



International Scientific Organization
<http://iscientific.org/>
 Chemistry International
www.bosaljournals.com/chemint/



Modeling and monitoring of treated wastewater based on water quality assurance parameters

Elias Barsenga Hassen^{1,*} and Abraham M. Asmare^{2,*}

¹Faculty of Chemical and Food Engineering, Bahir Dar Institute of Technology, Bahir Dar University, Ethiopia

²Institute of disaster Risk management and Food Security Studies, Bahir Dar University, Ethiopia

*Corresponding author's E. mail: elibarsa@gmail.com (E. B. Hassen), abraham.mebrat@aau.edu.et (A. M. Asmare)

ARTICLE INFO

Article type:

Research article

Article history:

Received June 2017

Accepted September 2018

January 2019 Issue

Keywords:

Artificial neural network

Wastewater treatment plant

Performance assessment

Modeling

ABSTRACT

Recently, process control in wastewater treatment plants (WWTPs) is, mostly accomplished through examining the quality of the water effluent and adjusting the processes through the operator's experience. This practice is inefficient, costly and slow in control response. A better control of WTPs can be achieved by developing a robust mathematical tool for performance prediction. Due to their high accuracy and quite promising application in the field of engineering, Artificial Neural Networks (ANNs) are attracting attention in the domain of WWTP predictive performance modeling. This work focuses on applying ANN with a feed-forward, back propagation learning paradigm to predict the effluent water quality of the Habesha brewery WTP. Data of influent and effluent water quality covering approximately an 11-month period (May 2016 to March 2017) were used to develop, calibrate and validate the models. The study proves that ANN can predict the effluent water quality parameters with a correlation coefficient (R) between the observed and predicted output values reaching up to 0.969. Model architecture of 3-21-3 for pH and TN, and 1-76-1 for COD were selected as optimum topologies for predicting the Habesha Brewery WTP performance. The linear correlation between predicted and target outputs for the optimal model architectures described above were 0.9201 and 0.9692, respectively.

© 2019 International Scientific Organization: All rights reserved.

Capsule Summary: The application of artificial neural network to wastewater treatment plant was proved to be excellent technique for the evaluation of effluent water quality parameters.

Cite This Article As: E. B. Hassen and A. M. Asmare. Modeling and monitoring of treated wastewater based on water quality assurance parameters. Chemistry International 5(1) (2019) 87-96.

INTRODUCTION

The aim of wastewater treatment process is to achieve a treated effluent and sludge quality that is environmentally safe for disposal and/or reuse (Fezzi, 2015). Considering the contemporary environmental problems and the global water scarcity, the development of efficient, sustainable and new

WWTPs is essential. At the same time, the loads on the existing plants are expected to increase due to the population growth, increasing industrial activity and expansion of residential areas. Moreover, the establishment of more stringent environmental regulations relative to WTP effluent quality and discharge puts more pressure on operators and decision makers for more efficient management and improved plant performance (Dairi et al., 2011).



Fig. 1: Habesha brewery wastewater treatment plant

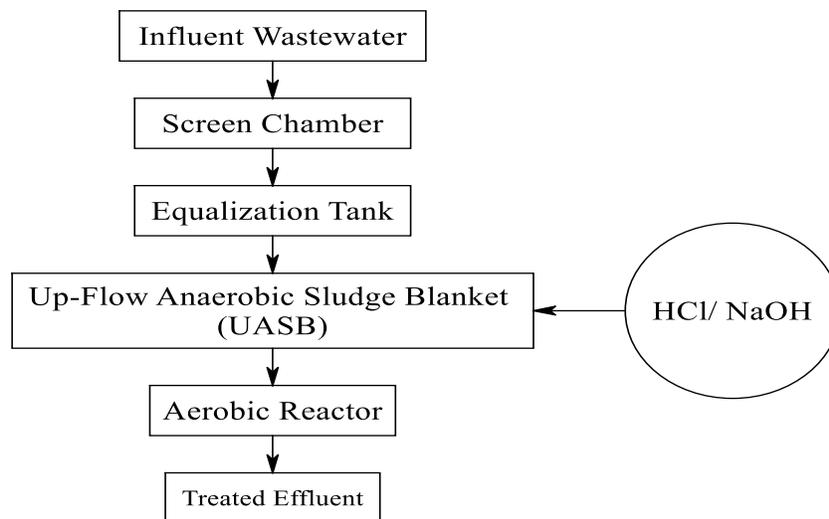


Fig. 2: Process flow diagram of Habesha brewery wastewater treatment plants

Currently, process control is generally accomplished through examining the quality of the water effluent and accordingly adjusting the processes based on the operators' experience. However, this practice is inefficient and slow in control response (Zhang and Stanley, 1999). The latter calls for more efficient and economic controls well as the development of mathematical tools describing the performance of WWTPs.

Model development is a sophisticated and cost-effective way to control WWTPs' operation and assess their performance (Vandejerckhove et al., 2008). These models can be divided into two main categories: linguistic and mathematical models. Linguistic models (such as expert systems) can relate cause to effect, without the construction of a mathematical model.

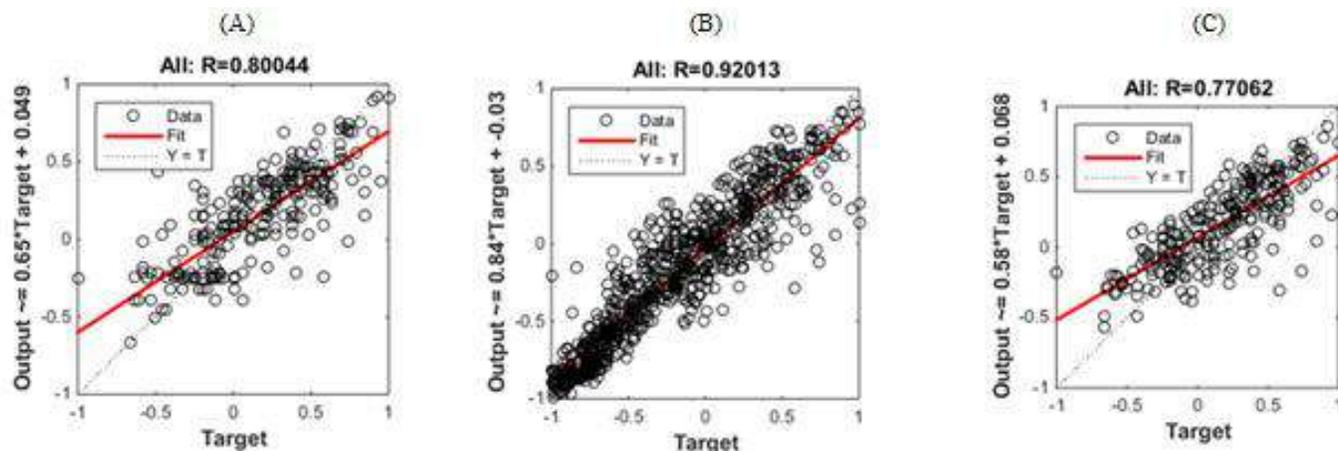


Fig. 3: Regression plot for the best performing pH prediction: (A) COD_{in}-pH_{out} (SISO) (B) MIMO and (C) MISO models

They are most suitable for describing phenomena in environmental systems that are very difficult to describe mathematically. Most of the expert systems that have been developed in the field of wastewater treatment are dedicated to diagnosis. Such models enable the identification of causes of poor performance, thus helping operators with the design of mitigation strategies (El-Din et al., 2004).

Mathematical models of wastewater treatment processes can be divided into two broad classes: white box (deterministic) and black box (empirical) models. White box models are most useful to understand the events occurring in a system (El-Din et al., 2004). Generally, deterministic models incorporate direct links between inputs and outputs through ordinary and partial differential equations describing mechanistic reactions. The development of such models which can accurately describe wastewater treatment processes requires a detailed knowledge and evaluation of the applied system, the factors that influence the system and the interaction among those factors. Although these models give a good insight into the dynamics of WWTPs, they require long and tiring calibration before customizing them to a specific plant (Moral, 2004). Activated Sludge Model No. 1 (ASM1) and Activated Sludge Model No. 2 (ASM2), both developed by the International Water Association (IWA), are among the basic mechanistic models developed in the field of wastewater treatment (Gernaey et al., 2004).

The other type of mathematical models (black box or empirical) is designed to describe highly complex systems. They are very useful to plant operators because they realistically mimic WWTP operation. Artificial neural network (ANN) is one of the most widely used black box modeling tools that are used to model wastewater treatment processes. The fact that black box models can be continually updated with minimal resource requirement has made them attractive in real-time control scenarios (El-Din et al., 2004). Since its introduction into the field of environmental engineering in the late 80's, black box application has increased at a steady rate (Zhang, 2002).

ANN is an artificial intelligence technique that mimics the human brain's biological neural network in the problem-solving processes. As we solve a new problem based on past experience, a neural network takes previously solved cases, looks for patterns in these examples, applies the pattern and develops the ability to correctly classify new patterns and predict and forecast process parameters (Zhang, 2002). Laboratory tests based on experimental data are often used for the determination of the effluent quality throughout the treatment process. However, this can be time-consuming, expensive and slow in response. This study aims at developing a robust predictive model for the determination of the quality of the treated effluent by using ANN.

WWTPs are dynamic, non-linear systems subjected to variation in the volume and composition of the incoming wastewater. Therefore, environmental regulations set restrictions on the effluent quality that must be met by any WTP. These stringent discharge standards along with the easily changing operational parameters render the proper WWTP management a highly challenging task.

Currently, the clear majority of industrial WTPs in Ethiopia, including Habesha Brewery, use the conventional experimental approach to determine the quality of the treated effluent water before discharge or reuse. However, this method can be expensive and time-consuming. In addition, this approach requires the engagement of experienced professionals to obtain high-quality results. Optimal WWTP operation and control can be achieved by the development of a robust mathematical tool that enables the prediction of the quality of the treated effluent based on past observations of certain key parameters. Hence, the aim of this work was to develop a robust model predicting the performance of Habesha Brewery WWTP by the use of the ANN technique

MATERIAL AND METHODS

WWTP Description

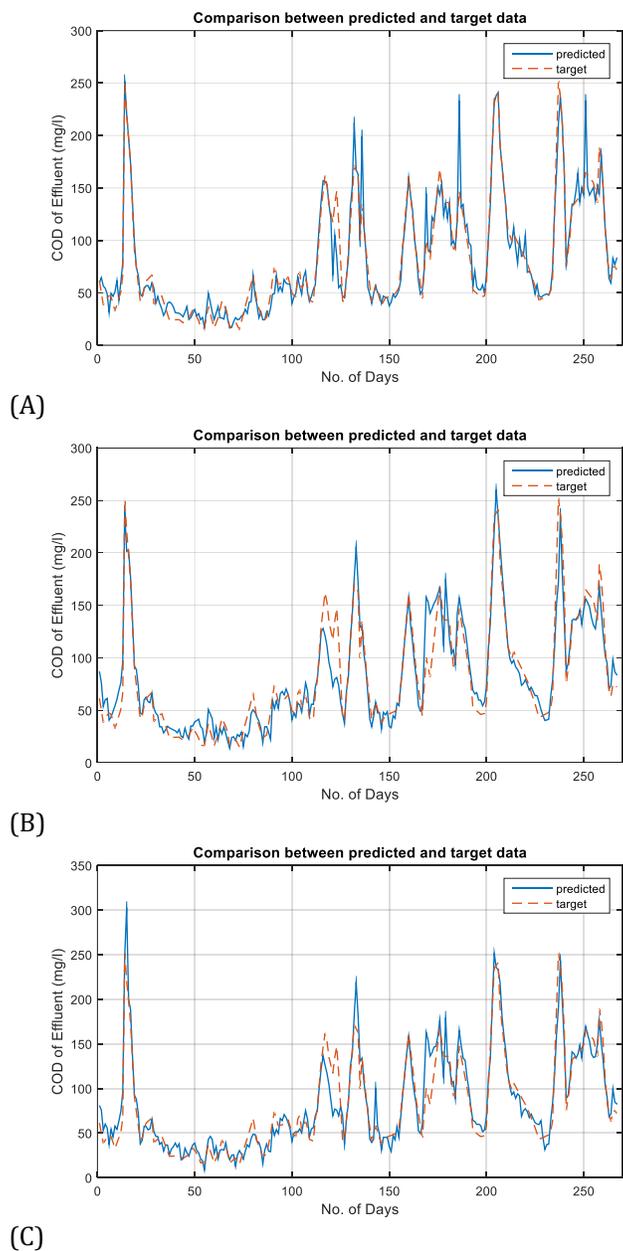


Fig. 4: Comparison between predicted and actual data for pH prediction using (A) COD_{in} - pH_{out} (SISO) (B) MISO and (C) MIMO models

Habesha Brewery WWTP, with a capacity of 700,000hl of beer annually, is located in the town of DebreBirhan 120km north of Addis Ababa City. The wastewater produced is biologically treated in a WWTP with a maximum capacity of 1,000 m³ at a time (Fig. 1).

Process Description of Habesha Brewery WWTP

The wastewater first passes through a screen chamber to remove large solid parts, then, it goes through an equalization tank where its pH is neutralized by the addition

of chemicals (HCl or NaOH depending on the initial pH of the wastewater). The next step is anaerobic biological treatment through an Up-flow Anaerobic Sludge Blanket (UASB). The anaerobically treated wastewater will be afterwards pumped to an aerobic reactor where it undergoes aerobic biological treatment before discharged to the nearby farms for irrigation. A simple process flow diagram of Habesha Brewery WWTP is presented in Figure 2.

Effluent quality standards and discharge guidelines

There is globally a growing political and social pressure to reduce the pollution related to industrial activities. The brewing industry is one such industry that generate large amount of wastewater; it has been documented that 3 to 10 l of wastewater can be generated per l of beer produced (Senthilraja et al., 2013).

The Environmental Protection Agency of Ethiopia (EPA) has developed general national pollutant discharge limits to control water pollution and enable wastewater reuse for agricultural purposes. EPA standard limits relevant to effluent physic-chemical parameters are shown in Table 1.

Historical operational data

The historical data used in this study were obtained out of the analyses carried out in the laboratory of the Habesha Brewery WWTP. For a period of about 11 months (from May 2nd 2016 to March 20th 2017) COD, pH and Total Nitrogen (TN) for both raw influent wastewater and treated effluent were recorded in the plant laboratory.

The obtained data were carefully investigated. After considering the available options for modeling the treatment plant performance, it was decided to relate the quality of the raw influent wastewater to the quality of the final treated effluent. The descriptive statistical analysis for the influent operational data and treated effluent are presented in Table 2.

Software

MATLAB® (R2014b) software neural network toolbox, which is a high performance interactive software package for scientific and engineering computation, was used for designing, building, training and testing the neural networks. MiniTab® v 17 and Microsoft Excel® 2016 were also used for the data organization and pre-processing tasks.

Data pre-processing

The quantity and quality of the available data sets will ultimately determine the performance and complexity of the ANN. Mostly, raw data collected from plant operations are noisy and incomplete by nature. Therefore, the raw data that was collected from the plant laboratory was examined for completeness, missing values and outliers by using MiniTab (statistical analysis software).

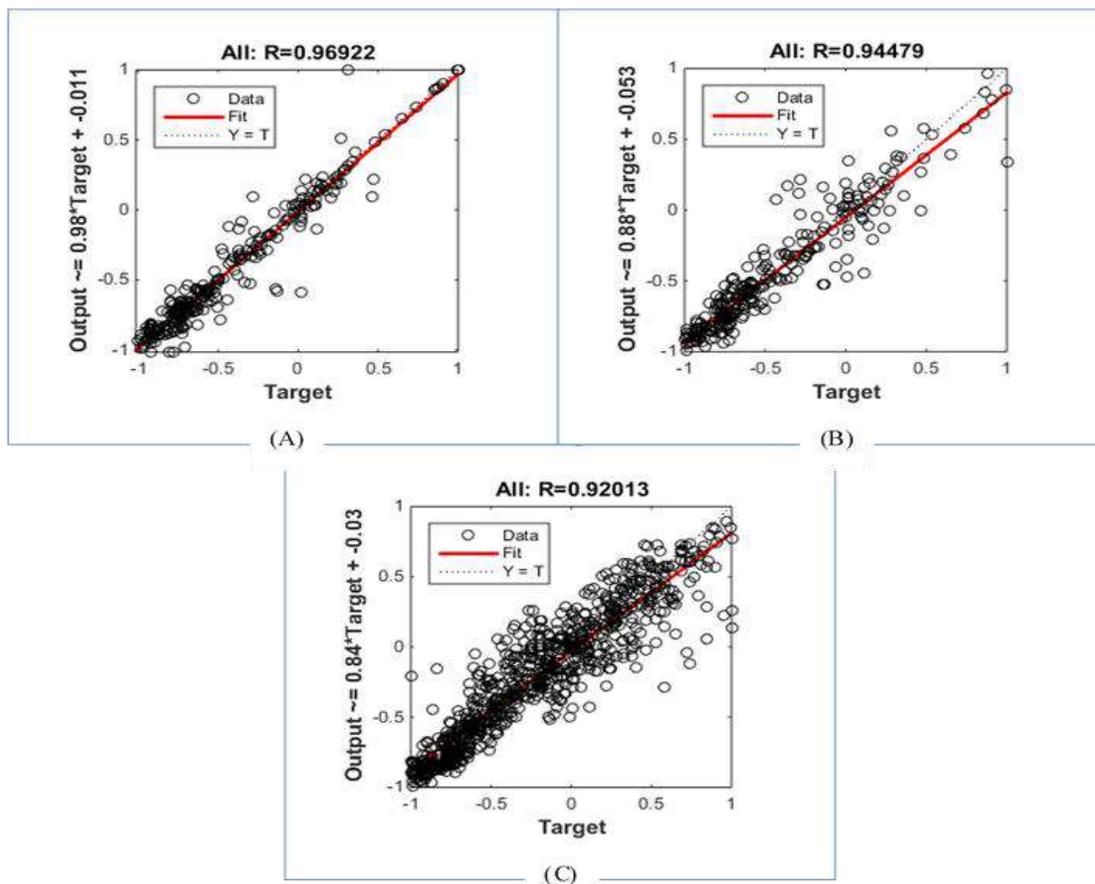


Fig. 5: Regression plot for the best performing COD prediction: (A) COD_{in}-COD_{out} (SISO) (B) MISO and (C) MIMO models

Missing values were estimated by using interpolation and outlier removal was accomplished by removing measurements that were not within the range of ± 3 standard deviations. The descriptive statistics of the pre-processed data is presented in the Tables 3.

To ensure that the statistical distribution of the values for each net input and output was roughly uniform, data scaling was carried out by using *mapminmax* function in MATLAB by which the data sets were scaled to a specified range of -1 to +1.

Data division

When training multilayer networks, the general practice is to first divide the data into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second set is the validation set. The error on the validation set is monitored during the training process. The validation error decreases during the initial phase of the training and the error tends to rise as the network begins to rise. Therefore, the validation set is used to decide when to stop training in order to avoid over-training. The test set is not used during

the training phase, but it is used to compare the performance of different models.

MATLAB provides four alternative functions for dividing data into training, validation and test sets. These are *dividerand* (the default), *divideblock*, *divideint*, and *divideind*. Normally the data division process is carried out automatically during the training process. In this study, all the four data division functions were used alternatively in an attempt to increase the network performance.

Model architecture selection

Model architecture determines the overall structure and direction of information flow in the model. Generally, ANN architectures are divided into feed-forward and recurrent networks. In feed-forward networks the information flows in one direction, i.e. from the input layer to the output layer. On the contrary, recurrent information moves in both forward and backward direction. Multilayer Perceptron (MLP) Networks, the most commonly used type of feed-forward neural network, were used in this thesis work.

Network structure selection

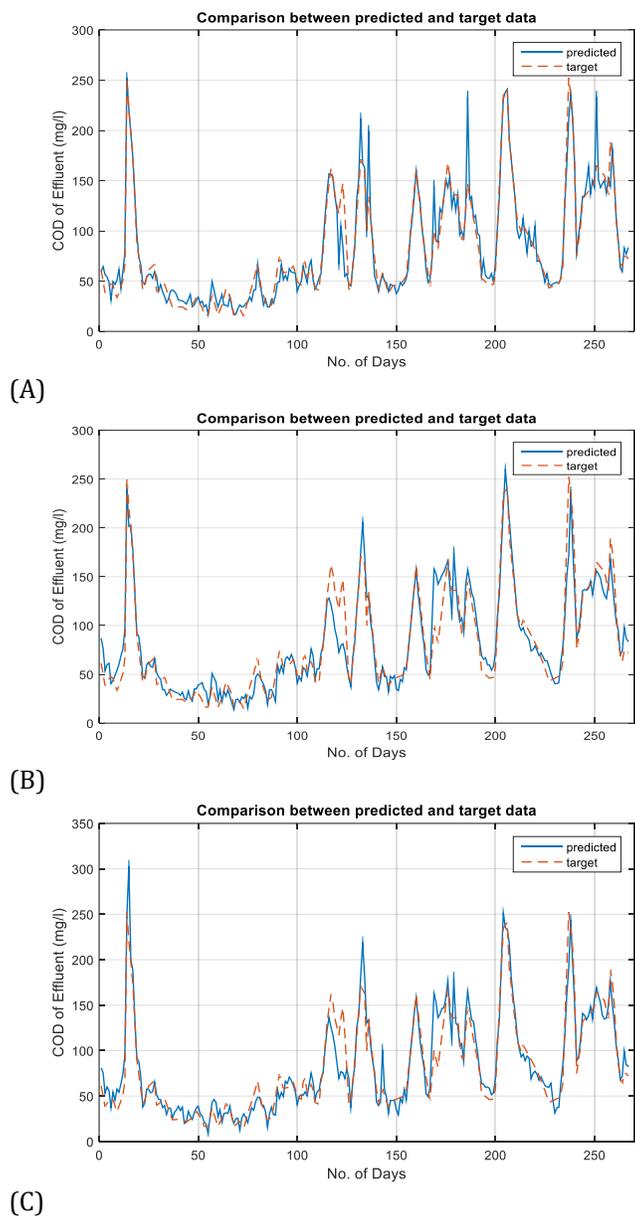


Fig. 6: Comparison between predicted and actual data for COD prediction using the (A) COD in-COD out (SISO), (B) MISO and (C) MIMO models

Table 1: EPA standard limits for the physicochemical parameters of the treated industrial wastewater (Fikresilasie, 2011)

Parameter	Limits
Temperature	Do not change ambient temperature by more than 1°C
pH	5-9
COD	≤ 125mg ^l ⁻¹
TN	≤ 30 mg ^l ⁻¹

Network structure, together with model network architecture, defines the functional form of the relationship between network inputs and outputs. The optimal network structure generally strikes a balance between generalization ability and network complexity. Optimal network structure was achieved by investigating the effect of the following network characteristics on the network performance.

- i. Network configuration: In this study, two types of configurations were developed based on the pre-processed data. pH, COD and TN of the raw influent wastewater were used as input variables to predict the quality of the treated effluent. A total of 13 models (9 SISO, 3 MISO and 1 MIMO configurations) were designed and evaluated in this study.
- ii. Selection of number of hidden neuron: Determination of optimal network structure involves the selection of the appropriate number of hidden neurons. Generally, a two-layer perceptron was developed in this study which was represented with a I-J-K topology, where I, J and K were the numbers of input, hidden and output neurons in the network, respectively. The input and output layer are easily known from the input and target data. Unfortunately, there is currently no universally accepted guideline to determine the number of hidden neurons. Therefore, the optimal number of neurons in this study was determined by using a trial and error method through adding one or two neurons at a time.
- iii. Selections of proper transfer function: Generally, the hyperbolic tangent and sigmoid functions are appropriate for most types of networks, especially for prediction problems. The hyperbolic tangent function was preferred over the log sigmoid function in this work for the following reasons:

The output was in the range of -1 to +1 for the hyperbolic tangent and only from 0 to 1 for the log sigmoid function. This means that the hyperbolic tangent function has a negative response for a negative input value and a positive response for a positive input value, while the log sigmoid function has always a positive response.

The slope of the hyperbolic tangent is much than the slope of the sigmoid function. The latter means the hyperbolic tangent function is more sensitive to small changes in the input.
- iv. Training algorithm selection: It is a challenging task to determine which training algorithm provides a fast learning experience for a given problem. Training algorithm selection depends on many factors including the complexity of the problem, the number of data points in the training set, the number of biases in the network, the error goal and for which purpose the network is used. MATLAB provides 13 training algorithms for training of a neural network among which the Bayesian Regularization algorithm (*trainbr*) was used in this work. Even if this algorithm takes more time for training compared to other algorithms, it has a good generalization capacity for difficult, small or noisy datasets (MathWorks, 2017).

Table 2: Descriptive statistics for raw operational data of influent wastewater and treated effluent

	Influent wastewater			Treated effluent		
	pH	COD (mg/l)	TN (mg/l)	pH	COD (mg/l)	TN (mg/l)
Min	6.030	920	8.44	7.560	14.6	2.69
Max	9.710	2764	49.6	8.330	422	27.6
Mean	7.208	1605	31.3	7.960	89.7	8.72
SD	0.551	437	6.82	0.140	67.4	5.30
VAR	0.303	190985	46.5	0.020	4545	28.1

Table 3: Descriptive statistics for pre-processed data of influent wastewater and treated effluent

	Influent wastewater			Treated effluent		
	pH	COD (mg/l)	TN (mg/l)	pH	COD (mg/l)	TN (mg/l)
Min	6.030	920	14.82	7.560	14.6	2.69
Max	8.750	2764	49.60	8.290	258	21.1
Mean	7.138	1602	31.56	7.951	83.5	8.36
SD	0.413	409	5.98	0.135	52.2	3.97
VAR	0.170	167036	35.76	0.0183	2724	15.75

Table 4: List of MATLAB data division functions (MathWorks, 2017)

Function	Algorithm
<i>dividerand</i>	Divide the data randomly (default)
<i>divideblock</i>	Divide the data into contiguous blocks
<i>divideint</i>	Divide the data using an interleaved selection
<i>divideind</i>	Divide the data by index

Model training

Model Training is an iterative process that seeks to modify the network through numerous presentations of data. Mainly, there are two type of model training; unsupervised and supervised. Unsupervised training occurs when the model is only fed with the input values and uses them to adjust its connection weights. In the case of supervised data, the model is fed with both input and target data values. Supervised mode of training was used in this study since it requires shorter time compared to the unsupervised training mode.

Model evaluation

In order to determine which network structure is optimal, the performance of the calibrated models was evaluated. In this work, the Mean Squared Error (MSE) and Correlation coefficient (R) were used to evaluate the performance of the networks, where MSE (given in Equation 1) is the average square difference between outputs and targets. Lower MSE values are considered better and zero means no error. R values measure the correlation between outputs and targets.

A R value of 1 means a close relationship, whereas 0 a random relationship.

$$MSE = \sqrt{\frac{\sum_{i=1}^n (Target-Output)^2}{n^2}} \quad (1)$$

RESULTS AND DISCUSSION

pH prediction models

Over all the 5 created networks (3 SISO, 1 MISO and 1 MIMO), models were developed to determine the optimum network topology for pH prediction for the treated effluent. Table 5 presents the statistical performance results of these models.

Based on the statistical performance of all the three configurations for the pH prediction, the MIMO model was observed to be the best by scoring a R value of 0.92. Thus, it can be suggested that the MIMO model generalizes the data well and is likely to make accurate predictions when new data (data that does not belong to the training or testing set) is provided compared to the other tested models.

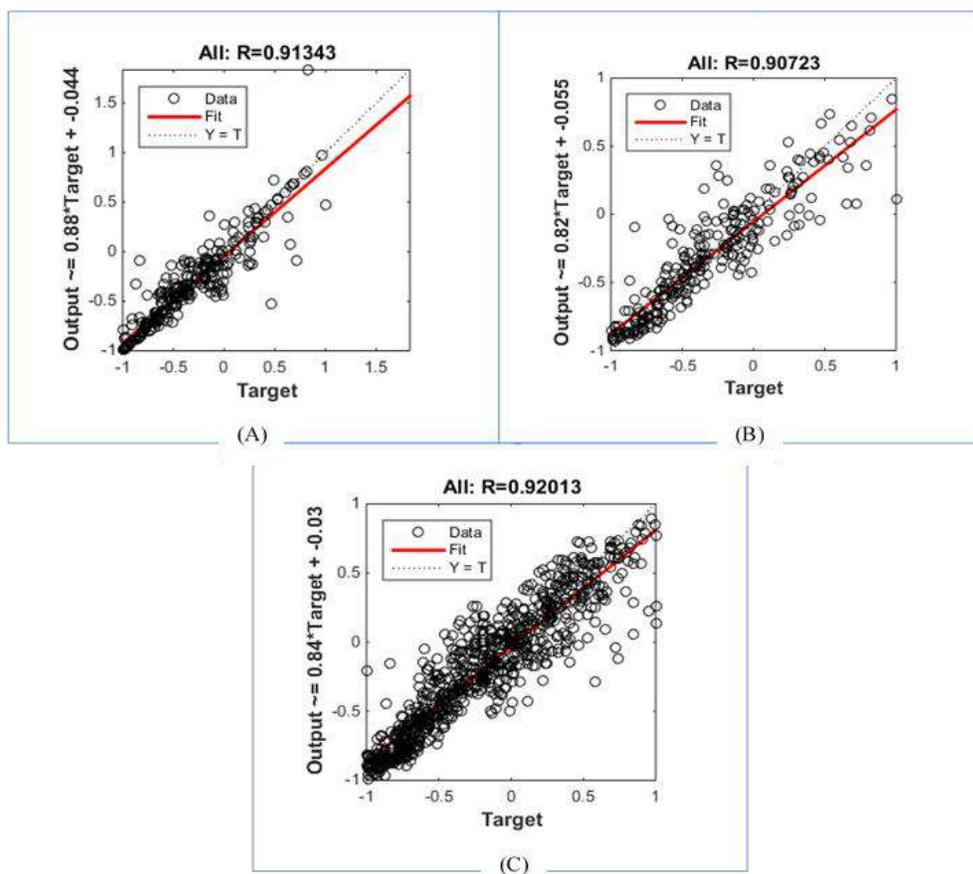


Fig. 7: Regression plot for the best performing TN prediction: (A) CODin-TNout (SISO), (B) MISO and (C) MIMO models

The linear regression plot for the best performing models from each configuration is presented in Figure 3. As shown in Figure 4, the comparison plot between the predicted and actual data (based on the MSE values), the MIMO model indicates a good fit compared to the other models. Therefore, the appropriate architecture for pH prediction was determined to consist of an input layer with 3 neurons, a hidden layer with 21 neurons and 3 output layer neurons.

COD prediction models

To predict the COD of the treated effluent 3 SISO configuration models with different input, 1 MISO and 1 MIMO configuration models were developed. The statistical performance for the best performing models from each configuration is presented in Table 6.

Among the models developed for COD prediction, the SISO model, where COD is used as an input has shown excellent generalization and predictive efficiency by scoring a R value of 0.97. Even if the MISO and MIMO models have scored a lower R value compared to the SISO model, both the configurations have shown a good accuracy by scoring R values greater than 0.9 for all the training and testing sets. These results were also supported by the regression plot and comparison plots presented in Figures 5 and 6, respectively.

Based on the R and MSE values of the models, the best topology for COD prediction is a SISO configuration with 1 neuron in the input layer, 76 neurons in the hidden layer and 1 neuron in the output layer.

TN prediction models

To predict the TN of the treated effluent, 3 SISO configuration models with different input, 1 MISO and 1 MIMO configuration models were developed. The statistical performance for the best performing models from each configuration is presented in Table 7.

As shown in Table 7, the best performing model for the TN prediction in the treated effluent was the MIMO model with topology of 3-21-3 by scoring a R value of 0.92. The result was also supported by the regression and comparison plot between the actual output data and predicted data that is presented in Figures 7 and 8.

CONCLUSIONS

Based on the performance statistical results, the MIMO model has shown a better predictive performance compared to the SISO and MISO configuration by scoring a R value of 0.92. In the case of the input variables, the

Table 5: Performance statistics for the pH prediction models

Configuration	Input	I-J-K	Training		Testing		All
			MSE	R	MSE	R	R
SISO	pH	1-52-1	0.120	0.276	0.161	0.421	0.455
	COD	1-41-1	0.054	0.805	0.057	0.747	0.804
	TN	1-76-1	0.135	0.254	0.178	0.050	0.172
MISO	pH, COD, TN	3-29-1	0.060	0.771	0.070	0.773	0.771
MIMO	pH, COD, TN	3-21-3	0.037	0.921	-	0.917	0.920

Table 6: Performance statistics for the COD prediction models

Configuration	Input	I-J-K	Training		Testing		All
			MSE	R	MSE	R	R
SISO	pH	1-33-1	0.202	0.136	0.181	0.247	0.305
	COD	1-76-1	0.007	0.982	0.057	0.885	0.970
	TN	1-48-1	0.111	0.305	0.092	0.079	0.275
MISO	pH, COD, TN	3-35-1	0.022	0.941	0.015	0.942	0.945
MIMO	pH, COD, TN	3-21-3	0.037	0.921	-	0.917	0.920

Table 7: Performance statistics for the TN prediction models

Configuration	Input	I-J-K	Training		Testing		All
			MSE	R	MSE	R	R
SISO	pH	1-96-1	0.133	0.340	0.184	0.729	0.330
	COD	1-68-1	0.024	0.937	0.131	0.779	0.913
	TN	1-51-1	0.145	0.229	0.172	0.024	0.208
MISO	pH, COD, TN	3-24-1	0.035	0.911	0.033	0.817	0.907
MIMO	pH, COD, TN	3-21-3	0.037	0.921	-	0.917	0.920

models with COD as input variables had a better quality of prediction and accuracy than the models where pH and TN were used as input variables. Model architectures 3-21-3 (MIMO configuration) for pH and TN, and 1-76-1 (SISO configuration with COD as input) for COD, were selected as the optimum topology for predicting the performance of Habesha Brewery WTP. The linear correlation between predicted and target outputs for the optimum model architecture described above were 0.92 and 0.97, respectively. Finally, this study has demonstrated the effectiveness of using neural networks to capture the non-linearity and complexity of the relationship between raw influent and treated effluent water quality data from industrial WTPs. In conclusion, the result showed that the trained and tested ANN models can potentially be employed for predicting the performance of Habesha Brewery WTP, thus providing a good supportive tool for the plant's control and monitoring.

REFERENCES

- Dairi, S., Mrad, D., Djebbar Y. and Hammar, Y., 2011. Dynamic Simulation for the requirements of oxygen about the Municipal Wastewater Treatment Plant- Case of Souk-Ahras/Algeria. *Journal of Material and Environmental Science* 2(S1): p. 507-512.
- El-Din, A.G., Smith, D.W., El-Din, M.G., 2004. Application of Artificial neural networks in wastewater treatment. *Journal of Environmental Engineering and Science* 3, S81-S95.
- Fezzi, M., 2015. A Pragmatic Approach to Wastewater Treatment Modelling: The Kallby Wastewater Treatment Plant as a Case Study, in Department of Chemical Engineering. Lund University: Lund, Sweden. p. 121.
- Fikresilasie, T., 2011, Impact of Brewery Effluent on River Water Quality: The Case of Meta abo Brewery Factory and Finchewa River in Sebeta, Ethiopia, in School of Graduate Studies. Addis Ababa University: Addis Ababa. p. 85.
- Gernaey, K.V., van Loosdrecht, M.C., Henze, M., Lind, M., Jørgensen, S.B., 2004. Activated sludge wastewater treatment plant modelling and simulation: state of the art. *Environmental Modelling & Software* 19, 763-783.

- MathWorks, Neural Network Toolbox User's Guide™. 2017. MathWorks Inc.: Massachusetts, USA.
- Moral, H., 2004. Modeling of Activated Sludge Process by using Artificial Neural Networks, in Department of Environmental Engineering. Middle East Technical University: Ankara, Turkey. p. 1126.
- Senthilraja, K., Jothimani, P., Rajannan, G., 2013. Effect of brewery wastewater on growth and physiological changes in maize, sunflower and sesame crops. International Journal for Life Science and Educational Research 1(1), 36-42.
- Vandekerckhove, A., Moerman, W., Hulle, S.W.H., 2008. Full-scale modelling of a food industry wastewater treatment plant in view of process upgrade. Chemical Engineering Journal 135, 185-194.
- Zhang, Q., Stanley, S.J., 1999. Real-time water treatment process control with artificial neural networks. Journal of Environmental Engineering 125(2), 153-160.
- Zhang, Q.J., 2002. Artificial neural network—advanced theories and industrial applications, Department of Civil and Environmental Engineering. University of Alberta: Edmonton, Canada. p. 183.

Visit us at: <http://bosaljournals.com/chemint/>
Submissions are accepted at: editorci@bosaljournals.com