Porosity Log Prediction Using Artificial Neural Network

Oki Dwi Saputro¹, Zulfikar Lazuardi Maulana¹ and Fourier Dzar Eljabbar Latief⁰

¹ Departement of Physics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Jalan Ganesa 10, Bandung 40132, Indonesia

Abstract. Well logging is important in oil and gas exploration. Many physical parameters of reservoir is derived from well logging measurement. Geophysicists often use well logging to obtain reservoir properties such as porosity, water saturation and permeability. Most of the time, the measurement of the reservoir properties are considered expensive. One of method to substitute the measurement is by conducting a prediction using artificial neural network. In this paper, artificial neural network is performed to predict porosity log data from other log data. Three well from ‘yy’ field are used to conduct the prediction experiment. The log data are sonic, gamma ray, and porosity log. One of three well is used as training data for the artificial neural network which employ the Levenberg-Marquardt Backpropagation algorithm. Through several trials, we devise that the most optimal input training is sonic log data and gamma ray log data with 10 hidden layer. The prediction result in well 1 has correlation of 0.92 and mean squared error of 5.67 × 10⁻⁴. Trained network apply to other well data. The result show that correlation in well 2 and well 3 is 0.872 and 0.9077 respectively. Mean squared error in well 2 and well 3 is 11 × 10⁻⁴ and 9.539 × 10⁻⁴. From the result we can conclude that sonic log and gamma ray log could be good combination for predicting porosity with neural network.

1. Introduction
Porosity is the fraction of pore space volume to the total volume of a rock. It is often related to the measure of fluid storage capacity inside rock. Total or absolute porosity includes all the pore space, while effective porosity describes only the interconnected pore space. Porosity could be an indicator of hydrocarbon presence. The porosity can be measured by mercury displacement, a bulk volume meter, mercury pump, caliper measurement, well logging and etc [1]. Well logging measures physical properties of reservoir directly in the borehole (in-situ). By well logging, we measures porosity by many method. The porosity measurement could be done by neutron porosity, formation density, nuclear magnetic resonance (NMR) and sonic [1] and none of the logs measures porosity directly. Density, NMR, and neutron logs are based on nuclear measurement. Sonic log uses acoustic measurement. Well logging or wireline measurement is one of important procedure in hydrocarbon exploration and most of the time it is very expensive.

Artificial neural network is one of the widely used method in science and engineering. It is a powerful tool that able to perform classification, feature extraction, diagnosis, function approximation and optimization. Nowadays, artificial neural network (ANN) has been widely used by scientist and engineer to predict data from other information data. In geophysics, ANN has been increasingly
applied to predict reservoir properties using well log data. ANN is used to predict permeability, normalized oil content, and any other reservoir properties [2,3,4]. ANN can also predict formation permeability even in highly heterogeneous reservoir condition [2].

The purpose of this study is to apply artificial neural network in predicting the porosity. The result of prediction is then compared to a real data. This study is also aimed to analyze the accuracy of the applied method when it involves only one data and when it uses combined data.

2. Background

A human brain is a very complex and non-linear system. It is capable of solving many kinds of problem such as pattern recognition, perception, faster than a computer. It consists of neurons that become stronger when they learn something new. Figure 1a shows biological neuron with its component. The connection between components becomes stronger when new electric signal comes through the neurons along dendrite and passes through a synapse. It is an important process of learning. Artificial neural network is inspired from the brain that made up from many neuron. It has a natural tendency for storing experience and the learning process thus helps in acquiring new knowledge.

![Figure 1a: Schematic of biological neuron](image1a)

![Figure 1b: Model of neural network](image1b)

**FIGURE 1.** (a) Schematic of biological neuron, (b) Model of neural network. [4]

Artificial neural network model has similar process to the biological neuron. As shown in Figure 1b, artificial neural network consist of various component. It has many synapse characterized by its own weights. The input connected to nucleus by this weights. With linear combination and activation function, the neuron learns something new. The elements of an artificial neural network are described as follows: a set of synapses $j$ connected to neuron $k$ and multiplied by synaptic weight $w_{kj}$, an adder for summing the input signals $x_j$ that weighted by respective synapses, and an activation function for limiting the amplitude of the output of a neuron [4]. The neuronal model also includes an externally applied bias, $b_k$ that increasing or lowering the net input of the activation function. The symbol $\phi(.)$ is the activation function, and $y_k$ is the output signal. The most common used of activation function is sigmoid function [4]. Sigmoid function has the range of $0 \leq \phi(v) \leq 1$ and it is differentiable and non-decreasing. An example of the sigmoid function is in equation (1), where $a$ is the slope parameter of the sigmoid function.

$$\phi(v) = \frac{1}{1 + \exp(-av)} \quad (1)$$
In this study, we used back-propagation-type artificial neural network. It has an input, an output, and one hidden layer. When an input is given to the machine, the forward propagation step begins and the backward-propagating step performs the required learning. The bias units of input and hidden layer improves the convergence. After we obtain the output from the process, error values are calculated and the weights are adjusted by backward propagation. The backward propagation starts at the output layer and moves backward until reaches the input layer[5]

In the neural supervised learning network, the data is split into training set, validation set, and test set. The training set consists of a set of examples used to improve learning of the network. It increases the accuracy of the weights. The validation set is a set of examples used to adjust the network parameters. The test set is used to examine the performance of a trained network. Mean square error used to measure the data’s errors as defined in Equation (2).

\[ MSE = \frac{1}{n} \sum_{i=1}^{n} (O_i - T_i)^2 \]  

where \( O_i \) is the real output for the training data \( i \) and \( T_i \) is the output of network from training data \( i \), and \( n \) is the number of data in the training data set. Lower MSE value means lower error and zero means no error. We used Levenberg-Marquardt (LM) training function in this study. LM is a network training function that updates weight and bias values according to LM optimization. It is very fast but required high amount of memory [3]. We use MATLAB to perform the neural network.

3. Case study
In this study, we use three well. As shown in Table 1, all well has gamma ray log, density log, and porosity log. We compare the porosity obtained from neural network to the one from the real data. Well no. 1 is chosen as an input training network. The resulted network is applied to well no. 2 and no. 3. If we compare all well data, the maximum value of porosity does not has any direct linear relationship to the gamma ray or sonic log.

| Well No. | Gamma Ray Log (api) | Sonic Log (μS/ft) | Porosity Log (%) |
|----------|---------------------|-------------------|-----------------|
| 1        | 34.20-180.62        | 35.8-113.7        | 0.1-37.12       |
| 2        | 31.089-204.7270     | 63.398-129.429    | 0.1-27.1998     |
| 3        | 37.779-194.935      | 65.1680-106.048   | 0.1-25.793      |

Statistically, relationship between sonic and porosity as shown in Figure 2a do not has good linear relation. R-square between porosity and sonic are 0.174. Figure 2b show that gamma ray and porosity are linearly good. The R-square are 0.7695. Gamma ray logs measure the natural radioactivity in formations and can be used for identifying lithology. As shale content increases, the gamma ray log response increases because of the concentration of radioactive material in shale. Gamma ray logs can be used for determining shale volumes. These is essential in calculating water saturations [6]. Water saturations has good relation with porosity. Sonic log measure interval transit time of a sound wave traveling through the formation along the axis of borehole. Sonic log has a dependency to the lithology and porosity [6]. Thus we can deduce that the gamma ray and sonic logs are suitable for our input training based on geological view.

4. Result and discussion
The ANN is built using back-propagation network trained by LM algorithm using MATLAB toolbox nftool. The sample is divided into three part: 70% for training, 15% for validation and 15% for testing.
Trial and error in determining the number of hidden layer are used to obtain the best and the fastest result. We first perform a test with gamma ray and sonic logs as training input and the result of our several tests are listed in Table 2. We decide to use 10 hidden because if we use more than 10, the MSE error or the R-square does not improve very much.

![Figure 2](image)

**FIGURE 2.** (a) Relation between sonic and porosity (b) Relation between gamma ray and porosity

We compared three network with different input: the first network used gamma ray logs only as data training, the second network used sonic logs only, and the third network used gamma ray and sonic logs. In Table 3 we can observe that the combination of training input gamma ray and sonic produced better result compared to the one using sonic. Thus gamma ray and sonic is considered the best to be used as training input.

**TABLE 2.** Relation between MSE, R-square and number of hidden layer

| Number of hidden layer | MSE (×10^-4) | R-square |
|------------------------|--------------|----------|
| 5                      | 7,13         | 0,922    |
| 10                     | 5,67         | 0,937316 |
| 20                     | 6,05         | 0,93352  |
| 30                     | 6,18         | 0,93373  |
| 40                     | 5,78         | 0,937108 |
| 50                     | 5,93         | 0,9355   |
| 60                     | 7,80         | 0,913852 |
| 70                     | 5,21         | 0,9433   |
| 80                     | 5,70         | 0,937    |

**TABLE 3.** Comparison MSE and R-square between different input training

| Input Training    | MSE (x10^-4) | R-square |
|-------------------|--------------|----------|
| Gamma Ray         | 9,58         | 0,894    |
| Sonic             | 33,5         | 0,5397   |
| Gamma Ray and Sonic | 5,67       | 0,937316 |

The network is then built with gamma ray log and sonic log as input using 10 hidden layer. Subsequently, we applied the network to other wells. The result as shown in Figure 3-5 tell that well...
number 1 give the best result as the correlation and R-square are the highest among all well. MSE of well number 1 also the smallest. The error in some area in well number 2 and 3 are high as seen in Figure 4 and 5. ANN output in well number 2 and 3 give underpredicted data in more area than well number 1. Well number 1 are chosen as the input training network. Thus, the network that have been built is fit to well number 1.

Overall, porosity from ANN are well matched with real data. The correlation from each ANN output also high. The correlation between output ANN and field data in well number 1 and 3 are more than 0.9. In well number 2 the correlation value are 0.87. R-square from well number 2 and 3 are 0.7572 and 0.8239. Moreover, the mean square error from output ANN also near zero. It means the ANN output is well matched with field data.

**FIGURE 3.** Neural network result comparison with real data in well no.1

**FIGURE 4.** Neural network result comparison with real data in well no.2
5. Conclusion

Comparison between the number of hidden layer, MSE and R-square, show that the result from the ANN cannot improve if its reach maximum performance. Increasing the number of hidden layer cannot increase the quality of the fit. We also compared different input training and the result show gamma ray and sonic give maximum performance. The result from ANN are considered good since the correlation, MSE and R-square from output ANN show that ANN with input gamma ray and sonic logs data works very well when predicting porosity.

Using this methodology, petrophysicist or geophysicist will be able to characterize reservoir with limited number of data. As future work, more combination of well log data could be used to predict porosity or another reservoir parameter and also could possibly improve the result.

Acknowledgments

This research is partly supported by the research program “Desentralisasi DIKTI-ITB” with contract number 310n/I1.C01/PL/2015.

6. References

[1] Norman Hyne 2014 Dictionary of Petroleum Exploration, Drilling & Production (Tulsa: Penwell Corporation) pp 394-395.
[2] M. Saemi, M. Ahmadi and A. Y. Varjani 2007 J. Petroleum Sci. and Eng. 59 pp 97-105
[3] A. Kadkodaie-Ikhchi, M. R. Rezaee and H. Rahimpour-Bonab 2009 J. Petroleum Sci. and Eng. 65, pp 23-32
[4] Bhatt A. 2002 Reservoir Properties from Well Logs using neural Networks (Norway: Norwegian University of Science and Technology)
[5] M. Bean and C. Jutten 2000 Geophysics 65 pp 1032-47.
[6] G. Asquith and D. Krygowski 2004 Basic Well Log Analysis (Tulsa: The American Association of Petroleum Geologists) pp 37-56

FIGURE 5. Neural network result comparison with real data in well no.3