PAPER • **OPEN ACCESS**

Stuck pipe optimization using duellist algorithm

To cite this article: T R Biyanto *et al* 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **672** 012104

IOP Conf. Series: Earth and Environmental Science **672** (2021) 012104 doi:10.1088/1755-1315/672/1/012104

Stuck pipe optimization using duellist algorithm

T R Biyanto[1], H Cordova[1], Matradji[1], T O Anggrea[1], H Suryowicaksono[1] and S Irawan[2]

[1] Teknik Fisika, Fakultas Teknologi Industri dan Rekayasa Sistem, Institut Teknologi Sepuluh Nopember (ITS)
[2] School of Mining and Geosciences, Nazarbayev University, Kazakhstan

*E-mail: trb@ep.its.ac.id

Abstract. Stuck pipes are one of the most serious drilling problems, stuck pipes can cost the oil industry hundreds of millions of dollars per year. One way to avoid the risk of a stuck pipe is to predict the condition of a stuck pipe with the available drilling parameters. Throughout the years, a lot research has been dedicated to finding the causes that lead to stuck pipe events. But it still not reached the study in the calculation of the optimization of drilling operation costs. In this final project, Artificial Neural Network (ANN) is used for prediction stuck pipe and optimized using the Duellist Algorithm (DA). As well as Increasing fewer data that will get easier to make it in to the model. In this model, it use 1 The input layer contains 12 input nodes, 14 hidden layers trained with 1 to 30 hidden nodes, and 1 output layer in 14 hidden layers with an RMSE value of 0. At the end of the optimization, the lowest cost is USD 17300 / hour at RPM 195, 69, and mudflow 722.28 GPM. As well as constraint conditions are maintained and not stuck.

1. Introduction

With the latest technological improvements in drilling operations, it has resulted in more varied wells such as deep water, HPHT, and borehole's which were initially less balanced and become more feasible and achievable in terms of economic and technical aspects. Although drilling operations with a high level of complexity can now be solved with new materials, methods, techniques, and equipment, the problem of stuck pipes or clogged pipes has always been a problem in drilling operations around the world. This clogged pipe event has an effect on non-productive time (NPT) where it can calculate additional drilling costs after considering the time and costs taken to clear the clogged pipe, the pipe caught in fish, or make other lines to get around the jammed pipe [1].

Throughout the years, much research has been dedicated to finding the causes that lead to stuck pipe events. In general, a stuck pipe can be caused by sticking a differential and a mechanical brake or a combination of both. By knowing the cause of clogged pipes, many drillers try to prevent pipe clogging using visual monitoring of changes in drilling parameters such as torque, standpipe pressure, and hook load. Currently, the work of detecting variations in drilling parameters can be done by software that can be used during drilling operations. The software can detect small changes in drilling parameters as well as evaluate long-term trends in drilling parameters which will then provide input to the software to assess the risk of pipe clogging. Before monitoring with this software was available, several researchers had tried to predict pipe clogging using statistical analysis such as regression and multivariate analyses. The drilling data entered into the analysis model included mud properties (PV, YP, gel strength, etc.), hole size, and BHA specifications. This initial statistical analysis has become the basis of the prediction of modern stuck pipes using machine learning algorithms [2]. With the development of modern computing, computing power has increased significantly over the last few decades. Historically, machine learning

IOP Conf. Series: Earth and Environmental Science **672** (2021) 012104 doi:10.1088/1755-1315/672/1/012104

was only a theoretical idea on paper because the computational power was not sufficient to turn algorithms into machine learning [3]. However, it can be implemented today using growing computing power. One of them is by using optimization with a duellist algorithm. The duellist algorithm is an algorithm that mimics from a fighter's behaviour on how to improve his ability to fight with his opponent. First, a population of registered fighters. Each fighter has properties that are coded into a binary array of ones and zeros. After that, each fighter was evaluated to determine their fighting ability, and then a duel schedule was arranged for each fighter which contained a set of duel participants, in this duel, each fighter would fight one on one with other fighters. In one-on-one combat, fighting gladiators was used to avoid optimum localization [4].

In line with the vision of industrial revolution 4.0, machine learning has been identified as one of the key areas that can drive the transition of traditional industries to highly automated and computerized industries. Today, machine learning has been applied in areas such as transportation, health care, marketing and sales, government services, and finance. However, the oil and gas industry has only recently started to adopt digitalization more recently, especially machine learning which is meant to build a highly digital work environment [5]. There are still many opportunities for optimization that can be realized through the application of machine learning in the oil and gas industry. One opportunity is the use of machine learning to predict clogged pipe events and then take action to reduce the risk of them occurring [6].

2. Material and methods

The research in this final project is designed on several stages, which are depicted in the flow chart as in the following. These stages include the following Figure 1.

In this research, the data collection stuck pipe which will be used for this study was collected from the Analysed Data software. This software is a proprietary Halliburton Company that can demonstrate access to data on drilling operations PSCB. This software can take and show systematic data taken from the drilling operation data database. The database consists of all types of data contained in drilling operations, including data on daily reports of drilling operations, specifications BHA, hole states such as slope, azimuth, hole diameter, MD, TVD; casing OD and ID size, operating conditions such as flow rate, rotational speed and ROP [7].

2.1. Data Analysis

After extracting data stuck and non-stuck pipe from the data analysis software, the drilling parameters have been obtained, collected, and converted into Excel format. Although some basic screening was carried out to eliminate coarse data and unidentified data, the data were not thoroughly checked. Therefore, the next step is to check the quality and completeness of the work. The percentage of data completeness is measured using the formula below:

After checking the input data for ANN, it was found that some data values such as "calcium (ppm)" and "Cl- (ppm)" had relatively low data completeness. Completeness of data was measured respectively 38.89% and 61.11% for sticking data, 19.78%, and 26.92% for non-stock data, respectively. Hence, it was decided that these two data columns will be deleted in the ANN training input because the drilling parameters are no longer representative of such high missing data points. The final step in preparing the dataset is changing qualitative data columns such as sludge types and stuck / not jammed condition to binary format so that the dataset can be read by programming software machine learning. For stuck / non-stuck conditions, 1 assigned to all data that is temporarily stuck 0 assigned to all non-stuck data [8].

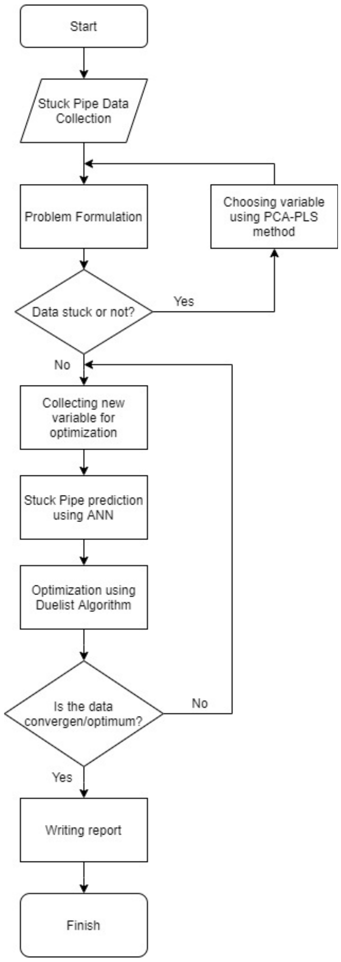

Figure 1 Research flowchart

2.2. Variable Selection Using PCA_PLS Method

In this research, PCA and PLS methods were used to obtain the most important variable that will be used for the next process, the stuck pipe data has been obtained, analysed the level of closeness of the relationship between the independent variable through the magnitude of the correlation coefficient [9]. In this research, PCA and PLS methods are used to obtain the most important variables that will be used for the next process, namely determining which variables have the most influence on stuck pipe and will be used in artificial neural network modelling. After that, the PCA - PLS method was carried out to obtain important variables. The PCA method is done by creating clusters of adjacent variables. Diversity tends to be collected in the first few main components and less and less in the last major components, so that the main components in the last sequence can be ignored without losing information. This diversity of data has a high percentage to represent other data.

2.3. Formulation of Optimization Problems

Problem formulation includes objective functions, limitations, and optimized variable determination. The objective function of this final project is to minimize the operational costs of the drilling process. The reason for choosing operational cost as an objective function is because every company wants maximum profit while maintaining pipe conditions so that a stuck pipe condition does not occur. Research that has been done on stuck pipes has yet to consider the economic side. Research that has been done on stuck pipes has yet to consider the economic side. While the limitation in this research is to avoid stuck pipe conditions that can be experienced in the drilling process. The optimized variables are RPM and mudflow rate.

2.4. Stuck Pipe Modeling Using Artificial Neural Network
In this research, the input variable is used is to use an artificial neural network. The model best of ANN is selected to be compared to determine the best machine learning model to predict stuck pipes. As a comparison, it is concluded that the ANN model can perform better in predicting congested pipes. The accuracy, sensitivity, and specificity of the best ANN models are consistently higher than the best SVM models. The best ANN models have a lower number of correctly predicted cases. In conclusion, as far as this study is concerned, the best machine learning model to apply in constructing a stuck pipe prediction [10].

2.5. Optimization Techniques Using Duelist Algorithm
The optimization done this time is optimization weights that affect the value of hidden nodes. This research uses 2 hidden nodes in the model artificial neural network. After training until obtained weight in the hidden layer (W) and weight of the output layer (V), then optimization is done using the Duellist Algorithm. The benchmark of optimization this time it is the resulting reduction in the RMSE price of this Duellist Algorithm optimization process.

2.6. Analysis of Data Optimization Results
The analysis will be carried out from the modelling results stuck pipe using an artificial neural network and optimized using the duellist algorithm this time is to compare the results of the data after optimization with the data before optimization.

3. Results and discussion
Stuck pipe data that have been obtained are analysed the degree of closeness of the relationship between the independent variables through the magnitude of the correlation coefficient. In this research, PCA and PLS methods were used. The stuck pipe data that had been obtained were analysed the level of closeness of the relationship between the independent variables through the magnitude of the correlation coefficient. In this study, the PCA and PLS methods were used to get the most important variable that will be used for the next process, namely determining which variable has the most influence on the stuck pipe and will be used in the artificial neural network modelling. However, before the PCA method, a zoning process was carried out between the MUD variable (as the x-axis) and the depth variable or ROP (as the y-axis). This zoning process aims to improve processing accuracy in getting the related variables. After that, the PCA - PLS method was carried out to obtain important variables. The PCA method is done by creating clusters of adjacent variables. The diversity tends to be collected in the first few main components and less and less in the last major components so that the main components in the last list can be ignored without losing information. This diversity of data has a high percentage to represent other data.

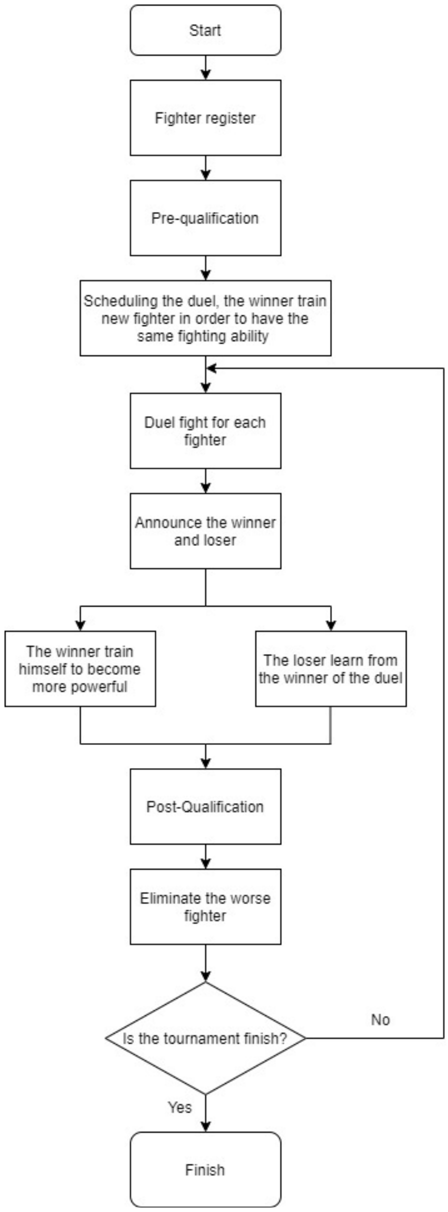

Figure 2 Optimization flowchart

IOP Conf. Series: Earth and Environmental Science **672** (2021) 012104 doi:10.1088/1755-1315/672/1/012104

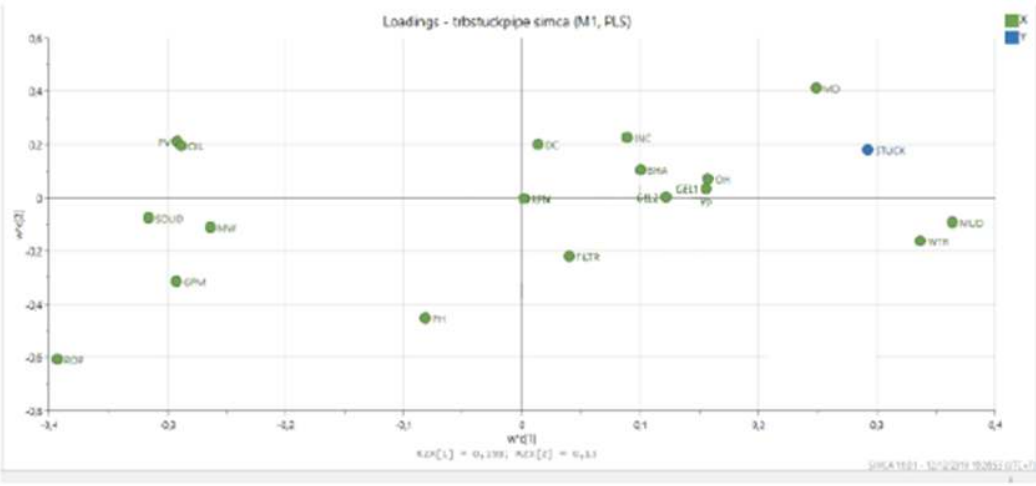

Figure 3 PCA Result

In the graph of the PCA results above, 2 clusters variables are not multicollinear and close to each other, namely PV-OIL and OH-GEL1-YP. The two clusters above show that these variables have high diversity so that they can represent variables others for further processing. From these data are taken each one variable that can be interpreted and can be measured. In this case, the variables taken are OIL and OH. So that the remaining variables are not processed.

The PLS method aims to determine the complexity of the relationship between variables. The variable that is used as output (y) is ROP. This is because the ROP log has the smallest component in the process compared to other logs. The y-axis in Figure 4 is Variable Importance for the Projection (VIP), which means the influence of a variable x on variable y and this is the result of the SIMCA-P software output. A VIP value of more than 1 indicates an important variable x, a value less than 0.5 indicates an unimportant variable x and a value between 0.5 - 1 is a variable in the grey zone, which means that the effect of the variable x on y depends on the data set used [33]. The minimum value limit of 1 in this final project is to determine the variables that influence the prediction of a stuck pipe condition in the drilling process.

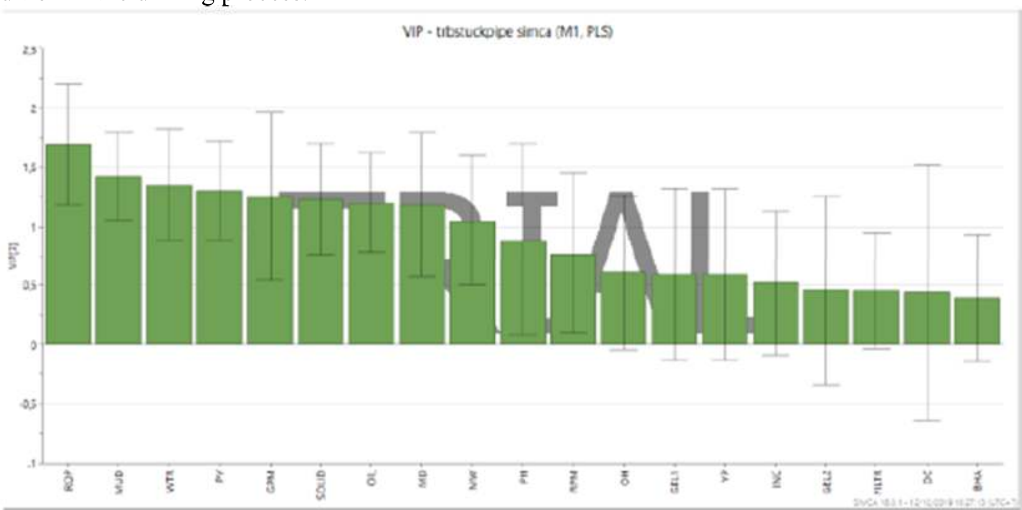

Figure 4 PLS Result

Obtained important variables from pre-processing data by the PLS method, namely ROP, MUD, WTR, PV, GPM, SOLID, OIL, MD, MW, PH, RPM, OH, GEL1, YP, INC, and DEPT, but after doing VIP analysis, it shows that the variables GEL2, FILTR, DC, and BHA, then it's cannot be included in the next process, to obtain the variables obtained the above variables which will be processed further.

In this research, a comparison between data that have not been done by the PCA-PLS method with data who have gone through the PCA-PLS method so that there is a difference in the number of inputs in the artificial neural network modelling. In the first model, there are 19 inputs whereas after removing redundant variables on PCA and selecting important variables in the PLS method using only variables that have a VIP value above 0.5 there are 12 inputs. The following is the result of neural network training by using 19 variables before using variables that have been through the PCA-PLS method.

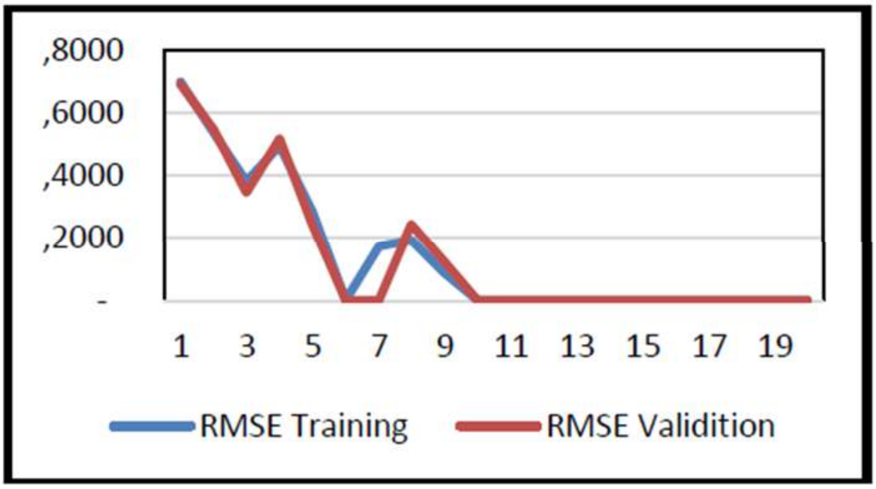

Figure 5 RMSE Result with 19 input data

Training results and validation of network training artificial neurons using 19 variables obtained the smallest RMSE value with the smallest hidden node with an RMSE value of 0 at hidden node 7. The blue line is the RMSE training value and the yellow line is the RMSE validation data.

The data used in predicting stuck pipe conditions this time the 12 variables were trained using an artificial neural network. The output of this training is operational status (1 = stuck, 0 = not stuck). In this model, this is done by using 1 input layer with 12 input nodes, 14 hidden layers trained from 1 to 30 hidden nodes, and 1 output layer in 14 hidden layers.

The data used in predicting stuck pipe conditions this time the 12 variables were trained using an artificial neural network. The output of this training is operational status (1 = stuck, 0 = not stuck). In this model, this is done by using 1 input layer with 12 input nodes, 14 hidden layers trained from 1 to 30 hidden nodes, and 1 output layer in 14 hidden layers.

Results of training on neural networks on assignment lately with is obtained the smallest RMSE with the smallest hidden node with the value of RMSE is 0 on the hidden node 14. The blue line is the result of RMSE and the yellow line is the RMSE validation data. The x-axis is the hidden layer and the y-axis is RMSE data as shown in Figure 6 below:

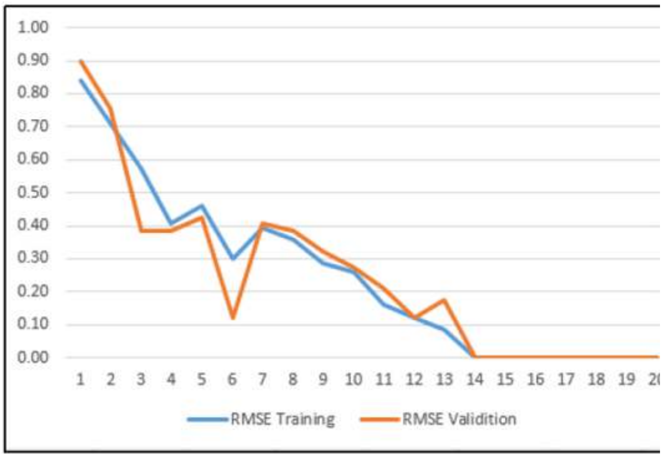

Figure 6 RMSE Result with 12 input data

The prediction results in the stuck pipe condition on this final project are as follows with a red graphic showing the results of the training with ANN and the blue graph is the actual data:

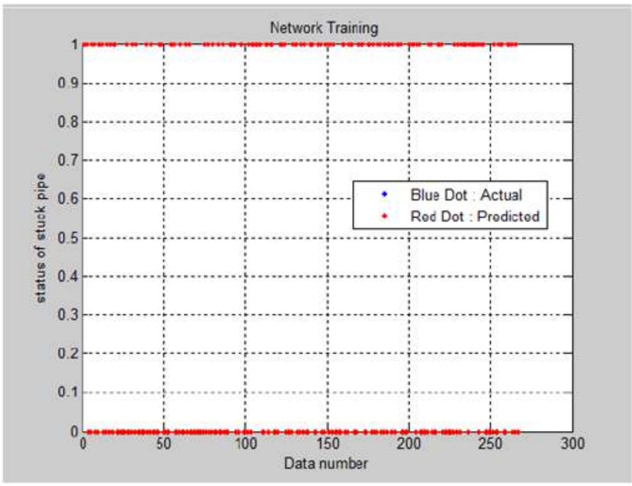

Figure 7 Prediction Result of the Stuck Pipe Condition

In the prediction results, the stuck pipe condition is above it was found that ANN predictions were the same as actual data where 1 indicates a stuck state, and 0 indicates a non-stuck state. In other words, the ANN model is ready for use in optimizing stuck pipe conditions.

The objective function of this optimization is to minimize the cost of drilling operations while maintaining a stuck pipe condition with the optimized variable being the vibration of the motor drilling, the mudflow rate used in the drilling process. The optimization technique used to solve this problem is by using a duellist algorithm.

The optimization process to get costs the lowest inverse must be the ability to compete in DA, so the cost function becomes 1 / cost to be maximized by DA. Iteration of competing abilities all the time can be seen in Figure 8 below.

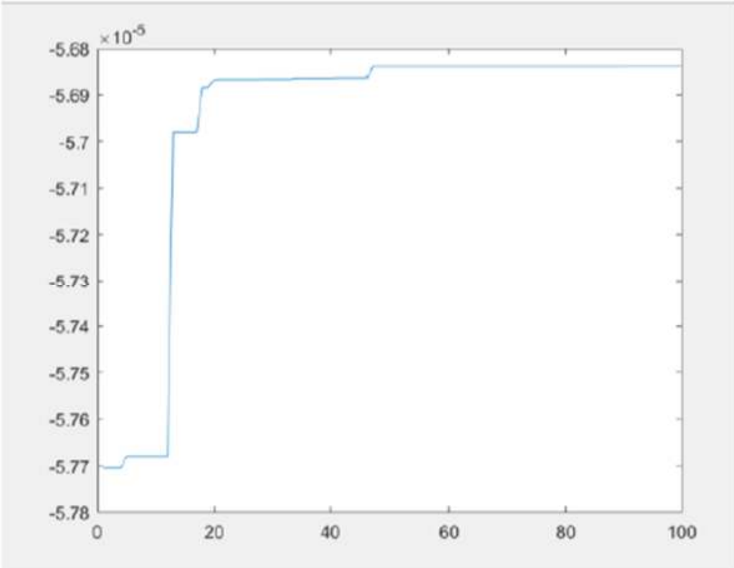

Figure 8 Duellist Algorithm Iteration Result

At the end of the optimization, the lowest cost is obtained of USD17300 / hour at 195,69 RPM and flow rate mud 722.28 GPM and the constraint conditions are maintained in the condition is not stuck. Comparison of operating conditions and costs, operations before and after optimization can be seen in the table below.

Table 1 Comparison Result

	Before Optimization	After Optimization
RPM	470	195.69
MUD	1150 GPM	722.28 GPM
The Lowest Cost	52599 /hour	17300 /hour

4. Conclusions and recommendations

There are several conclusions in this final project, that is:

- The stuck pipe has been modelled using Finite ANN Impulse Response (FIR) with two structures input. The 12 variable input structure is selected after going through PCA and PLS methods.
- ANN with less input structure is used as a model for due optimization requires less data collection. DA reaches the optimum value in the 48th iteration.
- The optimization results show the lowest cost at the end of the optimization of USD 17300 / hour at RPM 195.69 and mudflow rate of 722.28 GPM as well as conditions constraints are maintained in a condition that is not stuck.

Acknowledgments

The authors wishes to thank Institut Teknologi Sepuluh Nopember (ITS) Surabaya (Grant Number 1650/PKS/ITS/2020).

References

[1] Fox J and Andersen R 2005 Using the R Statistical Computing Environment to Teach Social
[2] Howard J and Glover S 1994 Tracking Stuck Pipe Probability While Drilling IADC/SPE Drilling

Conference
[3] Muqeem M A, Weeks A E and Al-Hajji A A 2012 Stuck Pipe Best Practices - A Challenging Approach to Reducing Stuck Pipe Costs SPE Saudi Arabia Section Technical Symposium and Exhibition
[4] Technical Symposium and Exhibition, Al-Khobar, Saudi Arabia, 2012. *Journal of Scientific & Industrial Research* **64**, 1:30-34.
[5] Gulati N 2012 Cost-Effectiveness in HVAC By Building Envelope Optimization Aus 14–7
[6] PCSB 2016, Wells Downtime Review Panel (WDRP) Terms of Reference Rev 1, PETRONAS, Kuala Lumpur.
[7] Jahanbakhshi R, Keshavarzi R, Shoorehdeli M A and Emamzadeh A 2012 Intelligent Prediction of Differential Pipe Sticking by Support Vector Machine Compared With Conventional Artificial Neural Networks: An Example of Iranian Offshore Oil Fields SPE Drilling & Completion **27** 586–95
[8] Annis M and Monaghan P 1962 Differential Pressure Sticking-Laboratory Studies of Friction Between Steel and Mud Filter Cake Journal of Petroleum Technology **14** 537–43. Venkataraman, P., 2002. Applied Optimization with Matlab Programming," Wiley-Interscience.
[9] Cakici M D 2003 Co-optimization of oil recovery and carbon dioxide storage *thesis*