Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2020.26043

201 (2020) 165–172
October

Regression modeling for rapid prediction of wastewater BOD_5

Ahmad Qasaimeh*, Ziad Al-Ghazawi

Department of Civil Engineering, Faculty of Engineering, Jordan University of Science and Technology, Irbid, 22110 Jordan, Tel. +962-2-7201000 22139; emails: argg22@yahoo.com (A. Qasaimeh), alghazawi@gmail.com (Z. Al-Ghazawi)

Received 23 October 2019; Accepted 25 April 2020

ABSTRACT

It is known that the biochemical oxygen demand (BOD_5) is not directly achieved because it takes a standard procedure of 5 d, and since simple variations in the biological processes in wastewater treatment plants will drastically change the required BOD_5 output; the main target in this work is to find measurable parameters at the effluent stream to find the effluent BOD_5 easily. Multi-linear regression is proposed here as a simple mathematic entity to predict the effluent BOD_5. In the first step, the data of the Irbid wastewater treatment plant was collected for 10 y. From the available measured data, it is aimed to find the effluent parameters that are correlated with the effluent BOD_5. The elected effluent quality parameters are dissolved oxygen (DO), pH, temperature, flow rate (Q), total suspended solids (TSS), and chemical oxygen demand (COD). These parameters are examined to check their correlation with the BOD_5. The worked data is 114 sets for each quality parameter; of which, 96 sets are used for training, and 18 sets are used for validation. By using the Pearson correlation, it is found that the BOD_5 is correlated with the following parameters: Q, TSS, DO, and COD. By using a series of multiple linear regression, it is found that the only significant correlated parameters are the COD, and the TSS. The evolved BOD_5 prediction model is a function of COD and TSS, and has the value of Pearson correlation R of (0.97), coefficient of determination R^2 of (0.94), P-values of (<0.05), and the significance level of (6.95E-59). It can be concluded that the obtained model can be applied as an automated system that predicts the effluent BOD_5 and tunes the parameters together with the required treatment efficiency.

Keywords: Biochemical oxygen demand; BOD_5; Prediction; Correlation; Multi-linear regression; MLR; WWTPs

1. Introduction

The treatment of municipal wastewater is of major concern in terms of health safety, environment protection, and resource savings. The wastewater treatment processes can be categorized as physical, chemical, and biochemical processes. The treatment design relies on the understanding of the basics governing the treatment processes [1]. The treatment of municipal wastewater entails primary, secondary, and tertiary treatment stages, of them, the secondary process is the main core of the treatment overall. The secondary treatment standards are based on the removal of bulk biodegradable organics (proteins, carbohydrates, etc.), total suspended solids (TSS), nutrients, and pH [2,3].

The 5 d biochemical oxygen demand (BOD_5) is considered as a conventional index of contaminated water that is produced from biodegradable organic materials [4]. The main disadvantages of the standard BOD_5 test; are the time needed for its accomplishment, and the variations of the results due to the experimental conditions and the microbial diversity of the samples used [5,6]. In addition, the BOD_5 differs according to the characterization of

* Corresponding author.

wastewater and differs during the time [7]. The wastewater characterization is usually based on measuring biochemical oxygen demand (BOD_5), chemical oxygen demand (COD), nitrogen, phosphorous, and TSS [8].

The variety in the equipment used in the wastewater treatment plants, and the range of wastewater strength; will imply variable operational techniques. Some mathematical relationships can be developed to enhance the operation and the management of the wastewater treatment plants. The relationships between wastewater quality parameters and treatment effectiveness may be entailed via monitoring and modeling processes such as those found in the literature [9,10]. Recently, intelligent computations such as neural networks, genetic algorithms, and fuzzy logic systems have been widely used for prediction in many domains of environmental engineering and wastewater treatment [11–18]. However, the computational modeling needs data parameterization, training, and validation processes that need experts with money and time-consuming process [19]. The prediction models are commonly directed to be simple parsimonious models that are plain models of high descriptive forecasting ability with a minimum number of variables [20]. Of parsimonious models, the regression models are simple statistical analyses that formulate an association between one dependent variable and other independent variables [21]. Multi-linear regression (MLR) is a simple and instant approach to evaluate correlated variables in the natural and environmental systems. Despite other techniques that may provide accurate results, the MLR is still able to provide simple, fast, and accurate results. Many researchers had widely used it for evaluation of quality parameters in urban runoff, reservoirs, surface water, and wastewater treatment plants [22–27].

In this study, the MLR will be used to predict BOD_5 from correlated variables. As the correlated variables are easily measurable, the BOD_5 prediction will be readily achieved by regression modeling using a data analysis tool available in the Excel program.

2. Materials and methods

The wastewater treatment plant that is being targeted in this study is the central Irbid WWTP. It is intended to check indicators at the effluent stream to tell the inspector about the level of the BOD_5 discharged. At the beginning, the candidate parameters are specified and then tested to find their correlation to the BOD_5. In the second stage, the correlated parameters to the BOD_5 are tested by means of MLR rounds to specify their significance. Finally, the significant correlated parameters will formulate the BOD_5 prediction model.

2.1. Study area

The wastewater data is taken from central Irbid wastewater treatment plant. It is located on the northern side of Jordan at 32°34′37.52″N latitude and 35°50′14.49″E longitude (Fig. 1). The area exhibits dry weather as the average temperature and rainfall pertain semi-arid conditions. The climate in Irbid is classified as warm and moderate in summer, cool, and wet in winter. The climate is classified

Fig. 1. Position of the central Irbid WWTP [28].

as Csa by the Köppen–Geiger system. The average annual rainfall volume is 428 mm. The temperature and rainfall variations throughout the year are detailed in Table 1.

Central Irbid wastewater treatment plant has been in process since 1986. The design hydraulic loading rate was initially 11,000 m^3/d. The plant is based on a hybrid trickling filtration and activated sludge process treatment. The flow is apportioned into two aerated grit chambers continuing to two rectangular primary clarifiers, then it proceeds to the first biological stage consisting of two trickling filters. The second biological stage consists of two aeration tanks. After the aeration, the wastewater undergoes clarification in two final circular sedimentation tanks, and finally, the chlorination is achieved.

There was a chronological increase in hydraulic loading rate after 2009. In 2016, the design loading rate had been raised to become 13,000 m^3/d, and the tertiary processes had been upgraded with final sand filtration and UV disinfection to utilize the treated effluent in irrigation.

2.2. Data sets

Irbid wastewater treatment plant holds over data records for main quality parameters as all other wastewater treatment plants. The main aim here is to find the correlated quality parameters that are easily measurable and can be used as indicators to predict the BOD_5 in the effluent.

In this work, the effluent quality parameters that can be used as indicators are: dissolved oxygen (DO), pH, temperature, flow rate (Q), TSS, and COD. The ranges of values and details of these parameters are illustrated in Table 2. The data used in this work is based on monthly records for 10 y (2007–2016). The data of the year 2010 and 2016 were excluded from the analyses, in order to use them later for the validation of the model.

Since the BOD_5 needs 5 d to be estimated; it is proposed here to find specific parameters to estimate effluent BOD_5 rapidly. These quality parameter indicators are examined to check their correlation with the BOD_5. Each quality

Table 1
Temperature and rainfall variations throughout the months [29]

	January	February	March	April	May	June	July	August	September	October	November	December
Average temperature (°C)	9.1	10	12.3	16.2	20.5	23.8	25.2	25.6	23.9	20.9	15.9	10.8
Minimum temperature (°C)	5.1	5.5	7.5	11	14.9	18.3	20.2	21	18.8	15.5	10.5	6.9
Maximum temperature (°C)	13.1	14.5	17.1	21.5	26.2	29.4	30.2	30.3	29	26.3	21.3	14.8
Average temperature (°F)	48.4	50.0	54.1	61.2	68.9	74.8	77.4	78.1	75.0	69.6	60.6	51.4
Minimum temperature (°F)	41.2	41.9	45.5	51.8	58.8	64.9	68.4	69.8	65.8	59.9	50.9	44.4
Maximum temperature (°F)	55.6	58.1	62.8	70.7	79.2	84.9	86.4	86.5	84.2	79.3	70.3	58.6
Rainfall (mm)	96	88	76	23	6	0	0	0	0	11	43	85

Table 2
Quality parameters for the effluent wastewater of Irbid WWTP (training sets)

Quality Parameter	Minimum value	Maximum value	Average value	Standard deviation	Number of data set
BOD_5, mg/L	27.5	89.4	66.88	14.66	96
pH	7.1	8.1	7.606	0.127	96
DO, mg/L	1.6	6.5	5.724	0.665	96
Temperature, °C	13.5	28	21.363	3.953	96
Q, m^3/d	5,800	8,409	7,051.984	629.813	96
TSS, mg/L	30	121.4	86.391	22.093	96
COD, mg/L	126	361.304	261.543	63.60	96

parameter has data records of 114 sets; 96 sets are used for training, and 18 sets are used for validation. The training data sets of the quality parameters are presented in Figs. 2–4. The validation sets will be discussed later after the building of the model.

2.3. Correlation and MLR modeling

Correlation is statistically defined as a criterion that specifies the degree to which two or more variables are associated together. The most common measure of correlation is Pearson's correlation, which is commonly referred to simply as the correlation coefficient (r), and is expressed in the following equation [30]:

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \sum_{i-1}^{n}\left(y_i - \bar{y}\right)}} \tag{1}$$

where x_i is the real observed values; y_i is the predicted values, n is the number of values; $\bar{x}$ is the average of observed values, and $\bar{y}$ is the average of predicted values.

In other terms, the multiple linear regression determines the fitness of a linear relationship between one dependent variable denoted by (Y) and other independent variables denoted by (X_i). The evolved MLR model can be represented as the following equation [31]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \tag{2}$$

where β_0 is the intercept of the regression line, β_i is the regression coefficient (slope), the Y is the dependent variable to be predicted, and the X_i is the independent variable.

Initially, the specified parameters are examined with Pearson correlation to check their connection to the BOD_5. The weakly correlated parameters will be excluded from the process. The correlated parameters will be used in MLR process to find a model of independent variables to predict the BOD_5. Excel program is being used here to find the Pearson correlation (R). In addition, the multiple linear regression model is being accomplished by using data analysis tool in Excel. The coefficient of determination (R^2), Pearson (R), and standard error are the governing factors for indicating the strength of the model. However, the P-value will rule the decision by choosing variables that have significant values ($P < 0.05$), and hence eliminating unnecessary variables. Finally, the significant correlated indicators will be used as a model for the BOD_5 prediction.

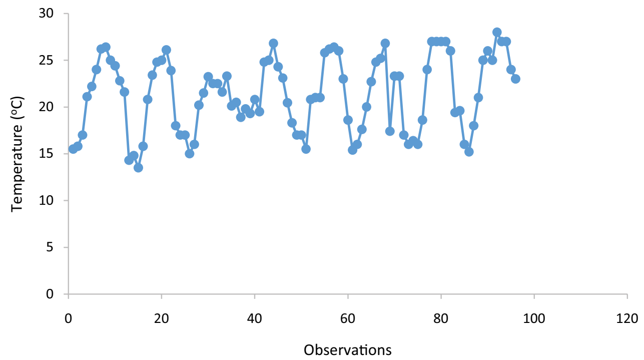

Fig. 2. Temperature variations of the effluent wastewater within 8 y.

3. Results and discussion

3.1. Determination of the correlated parameters

There are many parameters that affect the biochemical degradation process, however, in this work, the aim is to find parameters that are correlated with the effluent BOD_5. In the first stage, the Pearson correlation (r) is determined between the quality parameters and the effluent BOD_5 (Table 3).

Not all dedicated parameters are correlated with the effluent BOD_5; by inspecting Pearson correlation coefficient (r), it can be realized that the pH and the temperature have a very weak correlation with the effluent BOD_5. On the other hand, COD, TSS, and Q have a strong correlation with the BOD_5, whereas the DO has an inverse correlation. Based on

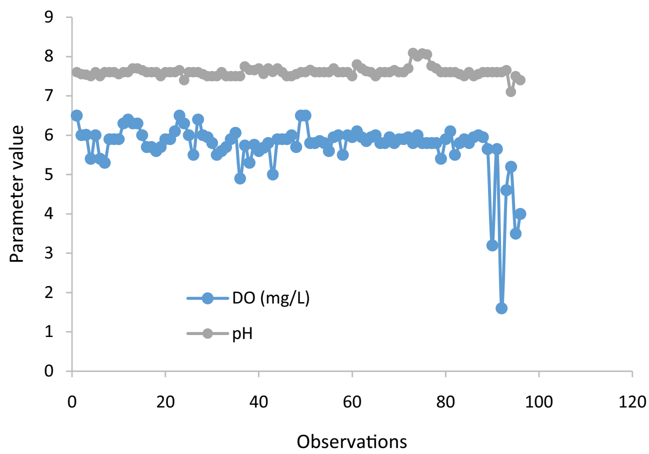

Fig. 3. Dissolved oxygen and pH variations of the effluent wastewater within 8 y.

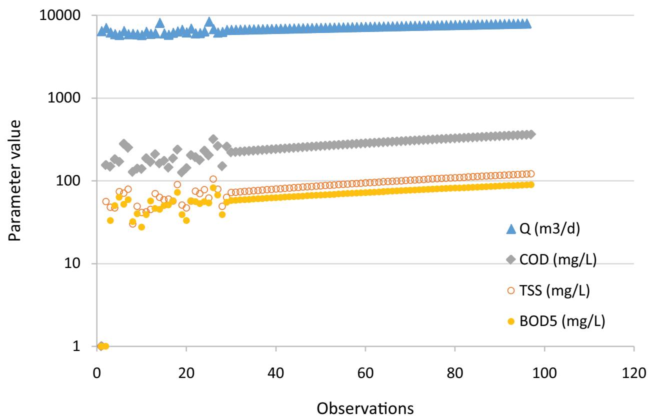

Fig. 4. Discharge, total suspended solids (TSS), COD, and BOD_5 variations of the effluent wastewater within 8 y.

the values of (r) illustrated in Table 3; only COD, TSS, Q, and DO will be taken to the next step of MLR.

3.2. Building MLR model

After choosing the parameters that have considerable correlation with the effluent BOD_5, here in this step, the only significant parameters will be retained, while the other parameters will be dismissed.

By using MLR between the effluent COD, TSS, Q, and DO as independent variables with the effluent BOD_5 as dependent variable; the obtained model has the following values: the multiple R is 0.97, R^2 is 0.94, and the significance level is as low as 5.17E-56 (Table 4). Though, the P-value for the DO and Q is larger than 0.05 showing insignificant role of these variables in the regression process. Thus, the DO will be omitted from the second round of regression, but the Q will still be kept for evaluation considering that it endured high Pearson correlation coefficient with BOD_5.

In the second round of regression, the parameters Q, TSS, and COD are analyzed as independent variables and the BOD_5 as dependent variable. The regression statistics of this round is shown in Table 5. It can be noticed that the multiple R is 0.97, R^2 is 0.94, and the significance level is low down to 2.09E-57 showing correlated and very accurate analysis. However, still the Q has high P-value, showing its insignificant role in regression.

In the third round of regression, MLR is held between the effluent parameters (COD and TSS) as the independent variables with the effluent BOD_5 as the dependent variable. The regression statistics of this round is shown in Table 6. It can be observed that the multiple R is 0.97, R^2 is 0.94, and that the significance level is down to 6.95E-59 showing very accurate analysis, and the P-value of all the components are below 0.05 pertaining significant role in regression analysis.

By summing up the previous analyses, the only significant independent variables are COD and TSS. The components of the prediction model and their statistical measures are illustrated in Table 7.

Based on the findings of the MLR analysis shown in Table 7, the prediction model can be written into the following equation:

$$BOD_5 = 0.498992(\text{TSS}) + 0.052154(\text{COD}) + 10.13197 \tag{3}$$

The evolved model will be involved in the monitoring system in the wastewater treatment plant by which the investigator can easily apply the TSS and the COD measured

Table 3
Effluent quality parameters for treated wastewater in central Irbid WWTPs and their Pearson correlation with the effluent BOD_5

Quality parameter	(r) Pearson correlation coefficient with BOD_5	Elected parameters
pH	0.095827	–
DO	–0.33938	DO (r = –0.34)
Temperature	0.015841	–
Q	0.8399	Q (r = 0.84)
TSS	0.96971	TSS (r = 0.97)
COD	0.951297	COD (r = 0.95)

Table 4
Model summary for the first round evaluation

Item	Value
Multiple R	0.971555
R^2	0.94392
Adjusted R^2	0.941454
Standard error	3.549468
Observations	96
Significance	5.17E-56
Variable	P-value
Intercept	0.231717
Q	0.753746
TSS	1.03E-10
COD	0.020304
DO	0.987441

Table 5
Model summary for the second round evaluation

Item	Value
Multiple R	0.971555
R^2	0.943919
Adjusted R^2	0.942091
Standard error	3.530129
Observations	96
Significance	2.09E-57
Variable	P-value
Intercept	0.148521
Q	0.75225
TSS	7.73E-11
COD	0.019504

Table 6
Model summary for the third round evaluation

Item	Value
Multiple R	0.971524
R^2	0.943858
Adjusted R^2	0.942651
Standard error	3.513011
Observations	96
Significance	6.95E-59
Variable	P-value
Intercept	2.12E-09
TSS	2.98E-12
COD	0.017689

Table 7
Prediction equation coefficients and its statistical and sensitivity measures

Variable	Coefficients	Standard error	*t* statistics	*P*-value
Intercept	10.132	1.527	6.635	0.000
TSS	0.499	0.062	8.026	0.000
COD	0.052	0.022	2.415	0.017

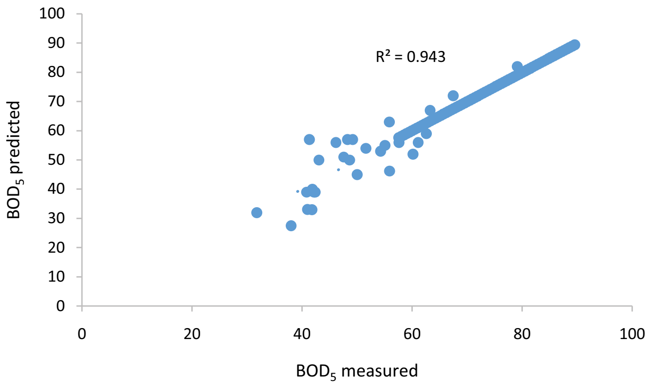

Fig. 5. Fitness of measured BOD_5 vs. predicted values for 96 observations.

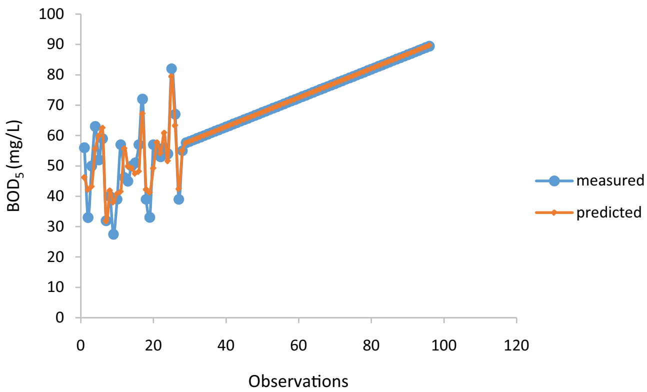

Fig. 6. Fitness of predicted and measured values of BOD_5 for 96 observations (trained model).

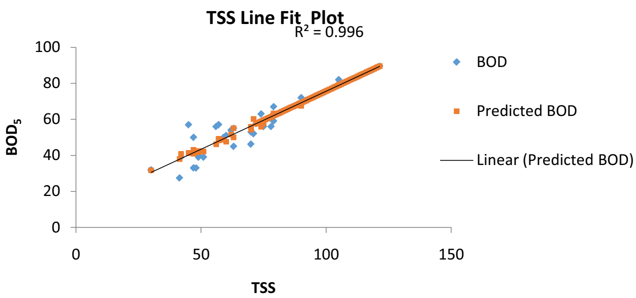

Fig. 7. Observed and predicted BOD_5 on TSS line fit plot.

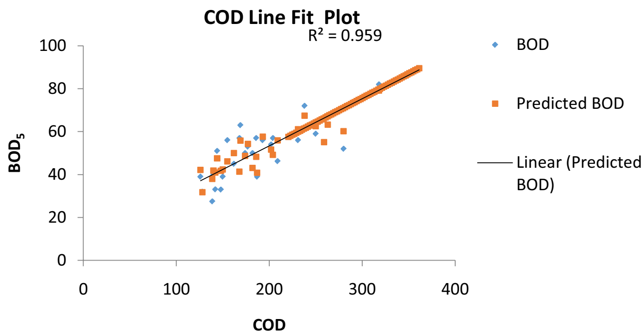

Fig. 8. Observed and predicted BOD_5 on COD line fit plot.

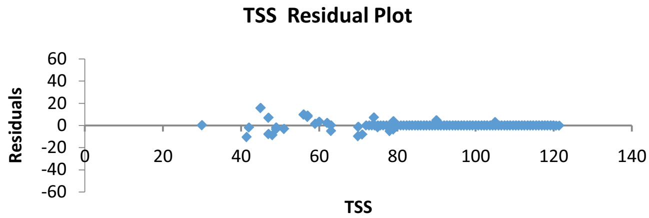

Fig. 9. Residual of predicted BOD_5 on TSS line fit plot.

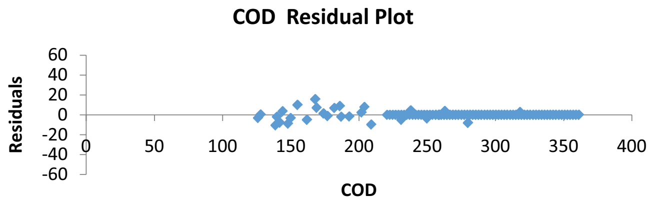

Fig. 10. Residual of predicted BOD_5 on COD line fit plot.

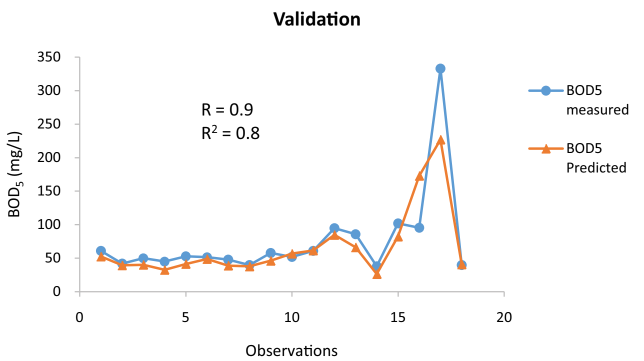

Fig. 11. Fitness of predicted and measured values of BOD_5 for 18 observations (validation of the model).

values into the model, and then the effluent biochemical oxygen demand will be estimated without the need of the 5 d period in the laboratory. Figs. 5–10 show how the observed values of the BOD_5 are fitting the output of the prediction model. The multi linearity of the independent

Table 8
Quality parameters for the effluent wastewater of Irbid WWTP (validation sets)

Quality Parameter	Minimum value	Maximum value	Average value	Standard deviation	Number of data set
BOD_5, mg/L	38	333	75	67.579	18
TSS, mg/L	26	361	93.713	92.646	18
COD, mg/L	54	707	185	146.459	18

variables with the predictable dependent variable and the minor residuals can be observed throughout the plots. Figs. 5 and 6 show the goodness of the fit (R^2 = 0.94) between observed and predicted values of the BOD_5.

The obtained multiple regression model is based on a linear relationship between the dependent variable and independent variables. Fig. 7 shows the linear relationship between TSS and BOD_5 with a high coefficient of determination (R^2 = 0.99) for the TSS line fit plot. Fig. 8 shows a linear relationship between COD and BOD_5 with a high coefficient of determination (R^2 = 0.96) for COD line fit plot. However, the model still has errors to appear in the deviated outputs. The standard error, as shown in Table 7, is 1.527 for the intercept, 0.062 for TSS, and 0.022 for COD. These errors lead to the residual values on the predicted outputs.

The residual is the deviation of the measured value from the predicted value by the model. In Figs. 9 and 10, the residual values are equally and randomly spaced around the horizontal axis with minor residual values.

3.3. Model validation

The evolved model is being validated in this section with different data than those used in training. An 18 data sets for each quality parameter are used for validation. The 18 data sets of each quality parameters are taken from the year 2010 and 2016. The data sets of the effluent quality parameters in the Irbid wastewater treatment plant are shown in Table 8. Fig. 11 shows that the measured BOD_5 is highly correlated with the predicted BOD_5 (R = 0.9), and that the output of the prediction model fits the observed measured data (R^2 = 0.8).

As per these findings, the inspector can rely on the two parameters, COD and TSS, as independent variables to predict the BOD_5. The unique association between BOD_5, COD, and TSS was also adapted by other researchers in their prediction models [32–35]. This approach enables engineers and technicians to measure the COD and TSS easily and consequently to find the expected BOD_5 from the prediction model. The WWTPs need such an approach to monitor, control, and automate the treatment processes into the required outputs and efficiencies.

4. Conclusion

BOD_5 is a major parameter that is used to indicate wastewater quality and treatment efficiency in wastewater treatment plants. However, it needs 5 d for the standard test to be accomplished. In this study, an MLR method is used to find a predictive model by easily regressing measurable independent variables (quality parameters) with the targeted dependent variable (BOD_5). It is obtained that the significant correlated independent variables are mainly; the effluent COD and the effluent TSS, and by simply measuring these parameters; the BOD_5 can be predicted from the evolved regression model. MLR model, in this research, shows high levels of correlation, fitness, and confidence. Therefore, the developed model can enhance the control and the automation of the investigated type of biological treatment, namely the hybrid trickling filters-activated sludge system. Similar approach is underway to examine this modeling technique in predicting BOD_5 from easily measured parameters in other types of treatment technologies.

References

[1] C.L. Grady Jr., G.T. Daigger, N.G. Love, C.D. Filipe, Biological Wastewater Treatment, CRC Press, Boca Raton, FL, 2011.
[2] Metcalf and Eddy, Wastewater Engineering: Treatment and Resource Recovery, 5th ed., McGraw-Hill, New York, NY, 2014.
[3] M. Clara, B. Strenn, O. Gans, E. Martinez, N. Kreuzinger, H. Kroiss, Removal of selected pharmaceuticals, fragrances and endocrine disrupting compounds in a membrane bioreactor and conventional wastewater treatment plants, Water Res., 39 (2005) 4797–4807.
[4] S. Velling, A. Mashirin, K. Hellat, T. Tenno, Non-steady response of BOD biosensor for the determination of biochemical oxygen demand in wastewater, J. Environ. Monit., 13 (2011) 95–100.
[5] S. Jouanneau, L. Recoules, M.J. Durand, A. Boukabache, V. Picot, Y. Primault, A. Lakel, M. Sengelin, B. Barillon, G. Thouand, Methods for assessing biochemical oxygen demand (BOD): a review, Water Res., 49 (2014) 62–82.
[6] K. Riedel, G. Kunze, A. König, Microbial Sensors on a Respiratory Basis for Wastewater Monitoring, N.N. Dutta, F. Hammar, K. Haralampidis, N.G. Karanth, A. König, S.H. Krishna, G. Kunze, E. Nagy, B. Orlich, A.E. Osbourn, K.S.M.S. Raghavarao, K. Riedel, G.C. Sahoo, R. Schomäcker, N.D. Srinivas, M. Trojanowska, Eds., History and Trends in Bioprocessing and Biotransformation, Vol. 75, Advances in Biochemical Engineering/Biotechnology, Springer, Berlin, Heidelberg, 2002, pp. 81–118.
[7] I. Skoczko, J. Struk-Sokołowska, P. Ofman, Seasonal changes in nitrogen, phosphorus, BOD and COD removal in Bystre wastewater treatment plant, J. Ecol. Eng., 18 (2017) 185–191.
[8] L. Metcalf, Wastewater Engineering: Treatment and Reuse, Metcalf & Eddy Inc., New York, NY, 2003.
[9] L. Benedetti, J. Langeveld, A. Comeau, L. Corominas, G. Daigger, C. Martin, P.S. Mikkelsen, L. Vezzaro, S. Weijers, P.A. Vanrolleghem, Modelling and monitoring of integrated urban wastewater systems: review on status and perspectives, Water Sci. Technol., 68 (2013) 1203–1215.
[10] H. Melcer, Monitoring and modeling VOCs in wastewater facilities, Environ. Sci. Technol., 28 (1994) 328A–335A.

[11] V. Ranković, J. Radulović, I. Radojević, A. Ostojić, L. Čomić, Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia, Ecol. Modell., 221 (2010) 1239–1244.
[12] S. Dahiya, B. Singh, S. Gaur, V.K. Garg, H.S. Kushwaha, Analysis of groundwater quality using fuzzy synthetic evaluation, J. Hazard. Mater., 147 (2007) 938–946.
[13] S. Areerachakul, Comparison of ANFIS and ANN for estimation of biochemical oxygen demand parameter in surface water, Int. J. Chem. Biol. Eng., 6 (2012) 286–290.
[14] H.M. Azamathulla, A.A. Ghani, S.Y. Fei, ANFIS-based approach for predicting sediment transport in clean sewer, Appl. Soft Comput., 12 (2012) 1227–1230.
[15] M. Savic, I. Mihajlovic, Z. Zivkovic, An ANFIS-based air quality model for prediction of SO_2 concentration in urban area, Serb. J. Manage., 8 (2013) 25–38.
[16] A.M. Ahmed, S.M.A. Shah, Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River, J. King Saud Univ., 29 (2017) 237–243.
[17] M. Bagheri, S.A. Mirbagheri, M. Ehteshami, Z. Bagheri, A.M. Kamarkhani, Analysis of variables affecting mixed liquor volatile suspended solids and prediction of effluent quality parameters in a real wastewater treatment plant, Desal. Water Treat., 57 (2016) 21377–21390.
[18] M. Bagheri, S.A. Mirbagheri, A.M. Kamarkhani, Z. Bagheri, Modeling of effluent quality parameters in a submerged membrane bioreactor with simultaneous upward and downward aeration treating municipal wastewater using hybrid models, Desal. Water Treat., 57 (2016) 8068–8089.
[19] J.P. Suen, J.W. Eheart, Evaluation of neural networks for modeling nitrate concentrations in rivers, J. Water Resour. Plann. Manage., 129 (2003) 505–510.
[20] Statistic Solutions, 2019. Available at: https://www.statisticssolutions.com/independent-and-dependent-variables/ (retrieved August, 2019).
[21] Regression Definition, 2019. Available at: https://www.investopedia.com/terms/r/regression.asp (retrieved August, 2019).
[22] A.E. Tümer, S. Edebalı, Prediction of Wastewater Treatment Plant Performance Using Multilinear Regression and Artificial Neural Networks, International Symposium on Innovations in Intelligent SysTems and Applications (INISTA), IEEE, Madrid, Spain, 2015, pp. 1–5.
[23] L. Qiuhua, S. Lihai, G. Tingjing, Z. Lei, O. Teng, H. Guojia, C. Chuan, L. Cunxiong, Use of principal component scores in multiple linear regression models for simulation of chlorophyll-a and phytoplankton abundance at a karst deep reservoir, southwest of China, Acta Ecol. Sin., 34 (2014) 72–78.
[24] V. Uyak, K. Ozdemir, I. Toroz, Multiple linear regression modeling of disinfection by-products formation in Istanbul drinking water reservoirs, Sci. Total Environ., 378 (2007) 269–280.
[25] M.C. Maniquiz, S. Lee, L.H. Kim, Multiple linear regression models of urban runoff pollutant load and event mean concentration considering rainfall variables, J. Environ. Sci., 22 (2010) 946–952.
[26] N. Basant, S. Gupta, A. Malik, K.P. Singh, Linear and nonlinear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water: a case study, Chemom. Intell. Lab. Syst., 104 (2010) 172–180.
[27] H.Z. Abyaneh, Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters, J. Environ. Health Sci. Eng., 12 (2014) 1–8.
[28] Google Earth, Pro 7.3.2.5776 (retrieved February, 2020).
[29] Climate Data. Available at: https://en.climate-data.org/ (retrieved August, 2019).
[30] V. Bewick, L. Cheek, J. Ball, Statistics review 7: correlation and regression, Crit. Care, 7 (2003) 451–459.
[31] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to Linear Regression Analysis, Vol. 821, John Wiley & Sons, Hoboken, NJ, 2012.
[32] A. Vijayan, G.S. Mohan, Prediction of effluent treatment plant performance in a diary industry using artificial neural network technique, J. Civ. Environ. Eng., 6 (2016) 1–4.
[33] F.S. Mjalli, S. Al-Asheh, H.E. Alfadala, Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance, J. Environ. Manage., 83 (2007) 329–338.
[34] O.T. Bakia, E. Arasb, U.O. Akdemirc, B. Yilmaza, Biochemical oxygen demand prediction in wastewater treatment plant by using different regression analysis models, Desal. Water Treat., 157 (2019) 79–89.
[35] Z. Al-Ghazawi, K. Bani-Hani, R. Alawneh, Testing and Evaluation of the Use of Digital Image Analysis and Artificial Neural Networks in Monitoring Wastewater Quality, 9th Annual International Conference on Civil Engineering, 24–27 June 2019, Athens, Greece.