The application of artificial neural network for quality prediction of industrial standard water

Y Muharni [1]*, Kulsum, A Denisa[2] and Hartono[3]

[1]Department of Industrial Engineering University of Sultan Ageng Tirtayasa, Banten 42435 Indonesia
[2]Vopak Terminal Merak, Banten Indonesia
[3]Department of Electrical Engineering University of Sultan Ageng Tirtayasa, Banten 42435 Indonesia

*Email: yusraini@untirta.ac.id

Abstract. Industrial Standard water is a source of water before being distributed in industry and household in a certain area. For the sake of health, quality water is a must to fulfill and control. Quality of water having several variables as indicators. One indicator, namely, Turbidity. It is defined as the level of cloudiness of water due to the presence of particles, mud or microorganisms. The highest of turbidity value meaning the index of water quality is low. In this study, we apply the Artificial neural network method for predicting the turbidity value. Three input variables are engaged, PH level, color spectrum, and electrical conductivity. As much of 827 data were collected during six months. Seventy percent are used for training and the rest thirty percent were used for testing. The ANN architecture consists of 3-6-1 configuration, 3 input variables, 6 hidden layers, and 1 output variable. The training was set into 1000 epoch and the MSE shows 0,0013, meaning that the ANN has the power of prediction. The prediction of turbidity level has a managerial implication as supporting information for purchasing decision of material in water processing

1. Introduction

As the growth of population increased along with the advancement of the industry, the presence of clean water becomes an important natural resource that need to be noticed. In a certain city, a water treatment plant is managing by a private company. A supplier of clean water with processing raw water into clean water. Most of the clean water produced is distributing for industrial needs and for the community. Water treatment plants process raw water to become clean quality water through some processings, physically, chemically, and biologically. Before being distributed to consumers, water was analysed and required testing of some level parameters of its quality. From several water quality parameters, turbidity is the most important parameter effects on water quality. The value is scored in Nephelometic Turbidity Unit (NTU). Refer to World Health Organisation (WHO), the score of turbidity in standard water can not exceed 1 before chlorination. [1] Score of tturbidity in raw water is very influential in determining the dosage of chemical raw materials (operating supplies) to be used. [2]

Water treatment instalation received water supplied from water basin in the outdoor area. There are several water supplies that will terminate in the basin. This situation will influenced the level of turbidity from time to time. In the rainy season, the presence of muds are increased significantly. Meaning that the score of turbidity will be high. In order to anticipated uncertainty of the peak of turbidity, the presence of prediction method could help in maintain the water treatment process. Artificial Neural

Network (ANN) is one of many computational methods frequently used in research related to prediction or forecasting. [3] ANN method was chosen because is a learning system that able to learn like the human brain based on knowledge base. In the context of water quality modelling, ANN bring the the current state-of-the-art of the integration of Artificial Intelligence.

Models used in ANN that is Backpropagation Neural network, used because of the ability to generalize which is good enough for prediction problems and its ability to lower gradient to minimize the square of error of network outputs [4]. To produce best prediction result with the smallest error value, it is necessary to know what factor is controlled the influence and how these factors must be set. To get the best ANN architecture, then the evaluation process carried out by the Taguchi method. Taguchi Method is used to avoid evaluation of the interaction of all factors thereby reducing the number of experiments involved. So, in this research, prediction of industrial standard water quality by applying Algorithm Backpropagation Neural Network using MATLAB software. The objective is to determine water quality predictions for upcoming period.

2. LITERATURE REVIEW

2.1 Water Quality

Water quality in general shows associated water condition with an activity or need certain. One of the parameters for water quality is called turbidity.[4] The term turbidity is used to describe the transparency of a liquid decreases due to the presence of insoluble substances. Turbidity of water can be caused if clay sand, organic substances, inorganic fine, plankton and other microorganisms exist, and usually will affect the level of water color. Turbidity is the most important parameter effect on water quality[4]. Score of turbidity called nephelometric turbidity units (NTU), in raw water is very influential in determining the dosage of chemical raw materials (operating supplies) to be used.

Quality water can be declare refer to its pH value. pH is the degree of acidity used to express the level of acidity or alkalinity possessed by a solution [5] The previous study conducted by Mandal analyzed how pH wastewater influenced water turbidity [6] The other physical parameter that having effect on water turbidity is Electrical conductivity. It is defined as the ability of a material to conduct electric current. [7] According to electrical conductivity in water, it is influenced by how big the ion content is, if the ion content is a condition for the occurrence of possible electrical conductivity, then these conditions can be fulfilled as a conductor of electric current. In the process of water treatment or water treatment, there is an important relationship between turbidity, PH, color and electrical conductivity. The levels between turbidity, PH, color and electrical conductivity are interrelated. For example the measurement of turbidity level is very useful for water treatment plants to ensure its cleanliness. In industrial processes, turbidity can be part of Quality Control to ensure efficiency in processing or related industrial processes.

2.2 Artificial Neural Network

Artificial neural networks are an information processing system that imitate how a human brain process the information in biological nerve cells. Artificial Neural Network was formed to solve a particular problem such as pattern recognition or classification since it has the ability of learning process.[8]
In mathematics forms, ANN derived biological neural networks, by assuming: [3]

1. The processing of information take place in many simple elements (neurons)
2. The signal is delivered between the neurons through the link
3. The connection between neurons has a bonded weight which in the network multiplies the transmission signal.
4. an activation function (usually nonlinear) is required in each neuron to determine the output signal.

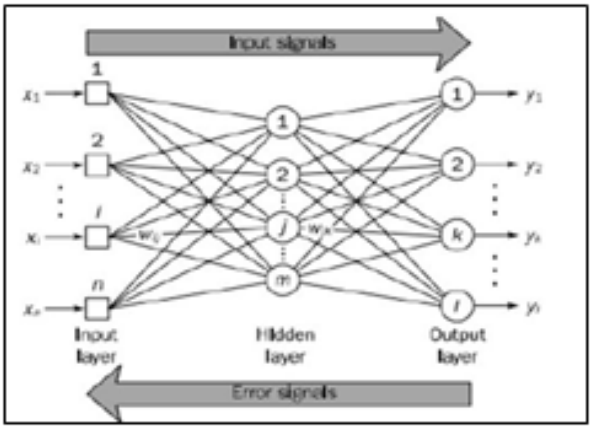

Figure 1. Backpropagation network architecture

2.3 Taguchi Method

New methodologies in engineering which aims to improve the quality products and processes and can suppress minimal costs and resources with an emphasis on achievement specific target and reduce the variation of a product or process by using parameter design. [9]

3. Research Method

3.1 Data

Data in this research were collected from water treatment plant located in Banten Province, Indonesia. Among the different water quality parameters, we choose four parameters there are turbidity, pH, color and electrical conductivity [5]. Totally there are 827 data for each parameter. The data for four water quality parameters are resumed in Table 1.

Table 1. Data for the water quality parameters

No	pH	Turbidity	Color	Electrical conductivity
1	7.35	25.1	350	93
2	7.37	19.2	259	96
3	7.44	17.9	246	93
4	7.38	22	298	94
5	7.41	19.8	231	97
…	…	…	…	…
823	7.52	46.5	385	171
824	7.49	39.8	387	181
825	7.61	39.5	361	180
826	7.67	42.3	394	179
827	7.66	51.5	477	177

Before being processed in the ANN, it is necessary to do data preprocessing through data normalization. The data normalization is a required before feed data into the network so as the network does not experience saturation and failure when do learning. The data normalized in the interval [0,1] [10], through a linear tranformation (Eq. 1)

$$x = \frac{x - x_{min}}{X_{max} - X_{min}} \quad (1)$$

Based on data acquisition, training and data testing requires 827 data input and target data. The input data that referred to is daily data consist of: pH of water, water color, and electrical conductivity, while the target data is turbidity data water. After normalization, the data can be processed for training and testing with ANN. As much as Seventy percent from 827 data are used for training and the rest thirty percent for testing.

3. 2 *Orthogonal arrays and design experiment*

This research makes a design experiment using matrix orthogonal array contained in Taguchi method, based on degrees of freedom, factor, and factor level. In this research using an orthogonal array L 9 (3 3), which means the experiment was carried out as much as 9 trials or nine combinations level with 3 columns, as shown in Table 2. After conducting the appropriate experiment orthogonal array, then we get 9 experimental results in the form of MSE values, with each experiment is replicated three times as shown in Table 3.

Table 2 orthogonal Array

Orthogonal Array Factors				Training Function and Design experiment		
	A	B	C	Training Function	*Layer*	momentum
1	0	0	0	Levenberg-Marquardt	3	0.1
2	0	1	1	Levenberg-Marquardt	6	0.5
3	0	2	2	Levenberg-Marquardt	9	0.9
4	1	0	1	Bayesian Regularization	3	0.5
5	1	1	2	Bayesian Regularization	6	0.9
6	1	2	0	Bayesian Regularization	9	0.1
7	2	0	2	Scaled Conjugate Gradient	3	0.9
8	2	1	0	Scaled Conjugate Gradient	6	0.1
9	2	2	1	Scaled Conjugate Gradient	9	0.5

3.3 Data Processing

ANN carried out for predicting the *turbidity* of water for upcoming period. The use of MATLAB is done by entering the *input* data, namely data pH, color and electricaal conductivity that has been normalized, and the target data i.e. data *turbidity* which has been normalized. The best architecture of ANN for prediction is 3-6-1 configuration (see Fig. 2) is achieved by apply Taguchi method.

Table 3 Mean squre error (MSE) value from the replications

No	**Replication 1**	**Replication 2**	**Replication 3**
1	0.002350	0.002542	0.002609
2	0.001744	0.002541	0.002477
3	0.002166	0.001437	0.002106
4	0.001973	0.003550	0.002315
5	0.001309	0.001313	0.001282
6	0.001107	0.001113	0.001590
7	0.002678	0.003253	0.002634
8	0.002507	0.002065	0.002441
9	0.002064	0.002695	0.003872

The architecture of ANN was built consist of 3 layer (Fig. 2). Three input node in the first layer have 3 variables involved. They are pH, color spectrum and electrical conductivity. In the second layer is hidden layer with six nodes. The hidden layer has activated function. That transformed the value from input variables. hidden layer outcome will be delivered to output layer. The best architecture of Artificial *neural* network to predict *turbidity* with the smallest error is using Bayesian training function Regularization with number of *hidden layers* 6, and a momentum of 0.9.

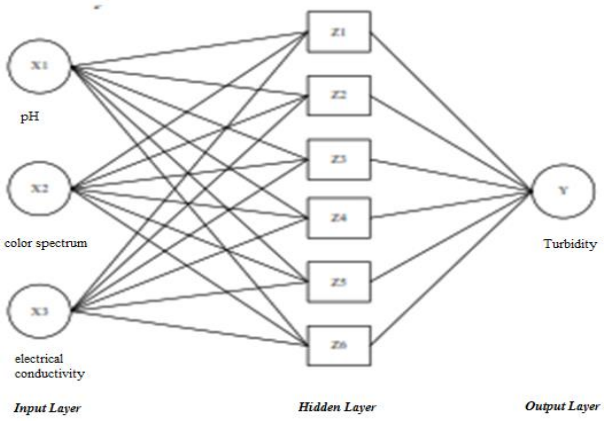

Figure 2. Best architecture of ANN for predicting turbidity level

4. Result and discussion

The ANN was set to run for 1000 epoch in order to experience the best training process. The training performance for predicting turbidity has its best at epoch 343th with mean square error 0,0013 which lower than 0,05. (Fig. 3). The error in training is very high in early epoch and it is decrease significantly as the number of epoch increase. This process shows that the network learn gradually and adaptation the weight in every network it has. The mean squarre error (MSE) has its lowest value at epoch 343th and ANN stop the training process. These MSE represents the average squared between outputs (x) and targets (y) during the training, see eq. 2 [11]

$$MSE = \frac{1}{n}\sum_{1}^{n}(x-y)^2 \qquad (2)$$

The other performance that show the ability of prediction is corelation value is 0. 944, meaning that the ANN has a good performance in prediction the level of turbiity.

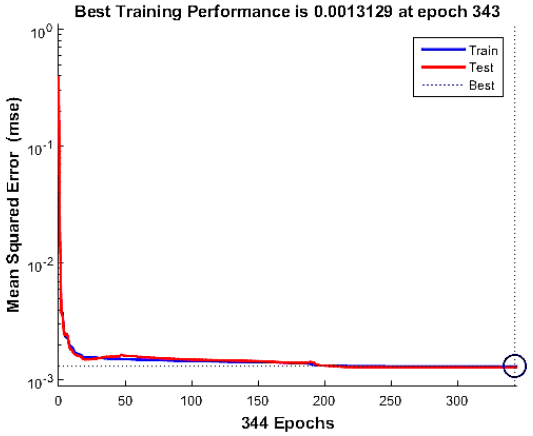

Figure 3 ANN Best Training Performancce for Turbidity prediction

As mention in the paragraph before that the ANN training process show a good performance, meaning that it is ready to start the testing to do the turbidity prediction. As shown in Fig. 4. The testing result lies very closely to the actual one. This result has the correlation value (testing) 0,94789 between outputs and targets during testing.

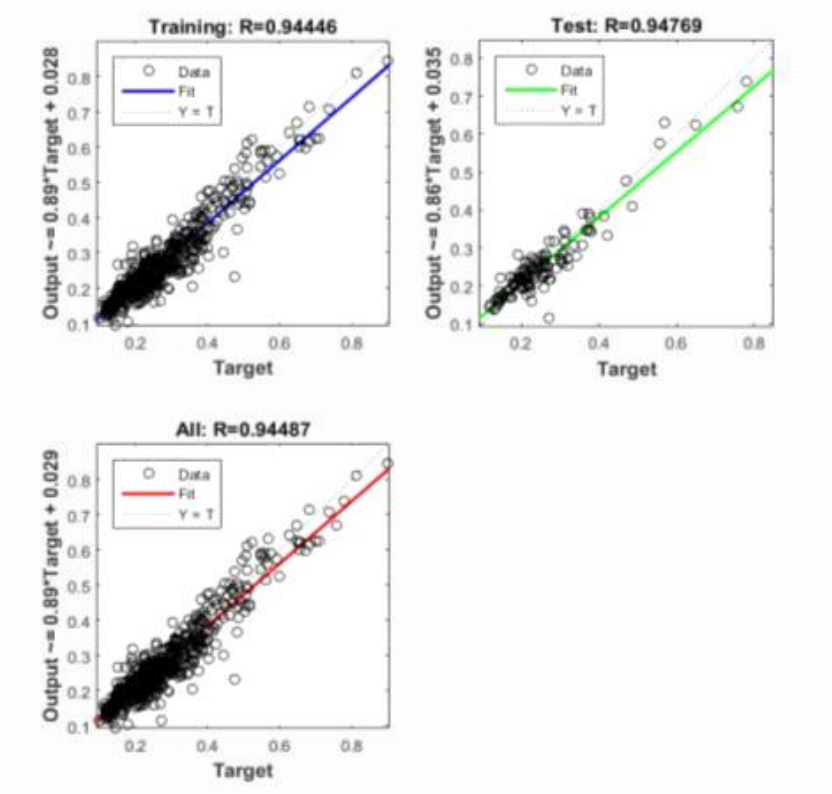

Figure 4. Turbidity Prediction Performance

The Artificial Neural Network was tried to predict turbidity of standard water in triple way, Training,Testing and Validity by dividing the data properly. Now, we can infere the ANN prediction results from figure 4. The corelation value during testing is 0.94769 not much differ from the value during training the networks 0.94446.Meaning that the overtraining phenomena did not exist. We can applied the ANN prediction results as information support of turbidity value in upcoming period. The application of neural network for prediction turbiditylevel has been widely applied with different input variables due to many factors affecting turbidity.[12]

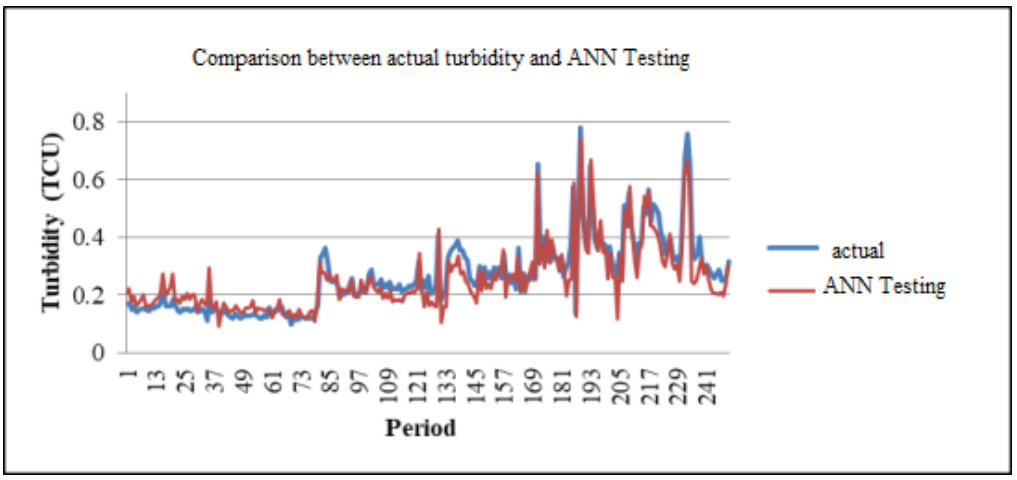

Figure 5. Comparison between the actual turbidity level and the result of prediction

5. CONCLUSION

In this study, the artificial Neural Network (ANN) was successfully develop to predict turbidity level in water treatment plant by involved pH, Color spectrum and electrical conductivity as input variable. The artificial Neural Network (ANN) has a good performance with MSE value is 0.0013. Prediction of turbidity level for the future period are 41.1 NTU in average.

REFERENCES

[1] M. Stevenson and B. Cristian, Decis. Support Syst. **02**, 1 (2019).
[2] Y. Muharni, K. Kulsum, and A. Denisa, Semiinar Nas. IENACO 2018 602 (2018).
[3] Y. Muharni, A. Irman, and M. Ilhamsyah, 1 (n.d.).
[4] M.S. Gaya, M.U. Zango, L.A. Yusuf, M. Mustapha, and B. Muhammad, **5**, 666 (2017).
[5] B.I. Gandiwa, L.B. Moyo, S. Ncube, T.A. Mamvura, L.L. Mguni, and N. Hlabangana, South African J. Chem. Eng. **34**, 158 (2020).
[6] Harasit Kumar Mandal, Int. J. Environ. Res. Dev. **4**, 105 (2015).
[7] A.A. Scandinavica, S.B. Soil, E. Skarbøvik, and R. Roseth, Acta Agric. Scand. Sect. B — Soil Plant Sci. **65**, 63 (2015).
[8] M.P. Bulletin, Mar. Pollut. Bull. **52**, 726 (2006).
[9] M. Sanjari, A.K. Taheri, and M.R. Movahedi, (2008).
[10] M. Khashei, S.R. Hejazi, and M. Bijari, **159**, 769 (2008).
[11] L. Wang, Z. Wang, H. Qu, and S. Liu, Appl. Soft Comput. J. (2018).
[12] C. Song and H. Zhang, Ecol. Modell. **432**, 109210 (2020).