PORE PRESSURE PREDICTION USING ARTIFICIAL NEURAL NETWORK BASED ON LOGGING DATA

Raka S.W. (1,a)*, Meredita Susanty(2,b) dan Hapsoro B.W. (1,c)

(1) Program Studi Teknik Perminyakan, Universitas Pertamina, Jakarta, Indonesia,
(2) Program Studi Ilmu Komputer, Universitas Pertamina, Jakarta, Indonesia,
Email: (a*) raka.sw@univesitaspertamina.ac.id, (b) meredita.susanty@univesitaspertamina.ac.id, (c) hapsorobagusw@gmail.com

Abstract. Pore pressure is a critical parameter in designing drilling operations. Inaccurate pore pressure data can cause problems, even incidents in drilling operations. Pore pressure data can be obtained from direct measurement methods or estimated using indirect measurement methods such as empirical models. In the oil and gas industry, most of the time, direct measurement is only taken in certain depth due to relatively high costs. Hence, empirical models are commonly used to fill in the gap. However, most of the empirical models highly depend on specific basins or types of formation. Furthermore, to predict pore pressure using empirical models accurately requires a good understanding in determining Normal Compaction Trendline. This proposed approach aims to find a more straightforward yet accurate method to predict pore pressure. Using Artificial Neural Network Model as an alternative method for pore pressure prediction based on logging data such as gamma-ray, density, and sonic log, the result shows a promising accuracy.

Keywords: formation pore pressure, artificial neural network

INTRODUCTION

Pore pressure is a critical parameter in designing drilling operations. Pore pressure will affect the design of the drilling mud, casing, cement, and tubing [1]. Errors in determining the pore pressure can cause problems in drilling operations such as lost circulation, kick, and blow out. These problems can result in increased drilling costs and even fatality.

Pore pressure can be defined as scalar hydraulic pressure acting on interconnected rock pores [2]. Pore pressure can be categorized into normal pore pressure, abnormal pore pressure (overpressure), and subnormal pore pressure (underpressure) [3].

The pore pressure, which has a pressure equal to the hydrostatic pressure of a fluid column with the height from the surface to a certain depth in the subsurface, is called the normal pore pressure. Therefore, Normal pore pressure can be calculated using the density of formation fluid and the depth of the observed point. Pore pressures whose value is higher than normal are called abnormal pore pressures while pore pressures whose values are below normal pressure are called subnormal formation pore pressures. Abnormal pore pressure is found almost in all sedimentary basins throughout the world and caused by the compacting effect, the effect of diagenesis, differences in density, and fluid migration [3].

Pore pressure can be determined using direct or indirect measurements. Direct measurement of pore pressure can be done using RFT (Repeat Formation Tester), DST (Drill Stem Test), or RCI (Reservoir Characterization Instrument) [4]. However, these methods are limited to a certain depth due to the issue of costs incurred. Besides, RFT and DST are only used in formations with high permeability whereas abnormal
pressures usually occur in formations that have low permeability such as shale [5].

The indirect measurement method can be done using empirical models. Some empirical models that can be used to calculate formation pore pressures using well logging data include Hotmann and Johnson [6], Eaton [7], and Bowers [8].

The Hotmann and Johnson model cannot be used in areas that have abnormal pressure (overpressure) caused by compaction as a result of overburden stress [6]. The Eaton model is the most widely used qualitative method [9]. The Eaton model is limited to be used in young sedimentary basins, where the cause of overpressure pressure is undercompaction [10]. The Bower method will give overestimation result when used in shallow formations that are unconsolidated [11].

As previously explained, most empirical models have their limitations. Empirical models are based on data in one basin so that it will be different from one basin to another. Justification and modification are needed to be able to use in another basin. An analysis of NCT (Normal Compaction Trendline) is also necessary to use empirical models. The accuracy of the method for determining the pore pressures based on rock porosity will depend on the determination of the Normal Compaction Trendline based on [12]. An understanding of resistivity compaction trends, the effect of lithology, formation fluid salinity, and temperature on the resistivity response to effective stress are also needed if the resistivity data is used [13]. Based on those limitations, more straightforward yet accurate method to predict pore pressure is necessary.

Artificial Neural Network (ANN) is one method that can help find the relationship between input parameters and output parameters without producing correlations. Artificial Neural Networks can be defined as an information-processing system that has similar characteristics to the biological nervous system [13]. The artificial neural network consists of connected input and output units that have associated weight in each connection [14]. The network is trained and adjusted until the output, and the target is matched. The schematic of the ANN network can be seen in Figure 1 [15].

With the ability to tolerate noisy data and classify new patterns, Artificial Neural Networks can be used as prediction, classification, data association, and data conceptualization tool [16]. Artificial Neural Network has been widely used in the oil gas industry, especially to solve complicated problems where conventional modeling is not a practical option [17].

This study aims to create an Artificial Neural Network Model as an alternative method to predict pore pressure. Data logging from several wells were used to develop an Artificial Neural Network model.

**Figure 1. Schematic of Artificial Neural Network [15]**

**METHODOLOGY**

The methodology of this research can be seen in figure Figure 2. Alyuda NeuroIntelligence software was used in this research.
Collecting and Analyzing Data

A sufficient amount of data is needed to build an artificial neural network. Logging data, daily drilling reports, lithology, and RFT data were collected from several wells. Logging data such as gamma-ray, density, and sonic log were collected and became input data. The pore pressure data were calculated using the Eaton equation and adjusted based on drilling event, mud density, and RFT. Pore pressure data were used as the output data. Then the data were analyzed where the outlier's data, typo, and missing values were eliminated from the dataset. Analyzed data were divided into 3 sets; training set, validation set, and test set. The training set is the dataset that used for neural network training, the validation set is the data that used to tune network parameters other than weights, and the test set is the data that will be used to test the performance of the neural network model. Simple Random Sampling (SRS) method was used to split the dataset. Using SRS method, the data were selected randomly with a uniform distribution. This method has the advantage of giving a low bias to the model's performance.

Pre-processing Data

Before the data were inputted to Artificial Neural Network, it needs to be normalized or scaled because ANN has a limited range of operating values and due to differences in the range of values in each parameter. The range value of each input parameter was scaled into the range of -1 to 1 using Equation 1 and Equation 2.

\[ x_p = SR_{\text{min}} + (x - x_{\text{min}}) \times SF \]  
\[ SF = \frac{(SR_{\text{max}} - SR_{\text{min}})}{(x_{\text{max}} - x_{\text{min}})} \]

Where \( x_p \) is preprocessed value, \( SR_{\text{min}} \) is lower scaling range limit, \( x \) is the actual value of the numeric column, \( x_{\text{min}} \) is the minimum actual value of the column, \( SF \) is scaling factor, \( SR_{\text{max}} \) is upper scaling range limit, \( x_{\text{max}} \) is the maximum actual value of the column, \( x_{\text{min}} \) is the minimum actual value of the column.

Developing ANN Model

The architecture of the ANN model used in this research consist of 3 layers; input layer, hidden layer, and output layer. The number of neurons in the hidden layer was selected based on iteration. The optimal number of neurons that give the highest correlation and the lowest error was chosen as the architecture of the ANN model. Training can be defined as the process of tuning the weights and biases from the input data. Each iteration of the training process consists calculating the predicted output, (feedforward) and tuning the weights and biases, known as backpropagation. The model was used the sigmoid function as the activation function.
function in the forward-propagation process. This function has a sigmoid curve and is calculated using Equation 3 and Equation 4, where \( W \) is the weight, \( x \) is the input layer, \( b \) is the bias, \( z \) is the input for the sigmoid function.

\[
z = Wx + b \tag{3}
\]

\[
\sigma(z) = \frac{1}{1 + e^{-z}} \tag{4}
\]

Back Propagation algorithm was used since this algorithm is the general-purpose training algorithm of choice. Several advantages of Backpropagation are simple and easy to program, no parameters to tune apart from the numbers of input, and this method is a standard method that generally works well\[^{21}\].

The training is conducted to find the best set of weights and biases that minimizes the loss function. Sum-of-squares error was used as the loss function. Sum-of-squares error can be calculated using Equation 5, where \( n \) is the amount of data, \( y_i \) is the actual value and \( \hat{y}_i \) is the predicted output.

\[
\text{Sum of squares error} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3}
\]

Testing ANN Model

The ANN model was tested to predict a new dataset that never been introduced to the model. The performance of the ANN model was checked by comparing the prediction pore pressure data with the actual pore pressure data. The coefficient of determination (\( R^2 \)) was used to see the correlation between the predicted result and the real data. The coefficient of determination (\( R^2 \)) was calculated using Equation 6 where \( n \) is the amount of data; \( m_i \) is predicted value at step \( i \), \( p_i \) is the observed value at step \( i \), \( \bar{m} \) is the mean of observed value and \( \bar{p} \) is the standard deviation of the observed value.

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (m_i - \bar{m}) (p_i - \bar{p})^2}{\sqrt{\sum_{i=1}^{n} (m_i - \bar{m})^2 \sum_{i=1}^{n} (p_i - \bar{p})^2}} \tag{4}
\]

RESULTS AND DISCUSSION

3 wells were selected in an onshore field. Those 3 wells were exploration wells and drilled directionally. The general stratigraphy of this field can be seen on Table 1. 1434 records were extracted. The data from logging tools such as True Vertical Depth (TVD), Gamma Ray (GR), Density, and Acoustic Log were collected and became the input data. The pore pressure data were selected as the input data.

Based on data analysis, 13 records were deleted (outliers), 995 records were selected for the training set (70%), 213 were selected for the validation set (15%), and 213 records were selected for the test set (15%). After the data were selected and classified, then each of parameter was scaled into the range of -1 to 1. The statistics of the data can be seen in Table 2. Statistic of The ANN Data.

The architecture of the ANN model consists of an input layer, one hidden layer, and output layer. One hidden layer is selected to make the network fast and efficient. The optimum number of neurons in the hidden layer was selected based on iteration. The 5 best architecture can be seen in Table 3. Ten neurons in the hidden layer were chosen because it showed the best model with the highest value of correlation (\( R = 0.974989 \)) above all. The
architecture of the model can be seen in Figure 3.

**Table 1. Stratigraphy**

| Formation | Description |
|-----------|-------------|
| A         | a sequence of claystones to siltstone, with interbeds to fine grained sandstones and traces of off white to light brown limestone stringer |
| B         | consists of calcareous claystones interbedded with fine to medium grained sandstones and calcareous siltstones |
| C         | comprises limestones with minor interbeds of claystones |
| D         | comprises calcareous claystones with interbeds of limestones and sandstones |
| E         | consists of claystones/shales interbedded with limestones |
| F         | consists of claystones/shales interbedded with packstones and medium grained |

**Table 2. Statistic of The ANN Data**

| Parameter               | Min.   | Max.    | Mean   | Standard Deviation |
|-------------------------|--------|---------|--------|--------------------|
| TVD (ft)                | 943    | 13250   | 6368.72| 2686.69            |
| Gamma Ray (API)         | 4.48   | 89.075  | 42.47  | 16.29              |
| Density (gr/cc)         | 1.72   | 2.738   | 2.38   | 0.17               |
| Acoustic (µs/ft)        | 48.75  | 175.17  | 103.88 | 23.85              |
| Pore Pressure (ppg)     | 8.055  | 14.406  | 9.92   | 1.375              |

As can be seen in Figure 3, the ANN model consists of 4 input layer, 1 hidden layer with 10 neurons and 1 output layer. Then it was trained using Back Propagation Algorithm. Back Propagation Coefficient was 1.75 and the learning rate was 0.1. Number of iterations was done until the lowest number of errors was achieved. After 487 iterations, the absolute error for training and validation was 0.19555 and 0.19893 respectively (Figure 4).

After ANN Model had been developed, it was tested to predict pore pressure using new data from the test set data. The result of this testing (ANN model output) was compared with the actual pore pressure data. This comparison can be seen in Figure 5. In Figure 5, the ANN output shows a good correlation with the actual data. Figure 6 shows the cross plot comparing actual value and ANN output. It shows the coefficient of determination (R²) of the actual data compared to ANN output is 0.927.

**Table 3. Hyperparameter Tuning**

| Architecture | Correlation |
|--------------|-------------|
| 4-4-1        | 0.921282    |
| 4-6-1        | 0.931653    |
| 4-7-1        | 0.936755    |
| 4-8-1        | 0.964859    |
| 4-10-1       | 0.974989    |
CONCLUSION

An Artificial Neural Network model has been developed using True Vertical Depth (TVD), Gamma-Ray, Density and Sonic log as the input parameter and pore pressure as the output parameter.

The model consists of one input layer with four nodes, one hidden layer with 10 nodes and one output layer with one node. The model uses sigmoid function as the activation function and sum of squares as the lost function.

The test's result of this model shows a good correlation with $R^2$ between the actual data and the ANN output is 0.927. Based on the result shown, ANN is a very potential alternative method or tool to predict pore pressure in Field X.

Figure 4. Training Graph

Figure 5. Actual vs. ANN Output Scatter
REFERENCES

[1] N. J. Adams, Drilling Engineering A Complete Well Planning Approach, Tulsa, Oklahoma: Penn Well Books, 1985.

[2] M. D. Zoback, Reservoir Geomechanics, New York: Cambridge University Press, 2007.

[3] A.T. Bourgoyne Jr, K.K. Millheim, M.E. Chenevert and F.S. Young Jr., Applied Drilling Engineering, Society of Petroleum Engineers, 1991.

[4] F. Jahn, M. Cook and M. Graham, Hydrocarbon Exploration And Production2nd Edition, Aberdeen: ELSEVIER, 2008.

[5] A. Kesarwani, A. Malhotra and V. Agarwal, "Pore Pressure: Causes, Methods of Determination and their Limitations," in 10th Biennial International Conference & Exposition, Kochi, 2013.

[6] C. Hottman and R. Johnson, "Estimation of Formation Pressures from Log Derived Shale Properties," in J. Pet. Tech., 1965.

[7] B. A. Eaton, "The Equation for Geopressured Prediction from Well Logs," in the 1975 SPE Annual Technical Conference and Exhibition, Dallas, Texas, 1975.

[8] G. L. Bowers, "Pore Pressure Estimation From Velocity Data: Accounting for Overpressure Mechanisms Besides Undercompaction," in IADC/SPE Drilling Conference, Dallas, 1995.

Figure 6. Actual vs. ANN Output Scatter Plot
NWEKE, I. FRANCIS, A. B. ORIJI and A. DOSUNMU, "Study of abnormally-high pore pressure prediction methods in regions with non-equilibrium compaction – Insight of Niger Delta fields," *The International Journal Of Engineering And Science (IJES)*, vol. 5, no. 2, pp. 01-10, 2016.

J. Lang, S. Li and J. Zhang, "Wellbore Stability Modeling and Real-Time Surveillance for Deepwater Drilling to Weak Bedding Planes and Depleted Reservoirs," in *SPE/IADC Drilling Conference and Exhibition*, Amsterdam, 2011.

J. J. Zhang, "Pore Pressure Prediction from Well Logs: Methods, modification, and new approaches," *Earth-Science Reviews*, vol. 108, pp. 50-63, 2011.

N. R. Ifeanyi, "Comparing methods of predicting pore pressure," *International Journal of Scientific & Engineering Research*, vol. 6, no. 7, pp. 1813-1822, 2015.

S. Mohaghegh, "Virtual Intelligence Applications in Petroleum Engineering: Part I-Artificial Neural Networks," *Society of Petroleum Engineers - Distinguished Author Series*, pp. 64-72, 2000.

J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, San Francisco: Elsevier, 2006.

H. Demuth and M. Beale, *Neural network toolbox for use with MATLAB: User's guide*, Massachusetts: MathWorks, Inc, 2004.

A. Abraham, "SECTION 3. AI SIGNAL PROCESSING TECHNIQUES," in *Handbook of Measuring System Design*, John Wiley & Sons, 2005, pp. 901-908.

H. H. Alkinani, A. T. T. Al-Hameedi, S. Dunn-Norman and R. E. Flori, "Applications of Artificial Neural Networks in the Petroleum Industry: A Review," in *SPE Middle East Oil and Gas Show and Conference*, Manama, 2019.

Z. Reitermanova, "Data Splitting," in *WDS'10 Proceedings of Contributed Papers*, 2010.

S. L. Lohr, *Sampling: Design and Analysis*, Duxbury Press, 1999.

J. Loy, "Towards Data Science," 14 May 2018. [Online]. Available: https://towardsdatascience.com/how-to-build-your-own-neural-network-from-scratch-in-python-68998a08e4f6. [Accessed 9 5 2020].

K. "guru99," [Online]. Available: https://www.guru99.com/backpropogation-neural-network.html. [Accessed 21 May 2020].