Available online at: http://jurnal.untirta.ac.id/index.php/jiss

JOURNAL INDUSTRIAL SERVICESS

Industrial Engineering Advance Research & Application

eISSN 2461-0631

An application of multiple regression for predicting turbidity of standard water quality for industrial and household consumption

Yusraini Muharni[1*], Natalia Hartono[2]

[1]*Industrial Engineering Department, Engineering Faculty, Universitas Sultan Ageng Tirtayasa*
[2]*Department of Mechanical Engineering, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

*Corresponding author: yusraini@untirta.ac.id

ARTICLE INFO

Received: 15 September 2021
Revision: 14 Oktober 2021
Accepted: 16 Oktober 2021

Keywords:
Water Quality
Turbidity
Multiple regression

ABSTRACT

A multiple regression approach was applied in this study with the aim of predicting Turbidity value of standard water in water treatment plant. Turbidity is a level of cloudiness in water due to the presence of particles or microorganisms. Turbidity in standard water did not affect human health in term of hazardous, even though it represent of poor quality water. Water treatment plant reduce the cloudiness in water by applying chlorination process. There are three independent variables of water quality involved to predict turbidity value. They are PH, color-spectrum and electrical conductivity. The correlation among variables were checked before conducting multiple regression. Color-spectrum has the highest correlation with the turbidity. The stepwise approach remain two independent variables involved in multiple regression equation, color-spectrum and electrical conductivity with the value of R-square equal to 0,97. Meaning that the two variables has the ability of explaining variances in turbidity up to 97 %.

1. INTRODUCTION

Water supply has important roles in human life either for personal purpose or for business activities. The standard water require special treatment before distributed into residential and industrial area. According to physical characteristic, standard water has several index set by the ministry of industry such as turbidity, color spectrum, electrical conductivity. Turbidity is defined as the "optical quality [of water] that causes light to be scattered and absorbed rather than transmitted in straight lines through a sample" [1, 2, 3]. Turbidity is used to indicate water quality [4, 5, 6, 7] and it is an important parameter to measure the water quality [8]. The aim of the research is to predict the turbidity of standard water in a water treatment plant using a multiple regression approach. Linear models have been used to predict water quality parameters which have provided useful insights into the behaviour of a natural system, however, have demonstrated relatively modest results [1].

2. LITERATURE REVIEW

Turbidity is a parameter for represent water quality. Turbidity is a term used to describe the clearness of a water decreases due to the presence of smallest substances, particles or microorganisms [9, 10, 11]. Refer to the World Health Organisation (WHO). The turbidity level of the water before the chlorination process is not more than 1 Nephelometric Turbidity Unit (NTU). The score of turbidity in raw water is very influential in determining the dosage of chemical raw materials (operating supplies) to be used in water treatment plant [12]. There are various studies focus on turbidity as one parameters of water quality [1], [2], [13, 14, 15]. The highest turbidity score in the water meaning that its quality is low.

Turbidity of water could be caused by the existence of other suspended particles, such clay and silts [16], or soluble colored organic compounds, and plankton and other microscopic organisms [17]. According to Mandal [17], various studies exclaimed that pH on the water system found to be significant showing positive and negative correlation on turbidity removal. The summary

http://dx.doi.org/10.36055/jiss.v7i1.12411

of previous research that focus on several parameter of water quality is presented in Table 1.

Table 1. Previous study related on water quality parameters

Quality water parameters	Authors (years)	Methods	Finding
Turbidity	LeChevallier et al. (1981) [16] Mustapha and Abdu (2012) [15] Stevenson et al (2019) [1] Song and Zhang (2020) [4]	Linear Regression Principal component analysis Multiple Linear regression Generelised Linear Model Random Forest Neural Network	Turbidity has a significant effect on water quality
pH	Mandal, H. E (2015) [17]	Experimental Design	no direct influence of wastewater pH on turbidity
Spectral	Read et al (2015) [18] V. Garg et al. (2017) [19]	Geospatial Technique	attempt to develop spectral library for water quality parameters and its application in spectral similarity analysis
Conductivity	Najah et al (2009) [20] Read et al (2015) E. Skarbøvik and R. Roseth (2015) [8]	Artificial Neural Network Random Forest Sensor	revealed the usefulness of sensor and real-time recording technoogy over traditional water sampling and laboratory analysis

3. RESEARCH METHOD

3.1 Data

Data in this research were collected from water treatment plant located in Banten Province, Indonesia. As much of 187 samples were taken for conducting multiple linear model. The samples are measured for four parameters which are turbidity, pH, color and electrical conductivity.

3.2. Descriptive statistics

Descriptive statistics of the water quality parameters have been summarized in Table 2. Table 2 describes statistical descriptive of parameters in this study. The pH mean is 7.6. it is explained clearly that the standard water in the water treatment plant is which is slightly above neutral level. However, far from acidity or alkalinity.

3.3. Correlation

It is required to check the correlation lying in between the parameters involved in this study. From Table 2, we can identify that Spectral has the strongest relation with Turbidity among other parameters.

Table 2. Statistical descriptive of parameters

Statistics	Turbidity	Spectro	DHL	pH
Mean	63,87	583,29	172,27	7,63
Std. Error	1,79	17,16	2,01	0,01
Median	56,9	515,5	169	7,63
Mode	49,2	450	209	7,61
Std. Dev	24,47	234,08	27,45	0,15
Variance	598,82	54793,7	753,69	0,02
Kurtosis	4,45	6,82	-0,84	0,03
Skewness	1,73	2,09	0,10	-0,19
Range	154,5	1712	118	0,79
Minimum	32,5	199	113	7,25
Maximum	187	1911	231	8,04
Sum	11880	108490	32043	1419,94
Count	186	186	186	186
Confidence Level (95,0%)	3,54	33,86	3,97	0,02

In the second row, with the correlation value of -0.515 electrical conductive has medium -relation to the turbidity with a different direction. Since the correlation has a minus sign. If the electrical conductivity value is increased, the Turbidity level is low.

From Table 2, we can identify that Spectral has the strongest relation with Turbidity among other parameters. In the second row, with the correlation value of -0.515 electrical conductive has medium -relation to the turbidity with a different direction. Since the correlation has a minus sign. If the electrical conductivity value is increased, the Turbidity level is low.

3.4. Multiple linear regression

Multiple linear regression is a statistical tool to analyze interrelationship between dependent variable with two or more independent variables. Mathematical model represent multiple linear regression is shown in equation (1) refer to the similar model in [8]

$$\boldsymbol{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon} \qquad (1)$$

where

y = Dependent variable (Turbidity in this case)

$\mathbf{x_n}$ = independent variable (parameters of water quality, n = 1...3)

$\boldsymbol{\beta_0}$= regression coeficient

$\boldsymbol{\varepsilon}$ = random error in term of regression model

By applying data analysis of 187 data, we can calculate the regression statistics as summarized in Table 4. Multiple R describes the correlation between actual and predicted values of the dependent variable; in this study refer to Turbidity. The R–square is the model's accuracy in explaining the dependent variable. Table 5 shows the result of ANOVA test for regression. The significance of F felt far below the F statistic with degree of freedom of 2. Meaning that, the multiple regression model developed is adequate for prediction turbidity purpose.

Table 3. Correlation matrix among variables

	Turbidity	Spectro	DHL	pH
Turbidity	1			
Spectro	0.984	1		
DHL	-0.515	-0.471	1	
pH	-0.212	-0.239	0.148	1

Table 4. Regression statistics of the water quality parameters for 187 observations

Multiple R	0,985525
R Square	0,971259
Adjusted R Square	0,970947
Standard Error	4,16417

Table 5. ANOVA for regression

	df	SS	MS	F	Sig. F
Regression	2	107822	53911	3109,01	1,5E-142
Residual	184	3190	17,34		
Total	186	111013			

Table 6. p-value of coefficient regression for independent variables

	Coefficients	P-value
Intercept	15,91223	1,846E-08
color spectral	0,09962	1,81E-131
electrical conductivity	-0,0588	5,569E-06

Table 6 describes the coefficient of variables in multiple regression model. The p-value of each variable is consider significant for using in the multiple regression model developed since the value below 0.05.

4. RESULT AND DISCUSSION

This study attempt to predict the standard water quality through the interrelationship among turbidity as the dependent variables, with three independent variables such as pH, Spectral and electrical conductivity. A multiple regression model was employed to find the mathematical model to predict turbidity level. Table 7 in the previous section shows the p-value for each coefficient of spectral (x1) and electrical conductivity (x2) are significant. Therefore, Turbidity level could be predicted by applying the equation below.

$$y = 15.92 + 0.0996\,x_1 - 0.05889\,x_2 + \varepsilon \qquad (2)$$

The multiple regression above is quite robust since the R square is 0.971259. Meaning that spectral (x1) and electrical conductivity (x2) can explained the variation of turbidity up to 97.13%, and the rest variation is explained by others. See Fig, 1 for the plot. In order to check that the assumption of multiple regression model are fulfilled, we plot the residual value. Figure 2 shows the residual plot of color spectral of standard quality water.

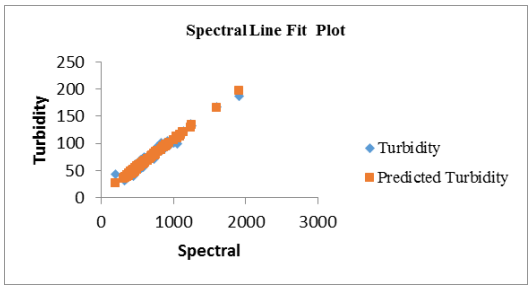

Figure 1. Turbidity vs predicted one

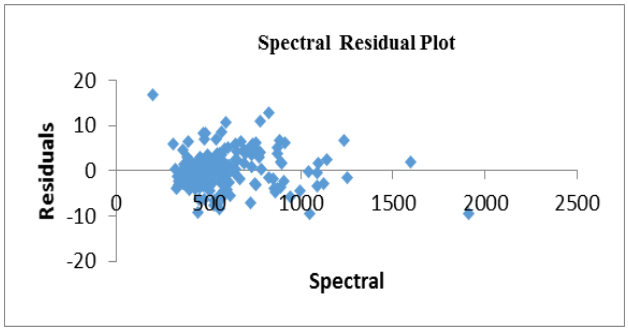

Figure 2. Spectral Residual Plot fulfill the assumption

5. CONCLUSION

This study has developed a turbidity prediction model by applying a multiple regression method with a stepwise approach. This model is very useful for managing water treatment plant. The level of turbidity has impact on purchasing the chemical stuff for improved the water quality.

ACKNOWLEDGEMENT

The authors would like to acknowledge the generous support from the Indonesia Endowment Fund for Education (LPDP); Industrial Engineering University of Pelita Harapan

REFERENCES

[1] M. Stevenson and C. Bravo, "Advanced turbidity prediction for operational water supply planning," *Decis. Support Syst.*, vol. 19, no. 4, pp. 72–84, 2019, doi: 10.1016/j.dss.2019.02.009.

[2] Y. Zhang, X. Yao, Q. Wu, Y. Huang, Z. Zhou, J. Yang, and X. Liu, "Turbidity prediction of lake-type raw water using random forest model based on meteorological data: A case study of Tai lake, China," *Journal of Environmental Management.*, vol. 290, no. 2, pp. 112657, 2021, doi: 10.1016/j.jenvman.2021.112657.

[3] S. Neupane, J. R. Vogel, D. E. Storm, B. J. Barfield, and A. R. Mittelstet, "Development of a turbidity prediction methodology for runoff–erosion models," *Water, Air, & Soil Pollution*, vol. 226, no. 12, pp. 415, 2015, doi: 0.1007/s11270-015-2679-9.

[4] C. Song and H. Zhang, "Study on turbidity prediction method of reservoirs based on long short term memory neural network," *Ecological Modelling*, vol. 432, no. 12, pp. 109210, 2020, doi: 10.1016/j.ecolmodel.2020.109210.

[5] C. Iglesias, J. M. Torres, P. J. G. Nieto, J. R. A. Fernández, C. D. Muñiz, J. I. Piñeiro, and J. Taboada, "Turbidity prediction in a river basin by using artificial neural networks: A case study in Northern Spain," *Water Resources Management*, vol. 26, pp. 319–331, 2014, doi: 10.1007/s11269-013-0487-9.

[6] J. Huang, R. Qian, J. Gao, H. Bing, Q. Huang, L. Qi, S. Song, and J. Huang, "A novel framework to predict water turbidity using Bayesian modeling," *Water Research*, vol. 202, pp. 117406, 2021,

doi: 10.1016/j.watres.2021.117406.

[7] C. M. Kim and M. Parnichkun, "Prediction of settled water turbidity and optimal coagulant dosage in drinking water treatment plant using a hybrid model of *k*-means clustering and adaptive neuro-fuzzy inference system," *Applied Water Science*, vol. 7, pp. 3885–3902, 2017, doi: 10.1007/s13201-017-0541-5.

[8] A. A. Scandinavica, S. B. Soil, E. Skarbøvik, and R. Roseth, "Use of sensor data for turbidity , pH and conductivity as an alternative to conventional water quality monitoring in four Norwegian case studies," *Acta Agric. Scand. Sect. B — Soil Plant Sci.*, vol. 65, no. 1, pp. 63–73, 2015, doi: 10.1080/09064710.2014.966751.

[9] M. E. Bote and W. M. Desta, "Removal of turbidity from domestic wastewater using electrocoagulation: Optimization with response surface methodology," *Chemistry Africa*, vol. 5, pp. 123–134, 2022, doi: 10.1007/s42250-021-00303-2.

[10] H. Yao, W. Zhuang, Y. Qian, B. Xia, Y. Yang, and X. Qian, "Estimating and predicting metal concentration using online turbidity values and water quality models in two rivers of the Taihu basin, Eastern China," *PLoS ONE*, vol. 11, no. 3, pp. e0152491, 2016, doi: 10.1371/journal.pone.0152491.

[11] Y. Muharni, K. Kulsum, and A. Denisa, "Prediksi kualitas air baku dengan pendekatan Adaptive Neuro Fuzzy Inference System," *Semiinar Nas. IENACO 2018*, pp. 602–606, 2018.

[12] M. S. Gaya, M. U. Zango, L. A. Yusuf, M. Mustapha, and B. Muhammad, "Estimation of Turbidity in Water Treatment Plant using Hammerstein-Wiener and Neural Network Technique," vol. 5, no. 3, pp. 666–672, 2017, doi: 10.11591/ijeecs.v5.i3.pp666-672.

[13] I. G. Ezemagu, M. I. Ejimofor, M. C. Menkiti, and C. C. Nwobi-okoye, "Modeling and optimization of turbidity removal from produced water using response surface methodology and artificial neural network," South African J. Chem. Eng., vol. 35, pp. 78–88, 2021, doi: 10.1016/j.sajce.2020.11.007.

[14] A. Mustapha and A. Abdu, "Application of Principal Component Analysis & Multiple Regression Models in Surface Water Quality Assessment," *Journal of Environment and Earth Science*, vol. 2, no. 2, pp. 16–23, 2012.

[15] M. W. Lechevallier, T. M. Evans, and R. J. Seidler, "Effect of Turbidity on Chlorination Efficiency and Bacterial Persistence in Drinking Water," Appllied Environ. Microbiol., vol. 42, no. 1, pp. 159–167, 1981.

[16] Harasit Kumar Mandal, "Influence of wastewater PH on Turbidity," *Int. J. Environ. Res. Dev.*, vol. 4, no. 2, pp. 105–114, 2015.

[17] E.K. Read, V.P. Patil, S.K. Oliver, A.L. Hetherington, J.A. Brentrup, J.A. Zwart, K,M. Winters, J.R. Corman, E.R. Nodine, R.I. Wollway, H.A. Dugan, "The importance of lake-specific characteristics for water quality accorss the continental United States," *Ecological Apps.*, vol. 25, no. 4, p 943-955, 2015, doi: 10.1890/14-0935.1.

[18] V. Garg, A.S. Kumar, S.P. Aggrawal, V. Kumar, P.R. Dhote, P.K. Thakur, B.R. Nikam, R.S. Sambare, A. Siddiqui, P.R. Muduli, G. Rastogi, "Spectral similarity approach for mapping turbidity of an inland waterbody," *J. of Hydrology*, vol 550, p. 527-537, 2017, doi: 10.1016/j.jhydrol.2017.05.039.

[19] A. Najah, A. Elshafie, O.A. Karim, O. Jaffar,"Prediction of Johor River water quality parameters using artificial neural networks," *European J. of Sci Research*, vol 28, no. 3, p. 442-435, 2009.

[20] Y. Wang, J. Chen, H. Cai, Q. Yu, and Z. Zhou," Predicting water turbidity in a macro-tidal coastal bay using machine learning approaches," *Estuarine, Coastal and Shelf Science*, vol 252, p. 107276, 2021, doi: 10.1016/j.ecss.2021.107276.