Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DETECTION OF SOIL FORMATION CHANGES USING A FUZZY INFERENCE SYSTEM-BASED OUTLIER

MOCHAMAD LUJENG PRATIKTAMA[1], SANI MUHAMAD ISA[2], HERMAN YOSEPH SUTARTO[3]

[1] Master Degree Candidate. Computer Science Department, Bina Nusantara University, Jakarta, Indonesia

[2] Lecturer. Computer Science Department, Bina Nusantara University, Jakarta, Indonesia

[3] Lecturer. Institut Teknologi Harapan Bangsa, Bandung, Indonesia

E-mail: [1]mochamad.pratiktama@binus.ac.id, [2]sani.m.isa@binus.ac.id, [3]hytotok@yahoo.com

ABSTRACT

Oil and gas drilling is a risky business because actual rock and soil formations may differ from the plans. The application of "artificial intelligence" (AI) is potent to detect soil-rock formation changes. "Fuzzy Inference System" (FIS) is one of many AI methods that may be used for outlier data detection from raw data sets produced by "Mud Logging Unit" (MLU) equipment from drilling activities. This research will be carried out to detect outliers using FIS. Then, detected outliers in drilling data are used to see if these outliers indicate changes in drilling parameters. While other research mainly detects outliers directly from the data set, this research will get outliers from the calculated standard statistical value that is fed to the FIS. Then the FIS results will be combined to determine whether the formation changes are indicated or not. The data sets being used are "rotation per minute" (RPM), "rate of penetration" (ROP), and "weight on bit" (WOB). From these three datasets, four (4) statistical standard values were calculated, namely DIST, N-POINT-DIST, MEM-DEG-DIST (distance from data point to center point cluster), and MEAN-DIST for each of the datasets mentioned above. The FIS method is used to detect outlier data with the input of the four (4) statistical standard values for each of the datasets above. There was outlier data detected at 6672 points for RPM, 8239 points for ROP, and 6783 points for WOB, respectively. Outlier data results are then used for further analysis to conclude whether there was a change in soil formation at a certain drilling depth. The result was obtained for 70 data points, indicating a change in soil-rock formation. Verification data was collected using the formation log report, and 0.93% of the log report was verified.

Keywords: *Drilling Operation, Fuzzy Inference System, Machine Learning, Outlier Detection, Rock-soil Formation*

1. INTRODUCTION

Oil and gas wells are drilled deep into the earth to access hydrocarbon reservoirs. Various types of "drilling rigs" are used to drill these wells, depending on the application and condition of the location. The rigs use drilling "bits" to drill the hole deeper that is attached to the drilling string (the pipe that connects and transfers "Weight on Bed" from the rig to the end of the bit). This drilling string allows drilling mud (which functions as a lubricant and drilled debris carrier) to be pumped down to the bit and bring the drilled material out from the wellbore.

Drilling operation is critical to the company's business and safety operation. Computing performance and logic techniques lead to machine implementation for reducing risk factors also while optimizing cost in many industries. To find solutions by reducing the risk of well stability problems and reducing costs, it is necessary to have a method of detecting changes in soil formation in drilling based on data from MLU.

Every drilled soil-rock formation is different with its physical properties such as hardness, elasticity, viscosity, and chemical properties that affect the dynamics and behavior of the drilling respective to the formation.

Fertl, et al. in [1] indicate in their article that drilling rate is a function of weight on the bit, rotary speed, bit type and size, hydraulics, the bottom-hole cleaning properties of drilling fluid, and formation characteristics. From those indications, it can be summarized that the drilling parameters that affect

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

drilling performance are weight on the bit (WOB), rotary speed (RPM), bit type, and size, hydraulic characteristics, and bottom-hole cleaning properties of drilling fluid. The drilling function of the respective parameter rate of penetration (ROP) or penetration speed of drilling to the formation is a function of the formation, RPM, and WOB, while drilling fluid and other parameters are assumed to be constant. when an outlier is detected and there is a change in ROP while RPM and WOB remain constant. So it can be concluded that there is a formation change. From that standpoint, this is a significant advance in detecting outliers in drilling parameters that represent formation change. A warning notice can be generated when the formation change occurs as quickly as possible so that the engineer can respond immediately to adjust and match the operation. This will make it safer and optimized.

The anomaly/outlier data detection system with reinforced learning used in this study is the Fuzzy Inference System (FIS) method. According to Cateni et al. [2] research, the FIS system has advantages and disadvantages when compared to other methods because this system is known to be a combined system to reduce detection weaknesses in other systems.

The data sets being used are rotation per minute (RPM), rate of penetration (ROP), and weight on bit (WOB). From these three datasets, four (4) statistical standard values were calculated, namely DIST, N-POINT-DIST, MEM-DEG-DIST (distance from data point to center point cluster), and MEAN-DIST for each of the datasets mentioned above. The FIS method is used to detect outlier data with the above four (4) statistical standard values as input to generate outlier data. This outlier data is then used for further analysis to produce a conclusion about whether there was a change in soil-rock formation at a certain drilling depth. Changes in soil-rock formation are then validated by comparing them to the formation log report.

2. REFERENCE

2.1 Fuzzy Application for Complex System

Tamaraat et al. [3] conclude that the fuzzy control system has advantages in controlling applications in systems with high nonlinearity or conditions that are difficult to predict and is very suitable for use in complex systems that are difficult to model mathematically in their research on fuzzy systems applied to the control system for wind turbines. Based on this research, the fuzzy method has a high possibility of being applied in drilling systems.

2.2 Drilling Data Parameter

Xue, et al. [4] conducted research by modeling a drilling system. Drilling failure is frequently caused by vibration in the drill string and drill bit. As shown in Figure 1, stick/slip is a function parameter value for rotation speed drill string that is directly related to spindle rotation speed (RPM) and torsional vibration of the drill string itself. The main parameters that are important in drilling are the spindle rotation speed (RPM), compressive force (derived as "weight on bit" / WOB) and soil-rock formation character.

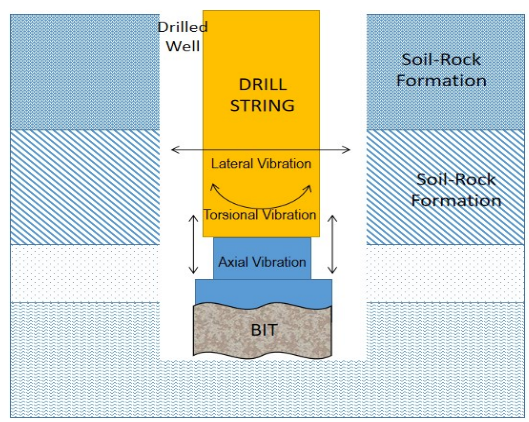


Figure 1: Vibration Direction on Drilling Dynamic

2.3 Membership Based – Fuzzy Logic Outlier Detection

Mazidi, et al. [5] conducted a study by observing several methods of detecting anomalous data. Each method has advantages and disadvantages, so certain anomaly detection methods are only suitable for use on certain systems or applications. He developed a new method for the detection of anomalies in data based on the weighting and the degree value of "membership".

2.4 Machine Learning Application on Wellbore Stability Prediction

In accordance with research conducted by Jahanbakhshi, et al. [6], wellbore stability is one of the most important and decisive factors for success in oil and gas drilling activities. Wellbore stability is one of the main challenges because it is related to geological conditions (formations), which are very complex and strongly influenced by formation pressure. The method that will be used in the research above is to utilize the artificial neural network (ANN) technique to make optimizations to prevent the instability of the drilling process.

2.5 Industrial application for Fuzzy Outlier Detection

Research conducted by Cateni, et al. [2] aims to create an anomaly/outlier data detection system by combining several approaches to obtain anomalous or outlier data. The methods used in this study include model-based outlier detection, the grubb test, fast outlier detection, outlier detection with RNN, the fuzzy clustering algorithm, and several other anomaly detection methods. The fuzzy inference system (FIS) is then used to combine the results of each method to produce a more precise detection of anomaly data while still maintaining the quality of the detection.

Fuzzy is a form of many-valued logic in which the truth value of variables may be any real number between 0 and 1. This fuzzy system is based on an approach just like people's observations when they make decisions based on imprecise and non-numerical information. FIS is a system that uses fuzzy set theory to map an input space to an output space. FIS uses a collection of fuzzy membership functions and rules to generate an output. A FIS is a way of mapping an input space to an output space using fuzzy logic. Instead of boolean logic, FIS uses a collection of fuzzy membership functions and rules to reason about data.

3.1 Methodology

In practice, this research will be carried out to detect outliers using FIS. Then, detected outliers in drilling data are used to see if these outliers indicate changes in drilling parameters, particularly data of RPM, ROP, and WOB. While other research mainly detects outliers directly from the data set, this research will get outliers from the calculated standard statistical value that is fed to the FIS. Then the FIS results will be combined to determine whether the formation changes are indicated or not.

For those above, raw data from the MLU is obtained, which is then prefiltered so that only the selected dataset is used. From these three (3) datasets, four (4) statistical standard values were calculated, namely DIST, N-POINT-DIST, MEM-DEG-DIST (distance from data point to center point cluster), and MEAN-DIST for each of the datasets mentioned above.

The FIS method is used to detect outlier data with input in the form of the four (4) statistical standard values for each ROP, RPM, and WOB dataset to produce three (3) outlier datasets for each data set. When anomalous data or outliers are detected, this indicates a significant change in parameters compared to the average data points and creates a batch outlier dataset. These outlier datasets were then processed again to indicate a variable named "change of formation/perubahan formasi tanah" according to the process in Figure 2.

Soil formation changes can be defined if there were changes in ROP (anomaly data), but the RPM and WOB are relative constants (normal). When all data produce outliers, then this condition cannot be determined to detect changes in soil formation.

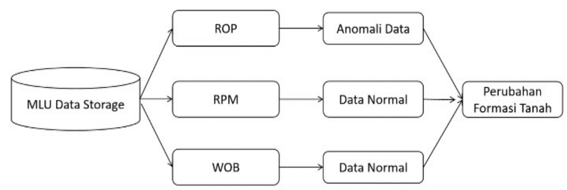


Figure 2: Method Obtaining "Perubahan Formasi Tanah"

3.2 Procedure

The research was carried out in the following steps, which are depicted in Figure 6.

a. Get data from the data storage and/or real-time MLU (mud logging unit) that generates and measures real time drilling data.

b. The model requires data pre-processing before it can be used. Data collected from a field has many parameters in the drilling system; for this study, we only use the types of data described in Table 1.

Table 1: MLU Data Used for Research

No	Physical Variable	Unit
1	Time	-
2	Hook load (avg)	N
3	Rate-Of-Penetration (ROP)	m/s
4	Revolution-Per-Minute (RPM)	rad/s
5	Depth Tracking	m

c. Filtering data of RPM, WOB, ROP.
This process will filter MLU data to only include the information needed for the calculation. RPM, WOB, ROP, and depth will be used as data.

d. Secondary data preparation.
In this study, each of these data sets will go through a calculation process according to the following four (4) statistical calculations as below.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Centroid-data point distance named as 'DIST',
- Neighboring Point Center-Data Point distance named as 'N-POINT-DIST',
- The K-Means Clustering method is used to determine the cluster center point. With the known cluster center, then the distance center cluster value to the data points can be calculated as 'MEM-DEG-DIST', and
- Mean distance to data points as named 'MEAN-DIST'.

e. Outlier detection.
FIS methods are used to detect outlier data, from which the input was four data points for the statistical calculation above.

f. Combining outlier.
The resulting outlier data is then processed to produce formation change data, as shown in Figure 7. Outlier data detected in the ROP data set will set a "formation change" value, but when the RPM data and WOB data are detected as outliers, the "formation change" variable is changed back to "no change" per the relation matrix seen in Table 2.

Table 2: Relation matrix for detected outlier – formation changes

Condition	ROP	RPM	WOB	Formation
1	normal	normal	normal	no-change
2	outlier	normal	normal	change
3	outlier	outlier	normal	n/a
4	outlier	outlier	outlier	n/a

g. Validation.
Finally, the formation change detection result is then compared to the drilling report to validate that the formation changes are in accordance with each depth point.

4.1 Data Processing

Data will be processed in the items listed below.

a. MLU raw data
Data can be obtained from MLU data storage sources or MLU real-time input data, which generates and measures drilling data in real-time. Real-time data is sourced from drilling instrumentation in the form of continuous data. MLU stored data is presented in ASCII format, as seen in Figure 3.

```
I_MD,MD,LG_ROP,PG_RPM,PG_HKLD,PG_TORQ,PG_WOB,PG_ECD,PG_PPRES,PG_MWIN,PG_GPM,PG_T_OUT
-188,188,117,62,43.6,978.2,10,9.75,8.95,9.25,622.2,110
-189,189,62,62,46.5,951.8,7.1,9.73,8.18,9.22,622.5,111
-190,190,132,62,43.9,976,9.8,9.75,8.9,9.22,622.3,111
-191,191,133,62,41.9,981,11.7,9.78,8.99,9.25,621.5,112
-192,192,109,62,42.2,977.3,11.5,9.76,8.55,9.21,622,112
-193,193,131,62,40.9,988.8,12.7,9.75,8.77,9.19,622,112
-194,194,106,62,41.3,982.2,12.4,9.76,8.32,9.18,622.4,112
-195,195,53,62,45.1,943.6,8.6,9.76,8.14,9.19,622.5,112
-196,196,73,62,43.8,956.2,9.8,9.75,8.16,9.18,622.4,112
-197,197,63,62,44.6,942.7,9,9.75,8.17,9.17,622.9,112
-198,198,109,62,40.7,970,12.9,9.76,8.31,9.18,622,112
-199,199,119,62,40.2,978.8,13.5,9.79,8.45,9.18,622.2,112
-200,200,142,62,39.3,980.3,14.3,9.77,8.77,9.15,621.8,112
-201,201,80,62,43.8,942.3,9.9,9.78,8.18,9.16,622.5,112
-202,202,96,62,42.8,950.1,10.9,9.77,8.37,9.14,622.6,112
-203,203,116,62,41.6,958,12.1,9.78,8.62,9.15,622,112
-204,204,114,62,39.9,962.8,13.8,9.8,8.31,9.16,621.8,112
-205,205,111,62,40.9,955,12.7,9.79,8.42,9.15,621.4,112
-206,206,48,62,45.2,913.7,8.5,9.81,8.15,9.15,621.5,112
-207,207,71,62,42.5,934.6,11.2,9.83,8.21,9.16,621.6,111
-208,208,77,62,43,928.4,10.6,9.84,8.1,9.17,621.6,111
-209,209,76,62,43.5,928.6,10.1,9.84,8.12,9.15,622.3,111
-210,210,72,62,42.3,931,11.4,9.81,8.14,9.11,621.8,111
```

Figure 3: Stored MLU Data

b. Platform
The google.colabs service can be used for various purposes of data simulation and data analysis using the Python language base. In addition, this service enables easy access to cloud hardware combined with many versions of open-source licenses and installation features.

c. Setup environment & variables
It is needed to import the Scikit-Fuzzy, Numpy, and Pandas libraries in order to set the environment and install the libraries required by the next process.

d. Data filtering
Filtering is performed to select data that is only related and required as input into the research model. This operation can be illustrated in Figure 4.

```
[ ] dff=df.filter(items=['MD','LG_ROP','PG_RPM','PG_WOB'])
    dff
```

	MD	LG_ROP	PG_RPM	PG_WOB
0	188	117.0	62.0	10.0
1	189	62.0	62.0	7.1
2	190	132.0	62.0	9.8
3	191	133.0	62.0	11.7
4	192	109.0	62.0	11.5
...	...	...	...	...
8543	8851	13.0	196.0	17.9
8544	8852	13.0	196.0	18.2
8545	8853	13.0	197.0	18.4
8546	8854	15.0	196.0	18.7
8547	8855	16.0	195.0	18.4

8548 rows × 4 columns

Figure 4: Filtering Data

e. Statistical attribute data calculation
There are four (4) variables in the statistical attribute that shall be calculated for each data point inside the dataset (RPM, WOB, and ROP). Those variables are described as follows.

Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Variable 'DIST' calculated with following function (1).

$$\mathrm{DIST_n} = \mathrm{X_n} - \frac{\sum_{i=0}^{n} \mathrm{X\ d\ A}}{\sum_{i=0}^{n} \mathrm{d\ A}} \quad (1)$$

- 'N-POINT-DIST' is the distance value from the neighboring point centroid to the data point. The specified neighboring points are 15 data points that follow the assumption of formation and will be the same for every 15 levels of depth. This value must come after function (2) below.

$$\mathrm{"N-POINT"_n} = \mathrm{X_n} - \frac{\sum_{i=n-1\,5}^{n} \mathrm{X\ d\ A}}{\sum_{i=n-1\,5}^{n} \mathrm{d\ A}} \quad (2)$$

- The K-Means Clustering method is run first to find the center cluster value for each dataset. The number of clusters is assumed to be 10 clusters for the entire depth data point. After knowing the center point of the cluster, the 'MEM-DEG-DIST" calculation is made according to the following function (3).

$$\mathrm{"MEM-DEG"_{n,c}} = \mathrm{X_{n,c}} - \mathrm{C_c} \cap \mathrm{c} \quad (3)$$

- "MEAN-DIST," calculated from the distance between data points to the "mean" value of the total data being processed, is the final variable.

f. FIS outlier detection

The initial process was determining the membership function of fuzzy input and output. The input set uses a Gaussian distribution function.

Before the fuzzy process is run, the program searches for max, min, mean, and median values for statistical attribute data in the previous steps. These maximum and minimum values are used as reference values in the formation of sets and membership functions. The following shows an example of a code snippet for the formation of sets and membership functions for the ROP, RPM, and WOB datasets and the output membership function "Outindx" below.

The FIS method is used as Mamdani FIS with six (6) rules applied, as seen on Figure 8. Those rules are described as in Table 3 below.

Table 3: Antecedent Membership List

No	Rules	Category
1	*rules* A:	IF (dist IS very high) AND (n-points IS very small) AND (memb-deg IS low) AND (mean-dist IS big) THEN (outindx IS very high).
2	*rules* B:	IF (dist IS medium) AND (n-points IS small) and (memb-deg IS quite low) and (mean-dist IS small) THEN (outindx IS quite high)
3	*rules* C:	IF (dist IS low) AND (n-points IS medium) and (memb-deg IS quite low) AND (mean-dist IS very small) THEN (outindx IS low)
4	*rules* D:	IF (dist IS medium) and (n-points IS very small) AND (memb-deg IS quite low) AND (mean-dist IS small) THEN (outindx IS high)
5	*rules* E:	IF (dist IS low) AND (n-points IS small) AND (memb-deg IS high) AND (mean-distance IS quite big) THEN (outindx IS low)
6	*rules* F:	IF (dist IS low) AND (n-points IS medium) AND (memb-deg IS high) AND (mean-dist IS small) THEN (outindx IS low).

Because the FIS Input membership function (antecedent) applies to each secondary data (statistical attribute), the total combination membership function was 12 antecedent membership functions, as shown in Table 4.

Table 4: Antecedent Membership List

No	Dataset	Statistical Attr.	Membership
1	ROP, RPM, WOB	DIST	dist_rop, dist_rpm, dist_wob
2	ROP, RPM, WOB	N-POINT	n_point_rop, n_point_rpm, n_point_wob
3	ROP, RPM, WOB	MEMB-DEG	mem_deg_rop, mem_deg_rpm, mem_deg_wob
4	ROP, RPM, WOB	MEAN-DIST	mean_dist_rop, mean_dist_rpm, mean_dist_wob

Below is a code snippet as an example for determining the parameter value of fuzzification membership for "dist_rop." This code also applies to and is modified for all twelve (12) antecedent related ROP, RPM, and WOB for each DIST, N-POINT, MEMB-DEG, and MEAN-DIST presented in Table 4 above.

```
#setting for fuzzification parameter setting
#-----------------Setting Parameter------------
#DIST ROP Parameter:
x=dffh.at['min','dist_rop']
y=dffh.at['max','dist_rop']
delta = abs(x-y)
```

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

```
dismean_rop_low = x
dismean_rop_mid = x+(delta*3/6)
dismean_rop_high = y
zigma_rop = delta/6

u_dist_rop['low'] = fuzz.gaussmf(u_dist_rop.universe,
dismean_rop_low, zigma_rop)
u_dist_rop['mid'] = fuzz.gaussmf(u_dist_rop.universe,
dismean_rop_mid, zigma_rop)
u_dist_rop['high']= fuzz.gaussmf(u_dist_rop.universe,
dismean_rop_high, zigma_rop)
```

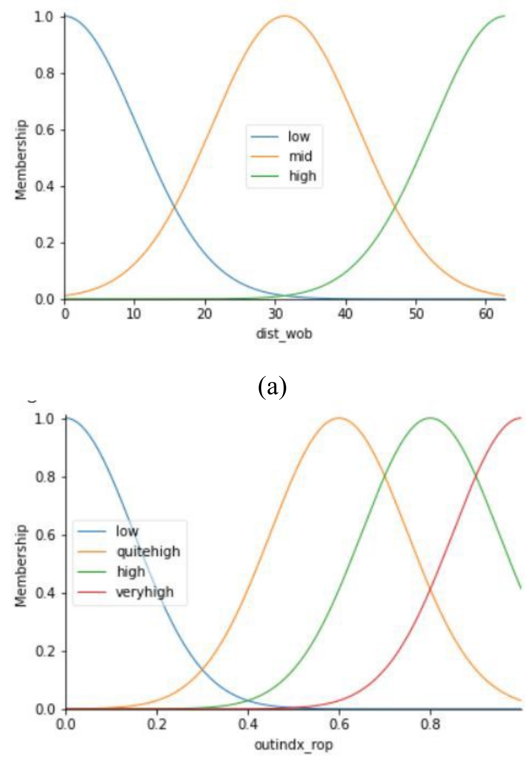
There are only three (3) output memberships in the FIS Output Membership function (consequent), which are outindx_rop, outindx_rpm, and outindx_wob. Figure 5 show plot example of "dist_wob" and "outindx_rop" membership functions.

Using the code snippet below, you can determine the next membership parameter based on the result.

```
#Output parameter (outindx)
   zigma_out = 0.15
   mean_out_low = 0
   mean_out_quitehigh = 0.6
   mean_out_high = 0.8
   mean_out_veryhigh = 1

   u_outindx_rop['low'] =
   fuzz.gaussmf(u_outindx_rop.universe, mean_out_low,
zigma_out)
   u_outindx_rop['quitehigh'] =
   fuzz.gaussmf(u_outindx_rop.universe, mean_out_quite
high, zigma_out)
   u_outindx_rop['high'] =
   fuzz.gaussmf(u_outindx_rop.universe, mean_out_high,
 zigma_out)
   u_outindx_rop['veryhigh'] =
   fuzz.gaussmf(u_outindx_rop.universe, mean_out_veryh
igh, zigma_out)

   u_outindx_rpm['low'] =
   fuzz.gaussmf(u_outindx_rpm.universe, mean_out_low,
zigma_out)
   u_outindx_rpm['quitehigh'] =
   fuzz.gaussmf(u_outindx_rpm.universe, mean_out_quite
high, zigma_out)
   u_outindx_rpm['high'] =
   fuzz.gaussmf(u_outindx_rpm.universe, mean_out_high,
 zigma_out)
   u_outindx_rpm['veryhigh'] =
   fuzz.gaussmf(u_outindx_rpm.universe, mean_out_veryh
igh, zigma_out)
   u_outindx_wob['low'] =
   fuzz.gaussmf(u_outindx_wob.universe, mean_out_low,
zigma_out)
   u_outindx_wob['quitehigh'] =
   fuzz.gaussmf(u_outindx_wob.universe, mean_out_quite
high, zigma_out)
   u_outindx_wob['high'] =
   fuzz.gaussmf(u_outindx_wob.universe, mean_out_high,
 zigma_out)
   u_outindx_wob['veryhigh'] =
   fuzz.gaussmf(u_outindx_wob.universe, mean_out_veryh
igh, zigma_out)
```



(a)

(b)

Figure 5. (a) Antecedent "dist_wob" membership function and (b) Consequent "outindx_rop" membership function

After the previous process of defining the rules and control settings, the FIS simulation is then run (with input in the form of data in each row of the dataset) according to the code snippet provided below.

```
#computing for every data row value
i=0
for row in dfg.itertuples():
 outindxing_rop.input['dist_rop']=dfg.at[i,'dist_rop']
```

```
 outindxing_rop.input['n_pt_dist_rop']=dfg.at[i,'n_pt_dist_rop']
 outindxing_rop.input['MEM_DEG_rop']=dfg.at[i,'MEM_DEG_rop']
 outindxing_rop.input['MEAN_DIS_DATA_rop']=dfg.at[i,'MEAN_DIS_DATA_rop']

 outindxing_rop.compute()
 dfg.at[i,'outindx_rop'] = outindxing_rop.output['outindx_rop']

 outindxing_rpm.input['dist_rpm']=dfg.at[i,'dist_rpm']
 outindxing_rpm.input['n_pt_dist_rpm']=dfg.at[i,'n_pt_dist_rpm']
 outindxing_rpm.input['MEM_DEG_rpm']=dfg.at[i,'MEM_DEG_rpm']
 outindxing_rpm.input['MEAN_DIS_DATA_rpm']=dfg.at[i,'MEAN_DIS_DATA_rpm']

  outindxing_rpm.compute()
  dfg.at[i,'outindx_rpm'] = outindxing_rpm.output['outindx_rpm']

 outindxing_wob.input['dist_wob']=dfg.at[i,'dist_wob']
 outindxing_wob.input['n_pt_dist_wob']=dfg.at[i,'n_pt_dist_wob']
 outindxing_wob.input['MEM_DEG_wob']=dfg.at[i,'MEM_DEG_wob']
 outindxing_wob.input['MEAN_DIS_DATA_wob']=dfg.at[i,'MEAN_DIS_DATA_wob']

 outindxing_wob.compute()
 dfg.at[i,'outindx_wob'] = outindxing_wob.output['outindx_wob']
i=i+1
```

4.2 Result

Detected outliers or anomalies at drilling depth for each dataset (ROP, RPM, and WOB) can be seen in Figures 9, 10, and 11 respectively. Then from that outlier data then the FIS Output (outindx) value of each dataset (ROP, RPM, and WOB) is seen at Figure 12.

The MLU datasets above have been processed with calculation into four (4) statistical values, and then those results are processed further with FIS to obtain outliers. The detected outliers for each dataset are then combined to indicate soil formation changes. The FIS method using input in the form of drilling data in this study has been obtained and resulted in some outlier data being detected. Based on the approach and results of the FIS operation, detected outliers indicating formation changes amounted to 70 data points, as seen in Figure 13.

According to the results, only 70 data points out of a total of 8548 data points (drilling depths ranging from 180 ft to 8855 ft) indicate change in soil formation. These 70-point results are then compared with the formation log report, which has 218 points indicating soil formation changes. It was found that only two data points were verified. The method conducted in this research has a calculated certainty level of 0.93%.

5. CONCLUSION

Reducing the risk of well stability problems during drilling is necessary. Detecting changes in soil formation in drilling based on data from MLU has a good opportunity to reduce the above-mentioned risk.

The MLU data that is processed is depth, ROP, RPM, and WOB. From these data sets, four (4) statistical parameters are calculated as input for FIS outlier detection. Then, detected outliers in drilling data are used to see if these outliers indicate changes in drilling parameters. Then the FIS results will be combined to determine whether the formation changes are indicated or not. In this study, the FIS method was obtained, and outlier data was detected using input in the form of drilling data.

Based on the approach and results of the FIS operation, detected outliers indicating formation changes amounted to 70 data points out of a total of 8548 data points (covering drilling depths from 180 ft to 8855 ft). These results are then compared with the formation log report, which has 218 soil formation changes. Only two data points were verified to match the formation log report. The method conducted in this research has a calculated certainty level of 0.93%.

This method has a low degree of certainty, and this result shall not be verified using a formation log report. This is because the data report comes from the MLU data storage that recorded and captured the measurement of mud in the "surface area". This data may be very different from what the actuals should be due to a delay in information caused by the time needed for cutting to reach surface ground level to be measured.

Unfortunately, the data that comes from MLU storage is stored data that has undergone an averaging and filtering process so that it becomes limited and has been "disturbed" from the actual data. Most likely, the developed system in this research was not able to detect any changes in soil formation due to this "disturbed" data.

Based on this research, the FIS method still needs to be optimized by adding more data to be used as feedback and training to increase the accuracy of the FIS model and to require further input from the expert to improve the quality of detection.

REFERENCES:

[1] Fertl, W. H., Chilingar, G. V., & Robertson Jr, J. O, Drilling parameters, In Developments in petroleum science, Vol. 50, 2002, pp. 151-167.

[2] Cateni, S., Colla, V., & Vannucci, M., A fuzzy logic-based method for outliers detection, In Artificial Intelligence and Applications, 2007, pp. 605-610.

[3] Tamaarat, A., Active and Reactive Power Control for DFIG Using PI, Fuzzy Logic and Self-Tuning PI Fuzzy Controllers, Advances in Modelling and Analysis C, 2019, pp. 95–102.

[4] Xue, Q., Leung, H., Huang, L., Zhang, R., Liu, B., Wang, J., & Li, L., Modeling of torsional oscillation of drill string dynamics, Nonlinear Dynamics, 2019, pp. 267–283.

[5] Mazidi, A., Roshanfar, F., Parvin Darabad, V., A Review of Outliers: Towards a Novel Fuzzy Method for Outlier Detection, Journal of Applied Dynamic Systems and Control, 2(1), 2019, pp. 7-17.

[6] Jahanbakhshi, R., & Keshavarzi, R., Intelligent prediction of wellbore stability in oil and gas wells: An artificial neural network approach. In 46th US Rock Mechanics/Geomechanics Symposium, 2012.

[7] Ahmed, M., Mahmood, A. N., & Islam, M. R., A survey of anomaly detection techniques in financial domain, Future Generation Computer Systems, ed 55, 2016, pp. 278-288.

[8] Baldwin, J. F., Fuzzy logic and fuzzy reasoning, International Journal of Man-Machine Studies, 11(4), 1979, pp. 465-480.

[9] Birdwell, J. D., Applications Of Artificial Intelligence In Control System Analysis And Design, In IECON'87: Automated Design and Manufacturing, Vol. 857, 1987, pp. 798-801.

[10] Bourgoyne, A., & Young, F. A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection, Society of Petroleum Engineers Journal, 14(04), 1974, pp. 371–384.

[11] Bowler, A., Harmer, R., Logesparan, L., Sugiura, J., Jeffryes, B., & Ignova, M., Continuous High-Frequency Measurements of the Drilling Process Provide New Insights Into Drilling-System Response and Transitions Between Vibration Modes, SPE Drilling & Completion, 31(02), 2016, pp. 106–118.

[12] Burrough, P. A., Macmillan, R. A., & Deursen, W., Fuzzy classification methods for determining land suitability from soil profile observations and topography, Journal of Soil Science, 43(2), 1992, pp. 193–210.

[13] Christine, M., Reinforcement learning: a powerful AI in ever more areas. June 11, 2022.

[14] Das Purkayastha, A., Rana, R., Talreja, R., & Rajaiah, N., Drilling Event Chart: A Kick Prevention Tool, In SPE Gas & Oil Technology Showcase and Conference, 2019.

[15] Dreyfus, H. L., & Dreyfus, S. E., Mind over machine: The power of human intuition and expertise in the era of the computer (First Edition), CRC Press (Oxford, UK), 1986.

[16] Duckstein, L., Bardossy, A., Fuzzy Rule-Based Modeling with Applications to Geophysical, Biological, and Engineering Systems, CRC-Press (UK), 1995.

[17] https://www.mathworks.com/help/fuzzy/fuzzy-inference-process.html, Fuzzy Inference Process - MATLAB & Simulink. (n.d.). June 11, 2022.

[18] Hu, J., Niu, H., Carrasco, J., Lennox, B., & Arvin, F., Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning, IEEE Transactions on Vehicular Technology, 69(12), 2020, pp. 14413-14423.

[19] Hu, J., & Lanzon, A., Distributed finite-time consensus control for heterogeneous battery energy storage systems in droop-controlled microgrids, IEEE Transactions on smart grid, 10(5), 2018, pp. 4751-4761.

[20] Jyothsna, V. V. R. P. V., Prasad, R., & Prasad, K. M., A review of anomaly based intrusion detection systems, International Journal of Computer Applications, 28(7), 2011, pp. 26-35.

[21] Kaur, H., Singh, G., & Minhas, J., A Review of Machine Learning based Anomaly Detection Techniques, International Journal of Computer Applications Technology and Research, 2(2), 2013, pp. 185–187.

[22] Macini, P., Mesini, E., Ferrari, M., Pisconti, G., Antoncecchi, I., & Terlizzese, F., Offshore Safety: Collection and Recording of Relevant Drilling Data in the Italian Offshore by Means

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of Virtual Black Boxes, In SPE Europec featured at 80th EAGE Conference and Exhibition, 2018.

[23] Mamdani, E. H., & Assilian, S., An experiment in linguistic synthesis with a fuzzy logic controller, International journal of human-computer studies, 51(2), 1999, pp. 135-147.

[24] Marsili-Libelli, S., Fuzzy Clustering of Ecological Data. Coenoses, 4(2), 1989, pp. 95–106.

[25] Mendel, J. M., & Zapalac, J. J., The application of techniques of artificial intelligence to control system design, In Advances in Control Systems, Vol. 6, 1968, pp. 1-94.

[26] Miller, D. J., & Browning, J., A mixture model and EM algorithm for robust classification, outlier rejection, and class discovery. In ICASSP (2), 2003, pp. 809-812.

[27] Odeh, I. O. A., McBratney, A. B., & Chittleborough, D. J., Soil pattern recognition with fuzzy-c-means: Application to classification and soil - landform interrelationships, Soil Science Society of America Journal, 56(2), 1992, pp. 505-516.

[28] Özkale, Aslıhan & Özer, Ceren & Kırbıyık, Selin & Eren, Tuna., A Case Study: Comparison of the Theoretical and Actual Drilling Hydraulic Pressures for Wells in Turkey, 2007.

[29] Paasch, J., Weiterentwicklung und Portierung des Fuzzy-Clustering-Systems ECOFUCS, Dipl. Inst. für Informatik und Praktische Mathematik, Universität Kiel, 1994.

[30] Russell, S. J., Artificial intelligence a modern approach, Pearson Education, Inc, 2010.

[31] Salski, A., & Kandzia, P., Modelling support system for ecological application based on fuzzy logic, Bouchon-Meunier, Valverde & Yager (Eds.): Uncertainty in Intelligent Systems, 1993, pp. 269-275.

[32] Sugeno, M., Industrial Applications of Fuzzy Control, New York: Elsevier Science Inc, 1985.

[33] Thomason, R., Logic and Artificial Intelligence. Stanford Encyclopedia of Philosophy, 2018.

[34] https://en.wikipedia.org/wiki/Artificial_intelligence, June 11, 2022.

[35] Zadeh, L. A., Fuzzy sets, Inform Control (8), 1965, pp. 338-353.

[36] Zielinski, K. M., Hendges, L. V., Florindo, J. B., Lopes, Y. K., Ribeiro, R., Teixeira, M., & Casanova, D., Flexible control of Discrete Event Systems using environment simulation and Reinforcement Learning, Applied Soft Computing, 111, 2021, 107714.

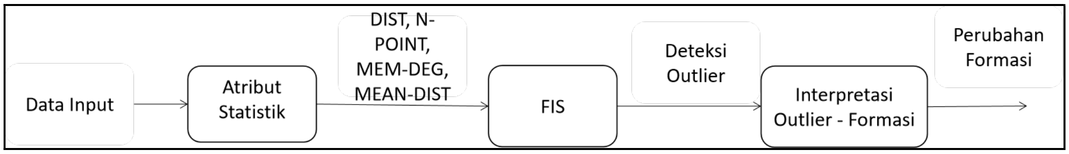


Figure 6: Steps Procedure

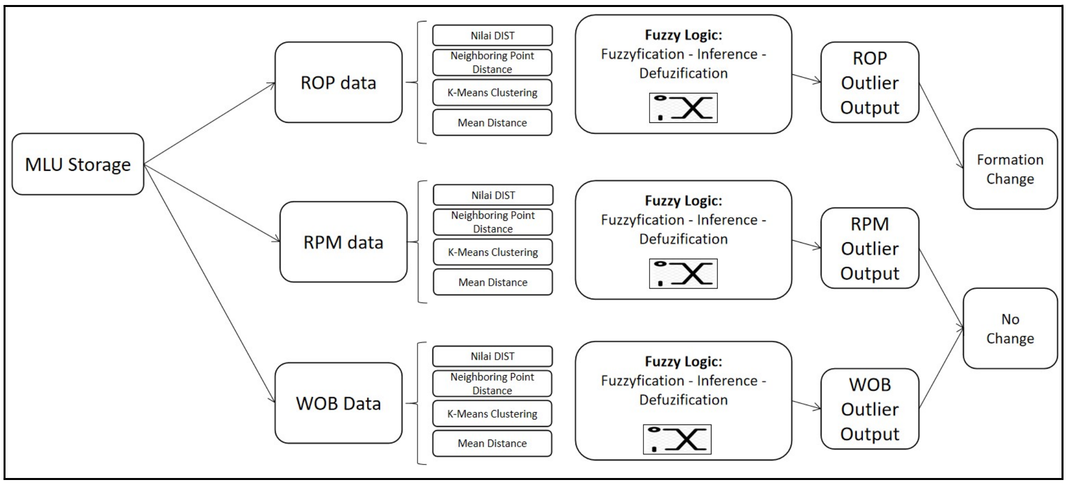


Figure 7: Detection Method of Soil Formation Changes

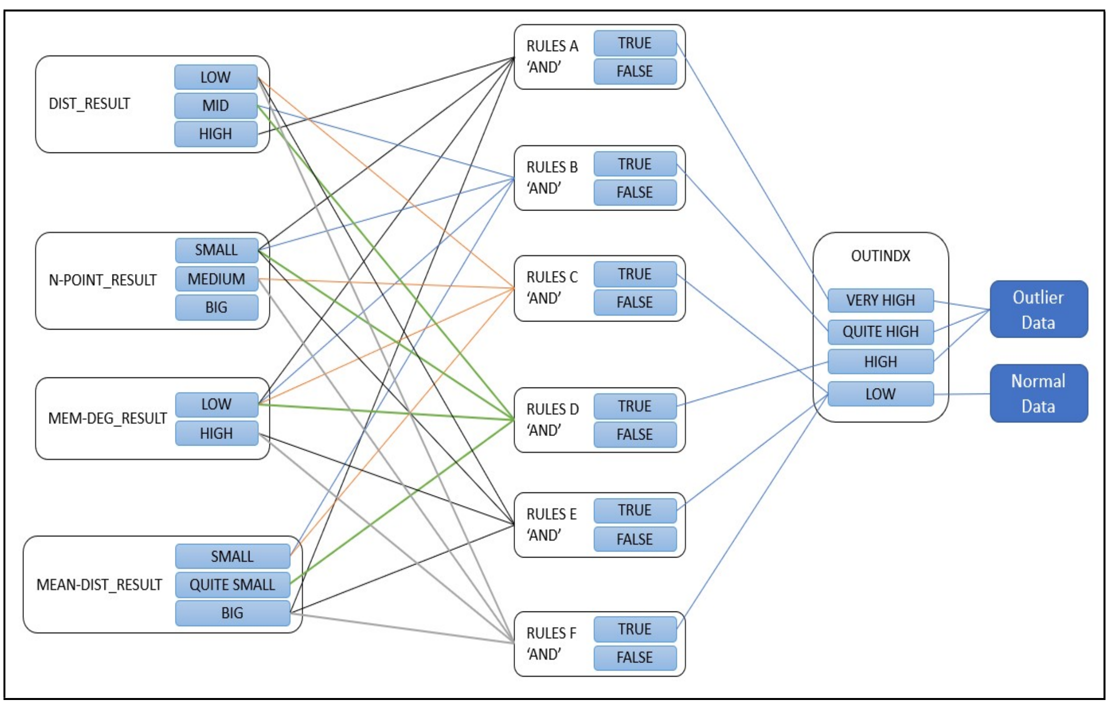


Figure 8: FIS Model (Input – Rules – Output)

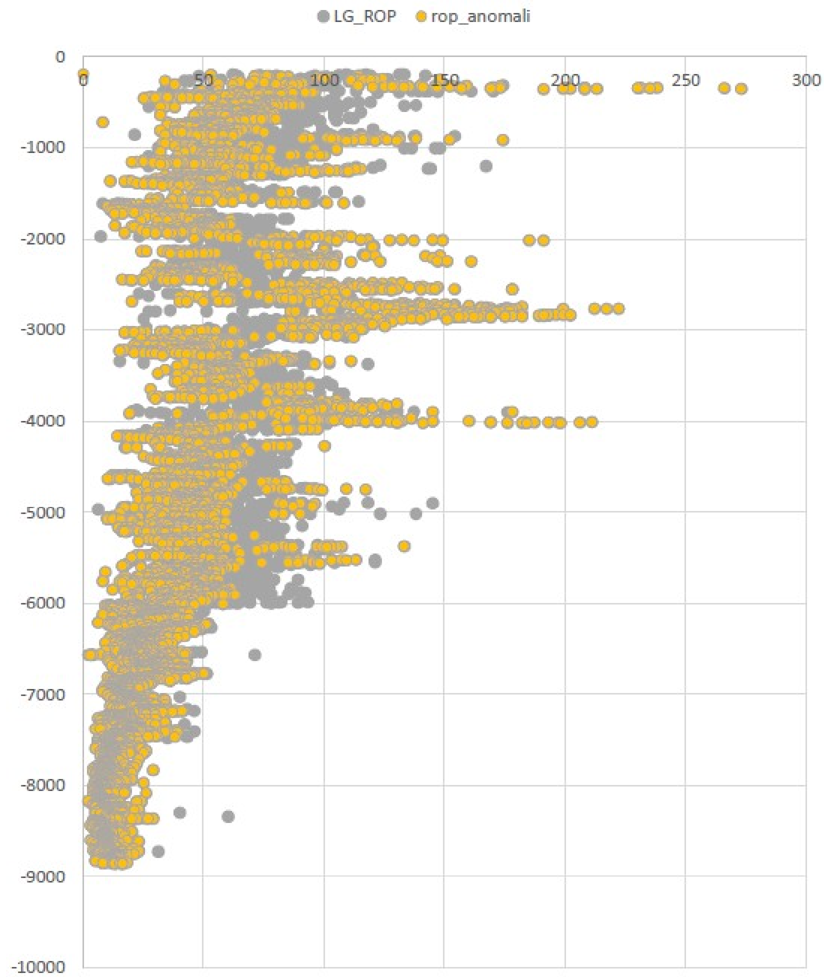


Figure 9: Outlier vs ROP Data "rop_anomali"

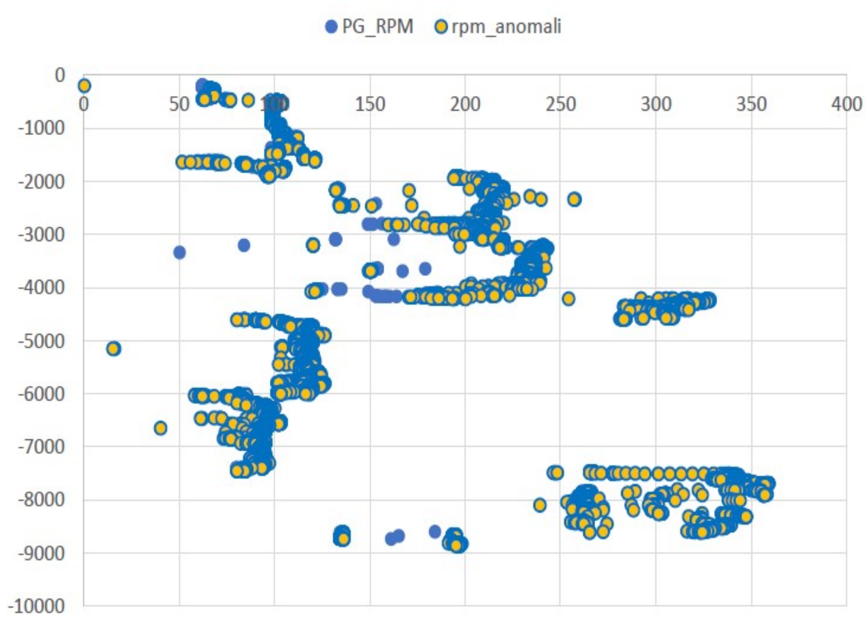


Figure 10: Outlier vs RPM Data "rpm_anomali"

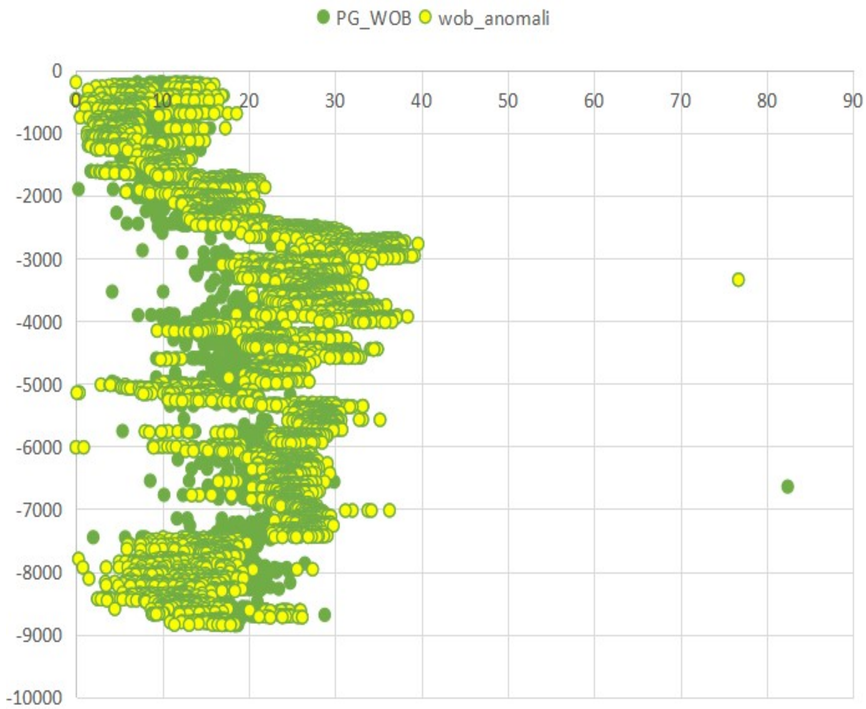


Figure 11: Outlier vs WOB Data "wob_anomali"

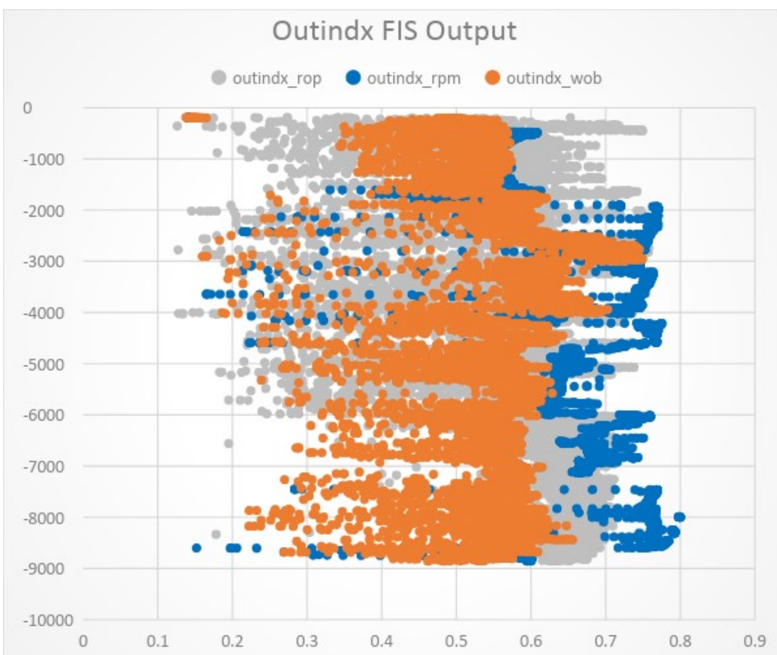


Figure 12: FIS Output Value vs Depth for WOB, RPM, and ROP

Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

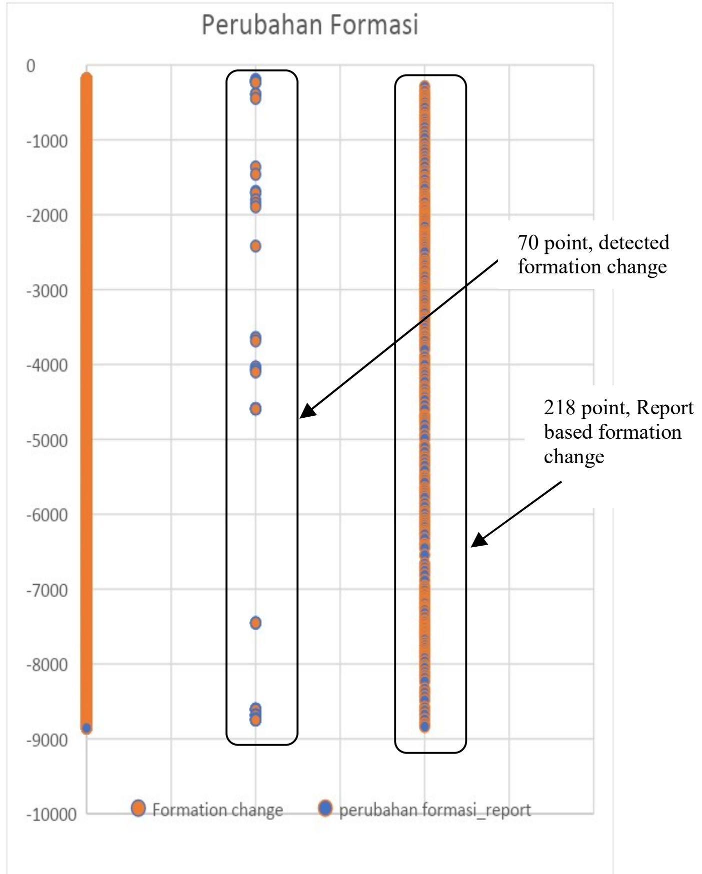


Figure 13: Detected Formation Change vs Actual Log Report "perubahan formasi report"