The Application of Neural Network for Predicting Corrotion Rate in Metal Pipe Installation

Zulkifli[1], Detak Yan Pratama[1], Dyah Sawitri[1], and Dotty Dewi Risanti[1]

Abstract—***Corrotion is one of the problems that must be considered in the metal pipe installation because it can disturb the operation of the plant. The possibility of the corrotion occurrence can be predicted using neural network system. The black box system in the neural network can be used to calculate several potential causes the corrotion and to predict the corrotion rate. This study had constructed the prediction system of corrotion rate using neural network. The input of the system are material compositions, pH, flow rate and temperature. The material compositions which are used are Carbon (C), Manganese (Mn), Silicon (Si), Phosphorus (P), Sulphur (S), Chromium (Cr), Molybdenum (Mo), Aluminium (Al), Nickel (Ni) and Iron (Fe). The corrotion rate prediction network is using one hidden layer and lavenberg marquardt for the learning algorithm. The Mean Square Error (MSE) which is used to analyze the network performance indicates that both of training and validation show excellence results. The MSE of training is 0,000338971 and the validation is 0,000493117.***

Keywords— Lavenberg Marquardt, Material Compositions, pH, Flow Rate and Temperature.

I. INTRODUCTION

One of the important factors on the pipe installation is the risk of corrosion. Corrosion is a unique physical phenomenon which often occurs in tropical region such as Indonesia. It is caused the tropical region has a different humidity in the day and night. If the pipes are installed on shore or off shore, the corrosion will be taken place faster. It is caused the saline concentration are higher than the ground. Also, the corruption can occur when the alloys are located in high sulfuric region.

The corrosions are very vital to be attend because it will influence performance of the system. Besides that, the higher risk gives high effect to the health safety and environmental system. Also, the corrosion can shut the system down which can give the business effects. These effects are persuaded by poor performance of the system, rejected product, lower safety, trouble system and high production cost.

Corrotion can be happened in all of the alloy materials. Almost all of industry use alloy for constructing their system. Thus, they must be attention to the corrotion effect. Installation of plant for producing oil and gases, ship, plane, bridge, car and many production process in industrial are utilized by alloy.

In order to construct the robust system for industry, it would be attention about the material which will be used and other external factors from the system. These external factors that must be attended are temperature of the surrounding and humidity.

Although, the corrosion effects still can be reduced by making the estimation rate of the corrotion. If the corrotion can be calculated, thus the effect can be overcomed. Thus, the possibility of the corrotion on the pipe installation need to be estimate to construct safety system. By calculating several cause of the corrotion, it can be used to predict the corrotion rate. If the corrotion rate has been known, it is possible to build robust system for industry. But, it is not the simple work to create a formula which correlate the cause and effect of the corrotion, because it can be a nonlinear correlation.

One of the several method for determining a problem which is influenced by several caused had been built. This is called Neural Network (NN). An NN is one of the statistical models which can be used to solve nonlinear problems. Like a biological system, NN is adopting human nervous system where sense the input signal and process it into neuron hereafter which called hidden nodes.

There are several neurons in hidden layer, which the number can be set manually or automatically depending on the type of the network model. Each of transmitted signal information from input layer is weighted before it is processed at the neurons in the hidden layer.

Based on the training method, the NN can be classified into two kind of methods that is supervised and unsupervised training. In the supervised NN, the most important process is to introduce the network on the number of data which are closed to the target output that is called as training process.

To train the network, it need several data which are indicate the true condition of the process. By using a proportional training data, the NN will identify a suitable pattern by training process and will attain a good accuracy of the prediction system. Thus, the NN process is defined as the black box processing. By using the black box processing, the NN can predict several outputs, which are correlated with multi inputs on the same network.

This study had been using neural network to predict the corrotion rate of the pipe installation. The data of the training are several factor which influence the corrotion of the pipe. They are material compositions, pH, and temperature and flow rate of the fluida into the pipe.

II. METHOD

NN model for predicting corrotion rate of the pipe installation had been developed. It needs several steps for creating a valid system. These steps include creating

[1]Zulkifli, Detak Yan Pratama, Dyah Sawitri, and Dotty Dewi Risanti are with Departement of Physcs Engineering, Faculty of Industrial Technology, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia. E-mail: zulab@ep.its.ac.id.

the database, modeling the network by using proportional data and evaluating the system.

This study is not include all of the factor which can be generate corrotion in pipe installation, but only several important factors, which are material compositions, pH, temperature and flow rate of the fluida. For the material compositions, this study uses Carbon (C), Manganese (Mn), Silicon (Si), Phosphorus (P), Sulphur (S), Chromium (Cr), Molybdenum (Mo), Aluminium (Al), Nickel (Ni) and Iron (Fe).

A. Database Construction

The fundamental factor of NN is the proportional and valid database. If the database are suitable for the main problems, thus the network can work accurately. As inform above, the network for this study use several factors which are thirteen material compositions, pH, flow rates and temperatures. These variable are used as the input in neural network. The scheme of network can be shown in figure 1.

To create the database, it needs to get data which are suitable to thirteen variables as shown in figure 1. Sometimes, it is difficult to find data which have the same form, but they can be devided as data point, table data or graphycal data. All of these data must be combined as a database. Because of that, all of data form will be modified as data point.

This study have collected 2091 data which are collected from several journal and books. These data will be divided by two parts as training and validation data. 75% of the data (or 1568 data) are used to train the network and the 25% data (or 523) are used to validate the network. These data must be statistically analyze before they are simulated to the network as shown in table 1.

From the table 1, it can be noticed that almost all of the data have low deviation. It indicates that the data are very sharp and it hopes can develop valid prediction network. Table 1 also shows that the biggest composition of the material is iron (Fe). It is suitable that pipe installations which are corroded easily are the pipes that are made from iron.

Before all of this database are trained to the network, they will be normalized based on mean and standard deviation method. It is usefull because the value of the database are very spread, thus they will be processed to produce closed value.

B. Neural Network Models

After the database are created, the next step of prediction system building is determined model of neural network. This study use feed forward back propagation method which has three layer, i.e. input, hidden and output layer as shown in figure 2.

Each layer of the network has a number of node which indicate the neuron in human neural. The input layer has thirteen nodes which are indicate the thirteen variables of the system. The data which enter from these nodes will be train to the next layer by using lavenberg marquardt as learning algorithm.

Whereas the number of nodes in hidden layer will be traced among 1 to 10 nodes. The exactness number of hidden nodes will be obtained if the network gives satisfied performance. The nodes of this layer will summing the all of the data coming from input layer and will be calculated by tangent sigmoid as an activation function. All of the output from this layer will be sent to the output layer and calculated again there.

All of the data from node in hidden layer will be calculated to the output layer to produce corruption rates. The calculation in this node is using activation function logarithmic sigmoid. Both of the input and hidden layer have bias node which will be also calculated to the network. Usually, the the value of each bias node is 1.

Each layer of the network will be correlated by weights. These weights will correlate all of the nodes from each layer. During the training, these weight will be updated until the network outputs has small different from the target outputs. In other words, MSE value of these conditions are small. For the initializing these weight, this study is applying nguyen widrow algorithm. This algorithm chooses values in order to distribute the active region of each neuron in the layer approximately evenly across the layer's input space. The values contain a degree of randomness, so they are not the same each time this function is called. Because of the randomly initializing weights, each training in this study will repeat five times. It is done for getting the best and accurate training results. The important and crusial key for building a network is the training part by using precise data.

C. Evaluation of Performance

For getting the suitable network, it needs performance indicator. In this study, performance indicator which is used is Mean Square Error (MSE). MSE is the error that shows the difference of output prediction and the target as shown in eq.(1) :

$$MSE = \frac{\sum_{j=1}^{t}(Y_j' - Y_j)^2}{t} \tag{1}$$

where, Y_j' is the predicted output, Y_j is the actual target output, t is the number of training data pairs, and j = 1,2,3,…,t. The best performance result is indicated by small value of MSE.

III. Result and Discussion

A. Determination of Network Architecture

For determining network architecture, it needs to train all of training data to the network. Inside of the training part, these training data needs to be simulated to the trained network. On the training simulation, it can be calculate the performance of the network by MSE. Also data for validation will be simulated to the trained network and are calculated the network performance by using MSE.

The results of MSE both of training and validation part had been shown in figure 3. These MSE values are computed as a function of number of hidden nodes. The performance of the network look fluctuating when the network had 1 to 4 hidden nodes. But, it will be seen as stable when the network had more than 4 nodes.

The MSE values on training part always get lower than validation. It can be done because the simulation for this part have the same data to the training network. Within simulation for validation part had different data to the training part. Even though simulation had different data from the training part, it will become a check to the network whether the network had the high-quality performance or not.

To verify the best performance based on the number of hidden nodes, it must be calculate among of training and validation. Best performance can be achieved when both of training and validation give small MSE values. Because of that, it must be noted to the performance in number of nodes from 4 to 10. Figure 3 is too difficult to analyze because the values are to close. Thus, figure 4 shows zooming of figure 3 especially in 5 until 10 nodes.

The MSE values for the network based on hidden nodes from 4 to 10 which had shown in figure 4 confirmed that both of training and validation in 9 nodes give the better performance compared others. For this condition, the MSE of training and validation part are 0.0008 and 0.002 respectively.

The MSE values of the training and validation part as shown in figure 4 are the average of five repeats from the network. The repetitions of this training are caused randomly initialization of the weights as said in methodology section above. Because these value still average of five repetition, it should be determined one of the value which indicate the truth small MSE value.

Figure 5 shows the MSE values on the 9 nodes based on the repetitions. From this figure, it can be analyze and specified when the network gives the better performance. Process of determining best repetitions is comparing both of training and validation part. The best repetition can be fulfilled when the network use 5^{th} repetition. For this condition, MSE training and validation have 0.0003 and 0.0004 respectively.

B. Weighting the Network

It can be said that the best architecture of the network will occur when the network has 9 nodes in its hidden layer. Then, the repetition of the network will be chosen is on 5^{th} repetition.

When the architecture of the network has been determined, thus it needs to record the weights from each layer connections to be used to predict corruption rate. The weight diantara input and hidden layer can be shown in table 2a and 2b. Table 2a shows the weights of each input node of network when connected to 1^{st} until 5^{th} nodes on hidden layer. While table 2b shows the weight of each input nodes to the 6^{th} to 9^{th} nodes in hidden layer.

The weights from the input bias node to the each nodes in hidden layer can be shown on table 3. Sedangkan weight from hidden node to the output and the bias hidden node to output node can be shown on table 4 and 5 respectively.

IV. CONCLUSION

From the methodology and results of this study, it can be concluded:

1. Neural Network can be used to train the database of corruption rate depend on chemical compositions, pH, temperature and flow rates.
2. The architecture of network is feed forward back propagation with 9 nodes in hidden layer
3. MSE values of training and validation simulation shows that NN can give an excellent results by 0.000338971 and 0.000493117 for training and validation part respectively.

ACKNOWLEDGEMENT

Equipment and facilities provided by the Department of Mechanical Engineering ITS, are acknowledged.

REFERENCES

[1]. Callan, R. 1999. The Essence of Neural Networks. London: Prentice Hall Europe.

[2]. Faussett, L. 1994. Fundamentals of Neural Networks: Architecture, Algorithms and Applications. New Jersey: Prentice Hall, Englewood Cliffs

[3]. Fang, Haitao,M.S.,November 2009. “Low temperature and high salt concentration effects on General co2 corrosion for carbon steel”. Chemical and Biomolecular hal. 18 – 22.

[4]. Sarver, Emily., Edwards, Marc., 2011. “Effects of Flow Brass Location, Tube Materials and Temperature on Corrosion of Brass Plumbing Devices”. Corrosion Science 53(2011)1813–1824 hal. 1815-1817

[5]. Aji, Gofar Ismail ., 2010. “Analisa Lju Korosi Berdasarkan Perbandingan Hasil Kupon, Corrotion modeling, dan Pengukuran Metal Loss Pada Sistem Perpipaan”. 361, hal 798–805

[6]. George, K.S ., Nesic,S.,2007. “Investigation of Carbondioxide of Mild Steel in the Presence of Acetic Acid Part 1: Masic Mechanism”. Corrotion Sciences Section , hal. 178 – 186

[7]. Adiyanti, Yuningtyas R.,Kurniawan,Budi Agung., 2011. “Pengaruh Temperatur dan pH Terhadap Karakterisasi Korosi Baja BS 970 di Lingkungan CO2. Journal of Corrotion Rate, hal. 1-7.

[8]. Kuroda,Takako., Takai, Ryuzo.,Kobayhasi,Yuki., Tanaka,Yoshiteru.,Hara,Soichi.,2008. “Corrosion Rate of Shipwreck Structural Steels under the Sea”. Journal of Corrotion rate , hal. 1 – 6.

[9]. Sulistyoweni, W., Ashadi,H.W., Wicaksono,A.K., 2002. “pengaruh unsur – unsur kimia korosif terhadap laju korosi tulangan beton di dalam air rawa”. Makara teknologi,Vol 6.No 2 , 1–5.

[10]. Malinov, S., Sha, W. & Guo, Z. 2000. Application of Artificial Neural Network for Prediction of Time-Temperature-Transformation Diagrams in Titanium Alloys. Materials Science and Engineering A283: 1-10.

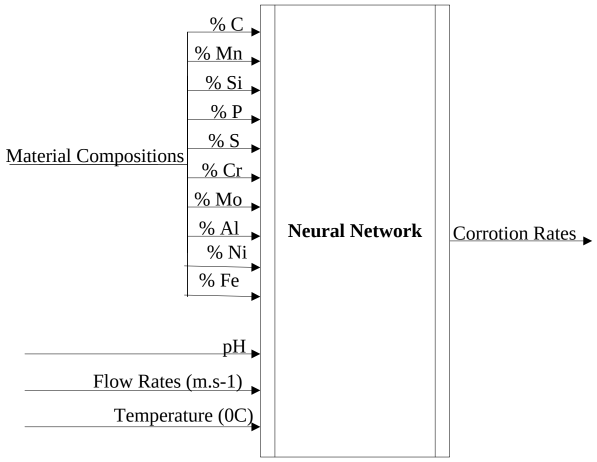

Figure 1. Prediction System

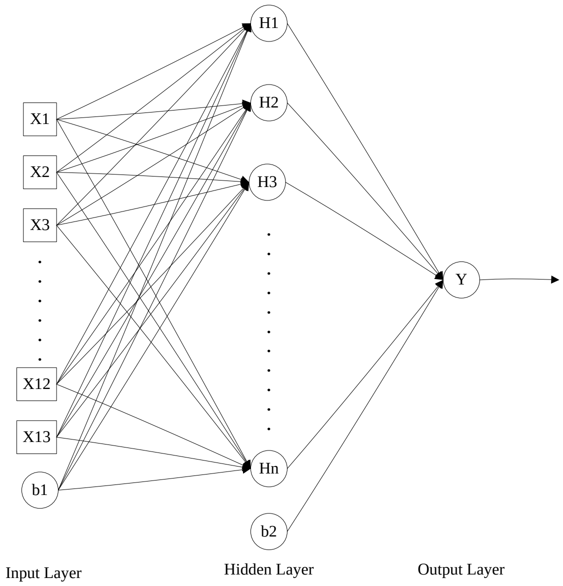

Figure 2. Neural Network Architecture

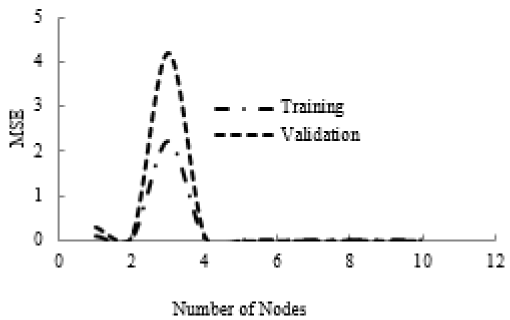

Figure 3. MSE for training and validation based on each hidden nodes

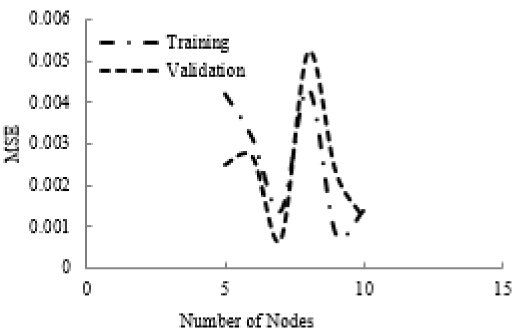

Figure 4. MSE for training and validation based on 5 to 10 hidden nodes

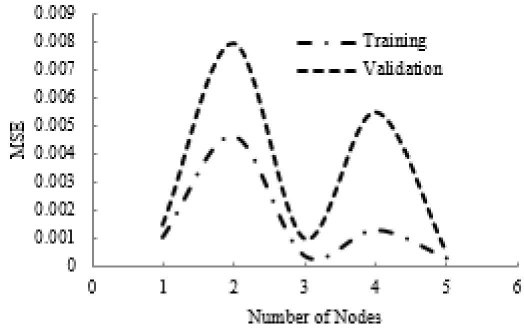

Figure 5. MSE value on 9 nodes based on repetition

TABLE 1
STATISTICALLY DATABASE ANALYSIS

Input	Number of Data	Minimum Value	Maximum Value	Mean	Deviation
C	2091	0. 14	0. 3	0. 212	0. 07640
Mn	2091	0. 662	1. 1	0. 97	0. 17739
Si	2091	0. 1	0. 175	0. 127	0. 03394
P	1597	0. 01	0. 035	0. 0240	0. 0124
S	1597	0. 032	0. 035	0. 033	0. 001437
Cr	2091	0. 0207	0. 4	0. 199	0. 1747
Mo	1597	0. 014	0. 15	0. 0904	0. 0670
Al	494	0	0. 047	0. 011	0. 0199
Ni	2091	0. 005	0. 4	0. 194	0. 179
Fe	2091	97. 52	99. 07	98. 17	0. 623
pH	2091	4. 87	7. 72	6. 21	1. 01
Flow rates(m/s)	494	0	1. 87	1. 03	0. 501
Temperature (^{o}C)	2091	25	90	42. 98	17. 70

TABLE 2A.
WEIGHTS AMONG INPUT AND FIRST TO 5TH HIDDEN NODES

Nodes	1	2	3	4	5
1	0.625465	-0.16853	0.230105	0.886662	0.497488
2	0.524094	0.754449	-0.00756	-0.66263	-0.30991
3	-0.23704	0.088346	-1.82527	-0.48477	-1.12168
4	-0.54216	0.079339	0.154572	0.51114	1.206474
5	0.317677	-0.52161	-0.94587	-0.44696	0.079656
6	0.297421	0.705587	0.078107	0.098603	1.628451
7	0.153082	0.093426	-0.29467	0.514411	0.93564
8	0.31934	-0.474	0.209221	0.531739	0.37707
9	-0.49859	0.431903	0.520904	0.74771	0.52619
10	-1.05174	-0.73037	0.504742	0.541078	-0.81489
11	-2.00411	0.339399	-0.82303	-0.68705	0.98655
12	0.402956	-0.82712	-3.18599	0.775512	-6.84827
13	-0.16031	-2.99736	-0.88296	0.137012	0.230036

TABLE 3.
WEIGHTS BETWEEN INPUT BIAS NODES AND EACH NODES IN HIDDEN LAYER

	Bias Nodes
1	-1.16369
2	0.82915
3	1.014161
4	0.052847
5	-1.02029
6	-0.75653
7	-3.18749
8	-1.24574
9	1.783606

TABLE 4.
WEIGHTS BETWEEN EACH NODES IN HIDDEN LAYER AND OUPUT NODE

	Bias Nodes
1	-1.37615
2	0.863291
3	-2.79165
4	1.55275
5	-0.40391
6	-1.9797
7	-0.31102
8	-0.64419
9	-2.02984

TABLE 5.
WEIGHT BETWEEN HIDDEN BIAS NODE AND OUTPUT NODE

	1
1	0.48496