Kaiser–Meyer–Olkin (KMO) is among the most common statistics used to assess the suitability of data in any factor analysis (FA) (Acikel et al., 2018). The classification of KMO coefficient can be demonstrated in Table 3.7 and KMO index is presented in equation 3.30. For more explanation of PCA refers to the studies of (Solgi et al., 2017; Singh et al., 2004).

Table 3.7: Classification of KMO coefficients

Relation of data with FA	KMO coefficient

Excellent	≥0.9
Very well	0.8-0.89
Well	0.7-0.79
Mediocre	0.6-0.69
Poor	0.5-0.59
Unacceptable	<0.5

$$KMO = \frac{\sum\sum r_{ij}^2}{\sum\sum r_{ij}^2 + \sum\sum r_{ij}^2} \quad (3.30)$$

Where r_{ij} is the correlation coefficient between the variable of i and j, and a_{ij} is the partial correlation coefficient between them.

3.11 Methodology for scenario III models (AR, HW, and NARX)

3.11.1 Autoregressive (AR) model

AR is commonly used in time-series simulation because of the stochastic process that was built with a degree of randomness and uncertainty (Adamowski et al., 2012). The AR model forecasts the value of a future process of any variable based on the prior values. In particular, the AR model is the regression of values based on the previous occurrence. Therefore, the AR model for an order *p* is defined as AR*(p)* and expressed as:

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots \epsilon_t \quad (3.31)$$

Where ϵ_t is white noise with E= (ϵ_t) and VAR (ϵ_t)=σ_e^2, the parameters $\beta_1, \beta_2, \ldots \beta_P$ are an AR coefficient (Hadi and Tombul, 2018).

3.11.2 Hammerstein- Weiner Model (HW)

A model in which a nonlinear block both precedes and follows a linear dynamic system is referred as Hammerstein- Wiener (HW) model (see, Figure. 3.14) (Bloemen et al., 2001; Wills et al., 2013, Gaya et al., 2017). HW, as a black-box model, is developed for identification of non-linear system (Zambrano et al., 2018). HW system comprises of series and parallel interconnected nonlinear dynamic and static blocks, as shown in Figure. 3.22 (Zambrano et al., 2018; Gaya et al., 2017). The block of HW model was characterized as an appropriate illustration with a clear and understandable relationship to the linear and nonlinear systems than the other traditional ANN. In addition, HW model involved a simple and flexible process of finding parametric specifications for non-linear models and functionally captured the physical knowledge about the system characteristics (Guo, 2004; Ababaei et al., 2013; Ljung, 1995; Guo, 2004).

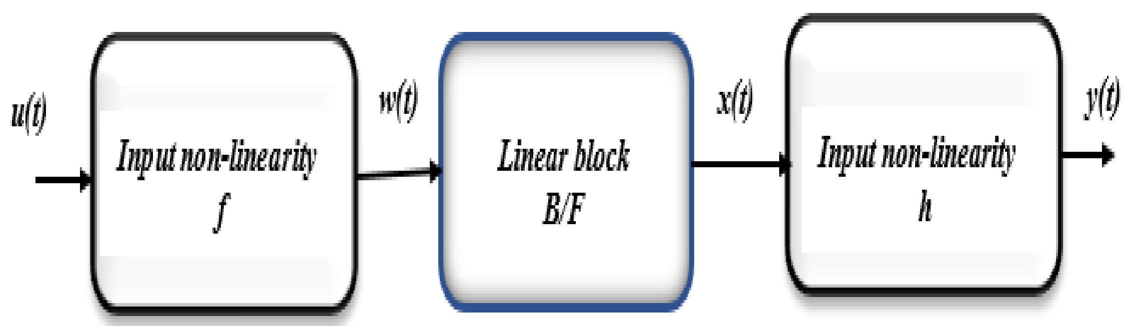


Figure 3.23: Schematic of Hammerstein-Wiener model

$w(t) = f(u(t))$ is a nonlinear function converting input data, $x(t) = w(t)B/F$ shows linear transfer function, f and h act on the input and output port of the linear block, respectively, the function *w(t)* and *x(t)* are variables that define the input and output of the linear block.

3.11.3 Nonlinear autoregressive with exogenous (NARX) neural network

NARX NN is a model of a nonlinear recurrent dynamic neural network, implemented with feedback connections and consisting of several layers (Boussaada et al., 2018). This NARX model is based on the linear ARX model, which is usually used in time series modelling. Therefore, NARX can accept dynamic inputs represented by time series sets. This represents the main advantage of the NARX over feedforward backpropagation neural networks (Di et al., 2016; Al-Sbou et al., 2017). As recurrent neural network possesses the network are quite suitable for nonlinear function approximation and control (Men et al., 2014). The configuration of the NARX model in both series and parallel can be shown in Figure 3.23. The expression for the NARX model is given as:

$$y(t) = f\big(y(t-1), y(t-2), \ldots y\big(t-n_y\big), u(t-1), (t-2), \ldots u(t-n_u)\big) \qquad (3.32)$$

where f is a nonlinear function to be approximated, n_y and $n_u$are the maximum lags input and output entering the model, respectively? The predicted output of future value $\widehat{y_p}(k+1)$of the series-parallel model is given by

$$\widehat{y_p}(k+1) = \emptyset\big[y_p(k), \ldots y_p(k), \ldots y_p(k-n+1); u(k), u(k-n+1)\big] \qquad (3.33)$$

where $\emptyset$ depicts the approximation provided by the series-parallel network identifier.

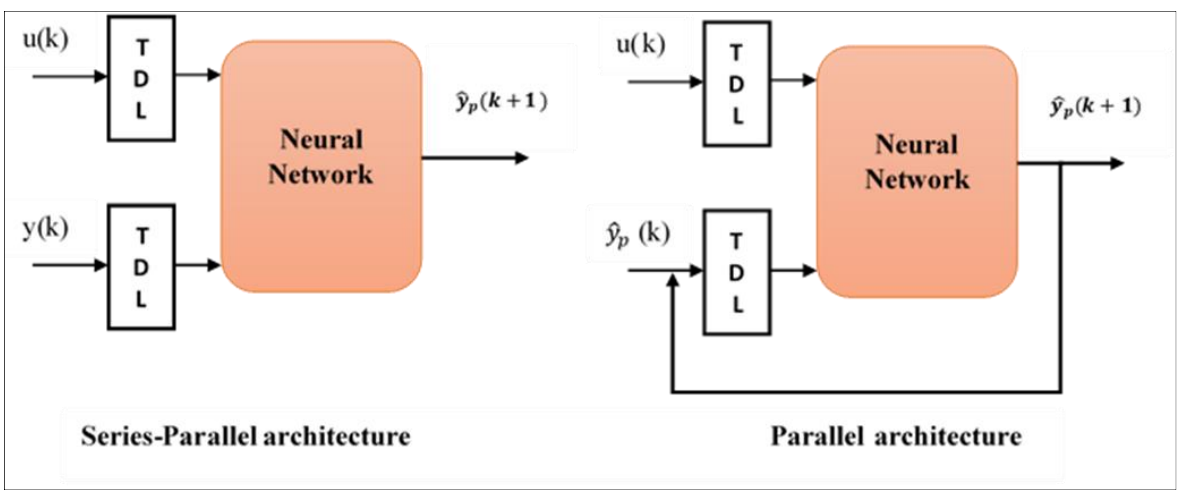


Figure 3.24: Architectures of the NARX neural network (Men et al., 2014)

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Results for the scenario I models (FFNN, SVM, ANFIS, MLR, and Ensemble)

4.1.1 Results of single model predictions of BOD_{eff}, $COD_{eff,}$ and TN_{eff}

It is worth mentioning that, determining the number of hidden neurons, training epoch number and transfer functions are essential aspects in designing the FFNN model. Lavenberg Marquardt was chosen and used in this study as the BP training algorithm due to its fast learning and high-performance accuracy. Four different FFNNs for each output were trained considering different input combinations and the best architecture was determined through the trial-error process. FFNN (5-1) which include 5 inputs and one output neurons found to be the best for all three simulated variables, as shown in Table 4.1. It is noticed that the obtained results of FFNN are satisfactory for predicting the NWWTP performance, which is supported by the values of DC and RMSE displayed in Table 4.1, except for the simulation of BOD_{eff}. For the determination of appropriate ANFIS model, different types of membership functions were examined by trial and error process. ANFIS models were trained for each model using hybrid algorithms (Table 4.1). According to Table 4, ANFIS (5-1) was found to be the best model, and the simulated results showed a good level of satisfaction. In addition, Table 4.1 confirms that the increase in performance is considerably due to the increase in input parameters. ANFIS performance was increased in the verification phase up to 10% for $BOD_{eff,}$ with no or negligible increase for COD_{eff} and TN_{eff} (Table 4.1). This shows that, for predicting the models of NWWTP, ANFIS is more recommended for $BOD_{eff,}$ modelling while both FFNN and ANFIS could be applied for the simulation of COD_{eff} and TN_{eff}.

For SVM modelling, the optimum model was obtained by adjusting two parameters, i.e. squared kernel and regularization constant parameter until the desired output set was achieved. Table 4.1 shows the simulated results of BOD_{eff}, $COD_{eff,}$ and TN_{eff} by SVM model. Table 4.1 proves that, for predicting the NWWTP performance, the use of SVM model is recommended for simulation of COD_{eff} and TN_{eff}, with regards to $BOD_{eff,}$ which the result was not so reliable. In the case of BOD_{eff}, SVM (4-1) was found to be the best while for COD_{eff} and TN_{eff}, SVM (5-1) could lead to better outcomes. MLR model was also applied as a conventional method to predict the performance of NWWTP. Four different models were used to simulate the BOD_{eff}, $COD_{eff,}$ and TN_{eff} (Table 4.1). The best model was determined according to the DC and RMSE criteria. The efficiency criteria of the BOD_{eff} modelling ranged between 0.5941 - 0.6212, 0.0077 - 0.0193 for DC and RMSE, respectively. However, for COD_{eff} those ranged between 0.7013 - 0.8669, 0.0012 - 0.0065 and for TN_{eff} ranged between 0.7055- 0.8392, 0.0006 - 0.0034 for DC and RMSE, respectively. The best model was found to be MLR (5-1) for BOD_{eff} and COD_{eff} while MLR (4-1) for TN_{eff}. Presented results indicate the improved performance of SVM in verification phase up to 10%, 4% and 11% for BOD_{eff}, COD_{eff} and TN_{eff} modelling, respectively over the MLR model. It is apparent that the SVM model slightly demonstrated the enhancement of prediction capability than the MLR model; the MLR also clearly indicates the extent of effluent removal efficiency and plant performance in NWWTP.

Table 4.1 also justifies that, when all the variables are fed into the models to simulate the outputs, the prediction turns to improve in terms of performance criteria. It can be observed from Table 4.1 that ANFIS is the best model among all applied models due to the capability of the fuzzy concept to handle uncertainty in the process. Meanwhile, COD_{eff} and TN_{eff} provide reliable accuracy, while BOD_{eff} is found to be the worst in all models. The fair BOD indicates the presence of organic matter and bacteria. The main focus of NWWTP is to reduce the BOD, COD, and TN in the effluent before discharging to natural waters. However, for the simulation of BOD_{eff} in the verification phase, ANFIS performance showed an increase up to 10%, 15% 16%, for FFNN, SVM, and MLR, respectively. For all models, the verification and calibration results were used to assess the accuracy and efficiency of the algorithm in terms of DC and RMSE. Nevertheless, in hierarchical comparison to other models, FFNN ranked second best followed by SVM and finally MLR model. It is also clear that the ratio between the various components in wastewater has a significant influence on the selection and functioning of NWWTP processes. In NWWTP, COD/BOD ratio is high, which indicates that a substantial part of organic matter will be difficult to degrade biologically, leading to a fair result of BOD_{eff}. The result of BOD_{eff} modelling also proved that the pollutions load is mostly contributed from the households and institutions with low significant contribution from the industrial catchment. Also, the variations and compositions of NWWTP are contributed by the amount of organic waste produced by domestic, institutional and commercial areas. Figures. 4.1 show the time series plot of the fitted models for three simulated outputs via ANN, SVM, ANFIS, and MLR. While Figure. 4.2. depicted the Observed vs predicted Scatter plot obtained by best single models for (a) $BOD_{eff,}$ (b) COD_{eff} and (c) $TN_{eff.}$ Interm of error Figure. 4.3 show the RMSE for the best model in the verification phase. This Figure also proved the capacity of the ANFIS model over the other models.

Table 4.1: Performance efficiency of single FFNN, ANFIS, SVM, and MLR models

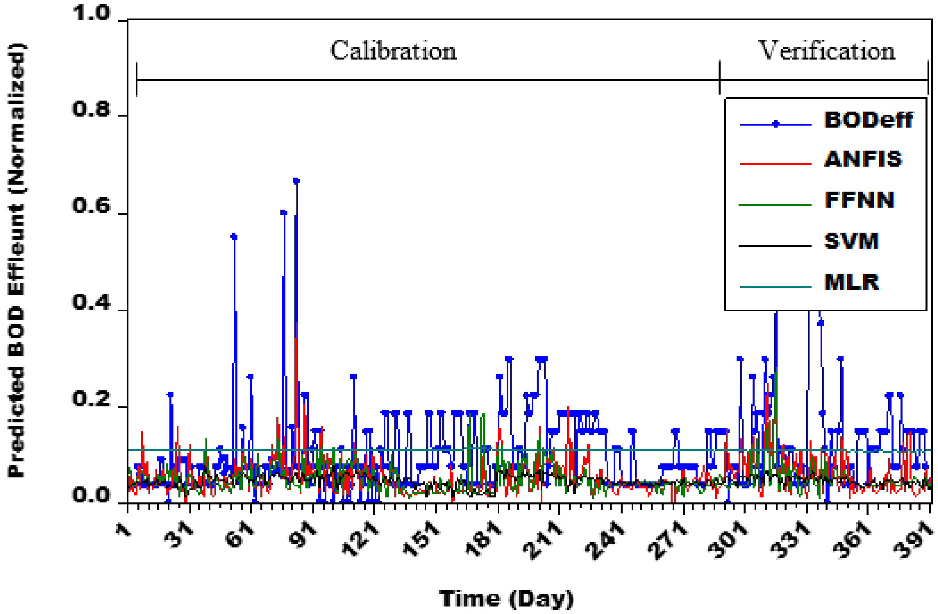
FFNN		Calibration DC	Calibration RMSE[a]	Verification DC	Verification RMSE[a]
BOD_{eff}					
	FFNN (2-1)	0.5889	0.0081	0.6721	0.0106
	FFNN (3-1)	0.6304	0.0076	0.6088	0.0105
	FFNN (4-1)	0.5788	0.0071	0.5046	0.0105
	FFNN (5-1)	**0.6779**	**0.0065**	**0.6600**	**0.0102**
COD_{eff}					
	FFNN (2-1)	0.9081	0.0014	0.9005	0.0062
	FFNN (3-1)	0.9297	0.0012	0.9104	0.0062
	FFNN (4-1)	0.9102	0.0011	0.9100	0.0059
	FFNN (5-1)	**0.9328**	**0.0014**	**0.9363**	**0.0053**
TN_{eff}					
	FFNN (2-1)	0.9367	0.0006	0.8943	0.0034
	FFNN (3-1)	0.9325	0.0006	0.8949	0.0034
	FFNN (4-1)	0.9258	0.0007	0.896	0.0033
	FFNN (5-1)	**0.9343**	**0.0004**	**0.9022**	**0.0034**

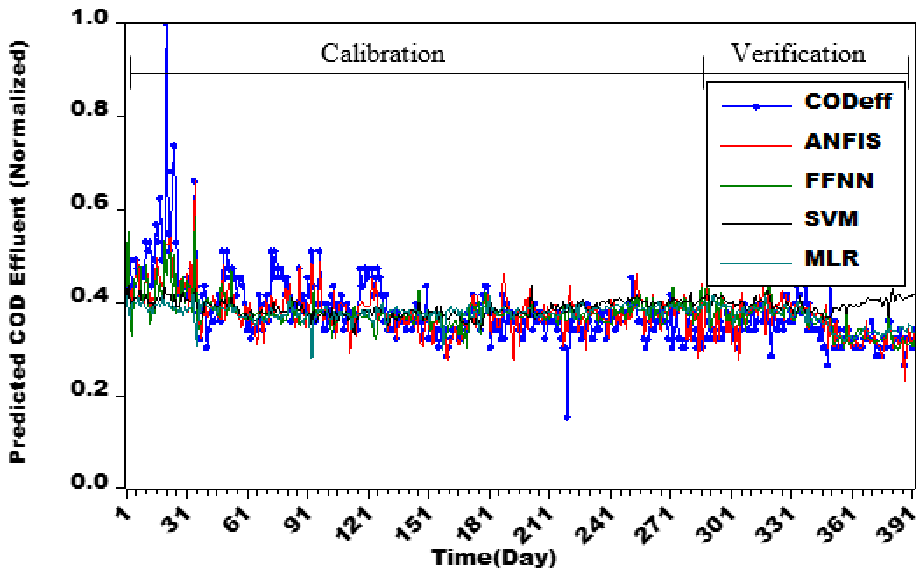
ANFIS		Calibration DC	Calibration RMSE[a]	Verification DC	Verification RMSE[a]
BOD_{eff}					
	ANFIS (2-1)	0.6068	0.0079	0.6766	0.0106
	ANFIS (3-1)	0.6806	0.0066	0.7454	0.0103
	ANFIS (4-1)	0.7512	0.0074	0.7183	0.0093
	ANFIS (5-1)	**0.7828**	**0.0053**	**0.7640**	**0.0083**
COD_{eff}					
	ANFIS (2-1)	0.9087	0.0013	0.9060	0.0062
	ANFIS (3-1)	0.9279	0.0014	0.9020	0.0059
	ANFIS (4-1)	0.9492	0.0013	0.9256	0.0054
	ANFIS (5-1)	**0.9388**	**0.0012**	**0.9260**	**0.0037**
TN_{eff}					
	ANFIS (2-1)	0.8365	0.0006	0.7946	0.0034
	ANFIS (3-1)	0.9389	0.0006	0.8985	0.0033
	ANFIS (4-1)	0.9383	0.0006	0.8113	0.0031
	ANFIS (5-1)	**0.9571**	**0.0005**	**0.9410**	**0.0010**

SVM		Calibration DC	Calibration RMSE[a]	Verification DC	Verification RMSE[a]
BOD_{eff}					
	SVM (2-1)	0.59607	0.0084	0.5635	0.0107
	SVM (3-1)	0.62809	0.0081	0.6393	0.0106
	SVM (4-1)	**0.6554**	**0.0080**	**0.6119**	**0.0106**
	SVM (5-1)	0.6013	0.6081	0.1739	0.0106
COD_{eff}					
	SVM (2-1)	0.9009	0.0039	0.9000	0.0063
	SVM (3-1)	0.9051	0.0053	0.9007	0.0061
	SVM (4-1)	0.9091	0.0048	0.9109	0.0061
	SVM (5-1)	**0.9096**	**0.0047**	**0.9018**	**0.0060**
TN_{eff}					
	SVM (2-1)	0.8966	0.0010	0.7055	0.0034
	SVM (3-1)	0.8929	0.0010	0.7957	0.0033
	SVM (4-1)	0.8892	0.0010	0.7958	0.0033
	SVM (5-1)	**0.8642**	**0.0013**	**0.8050**	**0.0032**

MLR		Calibration DC	Calibration RMSE[a]	Verification DC	Verification RMSE[a]
BOD_{eff}					
	MLR (2-1)	0.5999	0.0083	0.61003	0.0109
	MLR (3-1)	0.6146	0.0078	0.5941	0.0193
	MLR (4-1)	0.6007	0.0077	0.6000	0.0110
	MLR (5-1)	**0.6212**	**0.0077**	**0.6037**	**0.0109**
COD_{eff}					
	MLR (2-1)	0.8062	0.0012	0.7013	0.0065
	MLR (3-1)	0.8455	0.0013	0.8187	0.0064
	MLR (4-1)	0.8585	0.0013	0.7090	0.0064
	MLR (5-1)	**0.8669**	**0.0014**	**0.8591**	**0.0064**
TN_{eff}					
	MLR (2-1)	0.7375	0.0006	0.7029	0.0034
	MLR (3-1)	0.7377	0.0006	0.7117	0.0034
	MLR (4-1)	**0.8392**	**0.0006**	**0.7930**	**0.0034**
	MLR (5-1)	0.7956	0.0007	0.7227	0.0034

[a]since all data are normalized, the RMSE has no dimension.


Calibration
Verification
BODeff
ANFIS
FFNN
SVM
MLR
Predicted BOD Effleunt (Normalized)
Time (Day)
1.0
0.8
0.6
0.4
0.2
0.0
1
31
61
91
121
151
181
211
241
271
301
331
361
391


Calibration
Verification
CODeff
ANFIS
FFNN
SVM
MLR
Predicted COD Effluent (Normalized)
Time(Day)
1.0
0.8
0.6
0.4
0.2
0.0
1
31
61
91
121
151
181
211
241
271
301
331
361
391

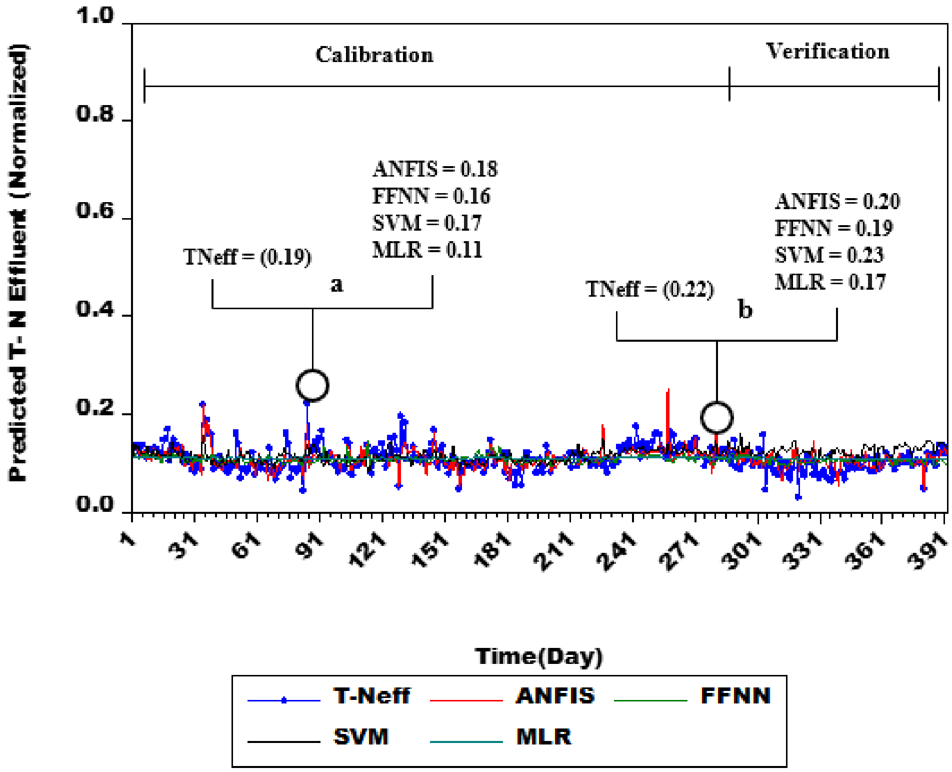
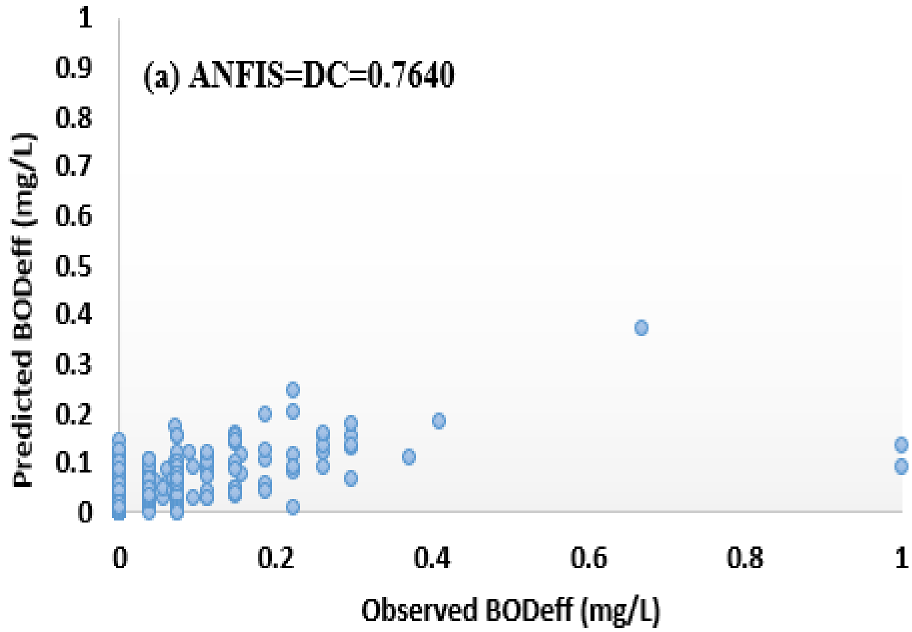
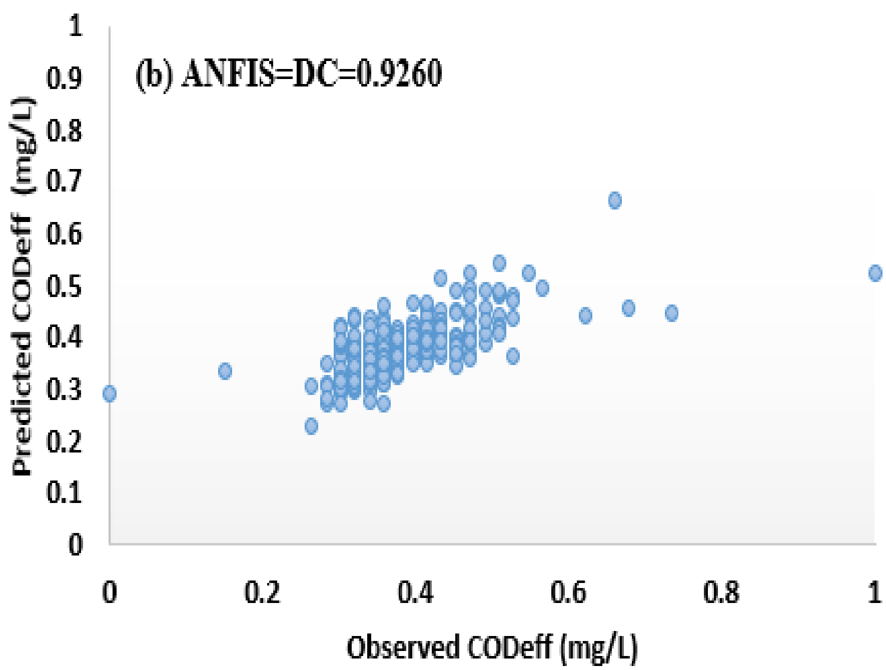


Figure 4.1: Observed vs predicted time series obtained by best single models for (a) BOD_{eff}, (b) COD_{eff} and (c) TN_{eff}

Figure. 4.1 depicts a plot of predicted TN_{eff} by different methods highlighting two sample points (a) and (b). From this Figure. it is clear that for sample point (a) ANFIS model could lead to a bit better performance than FFNN and SVM models. On the other hand, for sample point (b) FFNN and SVM models are better than the ANFIS model. Therefore, although the overall performance of one of the models may be better for the whole time series, at different spans of time series, the performance of the models may be different. As such, at different conditions, different methods may lead to different outcomes, and so it is a logical idea to ensemble the outcomes of different methods to get more accurate results for future predictions.



(a) ANFIS=DC=0.7640
Predicted BODeff (mg/L)
Observed BODeff (mg/L)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1



(b) ANFIS=DC=0.9260
Predicted CODeff (mg/L)
Observed CODeff (mg/L)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

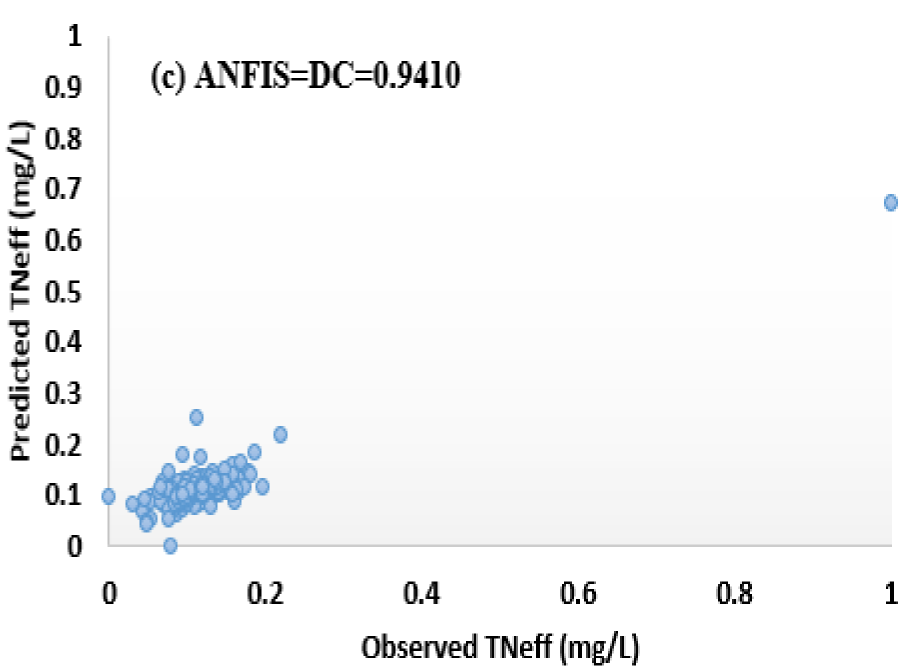


Figure 4.2: Observed vs predicted Scatter plot for the best single model obtained by best single models for (a) BOD_{eff}, (b) COD_{eff} and (c) TN_{eff}

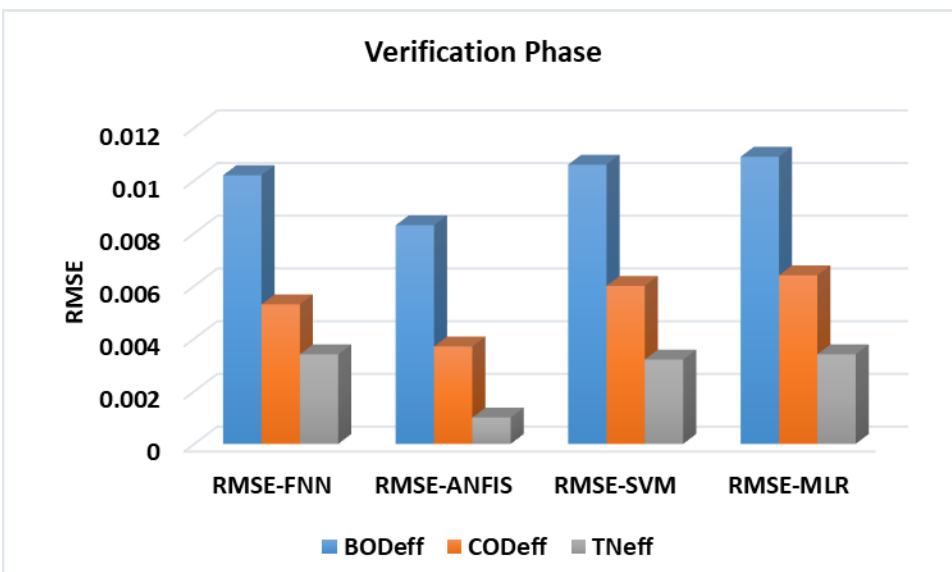


Figure 4.3: RMSE for the best model in the verification phase

4.1.2 Results of ensemble predictions for BOD_{eff}, COD_{eff}, and TN_{eff}

The ensemble of outputs from FFNN, ANFIS, SVM, and MLR were carried out based on proposed SAE, WAE and NNE to improve the overall prediction accuracy of the single models. Table 4.2 shows the obtained results by SAE, WAE, and NNE techniques. The obtained DC and RMSE values for both calibration and verification phases show improvement in the modelling efficiency with regards to the single models. The performance of ensemble techniques depends on the accuracy of each individual model as each model has its own drawback and merit in the modelling process. The results also proved that, for the prediction of BOD_{eff} in NWWTP, ensemble methods could lead to superior results with regard to single models (Table 4.2). This is because, the integration of a single model's outputs reduces the variance, bias and improves performance of the overall modelling.

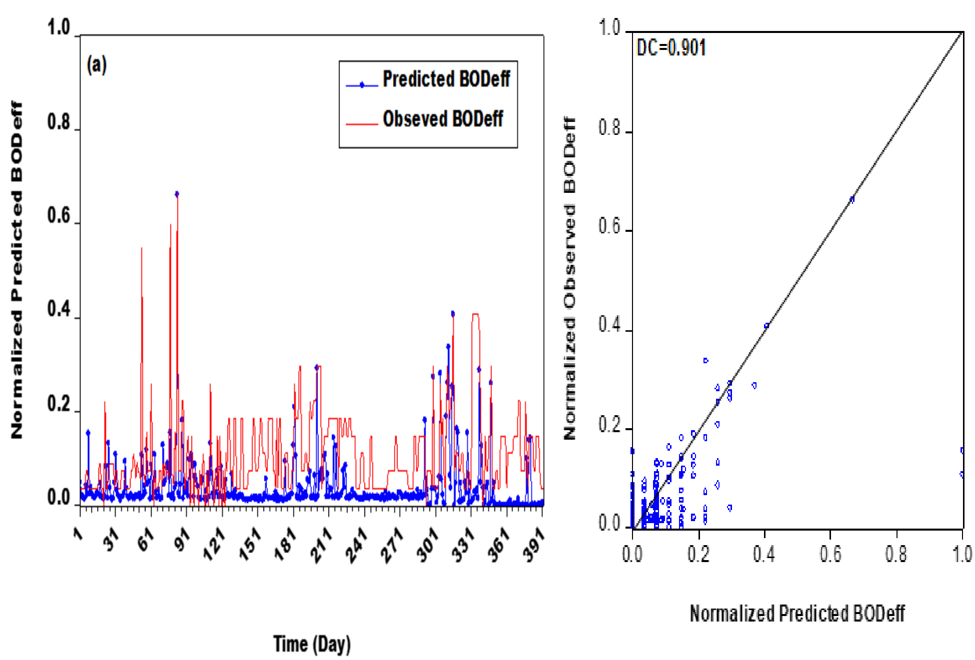
Despite a reliable result for all the ensemble techniques, NNE found to be more accurate, followed by WAE and lastly SAE. In the verification phase, SAE, WAE, and NNE increased the efficiency performance of AI modelling up to 14%, 20%, and 24%, respectively for predictions of BOD_{eff} and up to about 5% for modelling both COD_{eff} and TN_{eff} parameters. This proved a remarkable improvement in the prediction of BOD_{eff}, which was found poor using single models. Ensemble methods aimed primarily not only to integrate a set of models but also to decrease the weaknesses of every single model and come up with the enhance and composite model, which is feasible, reliable with high accuracy than single models. According to Table 4.2, the results of WAE slightly outperformed SAE due to the fact that weight is assigned to each parameter based on relative importance which is not in the case of simple averaging. The performance of NNE is better than two ensemble techniques in both calibration and verification steps, because of the robustness of NNE in handling nonlinear interactions, and able to backpropagate the produced error during calibration phase until the desired result is achieved. Figure. 4.4 shows the results obtained by NNE as the scatter plot and time series plots for BOD_{eff}, COD_{eff}, and TN_{eff} versus observed values.

Table 4.2: Results of the proposed ensemble techniques

		Calibration		Verification	
Ensemble technique[a]	Predicted variables	DC	RMSE[b]	DC	RMSE[b]
SAE	BOD_{eff}	0.884	0.006	0.860	0.008
	COD_{eff}	0.909	0.004	0.903	0.004
	TN_{eff}	0.897	0.009	0.873	0.002
WAE	BOD_{eff}	0.891	0.005	0.806	0.003
	COD_{eff}	0.919	0.004	0.900	0.009
	TN_{eff}	0.947	0.006	0.934	0.002
NNE	BOD_{eff}	0.902	0.085	0.899	0.053
	COD_{eff}	0.958	0.052	0.947	0.024
	TN_{eff}	0.979	0.020	0.968	0.015

[a] The result has been presented for the best structure.

[b] Since all data are normalized, the RMSE has no dimension



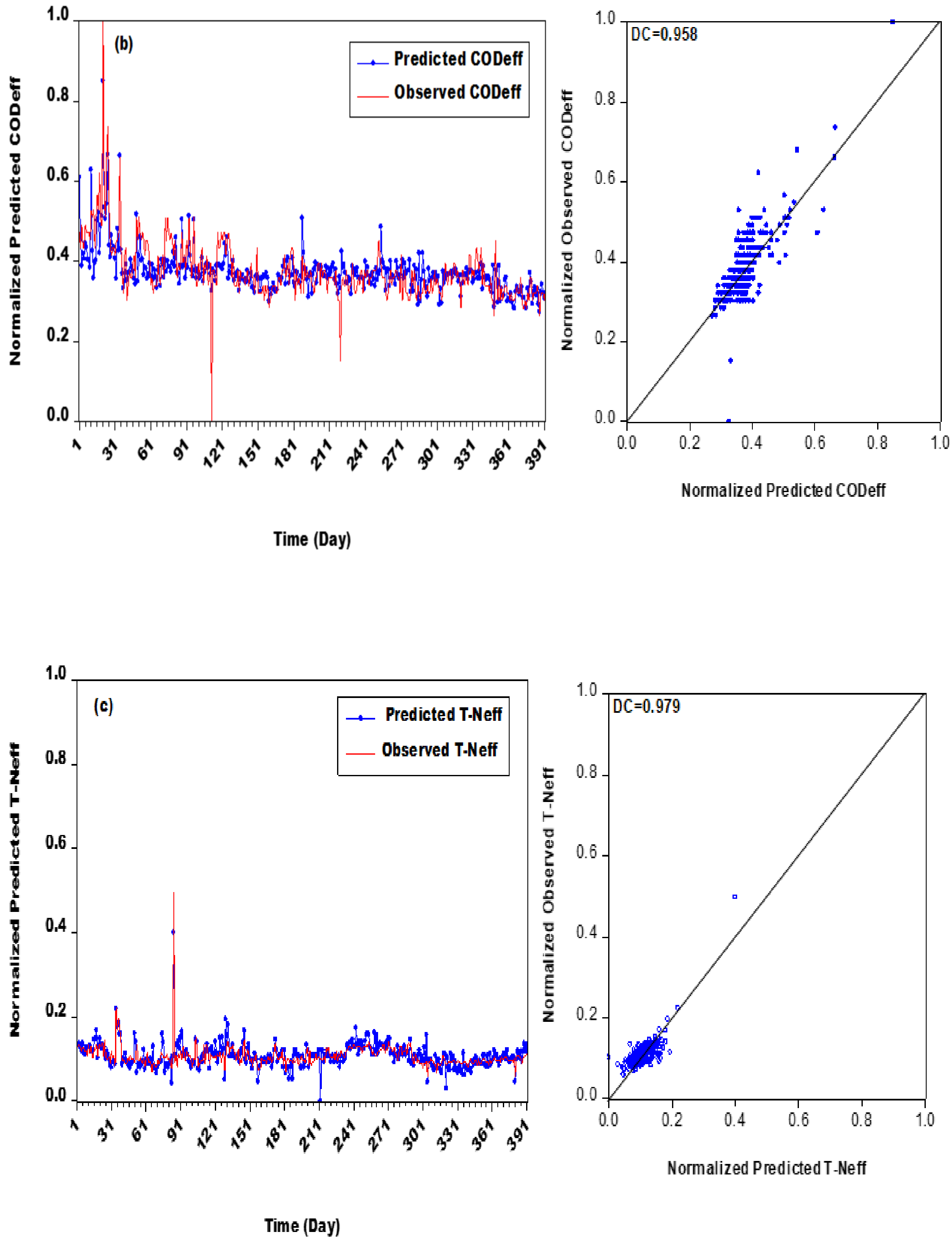
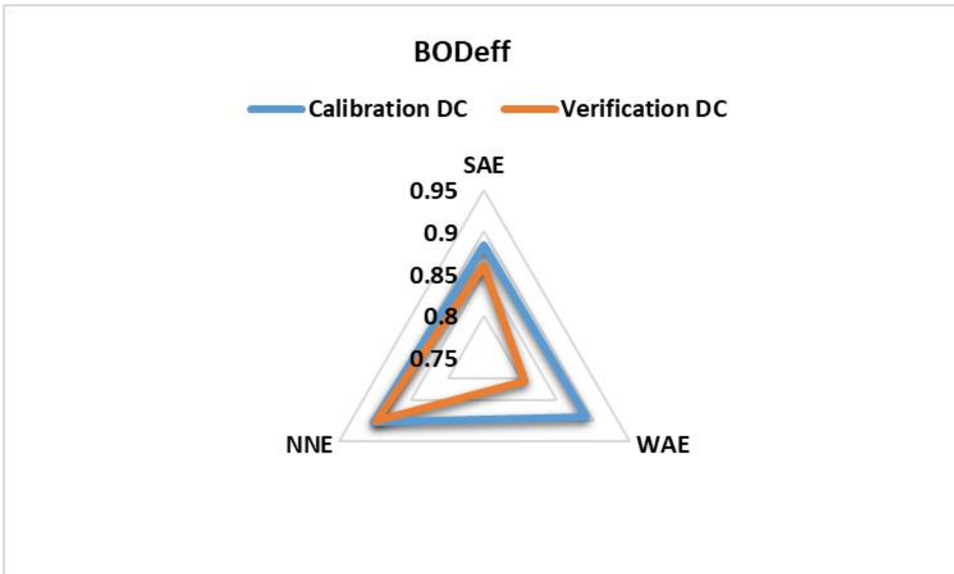
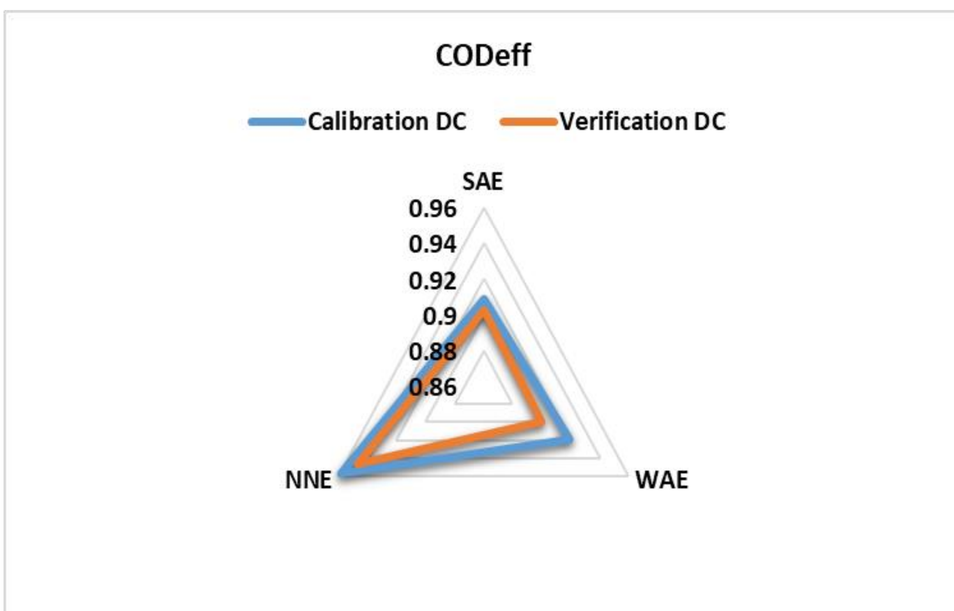


Figure 4.4: Scatter plot and time series plots of results obtained by NNE techniques for a) BOD_{eff}, (b) COD_{eff} and (c) TN_{eff}

In Figure. 4.5, Radar diagram was also used as the most widely recommended diagrams for accuracy comparison of SAE, WAE, and NNE for DC values for BODeff, CODeff, and TNeff. The performance of the models was compared during the calibration and verification phase. From Figure. 4.5 it can be seen that NNE outperformed other ensemble model techniques.





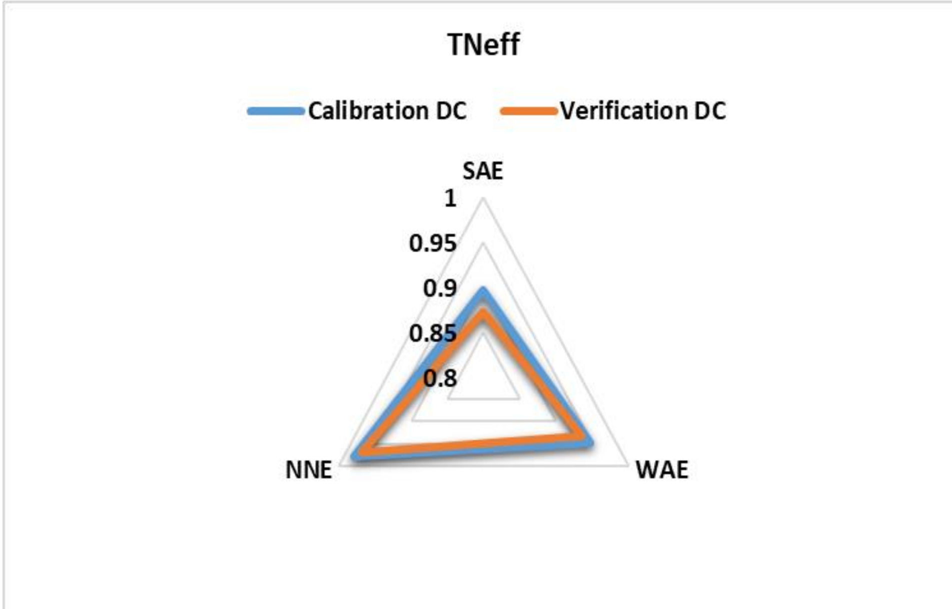


Figure 4.5: Radar chart of DC for SAE, WAE, and NNE for both calibration and verification phases (BODeff, CODeff, and TNeff)

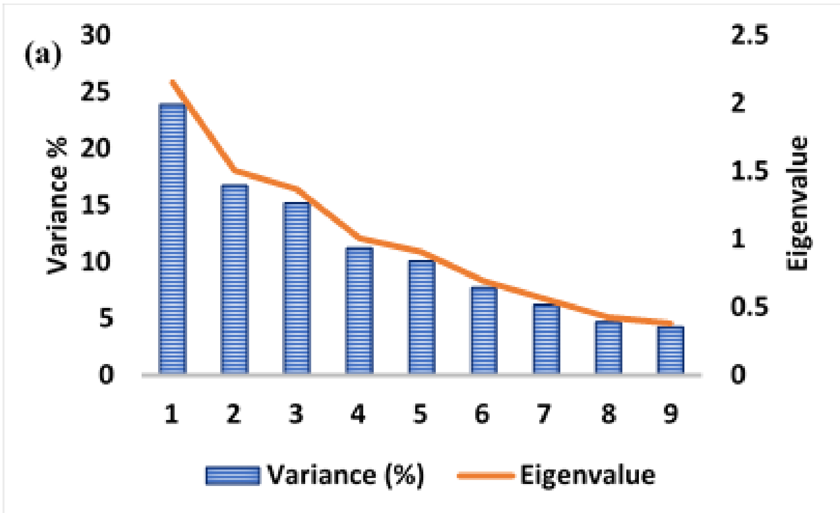
4.2 Results for scenario II models (ELM, MLP and PCA)

4.2.1 Implementation of scenario I

Various structure for ELM, 4PCs-ELM, and 6PCs-ELM was used to obtain the best structure of the model, the optimum hidden neurons were identified as the best optimal ELM structure for all the combination. PCA was employed for choosing the input variable in order to enhance the ELM prediction (Solgi et al., 2017). According to the obtained KMO=0.735, the PCA is suitable for all the output variables. In PCA different approaches were used for deciding which factors can noticeably affect the resulting pattern of the data, as such this research employed the approach of selecting the factors with eigenvalues equal or greater than 1.00 (see, Table 4.3). According to Holland, (2008), in any correlation matrix, eigenvalues are used to condense the variance where the highest eigenvalues (1 and above) are traditionally considered for any analysis by eigenvectors ranking. Figure. 4.6a shows the specific values and the percentage variance of each factor as a graph which demonstrates 9 inputs variable with the corresponding 9 eigenvectors and eigenvalues. Similarly, Table 4.3 shows the value of each factor and its percentage of separation from the primary variable. It can be seen from the table that more than 80% of the factors were explained by the first 6 variables. Likewise, the result indicated that up to 8 factors there exist a significant percentage of about 95%, this can be proved as in Figure. 4.6a of the obtained results.

Table 4.3: Eigenvalue and percentage of data explained by each factor.

Number	Eigenvalue	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.153019	0.646156	0.2392	2.153019	0.2392
2	1.506864	0.139714	0.1674	3.659883	0.4067
3	1.36715	0.358621	0.1519	5.027033	0.5586
4	1.008529	0.102154	0.1121	6.035561	0.6706
5	0.906374	0.213251	0.1007	6.941935	0.7713
6	0.693123	0.135041	0.077	7.635059	0.8483
7	0.558082	0.132987	0.062	8.193141	0.9103
8	0.425095	0.043332	0.0472	8.618236	0.9576
9	0.381764	---	0.0424	9	1



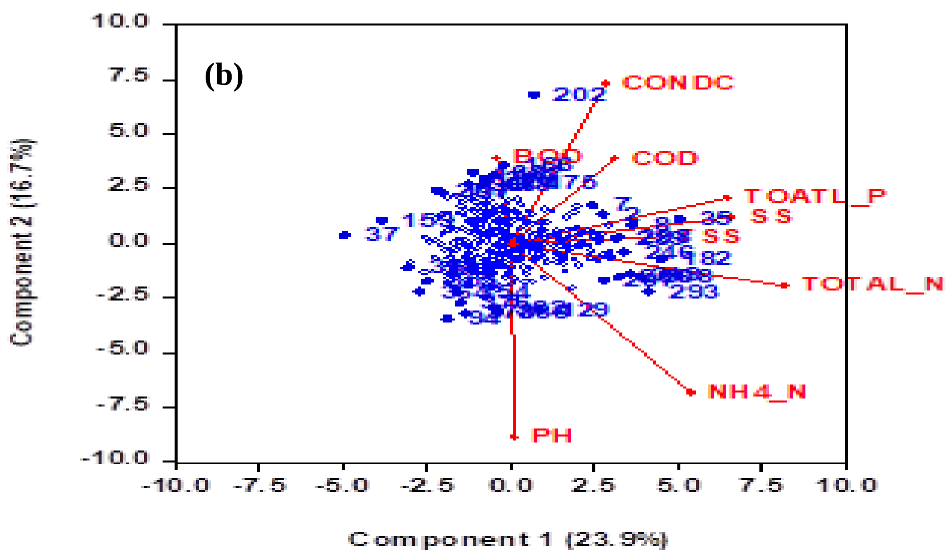


Figure 4.6: (a) shows the percentage variance vs the number of factors and eigenvalue vs the number of factors (b) orthonormal loadings biplot of the first two components of the PCA model

Figure. 4.6b examined the orthonormal loadings biplot relationship between the variables, the horizontal axis is the first PCA dimension representing 23.9%, and the vertical axis is the second PCA dimension. The long or short red vectors line indicates the suitability of the presentation or otherwise. From both the Figure. 4.6 (a) and (b) we can extract both the 4PCs and 6PCs accordingly. According to Table 4.4, the obtained results of the best model for BOD_{eff}, COD_{eff}, $TN_{eff,}$ and TP_{eff} were obtained using 4PCs-ELM, 6PCs-ELM, 6PCs-ELM, and 6PCs ELM, respectively. This can be proved by comparing the values of R^2, RMSE, and MAPE. Though PCs-ELM combination generates the most accurate results in all cases, using single ELM model also emerged to be reliable for the prediction, this is due to their promising ability to handle the highly complex and non-linear process.

Table 4.4: Results of ELM, and PCs-ELM for BOD_{eff}, COD_{eff}, $TN_{eff,}$ and TP_{eff}

		Calibration			**Verification**		
Parameter	**Model**	$\mathbf{R^2}$	**RMSE**	**MAPE**	$\mathbf{R^2}$	**RMSE**	**MAPE**
	All (9 inputs)	0.5439	0.0749	0.0126	0.5168	0.0803	0.0482
$\mathbf{BOD_{eff}}$	4PCs-ELM	0.5711	0.0714	0.0042	0.6341	0.0562	0.0143
	6PCs-ELM	0.5618	0.0727	0.0088	0.6285	0.0902	0.2009
	All (9 inputs)	0.9632	0.0101	0.0051	0.9541	0.0399	0.0191
$\mathbf{COD_{eff}}$	4PCs-ELM	0.9522	0.0268	0.0003	0.9545	0.0534	0.0452
	6PCs-ELM	0.9757	0.0208	0.0103	0.9742	0.0515	0.0403
	All (9 inputs)	0.8643	0.0424	0.0081	0.7651	0.0347	0.0837
$\mathbf{TN_{eff}}$	4PCs-ELM	0.9169	0.0387	0.0238	0.9128	0.0336	0.0561
	6PCs-ELM	0.9457	0.0983	0.0098	0.9656	0.0335	0.0522
	All (9 inputs)	0.8803	0.0819	0.0112	0.8159	0.0718	0.1019
$\mathbf{TP_{eff}}$	4PCs-ELM	0.8629	0.0191	0.0335	0.8509	0.0450	0.2542
	6PCs-ELM	0.9629	0.0312	0.0205	0.8807	0.0491	0.1303

A close examination shows that both ELM and PCs-ELM produced different performance accuracy, which signifies that the individual model type responds in a different way to the same or different input parameters. Table 4.4 also confirmed that, in both calibration and verification, PCs-ELM model achieved the lowest RMSE and MAPE for BOD_{eff}, COD_{eff}, $TN_{eff,}$ and TP_{eff} modelling. The result also shows the increase for PCs-ELM of about 12%, 2%, 20% and 6% for BOD_{eff}, COD_{eff}, TN_{eff} and TP_{eff} with regards to the novel ELM model. Box plots for observed and predicted models are shown in Figure 4.7. From the Figure, the PCs-ELM model was clearly found to obtain the best fit line between the observed and estimated values, hence, demonstrated high predicting ability in Nicosia MWWTPs which may be considered to serve as a valuable and reliable tool for identifying its performance analysis. The plot also demonstrated the closeness of all the models with the observed values, the plot contained (box and whisker median, mean and staples). According to the plot, the extent of spread values between the observed and predicted models indicates the superiority of the PCs -ELM models.

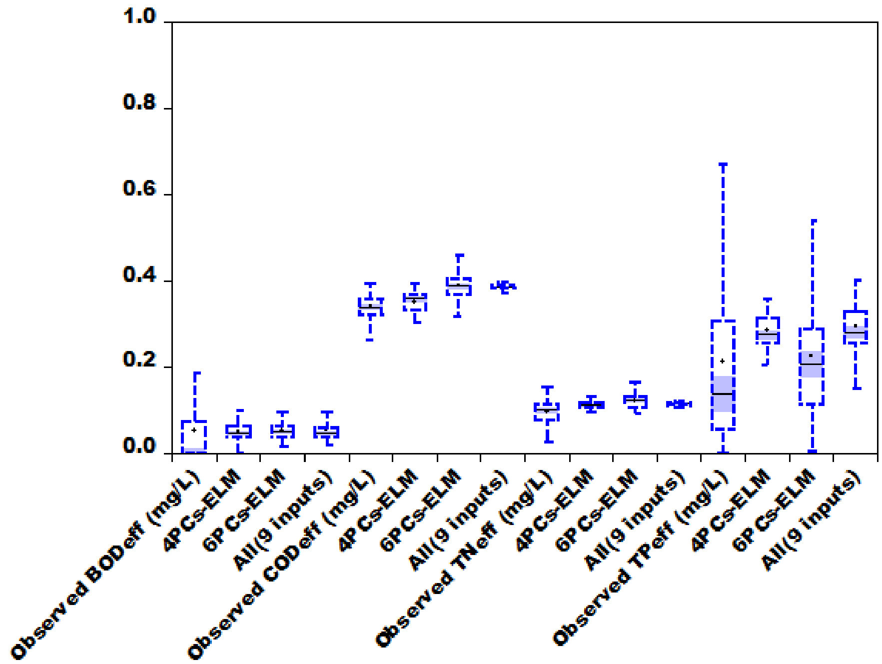
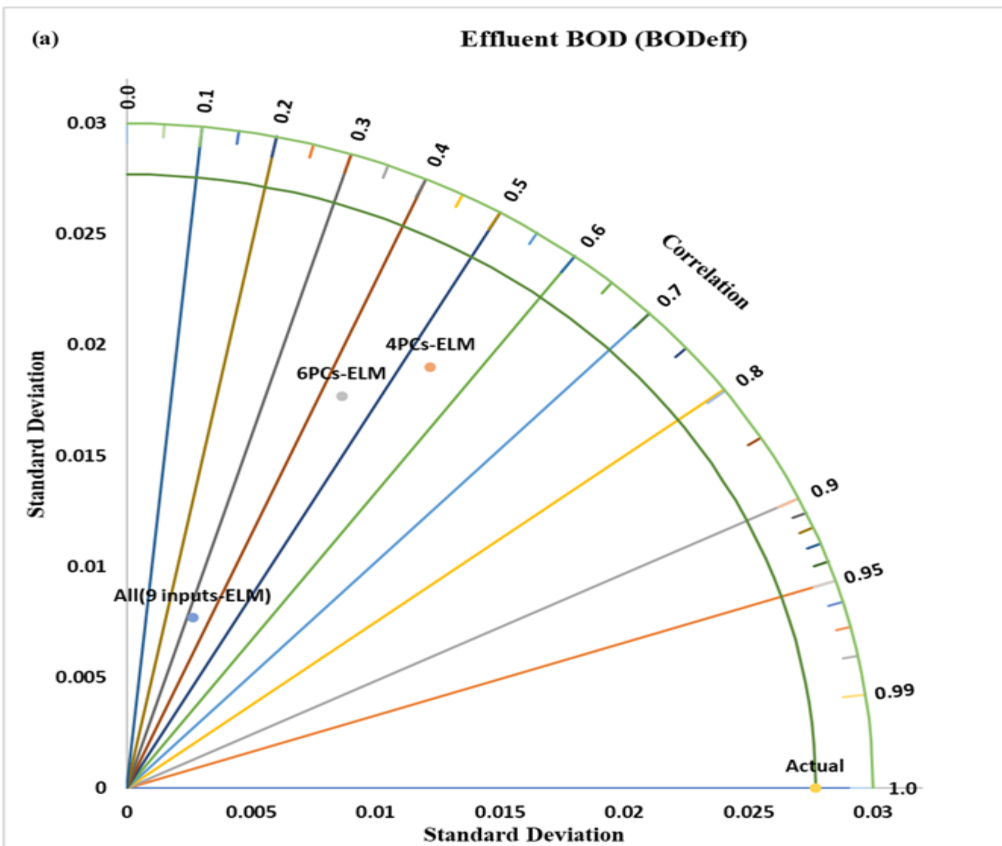
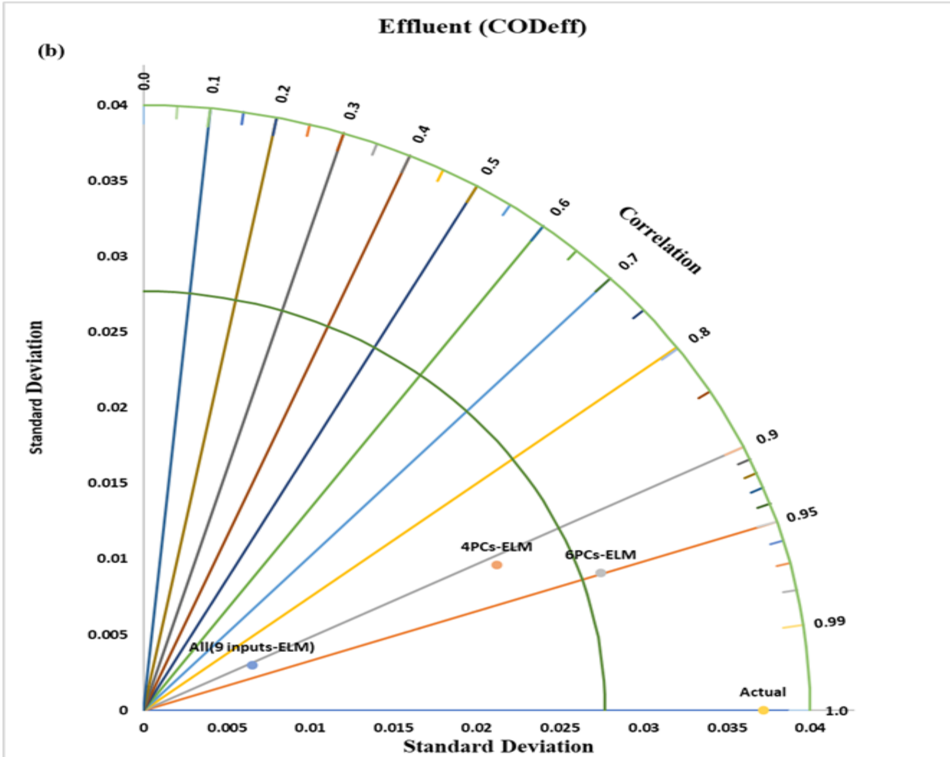
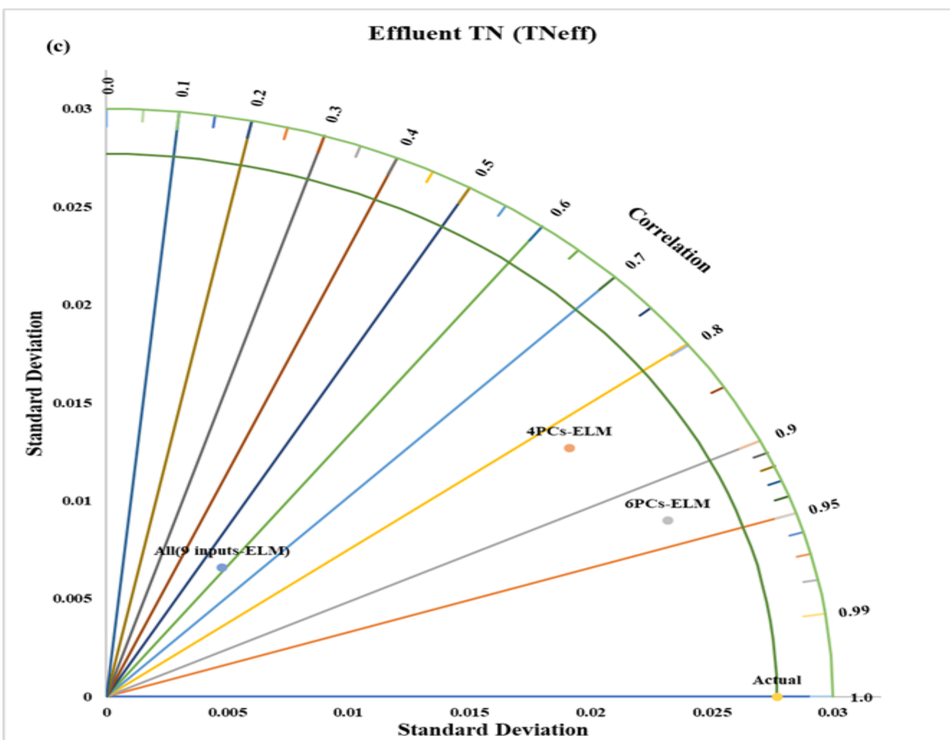


Figure 4.7: The comparison box-plot of the observed and all the predicted models

In the same way, the results of RMSE and MAPE depicts the performance indicator for the best model, and it was reported that the smaller the values of RMSE and MAPE the more accurate the predicting results (Gaya et al., 2017). Further examination of performance accuracy was also investigated using two- dimensional graphical diagram (i.e., Taylor diagram) as depicted in Figure. 4.8 (a-d). Taylor diagram is a graphical representation method that exhibits how closely a model or different model matches the observed and corresponding computed values. Moreover, the computed models and the observed data are described quantitatively in terms of their correlation coefficient (R) and standard deviations (SD).




(b)
Effluent (CODeff)
Correlation
Standard Deviation
4PCs-ELM
6PCs-ELM
All(9 inputs-ELM)
Actual


(c)
Effluent TN (TNeff)
Correlation
Standard Deviation
4PCs-ELM
6PCs-ELM
All(9 inputs-ELM)
Actual

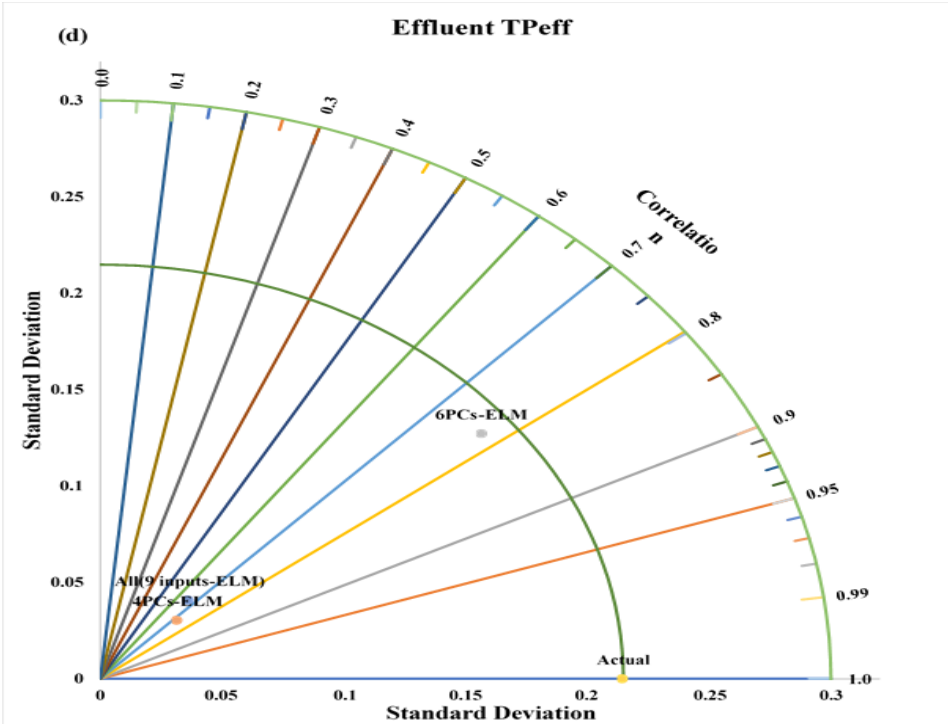
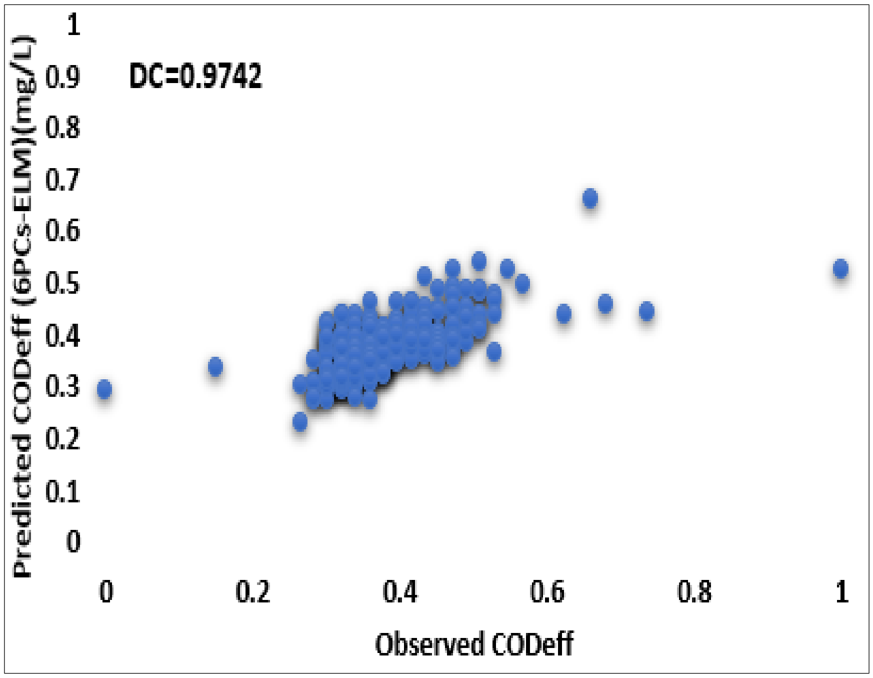
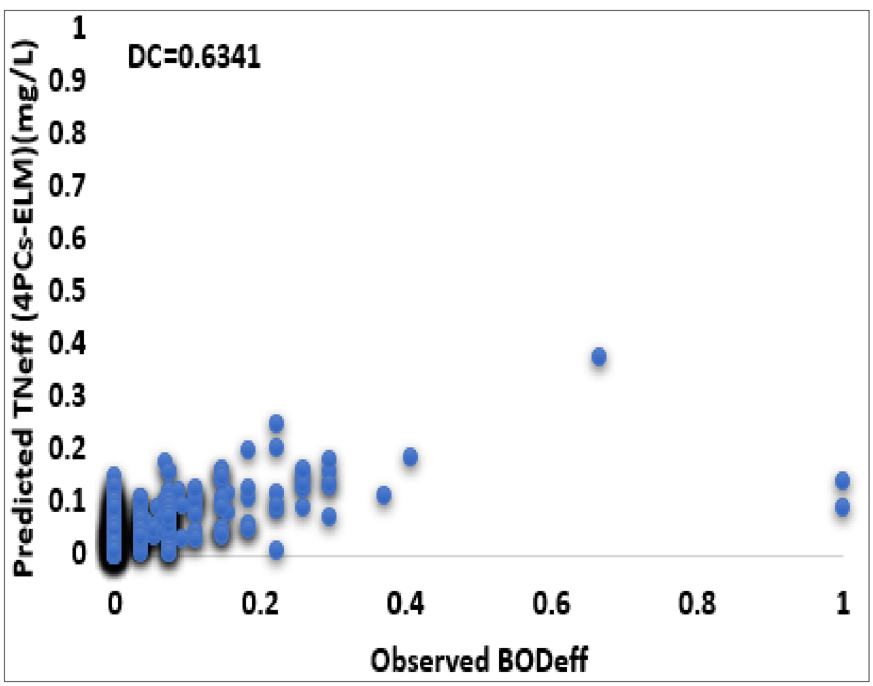


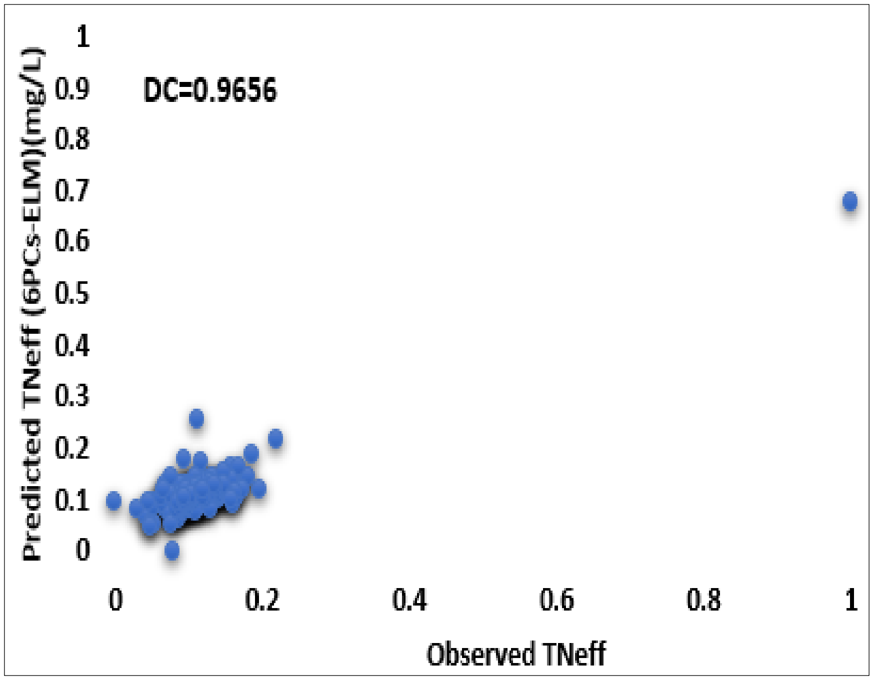
Figure 4.8: Taylor diagram showing the degree of prediction in terms of R and SD for (a) BODeff (b) CODeff (c) TNeff and (d) TPeff

Figure. 4.8a shows that the best predictive BODeff model is far from the actual (observed) data that signifies less performance accuracy which could be attributed to the small value of R and high dispersion between the observed and predictive model. Similarly, Figure. 4.8 (b-d), proved the results in Table 4.4, the discovered model (CODeff, TNeff, and TPeff) showed an outstanding performance in determining the performance of Nicosia MWWTP. According to the value of R and SD for Figure. 4.8 (b-d), the best models depicted the extent and degree of the prediction skills.

Moreover, the scatter diagram of the best-computed model is shown in Figure. 4.9. The plots indicate a closeness agreement between the observed and computed values for CODeff, TNeff, and TPeff during a fair agreement for BODeff. This conclusion is in line with that of Nourani et al., (2018a). Note that, data pairs closer to the 45° line represent better prediction cases in any scatter plots.







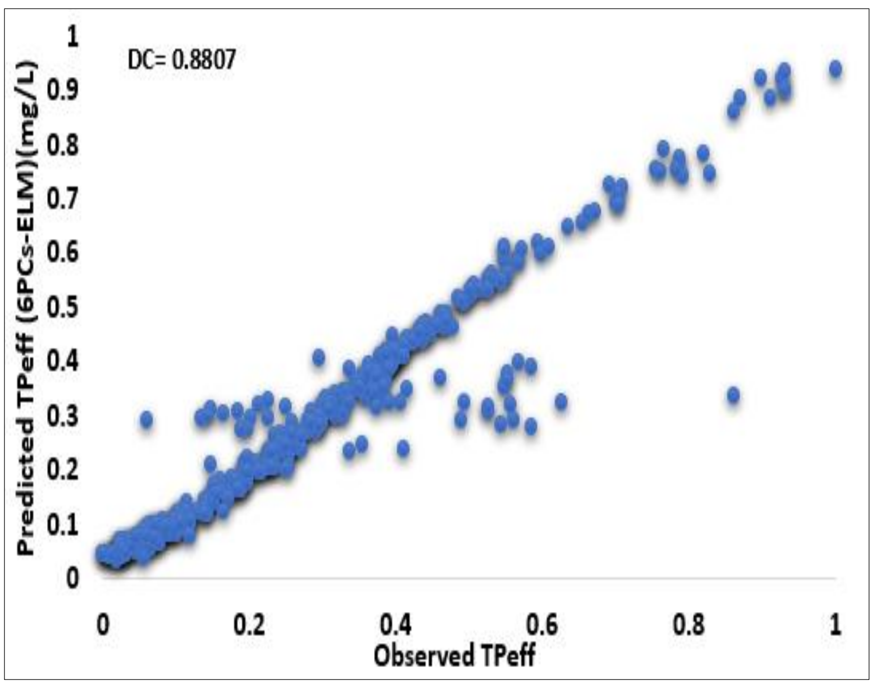
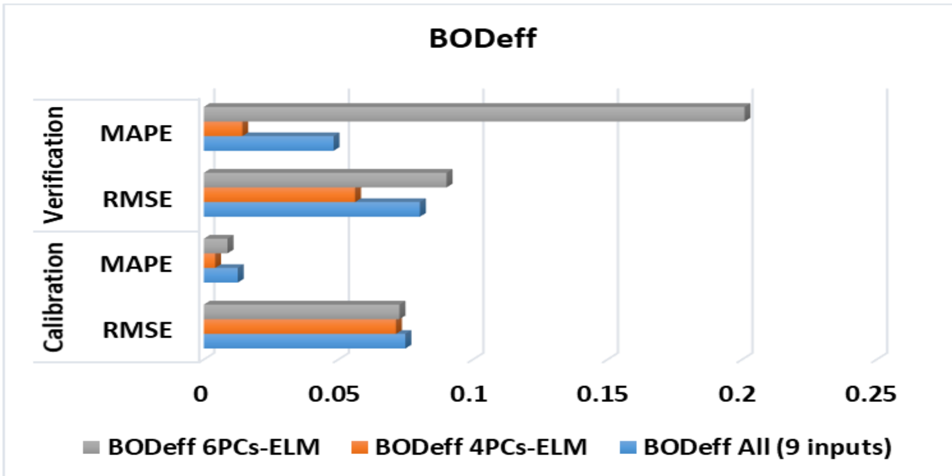
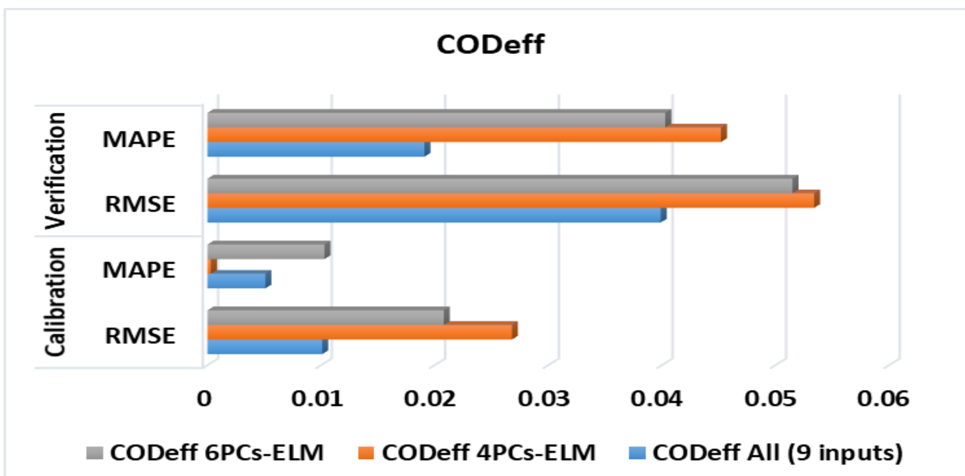
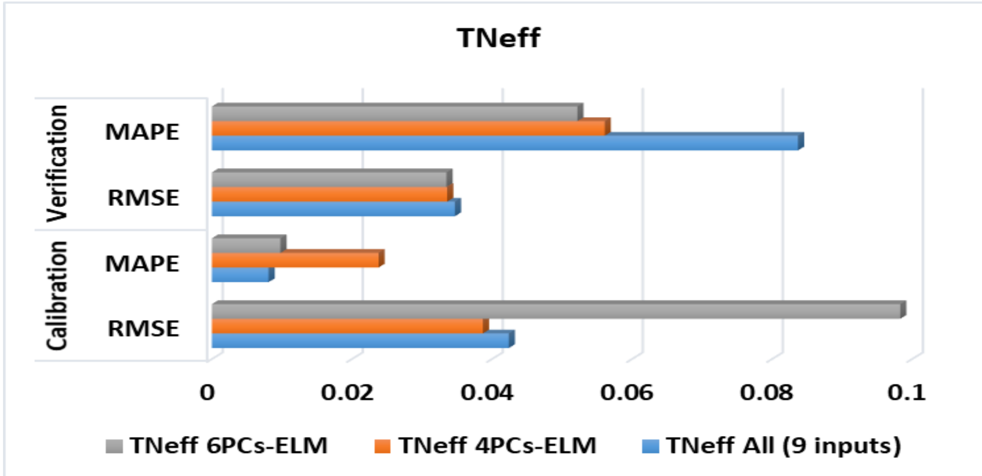


Figure 4.9: Scatter plots of observed and computed values for the best model of BODeff, CODeff, TNeff and TPeff

Figure 4.10 demonstrated the error bar chart plots for the simulated BODeff, CODeff, TNeff and TPeff in both calibration and verification phase. The from RMSE and MAPE proved the results of Table 4.4. from the Figure. 4.10 it can be seen that 4PCs-ELM has the lowest value of MAPE and RMSE for BODeff.







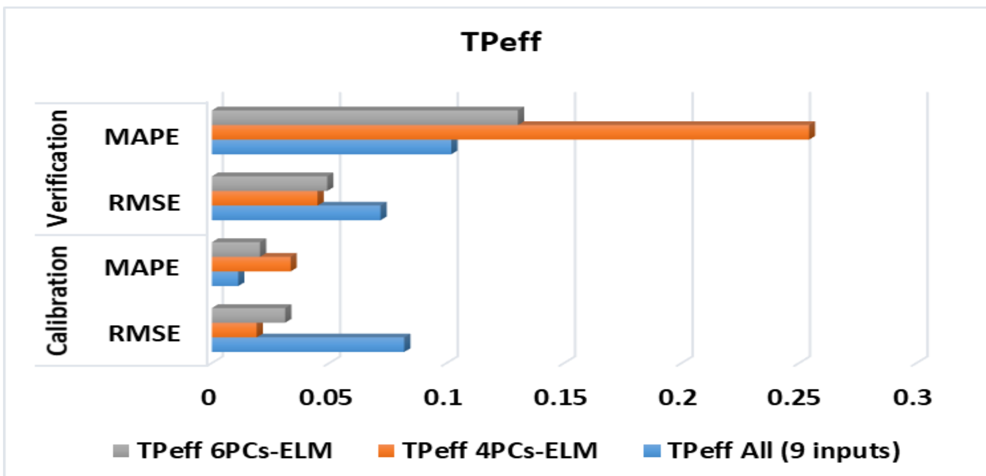


Figure 4.10: Error bar chart plots for the simulated BODeff, CODeff, TNeff and TPeff in both calibration and verification phase.

4.2.2 Implementation of scenario II

As stated above, different scenarios were constructed for multi-parametric prediction of MWWTP performance, in scenario II, MLP and MRL models were addressed according to the input variables stated in the equation above. As the tradition of any AI modelling, finding the optimal architecture is the main problems due to the fact that, there is no standard pattern for selecting the desired architecture prior to calibration phase (Kim and Singh 2014). As such, a different number of hidden neurons ranging from 1 to 30 were observed in MLP by trial and error procedure. Three different models were trained based on the scenario I in the section above. The model types were defined as MLP-M1 (4-6-1), MLP-M2 (6-6-1) and MLP-M3 (9-10-1) indicating the three-input combination set. In MLP-(4-6-1), 4 stands for a number of inputs imposed to the model, 6 indicates the hidden neuron and 1 stand for the target output of the model.

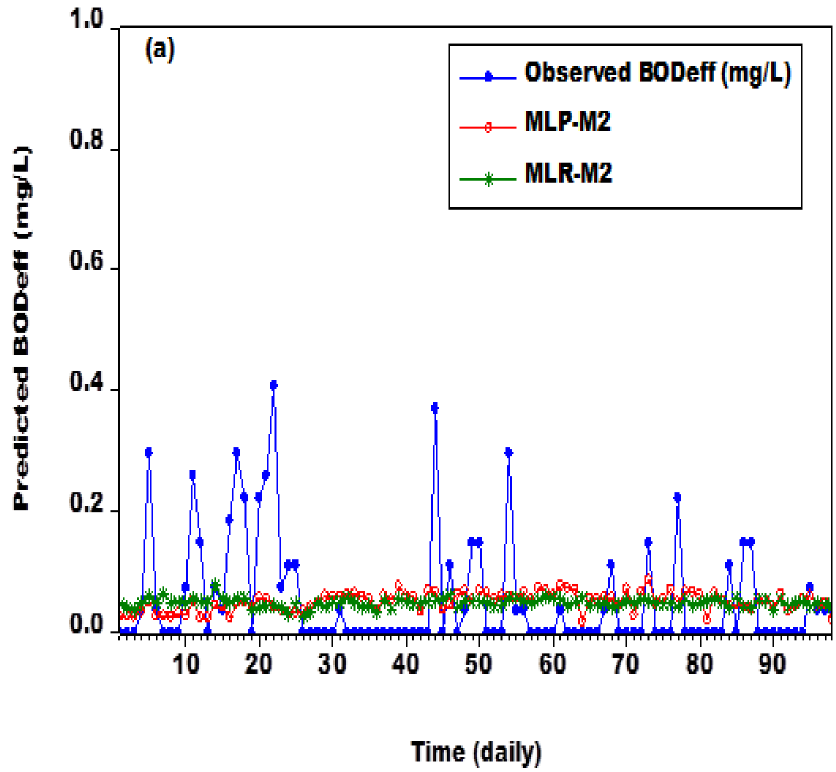
Similarly, for MLR, the model was defined as MLR-M1 (4-1). MLR-M2 (6-1) and MLR-M3 (9-1) indicating the model type, input, and output of the models. The performance indices of MLP and MLR are shown in Table 4.5. It is clearly noticed that MLP-M2 (6-6-1) and MLR-M2 (6-1) outperformed other models for modelling the BODeff while for modelling the CODeff, TNeff, and TPeff, MLP-M3 (9-10-1) and MLR-M3 (9-1) models types emerged to be the best combinations. The time series plots showing the relationship between the observed and computed values for the best of MLP and MLR models are shown in Figure.4.11.

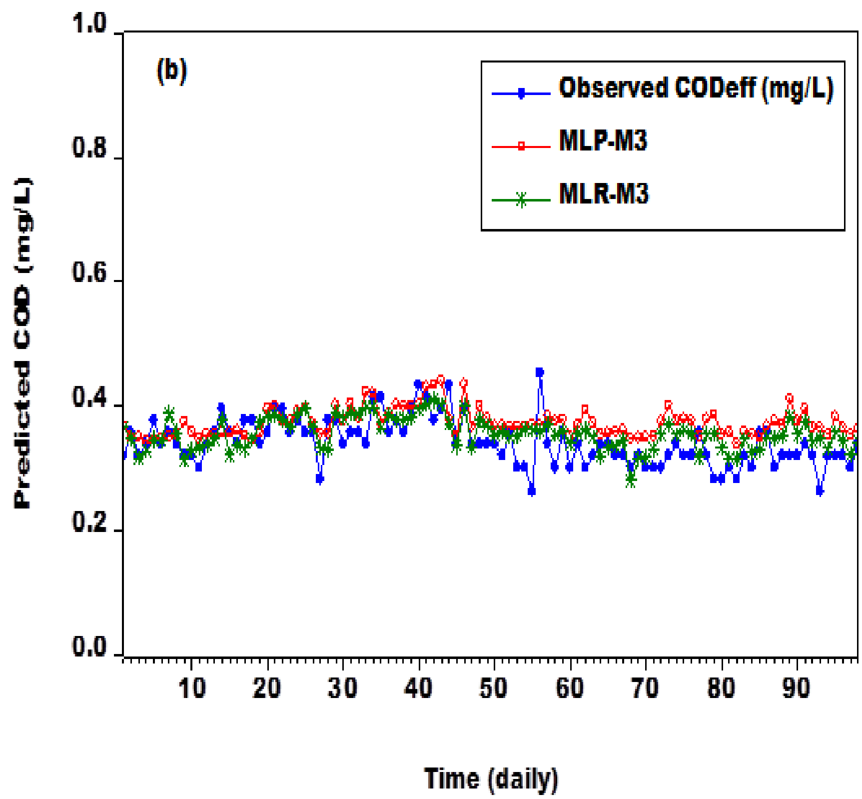
Table 4.5: Results of MLP and MLR models for BODeff, CODeff, TNeff, and TPeff

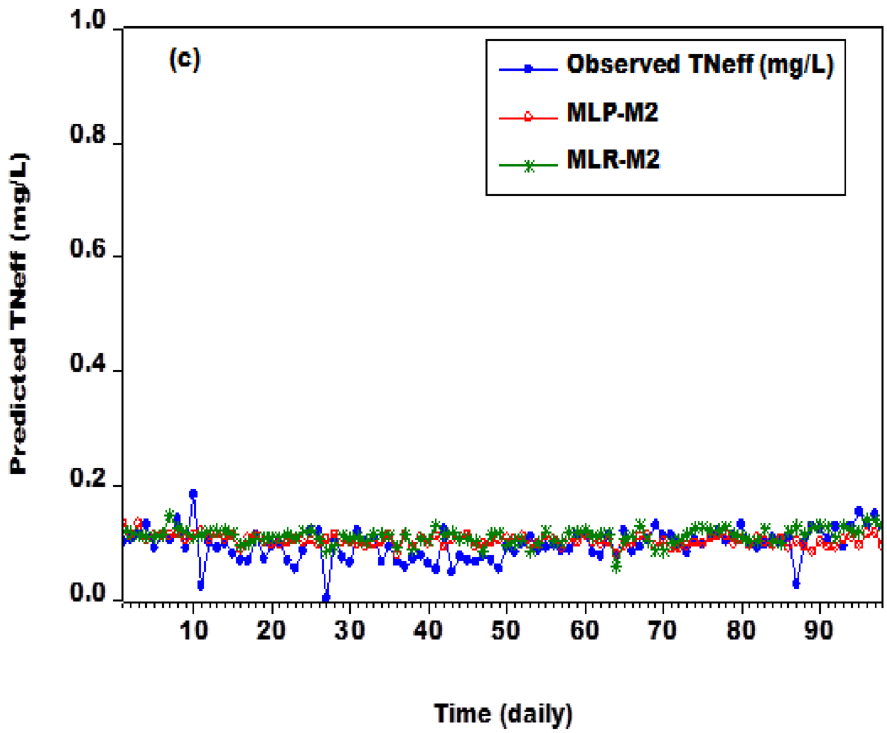
		Calibration			Verification		
	Model types	**R^2**	**RMSE**	**MAPE**	**R^2**	**RMSE**	**MAPE**
BODeff	MLP-M1 (4-6-1)	0.5473	0.1043	0.0564	0.4651	0.1093	0.0341
	MLP-M2 (6-6-1)	**0.5786**	**0.1024**	**0.0239**	**0.5776**	**0.1095**	**0.0468**
	MLP-M3 (9-10-1)	0.5331	0.1066	0.1445	0.5035	0.1091	0.1494
	MLR-M1 (4-1)	0.4775	0.1035	0.0093	0.4531	0.1093	0.0703
	MLR-M2 (6-1)	**0.5062**	**0.1034**	**0.0101**	**0.5020**	**0.1093**	**0.0757**
	MLR-M3 (9-1)	0.5005	0.1035	0.0291	0.4991	0.1091	0.2187
CODeff	MLP-M1 (4-6-1)	0.9516	0.0774	0.0116	0.9756	0.0646	0.0774
	MLP-M2 (6-6-1)	0.9599	0.0705	0.0051	0.9747	0.0747	0.0857
	MLP-M3 (9-10-1)	**0.9617**	**0.0689**	**0.0027**	**0.9555**	**0.0648**	**0.0960**
	MLR-M1 (4-1)	0.9505	0.0783	0.0094	0.9419	0.0549	0.0955
	MLR-M2 (6-1)	0.9505	0.0734	0.0088	0.9242	0.0547	0.0893
	MLR-M3 (9-1)	**0.9574**	**0.0727**	**0.0043**	**0.9552**	**0.0536**	**0.0437**
TNeff	MLP-M1 (4-6-1)	0.64026	0.08839	0.08961	0.63755	0.07613	0.55037
	MLP-M2 (6-6-1)	0.86359	0.08196	0.02441	0.81611	0.08324	0.16344
	MLP-M3 (9-10-1)	**0.87072**	**0.08028**	**0.03004**	**0.86662**	**0.08302**	**0.0799**
	MLR-M1 (4-1)	0.61499	0.08096	0.0187	0.52407	0.07361	0.21512
	MLR-M2 (6-1)	0.74987	0.08097	0.01867	0.75299	0.07358	0.21477

	MLR-M3 (9-1)	**0.76505**	**0.08044**	**0.01478**	**0.76181**	**0.07323**	**0.17002**
TPeff	MLP-M1 (4-6-1)	0.74479	0.20443	0.03042	0.73283	0.02207	0.39568
	MLP-M2 (6-6-1)	0.72657	0.20961	0.03789	0.73995	0.02087	0.30032
	MLP-M3 (9-10-1)	**0.74923**	**0.19557**	**0.00829**	**0.72544**	**0.01761**	**0.21034**
	MLR-M1 (4-1)	0.63768	0.20647	0.02897	0.29993	0.02253	0.39024
	MLR-M2 (6-1)	0.64319	0.20489	0.02841	0.35973	0.02155	0.38279
	MLR-M3 (9-1)	**0.63421**	**0.20746**	**0.0193**	**0.56072**	**0.01978**	**0.26007**

According to Table 4, the presented results indicated the improved performance accuracy of MLP with regards MLR up to 8%, 3%, 10% and 16% for BODeff, CODeff, TNeff, and TPeff, respectively. A similar conclusion was drawn by Nourani et al. (2018a) based on comparison of SVM and MLR models. Based on the employed performance indices, it is apparent that MLP demonstrated predictive skills than MLR models despite the promising ability of MLR to predict CODeff, TNeff, and TPeff. This finding was also in line with that of Zhu et al. (2019) which reported a slight performance increased of MLP over the MLR model. According to the proposed scenarios (I and II) the comparative results between Table 4.4 and 4.5, revealed that the best performance accuracy was obtained, ELM model. Hence, ELM yielded the best accuracy among all the models (MLP and MLR) in term of predictive skills.


(a)
Observed BODeff (mg/L)
MLP-M2
MLR-M2
Predicted BODeff (mg/L)
Time (daily)
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
60
70
80
90


(b)
Observed CODeff (mg/L)
MLP-M3
MLR-M3
Predicted COD (mg/L)
Time (daily)
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
60
70
80
90



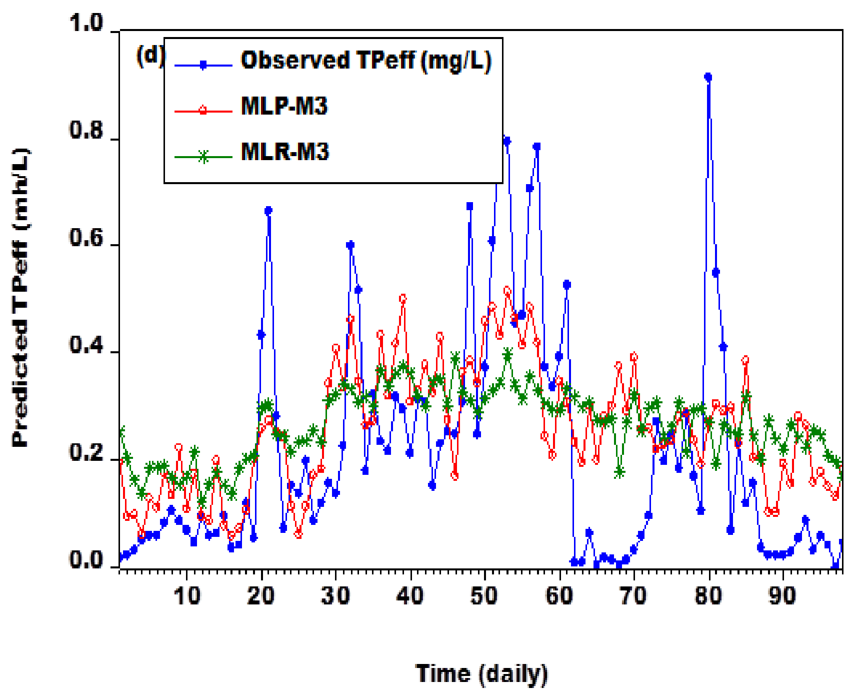
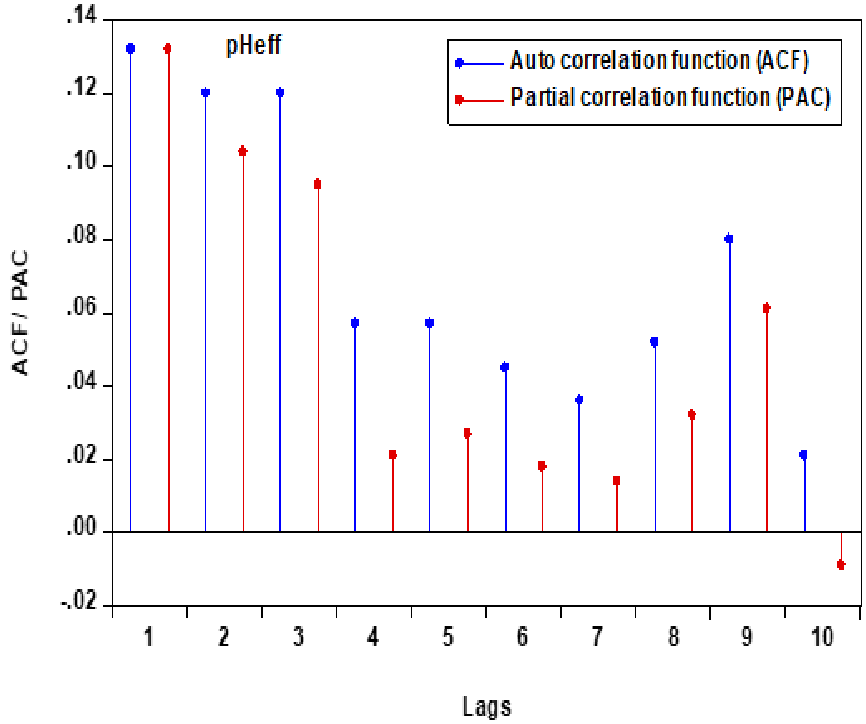


Figure 4.11: Scatter plots of observed and computed values for the best model of BODeff, CODeff, TNeff and TPeff

Further examination of the models proved that; it is obvious the ELM predicted values attained a high level of precision. The PCs-ELM increased the prediction accuracy of BODeff up 5% and 13%, CODeff up to an average of 2%, TNeff up to 10% and 20% and TPeff up to 15% and 32% with regards to MLP and MLR models, respectively. This served as the extra evidence on the capability of PCs-ELM for modelling the complex and uncertain system in MWWTP. Similarly, with the larger R^2 and smaller values of RMSE and MAPE, ELM ranked the best follow by MLP and lastly, MLR model.

4.3 Results for scenario II models (AR, HW, and NARX)

The HW and NARX models were developed using MATLAB 9.3 (R2017a) system identification toolbox based on the model configure-ration where the input and output nonlinearity estimators are both piecewise linear functions with a number of units equals to 10 as default, the complexity of the model increase by increases the number of units for HW model. Similarly, for NAXR model specify delay and number of terms in neural network regressors are chosen according to the input variables. The augmented Dickey-Full stationary test was conducted in order to meet the normality assumption of the AR model (Box and Cox 1964). Figure. 4.12 shows the Autocorrelation function (ACF), and Partial correlation function (PCF), the obtained ACF and PCF indicated the maximum number of lags (10) employed in the first analysis. Both the ACF and PCF are obtained to identify the number of the lags to be considered, the order of the AR lags is identified by using PCF. For the purpose of this research. The PCF for pH_{eff} and TSS_{eff} is considered as 4 and 6 lags. This is because the first 4 lags have the highest ACF, followed by the next two lags. Therefore, the number of developed models is equal to the lags considered (4 and 6) for each target outcome. For all the models, M1(4) represents the model with four input combination while in case of AR, it indicates the model with four number of lags.



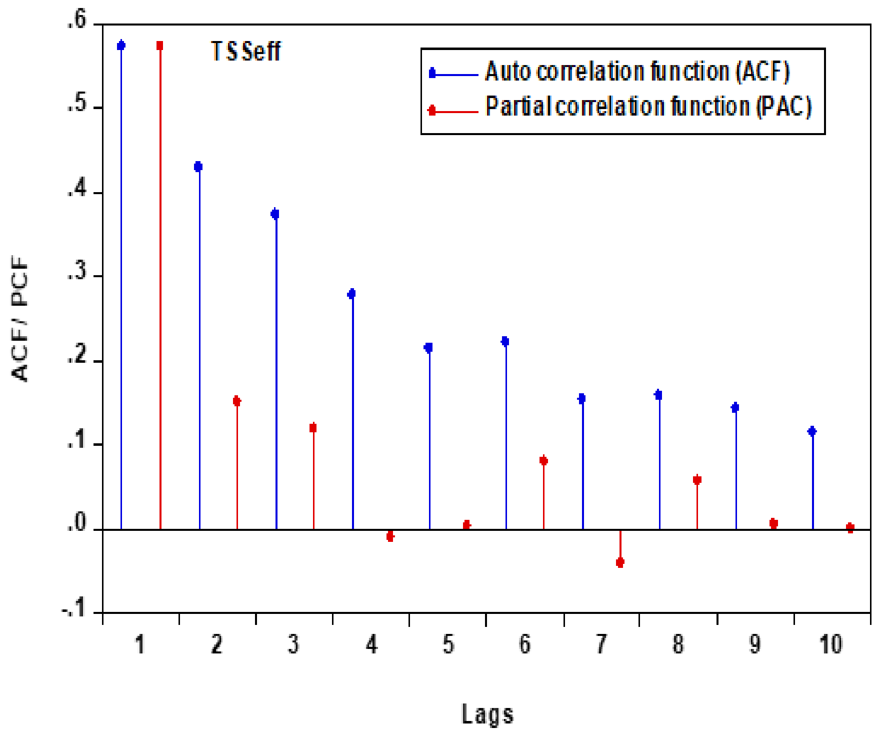


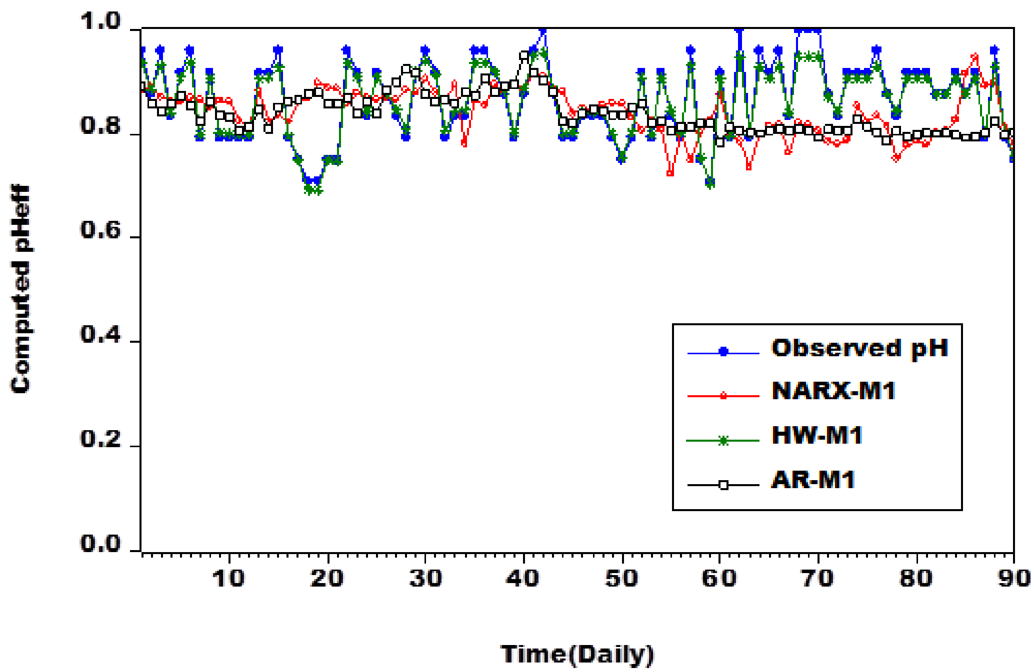
Figure 4.12: The autocorrelation and partial autocorrelation functions of pH_{eff} and TSS_{eff}

However, Table 4.6, shows the direct evaluation and comparison between the two models, it can be observed that HW and NARX model attained the highest accuracy in terms of performance indices for the estimation of pH_{eff} and TSS_{eff}, respectively. Among the model combination, M2(6) outperformed M1(4) in pH_{eff} estimation with approximately 9% and 2% in both calibration and verification, respectively. On the other hand, M1(4) emerged to be the best model for the estimation of TSSeff, with an average of 4% in both calibration and verification. The optimal AR model for both the pHeff and TSSeff is AR M1(4) consisting of 4 inputs variables and lags days. In general, NSI models are found to be close to each other, and the results are better than the linear AR model.

Table 4.6: Results of NSI models for pH_{eff} and TSS_{eff}

Effluents Parameter	**Models**	**DC**	**Calibration**			**Verification**	
			RMSE	**CC**	**DC**	**RMSE**	**CC**
pHeff	NARX-M1	0.6663	0.0136	0.8162	0.6293	0.0438	0.7932
	HW-1	**0.7416**	**0.0112**	**0.8611**	**0.7139**	**0.0322**	**0.8449**
	AR-M1	0.4187	0.0144	0.6471	0.3981	0.096	0.6310
	NARX-M2	0.5699	0.0143	0.7549	0.4812	0.0223	0.6936
	HW-M2	**0.8341**	**0.0130**	**0.9133**	**0.7355**	**0.1071**	**0.8578**
	AR-M2	0.3918	0.0144	0.6259	0.3654	0.095	0.6044
TSSeff	**NARX-M1**	**0.9864**	**0.0083**	**0.9932**	**0.9846**	**0.0093**	**0.9923**
	HW-M1	0.9540	0.0096	0.9762	0.9511	0.0073	0.9753
	AR-M1	0.9550	0.0093	0.9772	0.9306	0.0212	0.9647
	NARX-M2	**0.9852**	**0.0083**	**0.9926**	**0.9804**	**0.0049**	**0.9902**
	HW-M2	0.9758	0.0084	0.9878	0.9659	0.0097	0.9828
	AR-M2	0.9549	0.0093	0.9772	0.9199	0.0209	0.9591

The performance of the three models was also examined using some graphical presentations, such as time series, radar chart and Taylor diagrams. Figure.4.13 illustrates the time series and scatter plots of the observed versus the computed pH_{eff} and TSS_{eff} for the three best models in the verification phase. It is clear from Figure that the values are given by HW-M1, NARX-M1, AR-M1 for pH_{eff} and NARX-M1, HW-M2 and AR-M1 for TSS_{eff} are closer to the observed values and the other input combinations. It is clear also from Figure. 4.13 that the fitted values of all three models proved the superiority of 4 lags/input combinations (SS+TN+COD+NH4-N) over 6 lags/input combinations (SS+TN+COD+NH4-N+ pH+ TSS).



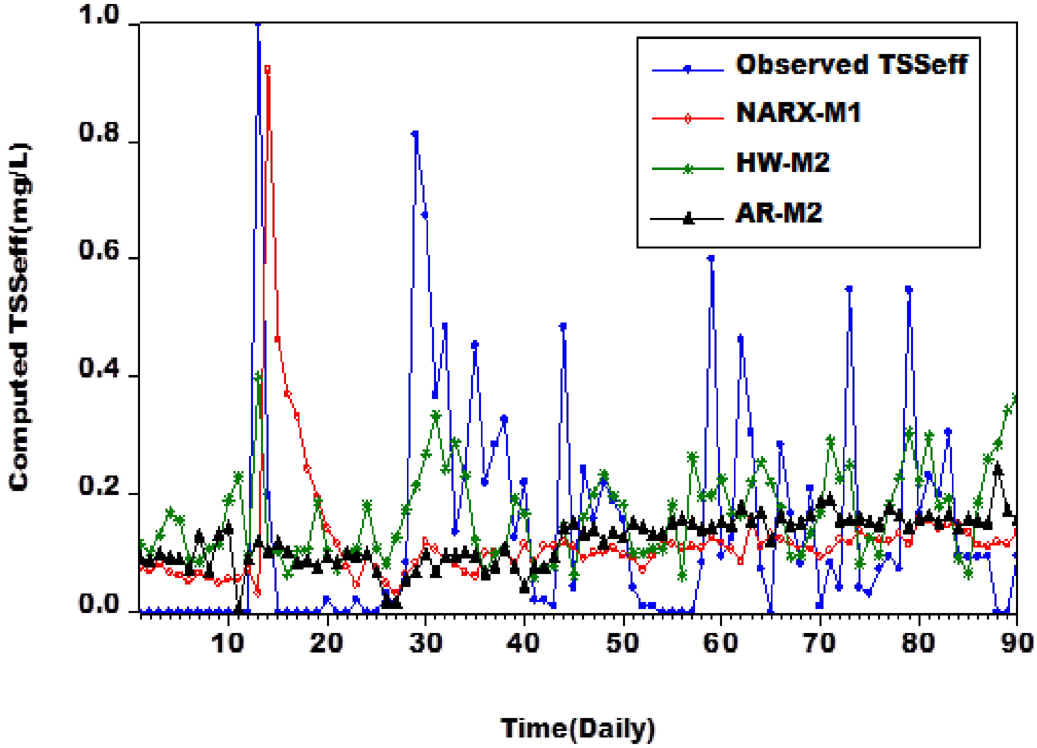


Figure 4.13: The time series plots for the best observed versus computed models for HW, NARX and AR models

A further method for diagnostic analysis of the models was employed using the Taylor diagram that has the capability to highlight the performance efficiency and accuracy of models based on the observed values. Taylor diagram provided a polar plot for acquiring a visual judgment of model performance and shows three different (i.e., correlation coefficient, normalized standard deviation, and RMSE) (Kim et al., 218). Figure 4.14 (a and b) provides the Taylor diagrams for pHeff and TSSeff, respectively.

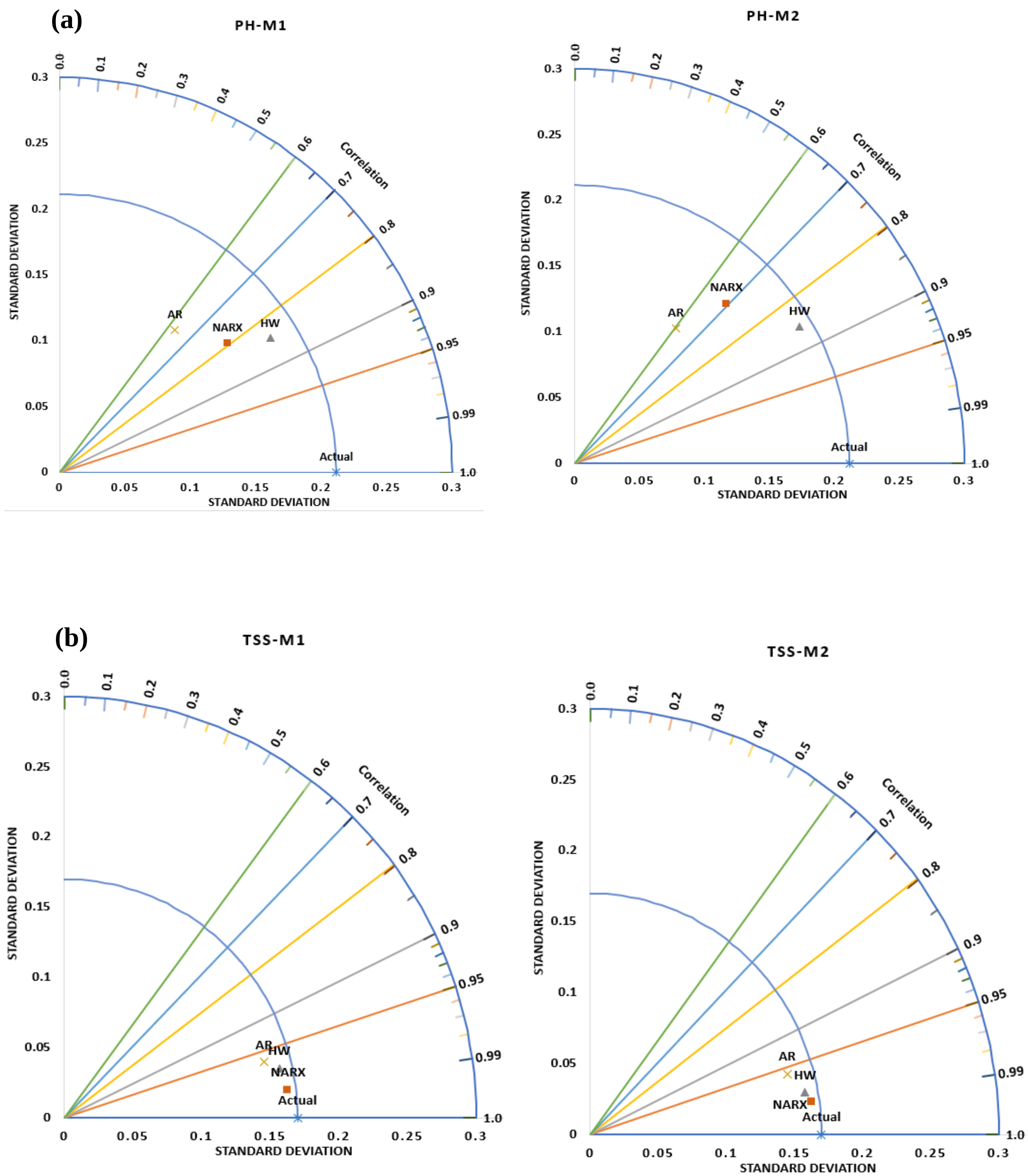
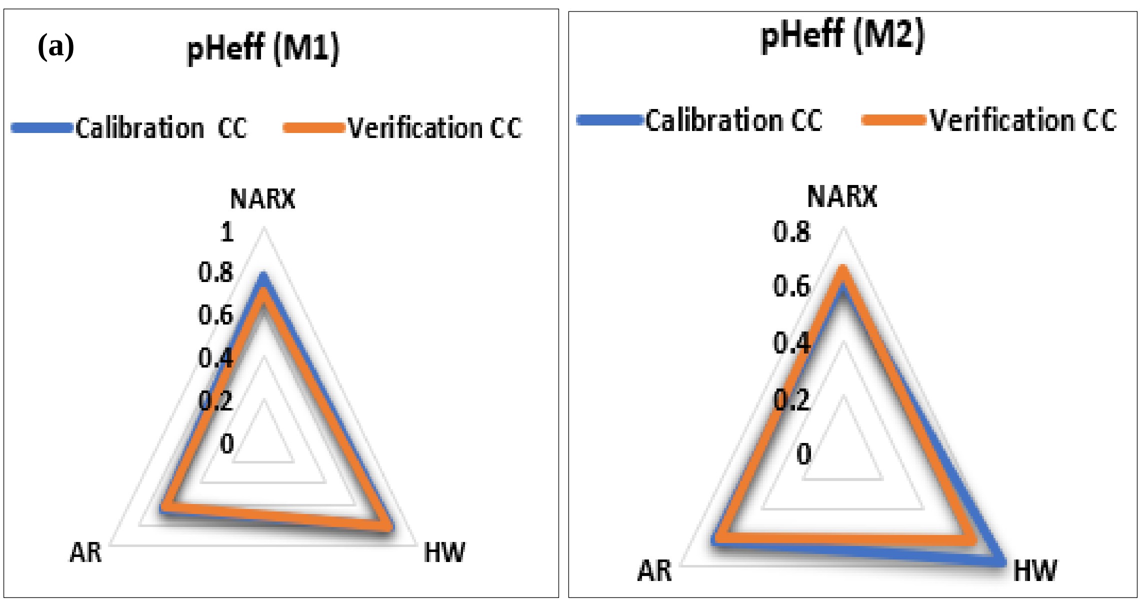


Figure 4.14: Taylor diagrams for evaluating the performances of best models for (a) pHeff (HW-M1 and M2) and (b) TSSeff (NARX-M1 and M2)

With reference to the observed values for pHeff and TSSeff, a perfect arrangement of predicted best model results can be identified for HW (M1 and M2) and NARX (M1 and M2) for pHeff and TSSeff, respectively. this plot strengthened the justification performance evaluations mentioned in Table 4.6. The predictive models are also compared in a radar chart as mentioned above to observe the high or low correlation value of each model combinations in order to perfectly display the performance of the model in terms of CC. Figure.4.15 (a and b) demonstrated the radar chart showing the different varieties of CC in both calibration and verification. From the Figure, it can be seen that 0.6044 and 0.9902 are the lowest and highest value of CC obtained from all the models in the verification phase. As it was reported in quite a lot of research that, the best performing model is attributed to the one with a high value of either CC.



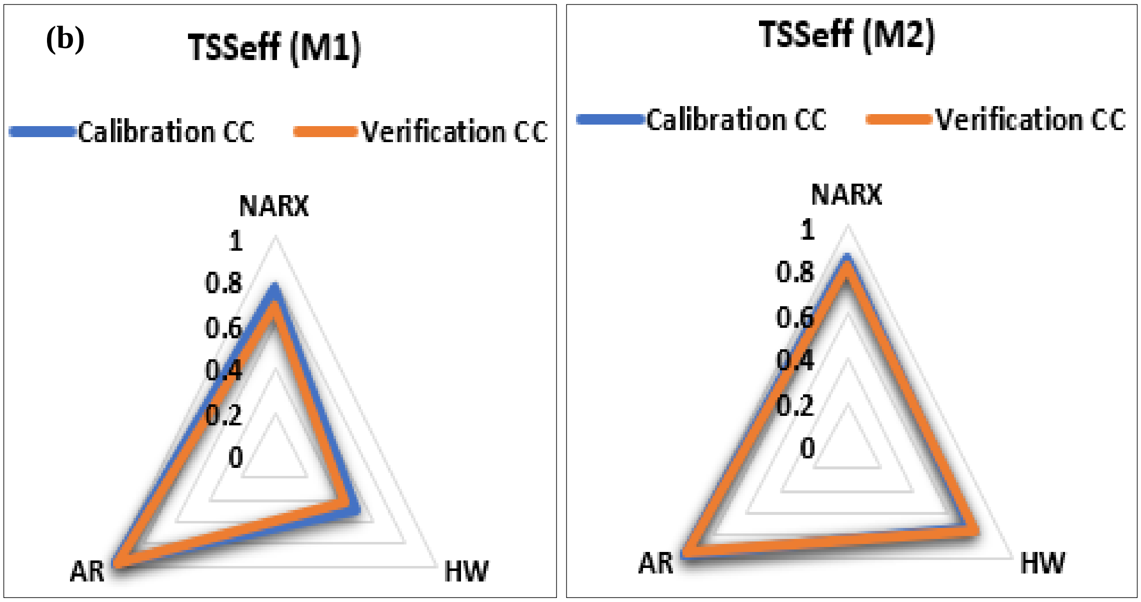


Figure 4.15: Radar chart for CC in both calibration and verification phase for (a) pHeff and (b) TSSeff

The exploratory analysis for HW and NARX models can also be justified and better visualized through boxplots (see, Figure. 4.16). Boxplots are a powerful graphical representation of data that gives an overview and a numerical summary of a data set. According Figure. 4.16, the closest of all the models to the observed values are selected to be the best model based on the mean value, the plot contained (box and whisker median, mean and staples). According to the plot, the extent of spread values between the observed and predicted models indicates that the pH_{eff} (HW-2) and TSS_{eff} (NARX-M1) ranked the best model among all the models.

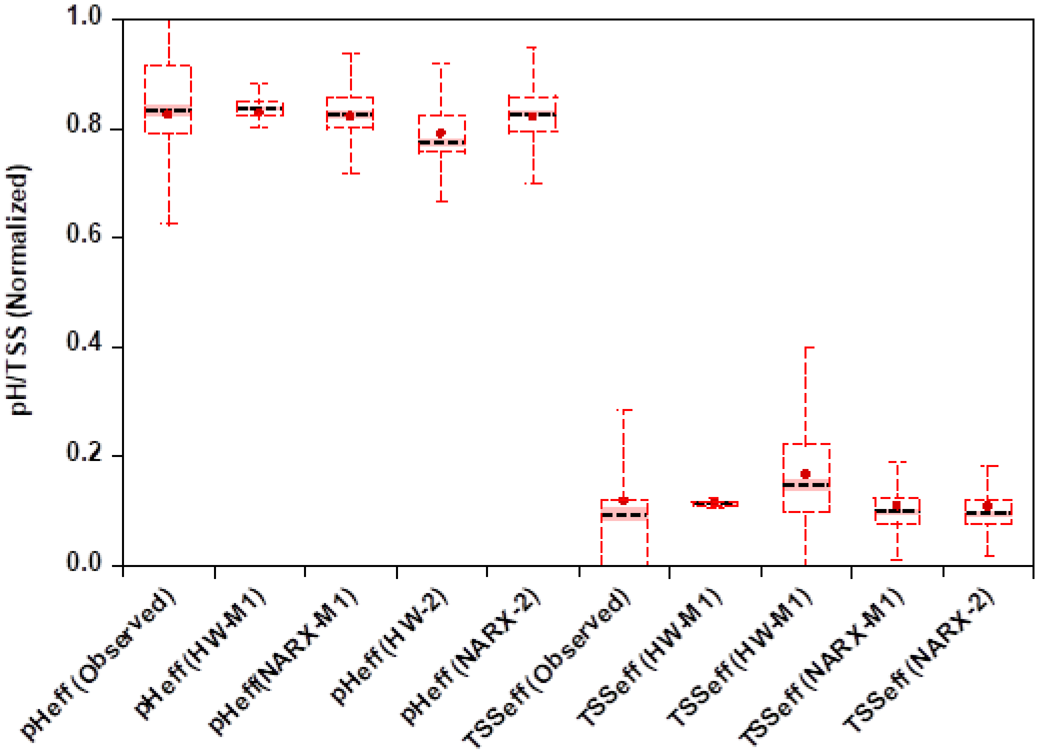
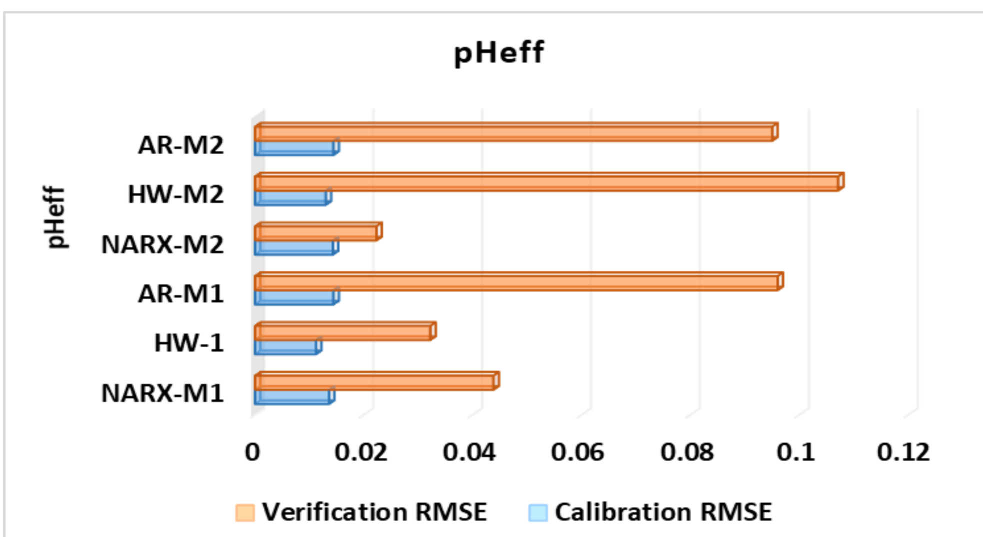


Figure 4.16: The box pots the observed and computed value of effluent pH_{eff} and TSS_{eff}



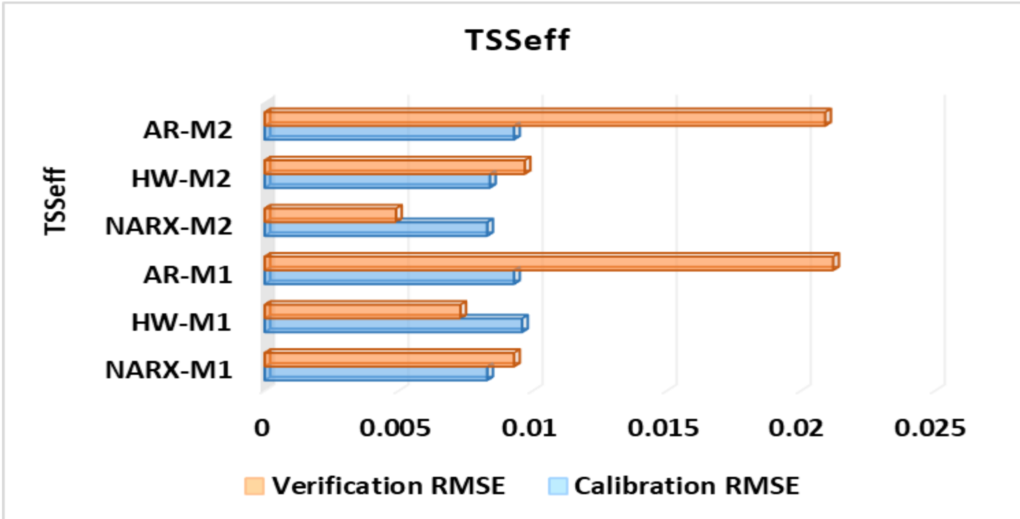


Figure 4.17: Error bar chart plots for the simulated pHeff and TSSeff in both calibration and verification phase.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion for scenario I models (FFNN, SVM, ANFIS, MLR and Ensemble)

In this paper, the performance of NWWTP was modelled by different AI models of FFNN, ANFIS, SVM, and a conventional MLR. Simple averaging, weighted and neural network ensemble techniques were subsequently employed to enhance the prediction performance of single models. For this purpose, daily data from NWWTP were obtained, and DC and RMSE were used in order to determine the prediction performance.

The comparison of single models showed that ANFIS was better than other single models in both calibration and verification phases. According to the results, SVM was found to be more reliable than the MLR model. Also, in the verification step of BOD_{eff} modelling, the models showed the more accurate performance of up to 10%, 15% and 16% with regards to FFNN, SVM, and MLR models, respectively. In the verification phase of ensemble predictions, SAE, WAE, and NNE increased the efficiency of AI modelling up to 14%, 20%, and 24%, respectively in the BOD_{eff} and about 5% for both COD_{eff} and TN_{eff} predictions. Among ensemble techniques, NNE was found to be a more robust and efficient method of combination and could improve the performance of AI modelling up to 24%.

The benefit of the NNE was due to the fact that the FFNN model has the ability to handle nonlinear behaviour in the system. According to the results obtained so far, firstly, single models should not be considered as a reliable model for the simulation of BOD_{eff} in NWWTP, as it proved fair results for all AI models. Secondly, all AI and classical models employed in this study were found to be satisfactory and therefore, recommended for the simulation of COD_{eff} and TN_{eff}. Thirdly, the NWWTP performance indicated the high quality of treated effluent, which can be used for irrigation and other re-use purposes and the ensemble results provide more reliable and promising results than the single models.

Finally, the study may serve as the background for researchers carrying out further studies in NWWTP. The outcomes also suggested that for the application of these models in the real world, the uncertainty involved in the process could be addressed. As such, the application of other AI tools may also be combined in the proposed ensemble approach in order to integrate a set of models so as to come up with a new model which could produce higher accuracy and more reliable estimates than the single models.

5.2 Conclusion for scenario II models (ELM, MLP and PCA)

In this research, two scenarios I and II were investigated for modelling the performance of Nicosia MWWTP in term of effluents BOD_{eff}, COD_{eff}, TN_{eff} and TP_{eff} using three different model input combinations. The extreme learning machine (ELM) as a newly emerged black-box model with a combined principal component analysis (PCA) was developed in the scenario I while in scenario II, traditional multilayer perceptron (MLP) neural network and multiple linear regression (MLR) models were established for comparison.

In the scenario I, PCA was employed in this study to understand whether it can be feasible to improve the accuracy of the emerging ELM algorithms. The PCA technique helps the ELM mapping by its orthogonal transformation of variables and reduction of system dimensionality. The obtained result showed the increase for PCs-ELM of about 12%, 2%, 20% and 6% for BOD_{eff}, COD_{eff}, TN_{eff} and TP_{eff} with regards to the novel ELM model. Nevertheless, the ELM model demonstrated accurate prediction capability and can also serve as a reliable tool. On the other hand, PCA algorithms can be employed to reduce the dimensionality of the input vectors, which may lead to achieving highly accurate prediction.

For scenario II, MLP and MRL models were addressed according to the input variables of first scenario and the results indicated the improved performance accuracy of MLP with regards MLR up to 8%, 3%, 10% and 16% for BODeff, CODeff, TNeff, and TPeff, respectively. According to the two scenarios, the comparative results revealed that the best performance accuracy was obtained by considering the inputs combination models ELM. Hence, ELM yielded the best accuracy among all the models (MLP and MLR) in term of predictive skills. The outcomes of the current study may contribute to the mentioned multi-parametric modelling of the treated effluents and provides a reference benchmark for wastewater management and control. It's suggested that other algorithms may be applied with the combination of PCs so as to come up with a new model which could produce higher accuracy and more reliable estimates.

5.3 Conclusion for scenario II models (AR, HW, and NARX)

A nonlinear system identification models have been found the promising tool for the estimation of highly nonlinear processes. The prime goal of this paper was to discover and employed two different NSI models viz: Hammerstein- Weiner Model (HW) and Nonlinear autoregressive with exogenous (NARX) neural network model with the classical linear method known as autoregressive (AR) model, for the estimation of effluents characteristic of total suspended solids (TSS_{eff}) and pH_{eff} from New Nicosia MWWTP. The results were evaluated in terms of widely used performance criteria (DC, RMSE, and CC).

The estimation results demonstrated that HW model statistically outperformed NARX in estimating the pH_{eff} while for TSS_{eff} NARX model performed better than the HW model. It was evident that the accuracy of the HW increased averagely up to 18% with regards to NARX model for pH_{eff}. Likewise, the TSS performance increased averagely up to 25% with regards to HW model. For comparison with the traditional AR, the results indicated that both HW and NARX are more accurate than the AR model. Hence, the outcomes determined that the NSI model (HW and NARX) are reliable modelling tool that could be adopted for the simulation of pH_{eff} and TSS_{eff}, respectively. The results also suggest that other non-linear techniques should also be considered in order to enhance the estimation accuracy of the model.

REFERENCES

Abba, S. I., And Elkiran, G. (2017). Effluent Prediction Of Chemical Oxygen Demand From The Wastewater Treatment Plant Using Artificial Neural Network Application. *Procedia Computer Science*, *120*, 156–163.

Abba, S. I., Hadi, S. J., And Abdullahi, J. (2017). River Water Modelling Prediction Using Multi-Linear Regression, Artificial Neural Network, And Adaptive Neuro-Fuzzy Inference System Techniques. *Procedia Computer Science*, *120*, 75–82.

Abdullahi, J., and Elkiran, G. (2017). Prediction of the future impact of climate change on reference evapotranspiration in Cyprus using an artificial neural network. *Procedia Computer Science*, *120*, 276-283.

Akrami, S. A., Nourani, V., and Hakim, S. J. S. (2014). Development of a nonlinear model based on wavelet-ANFIS for rainfall forecasting at Klang Gates Dam. *Water resources management*, *28*(10), 2999-3018.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000). Artificial neural networks in hydrology. II: Hydrologic applications. *Journal of Hydrologic Engineering*, *5*(2), 124-137.

Acikel, S., and Ekmekci, M. (2018). Assessment of groundwater quality using multivariate statistical techniques in the Azmak Spring Zone, Mugla, Turkey. *Environmental earth sciences*, *77*(22), 753.

Adamowski, J, Fung Chan, H, Prasher, SO, Ozga-Zielinski, B, Sliusarieva, A (2012) Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*, *48*(1).

Al-Sbou, YA, Alawasa, KM (2017) Nonlinear Autoregressive Recurrent Neural Network Model for Solar Radiation Prediction. *International Journal of Applied Engineering Research*, *12*(14), 4518-4527.

Anowar, F., and Sadaoui, S. (2019, May). Multi-class Ensemble Learning of Imbalanced Bidding Fraud Data. In *Canadian Conference on Artificial Intelligence* (pp. 352-358). Springer, Cham.

Areerachakul, S. (2012). Comparison of ANFIS and ANN for estimation of biochemical oxygen demand parameter in surface water. *International Journal of Chemical and Biological Engineering*, *6*, 286-290.

Baba, N. M., Makhtar, M., Fadzli, S. A., and Awang, M. K. (2015). Current Issues in Ensemble Methods and Its Applications. *Journal of Theoretical and Applied Information Technology*, *81*(2), 266.

Box, GEP, Cox, DR. (1964) An analysis of transformations. JR Statist. Soc. *E 26, 211, 252*.

Boussaada, Z, Curea, O, Remaci, A, Camblong, H, Mrabet Bellaaj, N (2018) A nonlinear autoregressive exogenous (narx) neural network model for the prediction of the daily direct solar radiation. *Energies*, *11*(3), 620.

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge university press.

Boyd, G., Na, D., Li, Z., Snowling, S., Zhang, Q., & Zhou, P. (2019). Influent forecasting for wastewater treatment plants in North America. *Sustainability (Switzerland)*, *11*(6), 1–14. https://doi.org/10.3390/su11061764

Chen, W. B., and Liu, W. C. (2015). Water quality modeling in reservoirs using multivariate linear regression and two neural network models. *Advances in Artificial Neural Systems*, *2015*, 6.

Civelekoglu, G., Yigit, N. O., Demopoulos, E., and Kitis, M. (2009). Modelling of COD removal in a biological wastewater treatment plant using adaptive neuro-fuzzy inference system and artificial neural network. *Water Science and Technology*, *60*(6), 1475-1487.

Chen, W.C., Chang, N.B., Shieh, W.K., 2001. Advanced hybrid fuzzy-neural controller for industrial wastewater treatment. *Journal of Environmental Engineering* 127 (11), 1048–1059.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Djeddou, M., and Achour, B. (2015). The use of a neural network technique for the prediction of sludge volume index in municipal wastewater treatment plant. *Larhyss journal issn 1112-3680*, (24), 351-370.

Dogan, E., Ates, A., Yilmaz, E. C., and Eren, B. (2008). Application of artificial neural networks to estimate wastewater treatment plant inlet biochemical oxygen demand. *Environmental Progress and Sustainable Energy*, *27*(4), 439-446.

Di Piazza, MC, Vitale, G (2016) Solar and wind forecasting by NARX neural networks. *Renewable Energy and Environmental Sustainability*, *1*, 39.

Elkiran, G., Nourani, V., Abba, S. I., and Abdullahi, J. (2018). Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river. *Global Journal of Environmental Science and Management*, *4*(4), 439-450.

Elkiran, G., Nourani, V., and Abba, S. I. (2019). Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach. *Journal of Hydrology*, 123962.

Eisinga, R, Grotenhuis, Pelzer, M (2013) The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown. *International journal of public health*, *58*(4), 637-642.

Erdirencelebi, D., & Yalpir, S. (2011). Adaptive network fuzzy inference system modeling for the input selection and prediction of anaerobic digestion effluent quality. *Applied Mathematical Modelling*, *35*(8), 3821-3832.

Frazão X., Alexandre L.A. (2014) Weighted Convolutional Neural Network Ensemble. In: Bayro-Corrochano E., Hancock E. (eds) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2014. Lecture Notes in Computer Science, vol 8827. Springer, Cham

Gaya, M. S., Wahab, N. A., Sam, Y. M., and Samsudin, S. I. (2014). ANFIS Modelling of Carbon and Nitrogen Removal in Domestic. *J. Teknol*, *67*(5), 29-34.

Gaya, M. S., Yusuf, L. A., Mustapha, M., Muhammad, B., Sani, A., Tijjani, A. T. A., ... and Khairi, M. T. M. (2017). Estimation of turbidity in water treatment plant using Hammerstein-Wiener and neural network technique. *Indonesian Journal of Electrical Engineering and Computer Science*, *5*(3), 666-672.

Gómez, T., Gémar, G., Molinos-Senante, M., Sala-Garrido, R., and Caballero, R. (2017). Assessing the efficiency of wastewater treatment plants: A double-bootstrap approach. *Journal of Cleaner Production*, *164,* 315-324.

Govindaraju, R. S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of hydrologic Engineering*, *5*(2), 115-123.

Granata, F., Papirio, S., Esposito, G., Gargano, R., and de Marinis, G. (2017). Machine learning algorithms for the forecasting of wastewater quality indicators. *Water*, *9*(2), 105.

Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J. P., and Cho, K. H. (2015). Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences*, *32*, 90-101.

Gozen, E, Turkman, A (2008) Water scarcity impacts on Northern Cyprus and alternative mitigation strategies. In *Environmental Problems of Central Asia and their Economic, Social and Security Impacts* (pp. 241-250). Springer, Dordrecht.

Guo, F (2004) A new identification method for Wiener and Hammerstein systems (p.98). For schungszentrum Karlsruhe.

Gulyani, B. B., and Fathima, A. (2017). Introducing Ensemble Methods to Predict the Performance of Waste Water Treatment Plants (WWTP). *International Journal of Environmental Science and Development*, *8*(7), 501.

Güçlü, D., & Dursun, Ş. (2010). Artificial neural network modelling of a large-scale wastewater treatment plant operation. *Bioprocess and biosystems engineering*, *33*(9), 1051-1058.

Güçlü, D., & Dursun, S. (2008). Amelioration of carbon removal prediction for an activated sludge process using an artificial neural network (ANN). *CLEAN–Soil, Air, Water*, *36*(9), 781-787.

Haghiabi, A. H., Azamathulla, H. M., and Parsaie, A. (2017). Prediction of head loss on cascade weir using ANN and SVM. *ISH Journal of Hydraulic Engineering*, *23*(1), 102-110.

Hamoda, M. F., Al-Ghusain, I. A., & Hassan, A. H. (1999). Integrated wastewater treatment plant performance evaluation using artificial neural networks. *Water Science and Technology*, *40*(7), 55-65.

Hanbay, D., Turkoglu, I., & Demir, Y. (2008). Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. *Expert Systems with Applications, 34*(2), 1038-1043.

Hamed, M. M., Khalafallah, M. G., and Hassanien, E. A. (2004). Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling and Software, 19*(10), 919-928.

Huang, Z., Luo, J., Li, X., & Zhou, Y. (2009). Prediction of effluent parameters of wastewater treatment plant based on improved least square support vector machine with PSO. *2009 1st International Conference on Information Science and Engineering, ICISE 2009*, 4058–4061. https://doi.org/10.1109/ICISE.2009.846

Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A., and Elshafie, A. (2017). Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications, 28*(1), 893-905.

Hassen, E. B., & Asmare, A. M. (2019). Predictive performance modeling of Habesha brewery wastewater treatment plant using artificial neural network.

Hoşkara, Ş. Ö., and Doratlı, N. (2007). A Critical Evaluation of The Issue Of "Conservation Of The Cultural Heritage" In North Cyprus.

Hong, Y., Fei, L., Yuge, X., and Jin, L. (2008, July). GA based LS-SVM classifier for waste water treatment process. In *Control Conference, 2008. CCC 2008. 27th Chinese* (pp. 436-439). IEEE.

Hong, Z., Bednarek, S. Y., Blumwald, E., Hwang, I., Jurgens, G., Menzel, D., ... & Verma, D. P. S. (2003). A unified nomenclature for Arabidopsis dynamin-related large GTPases based on homology and possible functions. *Plant molecular biology, 53*(3), 261-265.

Huang G, Bin, Song S, You K (2015) Trends in extreme learning machines: A review. Neural Networks 61:32–48. doi: 10.1016/j.neunet.2014.10.001

Holland, S. M. (2008). Principal components analysis (PCA). *Department of Geology, University of Georgia, Athens, GA*, 30602-2501.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. Extreme learning machine: Theory and applications. Neurocomputing 70 (1), 489–501. Neucom.2005.12.126.

Hou, M., Zhang, T., Weng, F., Ali, M., Al-Ansari, N., and Yaseen, Z. (2018). Global solar radiation prediction using hybrid online sequential extreme learning machine model. *Energies*, *11*(12), 3415.

Hornik K (1991) Approximation capabilities of multilayer Feedforward networks. *Neural Netw* 4(2):251–257.

Hadi, S J, Tombul, M (2018) Forecasting Daily Streamflow for Basins with Different Physical Characteristics through Data-Driven Methods. *Water resources management*, *32*, 3405-3422.

Hasanlou, H., Abdolabadi, H., and Aghashahi, M. (2015). Application of factor analysis in a large-scale industrial wastewater treatment plant simulation using principal component analysis–artificial neural network hybrid approach:(Case Study: Fajr Industrial Wastewater Treatment Plant, Mahshahr, Iran). *Environmental Progress and Sustainable Energy*, *34*(5), 1322-1331.

Jami, M. S., Husain, I. A., Kabashi, N. A., & Abdullah, N. (2012). Multiple inputs artificial neural network model for the prediction of wastewater treatment plant performance. *Australian Journal of Basic and Applied Sciences*, *6*(1), 62-69.

Jain, A. K,. and Punmia B. C. (2014). *Wastewater Engineering (Including Air Pollution)*. (B. Pumia, Ed.) (Second). India: Laxmi Publications (P) LTD.

Jeihouni, E., Eslamian, S., Mohammadi, M., and Zareian, M. J. (2019). Simulation of groundwater level fluctuations in response to main climate parameters using a wavelet–ANN hybrid technique for the Shabestar Plain, Iran. *Environmental Earth Sciences*, *78*(10), 293.

Kazienko, P., Lughofer, E., and Trawiński, B. (2013). Hybrid and ensemble methods in machine learning J. UCS special issue. *J. Univers. Comput. Sci*, *19*(4), 457-461.

Kajdanowicz, T., and Kazienko, P. (2013). Boosting-based Multi-label Classification. *J. UCS*, *19*(4), 502-520.

Khademi, F., and Behfarnia, K. (2016). Evaluation of concrete compressive strength using artificial neural network and multiple linear regression models. *Iran University of Science and Technology*, *6*(3), 423-432.

Khan, Y., and Chai, S. S. (2017). Ensemble of ANN and ANFIS for Water Quality Prediction and Analysis-A Data Driven Approach. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2-9), 117-122.

Kim S, Singh VP (2014) Modeling daily soil temperature using data driven models and spatial distribution. *Theor Appl Climatol* 118(3):465–479.

Kişi, Ö (2008) Stream flow forecasting using neuro-wavelet technique. *Hydrological Processes: An International Journal*, *22*(20), 4142-4152.

Kim, S, Seo, Y, Rezaie-Balf, M, Kisi, O, Ghorbani, MA, and Singh, VP (2018) Evaluation of daily solar radiation flux using soft computing approaches based on different meteorological information: peninsula vs continent. *Theoretical and Applied Climatology*, 1-20.

Lermontov, A, Yokoyama, L, Lermontov, M, Machado, MA S (2009) River quality analysis using fuzzy water quality index: Ribeira do Iguape River watershed, Brazil. *Ecological Indicators*, *9*(6), 1188-1197.

Ljung, L (1995) System identification toolbox*: User's guide*. Natick, MA: MathWorks Incorporated.

Li-juan, W., & Chao-bo, C. (2008, December). Support vector machine applying in the prediction of effluent quality of sewage treatment plant with cyclic activated sludge system process. In *2008 IEEE International Symposium on Knowledge Acquisition and Modeling Workshop* (pp. 647-650). IEEE.

Liu, H., Huang, M., & Yoo, C. (2013). A fuzzy neural network-based soft sensor for modeling nutrient removal mechanism in a full-scale wastewater treatment system. *Desalination and Water Treatment*, *51*(31-33), 6184-6193.

Memon, N A., Unar, M A, Ansari, A K. (2012) pH prediction by artificial neural networks for the drinking water of the distribution system of Hyderabad City. *Mehran University Research Journal of Engineering and Technology, 31(1) 137-146.*

Men, Z, Yee, E, Lien, FS, Yang, Z, Liu, Y (2014) Ensemble nonlinear autoregressive exogenous artificial neural networks for short-term wind speed and power forecasting. *International scholarly research notices*, *2014*.

Mehr, A D, Nourani, V, Kahya, E, Hrnjica, B, Sattar, AM., Yaseen, ZM (2018) Genetic programming in water resources engineering: A state-of-the-art review. *Journal of hydrology*.

Mamdani, E. H. (1974, December). Application of fuzzy algorithms for control of simple dynamic plant. In *Proceedings of the institution of electrical engineers* (Vol. 121, No. 12, pp. 1585-1588). IET.

Maleki, A., Nasseri, S., Aminabad, M. S., and Hadi, M. (2018). Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics. *KSCE Journal of Civil Engineering*, 1-13.

Mjalli, F. S., Al-Asheh, S., & Alfadala, H. E. (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management*, *83*(3), 329-338.

Mesdaghinia, A., Nasseri, S., Mahvi, A. H., Tashauoei, H. R., and Hadi, M. (2015). The estimation of per capita loadings of domestic wastewater in Tehran. *Journal of Environmental Health Science and Engineering*, *13*(1), 25.

Mikosz, J. (2016). Analysis of greenhouse gas emissions and the energy balance in a model municipal wastewater treatment plant. *Desalination and Water Treatment*, *57*(59), 28551-28559.

Moreno-Alfonso, N., & Redondo, C. F. (2001). Intelligent waste-water treatment with neural-networks. *Water policy*, *3*(3), 267-271.

Mushiri, T., Manjengwa, K., & Mbohwa, C. (2014). Advanced Fuzzy Control in Industrial Wastewater Treatment (pH and Temperature Control). In *Proceedings of the World Congress on Engineering* (Vol. 1).

Nazir, H. M., Hussain, I., Faisal, M., Shoukry, A. M., Gani, S., and Ahmad, I. (2019). Development of Multidecomposition Hybrid Model for Hydrological Time Series Analysis. *Complexity*, *2019*.

Nasr, M. S., Moustafa, M. A., Seif, H. A., and El Kobrosy, G. (2012). Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. *Alexandria engineering journal*, *51*(1), 37-43.

Nourani, V. (2017). An Emotional ANN (EANN) approach to modelling rainfall-runoff process. *Journal of Hydrology, 544*, 267-277.

Nourani, V., Hakimzadeh, H., and Amini, A. B. (2012). Implementation of artificial neural network technique in the simulation of dam breach hydrograph. *Journal of Hydroinformatics, 14*(2), 478-496.

Nourani, V., Khanghah, T. R., and Sayyadi, M. (2013). Application of the Artificial Neural Network to monitor the quality of treated water. *International Journal of Management and Information Technology, 3*(1), 39-45.

Nourani, V., Kisi, Ö., and Komasi, M. (2011). Two-hybrid artificial intelligence approaches for modelling rainfall-runoff process. *Journal of Hydrology*, *402*(1-2), 41-59.

Nourani, V., Andalib, G., and Sadikoglu, F. (2017). Multi-station streamflow forecasting using wavelet denoising and artificial intelligence models. *Procedia Computer Science*, *120*, 617-624.

Nourani, V., Elkiran, G., and Abba, S. I. (2018a). Wastewater treatment plant performance analysis using artificial intelligence–an ensemble approach. *Water Science and Technology*.

Nourani, V., Elkiran, G., Abdullahi, J., and Tahsin, A (2018b). Multi-region Modeling of Daily Global Solar Radiation with Artificial Intelligence Ensemble. *Natural Resources Research*, 1-22.

Noori, R., Abdoli, M. A., Ghasrodashti, A. A., and Jalili Ghazizade, M. (2009). Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: a case study of Mashhad. *Environmental Progress and*

Sustainable Energy: An Official Publication of the American Institute of Chemical Engineers, *28*(2), 249-258.

Nourani, V., Elkiran, G., and Abdullahi, J. (2019). Multi-station artificial intelligence-based ensemble modelling of reference evapotranspiration using pan evaporation measurements. *Journal of Hydrology*, 123958.

Oliveira-Esquerre, K. P., Mori, M., & Bruns, R. E. (2002). Simulation of an industrial wastewater treatment plant using artificial neural networks and principal components analysis. *Brazilian Journal of Chemical Engineering*, *19*(4), 365-370.

Parmar, K. S., and Bhardwaj, R. (2015). River water prediction modelling using neural networks, fuzzy and wavelet coupled model. *Water resources management*, *29*(1), 17-33.

Pai, T. Y., Wan, T. J., Hsu, S. T., Chang, T. C., Tsai, Y. P., Lin, C. Y., ... & Yu, L. F. (2009). Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent. *Computers & Chemical Engineering*, *33*(7), 1272-1278.

Pai, T. Y., Wan, T. J., Hsu, S. T., Chang, T. C., Tsai, Y. P., Lin, C. Y., ... & Yu, L. F. (2009). Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent. *Computers & Chemical Engineering*, *33*(7), 1272-1278.

Pisa, I., Santín, I., Lopez Vicario, J., Morell, A., & Vilanova, R. (2018). A recurrent neural network for wastewater treatment plant effuents' prediction. *Actas de las XXXIX Jornadas de Automática, Badajoz, 5-7 de Septiembre de 2018*.

Perendeci, A., Arslan, S., Tanyolaç, A., & Çelebi, S. S. (2007). Evaluation of input variables in adaptive-network-based fuzzy inference system modeling for an anaerobic wastewater treatment plant under unsteady state. *Journal of Environmental Engineering*, *133*(7), 765-771.

Pakrou, S., Mehrdadi, N., & Baghvand, A. (2014). Artificial Neural Networks Modeling for Predicting Treatment Efficiency and Considering Effects of Input Parameters in Prediction Accuracy: a Case Study in Tabriz Treatment Plant. *Indian Journal of Fundamental and Applied Life Sciences*, *4*, 2231–6345.

Pakrou, S., Mehrdadi, N., & Baghvand, A. (2015). ANN Modeling to Predict the COD and Efficiency of Waste Pollutant Removal from Municipal Wastewater Treatment Plants. *Current World Environment*, *10*(Special-Issue1), 873–881. https://doi.org/10.12944/cwe.10.special-issue1.106

Raduly, B., Gernaey, K. V., Capodaglio, A. G., Mikkelsen, P. S., & Henze, M. (2007). Artificial neural networks for rapid WWTP performance evaluation: Methodology and case study. *Environmental Modelling & Software*, *22*(8), 1208-1216.

Rahmat, M. F., Samsudin, S. I., Wahab, N. A., Najib, S., Salim, S., And Gaya, M. S. (2011). Control Strategies Of Wastewater Treatment Plants, Australian Journal of Basic and Applied Sciences,*5*(8), 446–455.

Ribeiro, D., Sanfins, A., and Belo, O. (2013, July). Wastewater treatment plant performance prediction with support vector machines. In *Industrial Conference on Data Mining* (pp. 99-111). Springer, Berlin, Heidelberg.

Rahimzadeh, A., Ashtiani, F. Z., & Okhovat, A. (2016). Application of adaptive neuro-fuzzy inference system as a reliable approach for prediction of oily wastewater microfiltration permeate volume. *Journal of environmental chemical engineering*, *4*(1), 576-584.

Sirhan, H., and Koch, M. (2013). Prediction of Dynamic Groundwater Levels in the Gaza Coastal Aquifer, South Palestine, using Artificial Neural Networks. *Journal of Hydrocarbons Mines and Environmental Research*, *4*(1), 42-56.

Solgi, A., Zarei, H., Nourani, V., and Bahmani, R. (2017). A new approach to flow simulation using hybrid models. *Applied Water Science*, *7*(7), 3691-3706.

Solgi, A., Pourhaghi, A., Bahmani, R., and Zarei, H. (2017). Improving SVR and ANFIS performance using wavelet transform and PCA algorithm for modeling and predicting biochemical oxygen demand (BOD). *Ecohydrology and Hydrobiology*, *17*(2), 164-175.

Singh, K. P., Malik, A., Mohan, D., and Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water research*, *38*(18), 3980-3992.

Şimşek, M. (2016). Nonlinear free vibration of a functionally graded nanobeam using nonlocal strain gradient theory and a novel Hamiltonian approach. *International Journal of Engineering Science*, *105*, 12-27.

Sargent, R. G. (2009). Verification and validation of simulation models. In *Simulation Conference (WSC), Proceedings of the 2009 Winter* (pp. 162-176). IEEE.

Soltaninezhad, K., Beyranvand, M. R., Momenzadeh, S. A., & Shadnia, S. (2012). Electrocardiographic findings and cardiac manifestations in acute aluminum phosphide poisoning. *Journal of forensic and legal medicine*, *19*(5), 291-293.

Tiantian, W., Dewu, J., Jian, Y., Ji, L., and Qiangmin, W. (2019). Assessing mine water quality using a hierarchy fuzzy variable sets method: a case study in the Guojiawan mining area, Shaanxi Province, China. *Environmental Earth Sciences*, *78*(8), 264.

Takagi, T., and Sugeno, M. (1993). Fuzzy identification of systems and its applications to modeling and control. In *Readings in Fuzzy Sets for Intelligent Systems* (pp. 387-403).

Takagi, T. Sugeno, M. (1974). Application Of Fuzzy Algorithm For Control Of Simple Dynamic Plant. *Ieee Proc.*, *12*, 1585–1588.

Tumer, A. E., and Edebali, S. (2015). An Artificial Neural Network Model for Wastewater Treatment Plant of Konya. *International Journal of Intelligent Systems and Applications in Engineering*, *3*(4), 131-135.

Tsioptsias, N., Tako, A. A., and Robinson, S. (2016). Model validation and testing in simulation: a literature review.

Tay, J. H., & Zhang, X. (2000). A fast predicting neural fuzzy model for high-rate anaerobic wastewater treatment systems. *Water Research*, *34*(11), 2849-2860.

UNDP (2014) New Nicosia Waste Water Treatment Plant. Nicosia, Northern Part of Cyprus.

UNESCO (2015). Water Supply, Sanitation And Health. Available At: Http://Www. Unesco.Org/New/En/Natural-Sciences/Environment/Water/Wwap/Facts-Andfigures/ Water-Supply-Sanitation-And-Health/.

Verma, A., Wei, X., and Kusiak, A. (2013). Predicting the total suspended solids in wastewater: a data-mining approach. *Engineering Applications of Artificial Intelligence*, *26*(4), 1366-1372.

Wan, C., & Li, Y. (2011). Effect of hot water extraction and liquid hot water pretreatment on the fungal degradation of biomass feedstocks. *Bioresource Technology*, *102*(20), 9788-9793.

World Health Organization (WHO); United Nation Children Fund (UNICEF) (2012) Progress on drinking water and sanitation WHO, Geneva and UNICEF, New York, pp 122–125.

Wang, W. C., Chau, K. W., Qiu, L., and Chen, Y. B. (2015). Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. *Environmental research*, *139*, 46-54.

Yetilmezsoy, K., Ozkaya, B., and Cakmakci, M. (2011). Artificial intelligence-based prediction models for environmental engineering. *Neural Network World*, *21*(3), 193.

Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J., and El-Shafie, A. (2016). Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. *Journal of Hydrology*, *542*, 603-614.

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., and Chau, K. W. (2018). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of hydrology*.

Yaseen ZM, Sulaiman SO, Deo RC, Chau K-W (2019) An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. Journal of Hydrology 569:387–408. doi: 10.1016/j.jhydrol.2018.11.069

Yel, E., & Yalpir, S. (2011). Prediction of primary treatment effluent parameters by Fuzzy Inference System (FIS) approach. *Procedia Computer Science*, *3*, 659-665.

Zhu, S., Heddam, S., Wu, S., Dai, J., and Jia, B. (2019). Extreme learning machine-based prediction of daily water temperature for rivers. *Environmental Earth Sciences*, *78*(6), 202.

Zheng, Z., and Padmanabhan, B. (2007). Constructing ensembles from data envelopment analysis. *INFORMS Journal on Computing*, *19*(4), 486-496.

Zhang, Q., & Stanley, S. J. (1999). Real-time water treatment process control with artificial neural networks. *Journal of Environmental Engineering*, *125*(2), 153-160.

Zhu, X., Ni, J., & Lai, P. (2009). Advanced treatment of biologically pretreated coking wastewater by electrochemical oxidation using boron-doped diamond electrodes. *Water research*, *43*(17), 4347-4355.