Mid-Semester Progress Report

DSA5900 – Spring 2021

Daniel Addokwei Tetteh

June 20th, 2022

Development of a Data-driven Model for Estimation of Reservoir Properties (Shale Volume, Porosity, And Fluid Saturations) Using Well Logs: A Case Study of the Volve Field; Norwegian North Sea

1. Introduction

Well logging is the discrete or continuous recording of measurements within a wellbore with a well logging tool or a probe (Hearst and Nelson, 1985). The logging tools are usually placed at the end of a wireline tool and lowered into a petroleum well to measure rock and fluid properties. Standard well logs in the oil and gas industry include resistivity, density, sonic, gamma-ray, caliper, and neutron porosity logs. Well logs provide essential information about geologic formations, which are used in the analysis of geophysical properties of wells, evaluation of formation rock characteristics, estimation of hydrocarbons initially in place, investigation of reservoir pressures, and estimation of in-situ petrophysical, geomechanical and geochemical reservoir properties which is essential for reservoir modeling and production forecasting amongst others (Mondol, 2015).

Despite the numerous advantages of using well logs, well logging tools do not directly measure these properties. On the contrary, these properties are obtained by processing, interpreting, and calibrating the well logs and thus are associated with several uncertainties that may affect the overall outcome (Moore et al., 2011). Furthermore, petroleum engineers and geophysicists use statistical and empirical methods together with well logs to build three-dimensional reservoir models for reserve estimation and improvement of production. However, the process is usually costly, cumbersome, and associated with human errors (Wang et al., 2021).

Data-driven models have recently been a better alternative in reservoir property estimation and characterization using well logs (Fajana et al., 2018; Saproetti et al., 2018). Machine learning has tremendous potential in predicting reservoir properties better than conventional methods using the large volumes of well log data presently available in the oil and gas industry. It is a less cumbersome approach and can be done at a minimal cost.

2. Objectives

1. The primary objective of this project is to develop a machine learning data-driven model to estimate reservoir properties, including shale volume, porosity, and fluid saturation, based on a standard set of well logs, including gamma-ray, bulk density, neutron porosity, resistivity, and sonic.
2. A secondary objective is to cluster well logs from the training data into various electrofacies (i.e., clusters showing rock intervals with specific characteristics). These electrofacies would be depth-matched and compared with lithological plots obtained from the original database as an extra step in validating the performance of the clustering algorithm.

2.1.Learning Objectives

This project would promote a thorough exploration of the Exploratory Data Analysis (EDA), Imputation techniques, and outlier detection and removal techniques. Also, feature engineering, model development, and model performance improvement techniques like cross-validation and data stratification will be investigated. Overall deployment of data-driven models as a cost-effective solution to pertinent petroleum engineering problems is investigated.

3. Data Source and Acquisition

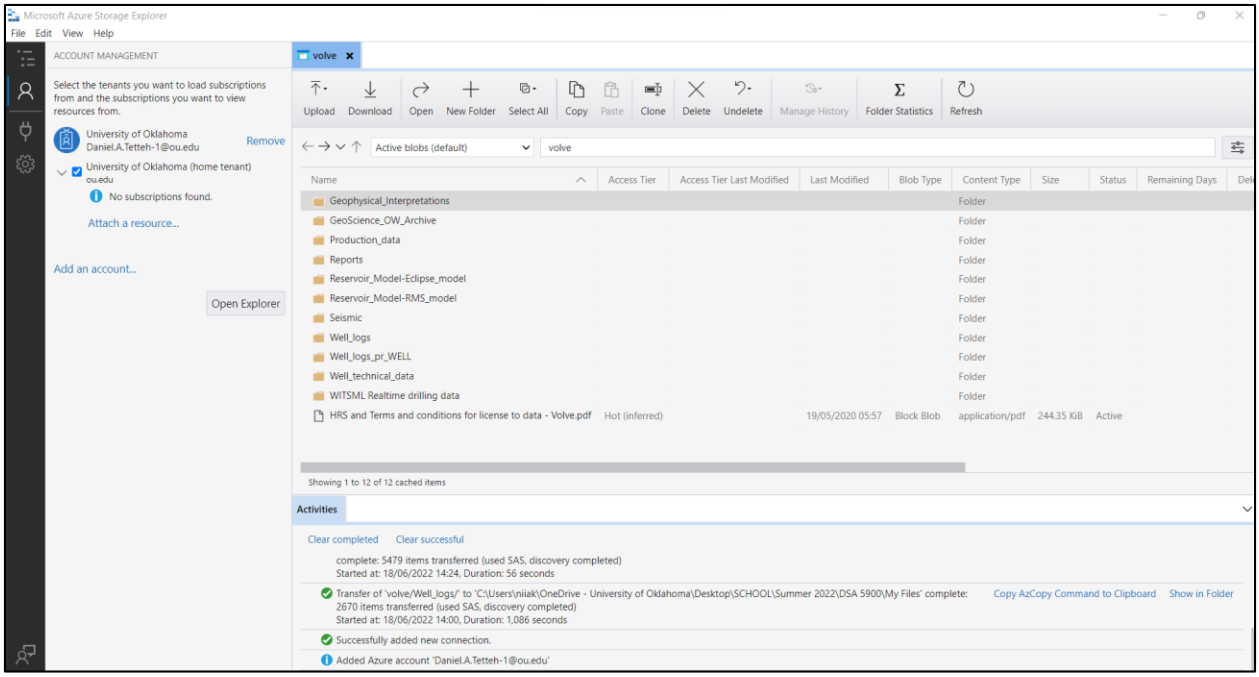
Data employed in this work is from the Volve Field owned by Equinor.

Field Description: The Volve is a Norwegian offshore petroleum field discovered in 1993. The field is located about 5 km from the north Sleipner Ost field, where the water depth is about 85m deep. The Hugin sandstone formation was the main producer from the field, with a permeability of about 2913 millidarcy (Deepak). The field has a porosity of about 0.226 and showed a peak oil production of 56,000 barrels per day with an overall production of about 63 million barrels (Sen et al., 2019). The data was made available for public consumption in 2018.

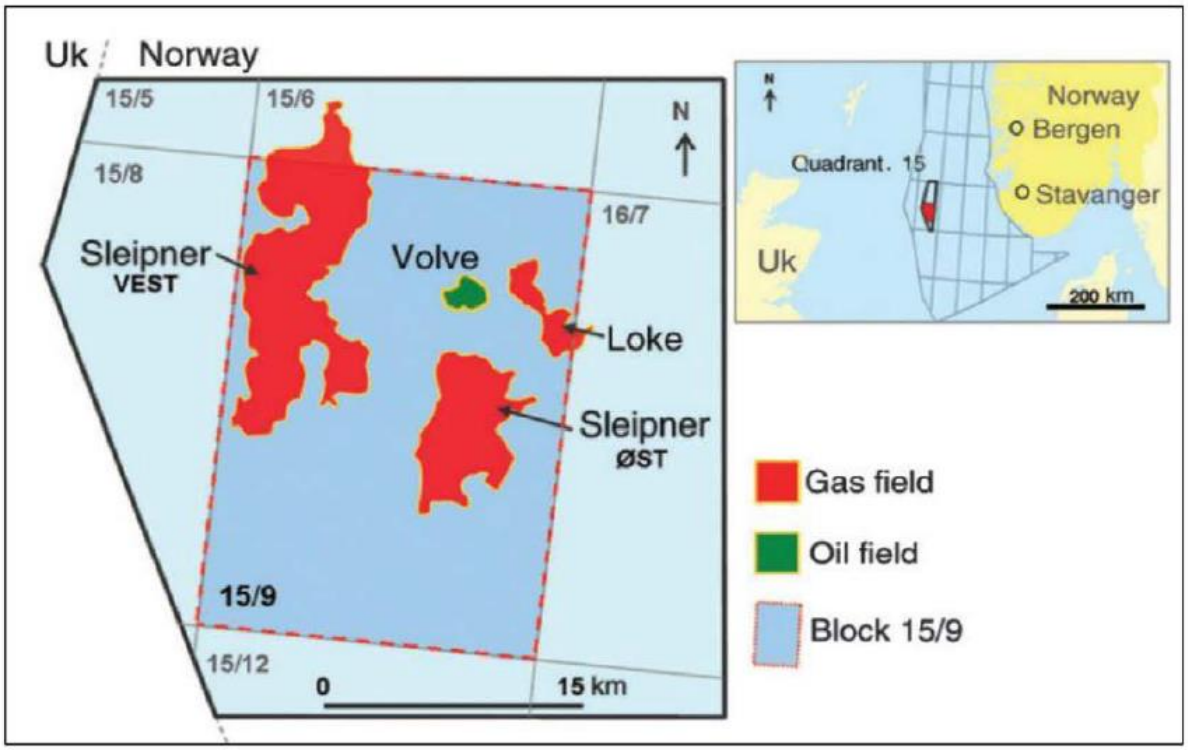
Data Acquisition: Microsoft Azure Storage Explorer was used to acquire data from the Equinor official website (https://www.equinor.com/energy/volve-data-sharing.

Available data include, Well log, reservoir modeling, seismic, production, and geophysical data, amongst others. For this work, data of interest include well logs, lithology and reservoir properties, field maps, and geological maps.

Key Challenges: Key challenges faced in the data acquisition process were extracting pertinent information from cumbersome oil field data formats and converting them into usable formats in a python coding environment. Oil and gas datasets usually have numerous inconsistencies and errors, making data mining very slow and unreliable. A .csv format form of the same data was obtained from the Society of Petrophysics and Well Log Analysts (SPWLA) PDDA (Petrophysical Data-Driven Analytics) division (https://github.com/pddasig/Machine-Learning-Competition-2021). The data has about 10 different wells with well log data and target petrophysical data to be predicted. i.e., the volume of shale (VSH), porosity (PHIF), and water saturation (SW). The training and test data sets obtained were distinct from each other. The data obtained from SPWLA was cross-checked and validated with the data acquired from the official equinor website.



Volve Field Data on Microsoft Azure Storage Explorer



Volve Field Location in the North Sea (Ravasi et al., 2015).

4. Workflow

The workflow for this project is described as follows.

1. Description of data source and acquisition
2. Explanation of fundamental concepts in estimating petrophysical properties from well logs.
3. **Exploratory Data Analysis**: Data description, univariate analysis, bivariate analysis, handling of missing values and outliers, and final data preparation for modeling. The process would be done for both the training and test datasets.
4. **Unsupervised Learning**: Well logs from the training datasets would be clustered into electrofacies using unsupervised classification algorithms; K-Means and Agglomerative Clustering. The clusters would be depth-matched and compared to lithological plots from the original dataset to validate the clustering algorithm.
5. **Supervised Learning**: Four main regression algorithms, Linear Regression, Support Vector Machine, K-Nearest Neighbours, and Random Forest models, would be explored in developing a single best model for predicting the target features. Optimization techniques like feature engineering and boosting algorithms would be employed to optimize the model's performance where applicable.
6. **Train-test split and Model Validation**: The training dataset would be divided into a training dataset and a validation data set. Standard methods like data stratification and cross-validation validation would be employed to avoid data imbalance and over-fitting or under-fitting the regression models. The performance of the models would be determined using R^2 and adjusted R^2 values and root-mean-square-error values on training and validation datasets.
7. Finally, the best model would be used to predict the test dataset's target variables. As an extra step to validate the model performance, the well logs of the test dataset would be clustered into elctrofacies and depth-matched in a similar manner to the training dataset. Also, the predicted values (i.e., Volume of Shale, Porosity, and Water saturation) would be plotted against depth and compared to their respective values for a given formation depth as seen in lithology plots from the original volve dataset from Equinor.

5. Fundamentals of Petrophysical Analysis

Before describing the exploratory data analysis (EDA) process, this section explains some primary petrophysical analysis concepts are explained showing meaningful relationships between well logs and the petrophysical properties to be predicted.

5.1.Estimation of Formation Water Saturation

The famous Archie’s equation, as shown below, is used to determine the saturation of water (Sw) in an uninvaded zone in a formation. The equation shows the dependency of water saturation on formation resistivity, obtained from resistivity logs.

$$S_w^{\,n} = \frac{R_w}{(\Phi^m \times R_t)}$$

where:

- S_w = water saturation of the uninvaded zone
- n = saturation exponent, which varies from 1.8 to 4.0 but normally is 2.0
- R_w = formation water resistivity at formation temperature
- Φ = porosity
- m = cementation exponent, which varies from 1.7 to 3.0 but normally is 2.0
- R_t = true resistivity of the formation, corrected for invasion, borehole, thin bed, and other effects

The Archie's Equation ((Archie, 1952); AAPG, 2022)

5.2. Porosity Estimation from Well Logs

Formation bulk density, neutron, and sonic logs (which usually depict shear and compressional wave travel times) are used in estimating formation porosity. The associated equations are shown below;

$$\Phi = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f}$$

where:

- Φ = porosity
- ρ_{ma} = matrix density (see table below)
- ρ_b = formation bulk density (log value)
- ρ_f = density of the fluid saturating the rock immediately surrounding the borehole—usually mud filtrate (use 1.0 for freshwater and 1.1 for saltwater mud)

Use the lithology matrix densities to determine porosity and average P_e to determine lithology listed in the table below.

Lithology	Density, g/cc	Average pe
Sandstone	2.65	1.8
Limestone	2.71	4.8
Dolomite	2.876	3.0
Anhydrite	2.977	5.05
Salt	2.032	4.6

Porosity estimation from formation density (AAPG, 2021; Alberty, 1992)

A standard early method for estimating porosity from neutron logs was proposed by (Brown and Bower; 1958) using neutron counts.

$$\log \phi = -mN_d + K$$

ϕ = porosity
N_d = neutron count
m = slope of best-fit line
K = a constant

Porosity estimation from neutron logs (Brown and Bower; 1958)

Also, in modern petroleum engineering, neutron and formation bulk density porosities are combined to estimate the true porosity of a formation.

Calculate porosity using the equation

$$\Phi = \left(\frac{\Phi_N^2 + \Phi_D^2}{2} \right)^{1/2}$$

where Φ is percent porosity, Φ_N is neutron percent porosity, and Φ_D is density percent porosity.

Formation porosity estimation equation from neutron and bulk density porosities (AAPG, 2021)

$$\frac{1}{v} = \frac{\phi}{v_f} + \frac{(1-\phi)}{v_{ma}}, \quad \text{.....(1)}$$

where

- ϕ = fractional porosity of the rock
- v = velocity of the formation (ft/sec)
- v_f = velocity of interstital fluids (ft/sec)
- v_{ma} = velocity of the rock matrix (ft/sec)

In terms of transit time (Δt):

$$\Delta t = \phi \Delta t_f + (1-\phi)\Delta t_{ma}, \quad \text{.....(2)}$$

or

$$\phi = \frac{\Delta t - \Delta t_{ma}}{\Delta t_f - \Delta t_{ma}}, \quad \text{.....(3)}$$

where

- Δt = acoustic transit time (µsec/ft)
- Δt_f = acoustic transit time of interstitial fluids (µsec/ft)
- Δt_{ma} = acoustic transit time of the rock matrix (µsec/ft)

Porosity estimation from acoustic logs (AAPG, 2021; Tenchov, 2016)

5.3 Estimation of Volume of Shale from GR logs

$$I_{GR} = \frac{GR_{log} - GR_{min}}{GR_{max} - GR_{min}} \qquad V_{Sh} = \frac{I_{GR}}{3 - 2\, I_{GR}}$$

The volume of shale estimation from GR logs (Moradi et al., 2016)

IGR = gamma ray index
GRmin = gamma ray response for the cleanest formation
GRmax = gamma ray response in shale layer
GRlog = gamma ray log value in zone of interest
Vsh= volume of shale

It must be mentioned that these estimations are not final. They are combined with outcomes from geological and reservoir models, laboratory core analysis, and tests to determine the appropriate reservoir petrophysical properties.

6. Data Exploration

The training data set has 17 features, as shown below;

- WELLNUM - Well, number
- DEPTH - Depth, unit in feet
- DTC - Compressional Travel-time, unit in nanosecond per foot
- DTS - Shear Travel-time, unit in microseconds per foot
- BS - Bit size, unit in inch
- CAL - Caliper, unit in Inc.
- DEN - Density, unit in Gram per cubic centimeter
- DENC - Corrected density, unit in Gram per cubic centimeter
- GR - Gamma Ray, unit in API
- NEU - Neutron, unit in dec
- PEF - Photo-electric Factor, unit in barns/e
- RDEP - Deep Resistivity, unit in Ohm.m
- RMED - Medium Resistivity, unit in Ohm.m
- ROP - Rate of penetration, unit in meters per hour
- PHIF - Porosity, a unit equals to the percentage of pore space in a unit volume of rock.
- SW - Water saturation
- VSH - Shale Volume

The depth unit was changed to meters to correlate the log and lithological plots from model development to that obtained from the field data. The original training dataset had 318967 rows with 17 columns. The statistics of the "raw" training data are shown in Table 1.

Table 1: Statistics of rawdata for training

	WELLNUM	DEPTH	DTC	DTS	BS	CALI	DEN	DENC	GR	NEU	PEF	RDEP	RMED	ROP	PHIF	SW	VSH
count	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000	318967.000000
mean	4.081012	2133.045195	-7791.051225	-8605.136709	-958.267420	-6982.041578	-7000.762197	-7205.371313	-43.026647	-7007.119486	-7221.969013	-949.836141	-804.023955	-989.366887	-8515.777016	-8515.693410	-8585.158486
std	2.462805	1157.611118	4168.057493	3491.479107	2969.159669	4592.475975	4582.288708	4486.572798	990.102434	4578.760758	4480.143550	2944.074718	4177.156020	3028.429856	3554.017382	3554.217716	3484.038870
min	0.000000	102.156797	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000	-9999.000000
25%	2.000000	1145.399959	-9999.000000	-9999.000000	8.500000	-9999.000000	-9999.000000	-9999.000000	23.330000	-9999.000000	-9999.000000	0.648200	0.669800	15.197600	-9999.000000	-9999.000000	-9999.000000
50%	4.000000	2104.700016	-9999.000000	-9999.000000	17.500000	-9999.000000	-9999.000000	-9999.000000	55.508400	-9999.000000	-9999.000000	1.102400	1.146000	24.969900	-9999.000000	-9999.000000	-9999.000000
75%	6.000000	3063.998854	-9999.000000	-9999.000000	26.000000	8.556900	2.246500	0.030300	78.978000	0.085100	0.058400	2.020650	2.154900	30.633700	-9999.000000	-9999.000000	-9999.000000
max	8.000000	4770.601431	181.813900	388.839700	36.000000	20.330400	3.089600	0.334158	1124.440000	1.463474	13.840700	80266.800000	97543.400000	208.633000	0.403294	1.000000	3.654300

A significantly large number of the individual column data values was "-999," which is quite unexpected for the features in context. A log plot and box plot were thus plotted to visualize the logs and target variables, as shown in fig 1 and 2, respectively. From the log plot, we can infer that the '-999' values are possible outliers as no significant deflections are seen from the surface to a depth of about 2500m for all logs beside the gamma ray (GR) log. From domain knowledge of petroleum well logging, well logging tools are usually deployed at depths of interest (usually the pay zone). The Hugin sandstone, the primary producing formation in the Volve field, is found at a depth of about 2500m from the surface across all wells (Sen et al., 2019). The basic boxplot also showed many outliers.

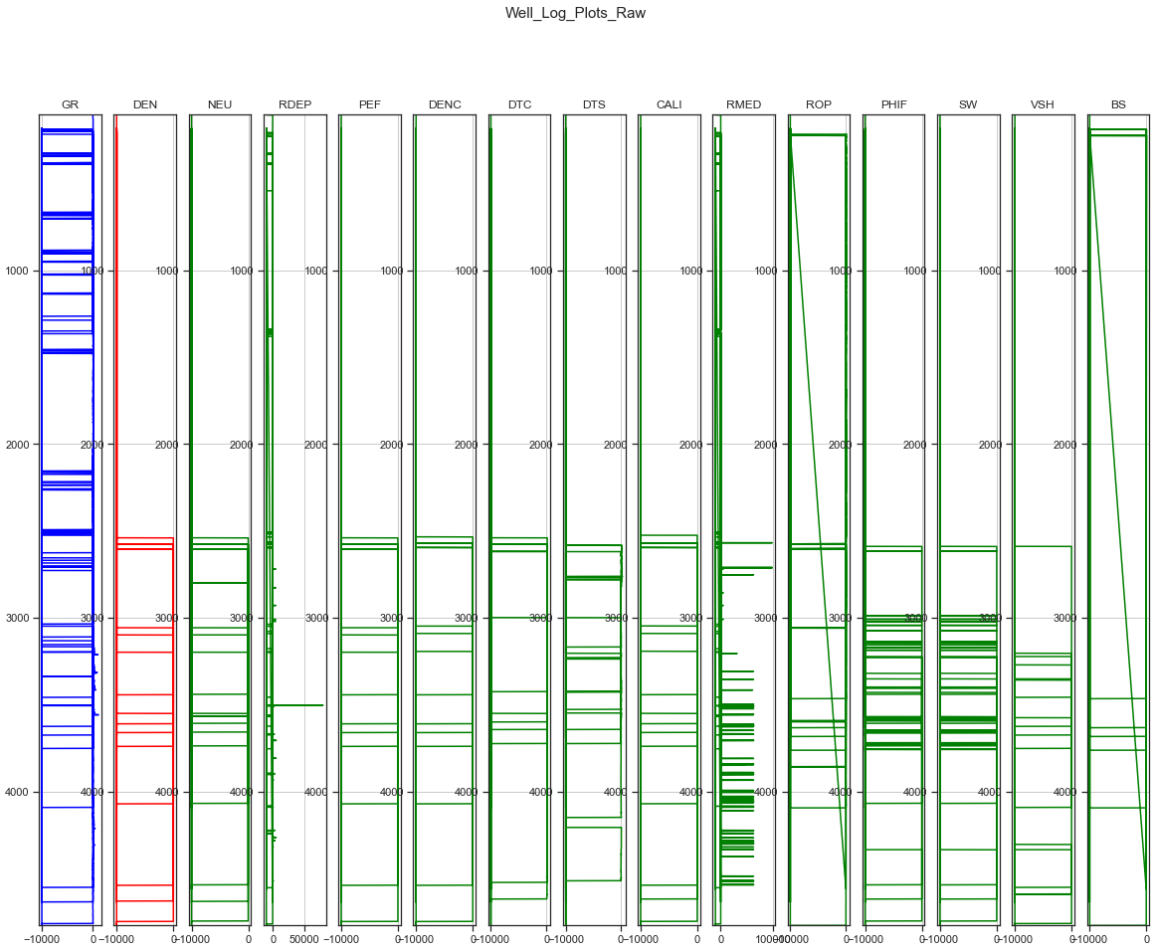


Fig 1. log plot of rawdata

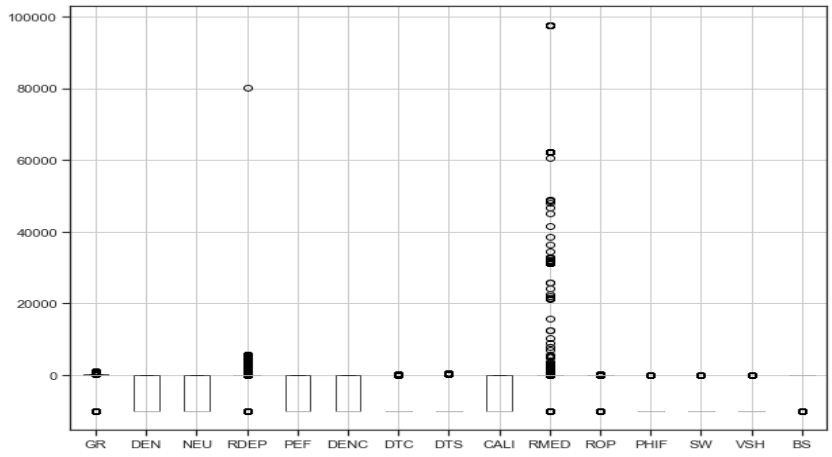


Fig 2. Boxplot of rawdata

6.1.Univariate Analysis

The following inferences were made from the univariate analysis of the raw data set.

1. Well number and well depth are uniformly and normally distributed, respectively, which is expected. This is because 'well number' is nominal. Also, we expect more data to be collected at a depth of interest, around 5000 to 10000ft.
2. Bit size is also nominal as it depends on the hole size being drilled.
3. ROP, DTS, and GR logs are right-skewed, while PEF showed some bimodality.
4. DTC, DTS, DEN, DENC, NEU, PEF, and PHIF have a more significant proportion of their distribution with average values of -999. This is unusual using domain knowledge as compression travel time, shear travel time, density, and neutron porosity logs do not usually provide such high negative values in oil and gas operations.

   Assumption: We would assume that this data was obtained from depths where the logging tools were not deployed, which are not of interest to the study.

5. RDEP and RMED values have averages of -948 and -804, respectively which is unusual for a conventional oil field like the Volve field. Resistivity logs are usually presented on logarithmic scales from 0.2 to 2000 ohms (https://glossary.oilfield.slb.com/en/terms/r/resistivity_log).

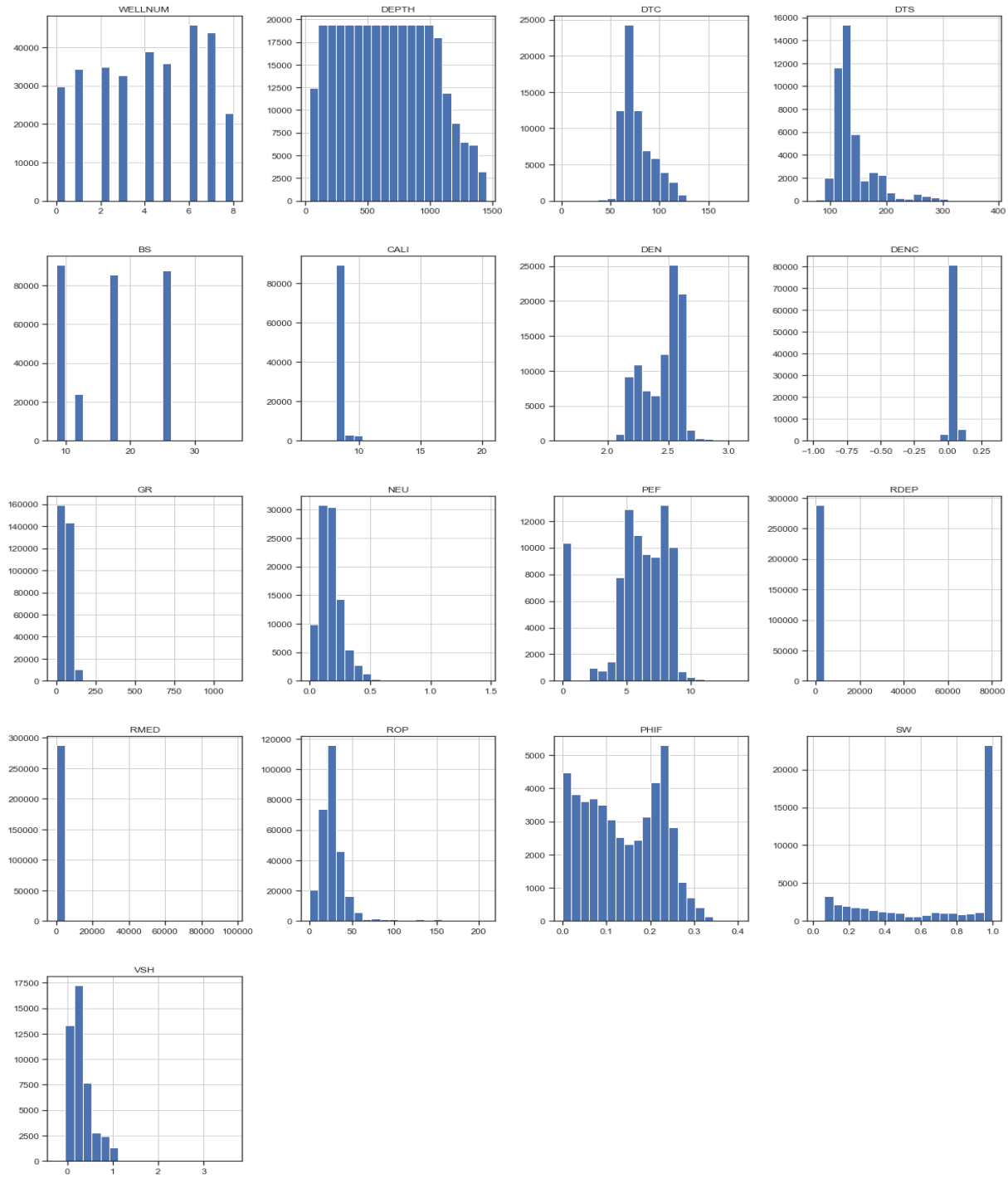


Fig 3. Histogram of rawdata features

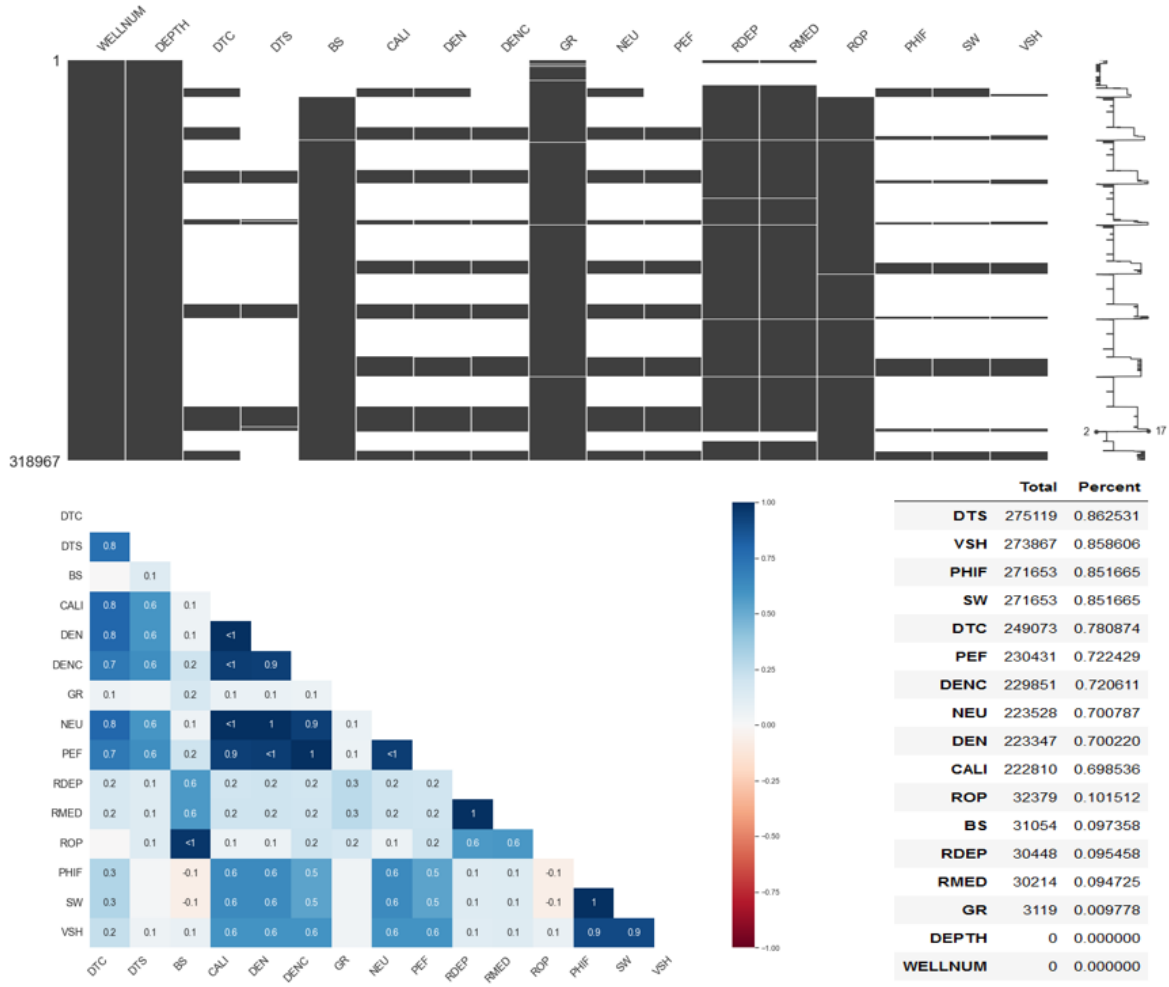
6.2.Missing Data

Before determining the missing data, all "-999" values were replaced with NAN values. Again, this was an assumption based on domain knowledge that logging tools are only deployed at depths of interest. Thus, no values (represented by -999 in this case) were recorded for depths from the surface to about 2400m. The statistics of the features after replacing '-999' values with NAN are shown in Table 2.

Table 2: rawdata statistics after replacing '-999' with NAN values.

	WELLNUM	DEPTH	DTC	DTS	BS	CALI	DEN	DENC	GR	NEU	PEF	RDEP	RMED	ROP	PHIF	SW	VSH
count	318967.000000	318967.000000	69894.000000	43848.000000	287913.000000	96157.000000	95620.000000	89116.000000	315848.000000	95439.000000	88536.000000	288519.000000	288753.000000	286588.000000	47314.000000	47314.000000	45100.000000
mean	4.081012	2133.045195	77.155278	140.490795	16.856695	8.697590	2.452806	0.050244	55.288621	0.173838	5.757870	5.140624	158.102867	28.550509	0.137371	0.701001	0.307895
std	2.462805	1157.611118	15.387921	36.085217	7.071820	0.384869	0.156333	0.020949	38.602323	0.095899	2.533365	168.911711	3082.521918	18.161868	0.085265	0.350021	0.254315
min	0.000000	102.156797	1.025100	74.822400	8.500000	6.000000	1.626600	-0.982700	0.148800	-0.003400	-0.023200	0.065000	0.064900	0.000000	0.000000	0.013000	-0.248000
25%	2.000000	1145.399959	66.363775	119.019775	8.500000	8.578100	2.311100	0.043900	24.329750	0.106300	4.891800	0.743400	0.769700	19.451200	0.060000	0.336000	0.118375
50%	4.000000	2104.700016	72.396000	130.534100	17.500000	8.625000	2.506800	0.053100	55.983000	0.156900	6.143150	1.208590	1.255200	25.916100	0.133000	0.933496	0.246509
75%	6.000000	3063.998854	85.584000	144.340900	26.000000	8.687500	2.576100	0.060600	79.240600	0.218700	7.737325	2.183900	2.348500	32.220200	0.216000	1.000000	0.390100
max	8.000000	4770.601431	181.813900	388.839700	36.000000	20.330400	3.089600	0.334158	1124.440000	1.463474	13.840700	80266.800000	97543.400000	208.633000	0.403294	1.000000	3.654300

The missing data in the data afterward are summarized in Fig 4. The white spaces in the first plot represent the missing values. The percentage of missing values in the individual features and the fraction of the missing values between any two features are also depicted.

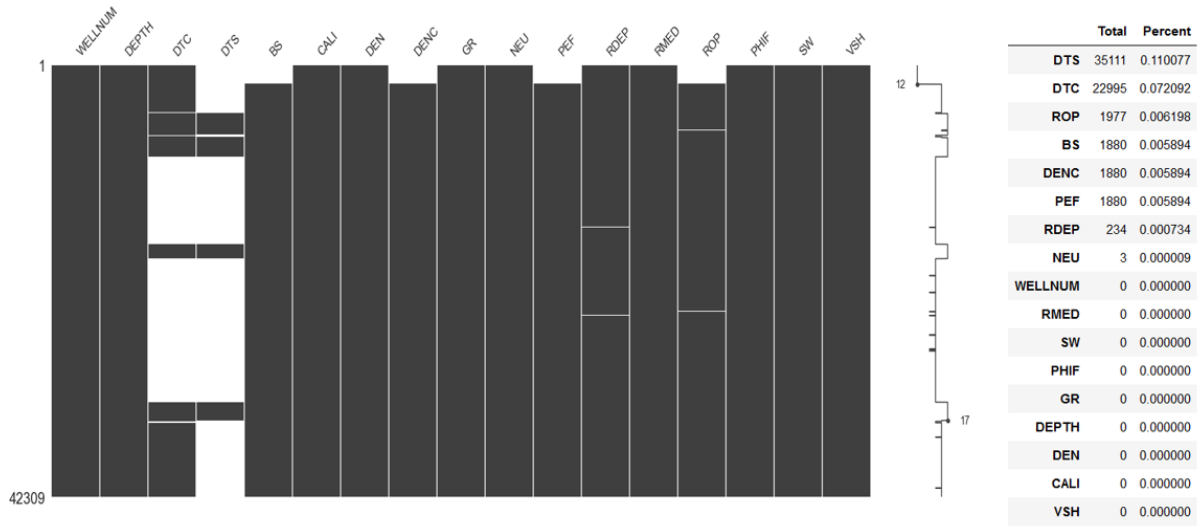


	Total	Percent
DTS	275119	0.862531
VSH	273867	0.858606
PHIF	271653	0.851665
SW	271653	0.851665
DTC	249073	0.780874
PEF	230431	0.722429
DENC	229851	0.720611
NEU	223528	0.700787
DEN	223347	0.700220
CALI	222810	0.698536
ROP	32379	0.101512
BS	31054	0.097358
RDEP	30448	0.095458
RMED	30214	0.094725
GR	3119	0.009778
DEPTH	0	0.000000
WELLNUM	0	0.000000

Fig 4: Missing value plots in rawdata

6.3.Data Cleaning

First, all missing values in the target variables, PHIF, SW, and VSH, were removed. This was done because the originality of the data in these features cannot be changed. The new dataset has a shape of 42309 rows and 17 columns which is good enough to build a model. Also, the statistics of this new dataset showed that they are within depths of interest. A minimum and a maximum depth of 2588.97m to 4744.80m, respectively, with an average of 3732.46m. The data still had missing values, as shown in Fig 5, although it reduced significantly. This dataset was referred to as rawdata1.



	Total	Percent
DTS	35111	0.110077
DTC	22995	0.072092
ROP	1977	0.006198
BS	1880	0.005894
DENC	1880	0.005894
PEF	1880	0.005894
RDEP	234	0.000734
NEU	3	0.000009
WELLNUM	0	0.000000
RMED	0	0.000000
SW	0	0.000000
PHIF	0	0.000000
GR	0	0.000000
DEPTH	0	0.000000
DEN	0	0.000000
CALI	0	0.000000
VSH	0	0.000000

Fig 5: Missing Values in rawdata1

6.4.Handling Missing Data

Two techniques will be explored in handling the missing data.

1. Imputation of missing values
2. The use of algorithms that accommodate missing values.

Imputation Technique: Petroleum well logs and petrophysical data are measured on different scales and are highly variable. This is due to the significant differences in mineralogical constituents of different rock formations encountered during well logging. The use of conventional machine learning imputation techniques like imputing mean or any other measure of central tendency would significantly affect the integrity of the data and model accuracy. In this work, we predicted the missing data using rows within rawdata1 without missing values with a K-Nearest Neighbors (KNN) linear regression model. The method was first used by (Akinnikawe et al., 2018) in generating synthetic Photoelectric and Unconfined compressive strength logs from other wireline logs. The missing data for DTS, DTC, PEF, DENC, and ROP were predicted and imputed as they were significantly large. Multicollinearity between the individual predictors was determined using the variance inflation factor (VIF) method, and a threshold value was determined

based on the range of values of VIF. Generally, all the well logs were highly correlated hence very high VIF values. The defined threshold value was thus higher than the conventional value of 10. All other missing values were removed. The final dataset after imputation and cleaning was called 'rawdata_new' with 42072 rows and 15 columns. Log plots, distribution plots, and boxplots were done for the new data set rawdata_new.

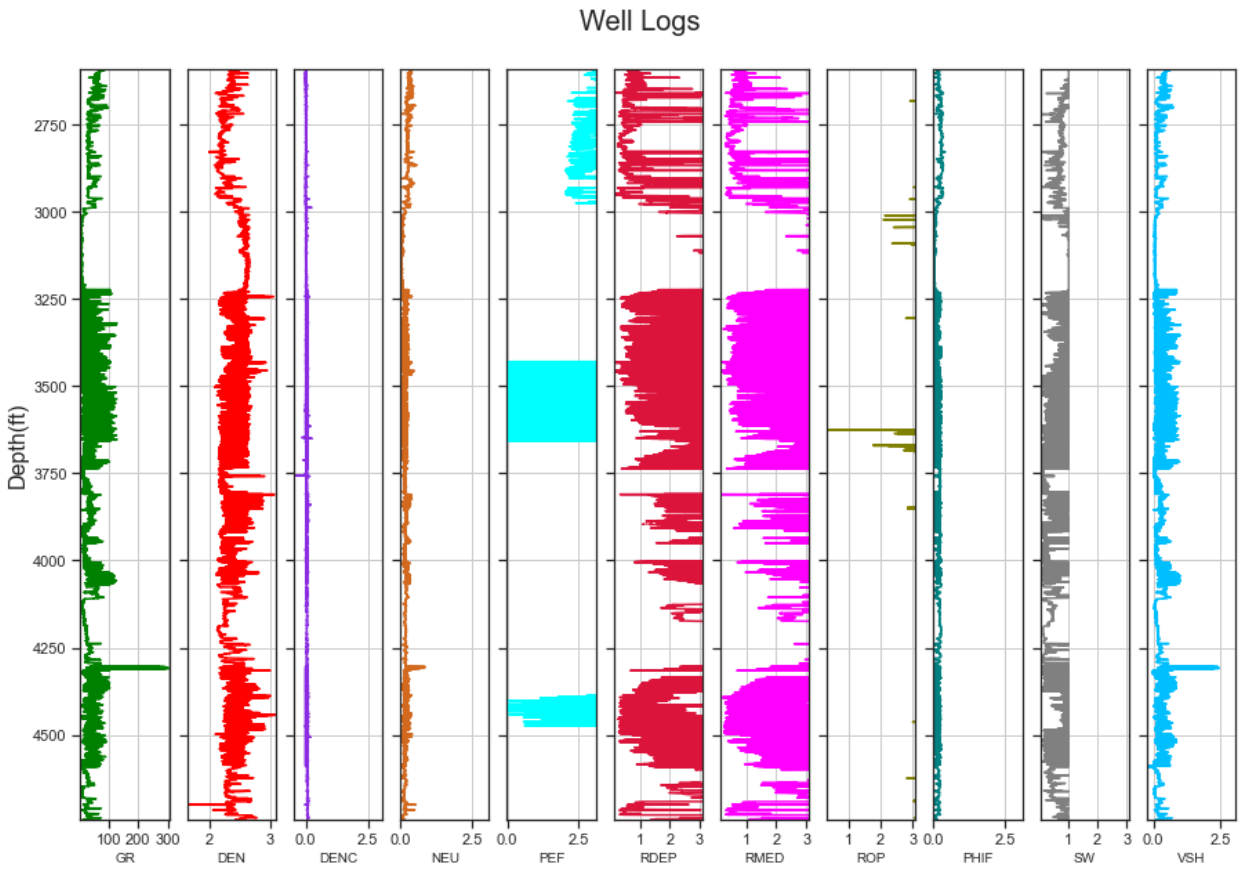


Fig 6. Log plot for rawdata_new

6.5.Handling Outliers

Outliers within the dataset were detected using the isolation forest algorithm, as shown in Fig 8. below. The boxplots also confirmed the presence of outliers. The z-score transform was used to handle the outliers, where all data that fell beyond a threshold value of 2.5 for all the features were removed. The new data set was named "rawd_zscore." Other outlier handling techniques will be explored based on the model's performance. Significant number of outliers within the individual features were removed. The pair plot and boxplot below in figs 9 and 10 show how the data looked after outlier removal.

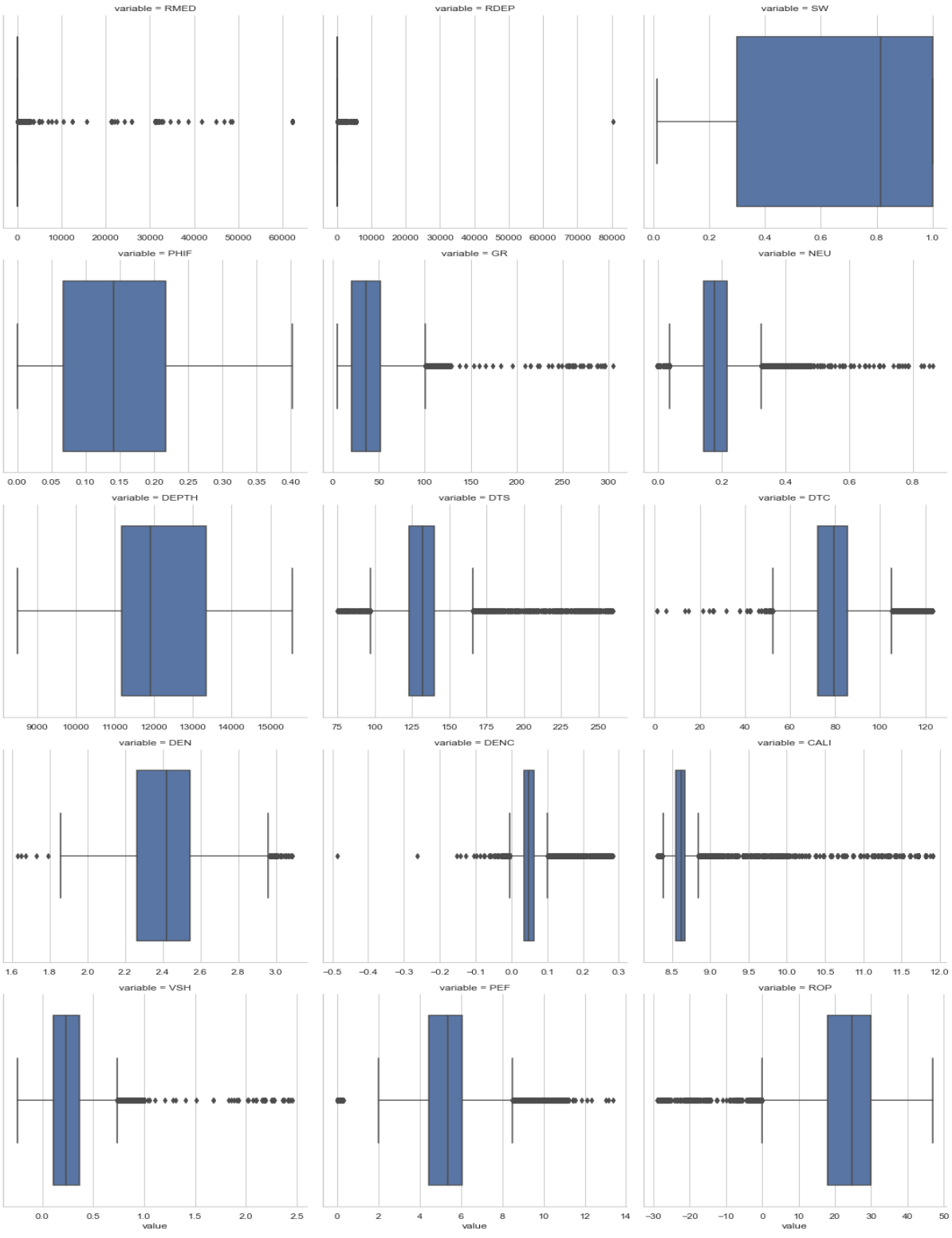


Fig 7. Boxplot for rawdata_new

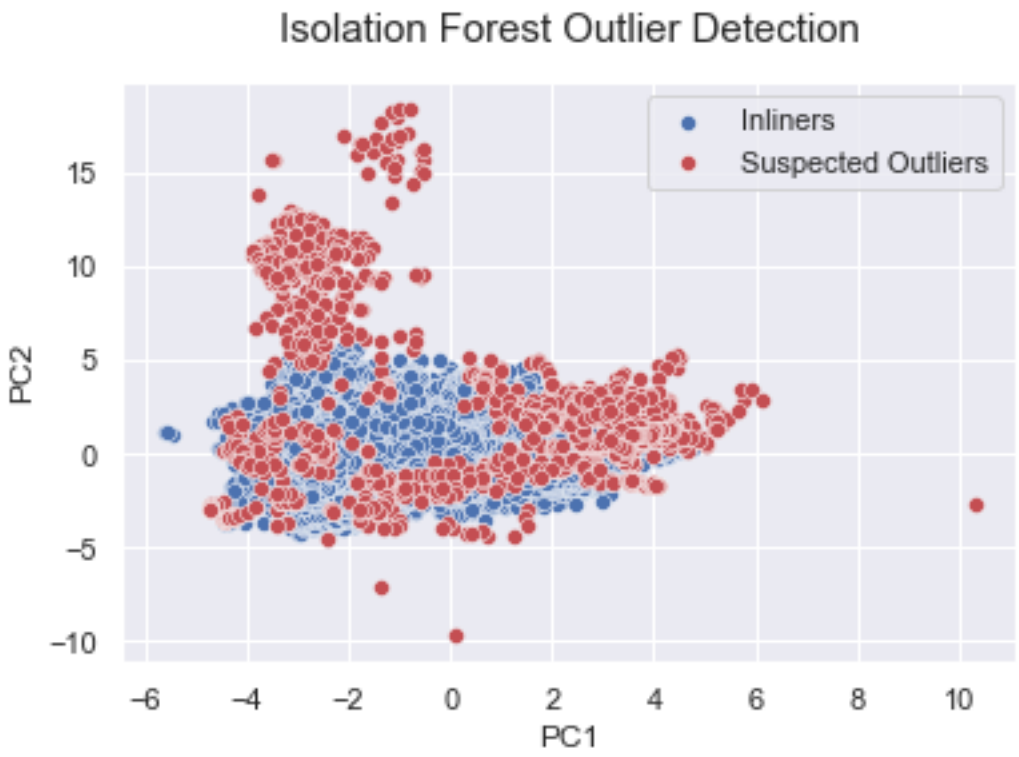


Fig 8: Outlier detection in rawdata_new

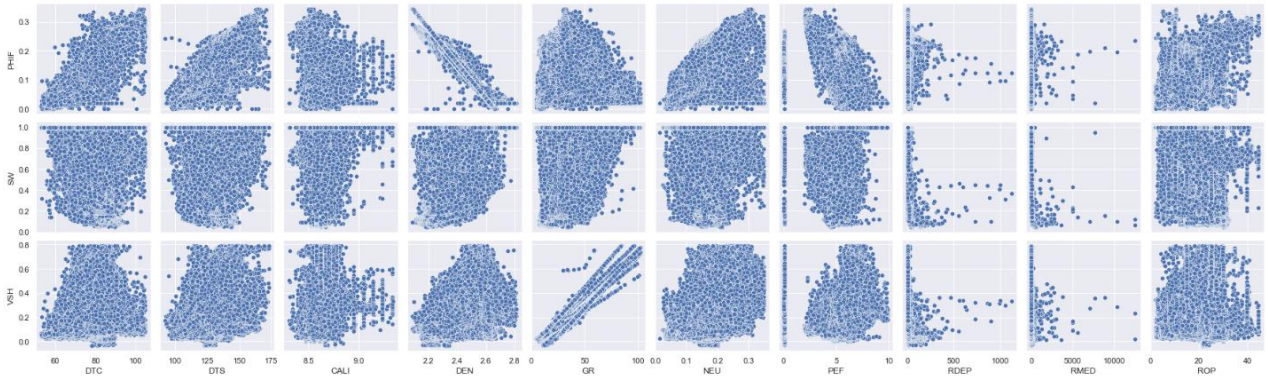


Fig 9: Pairplot after outlier removal

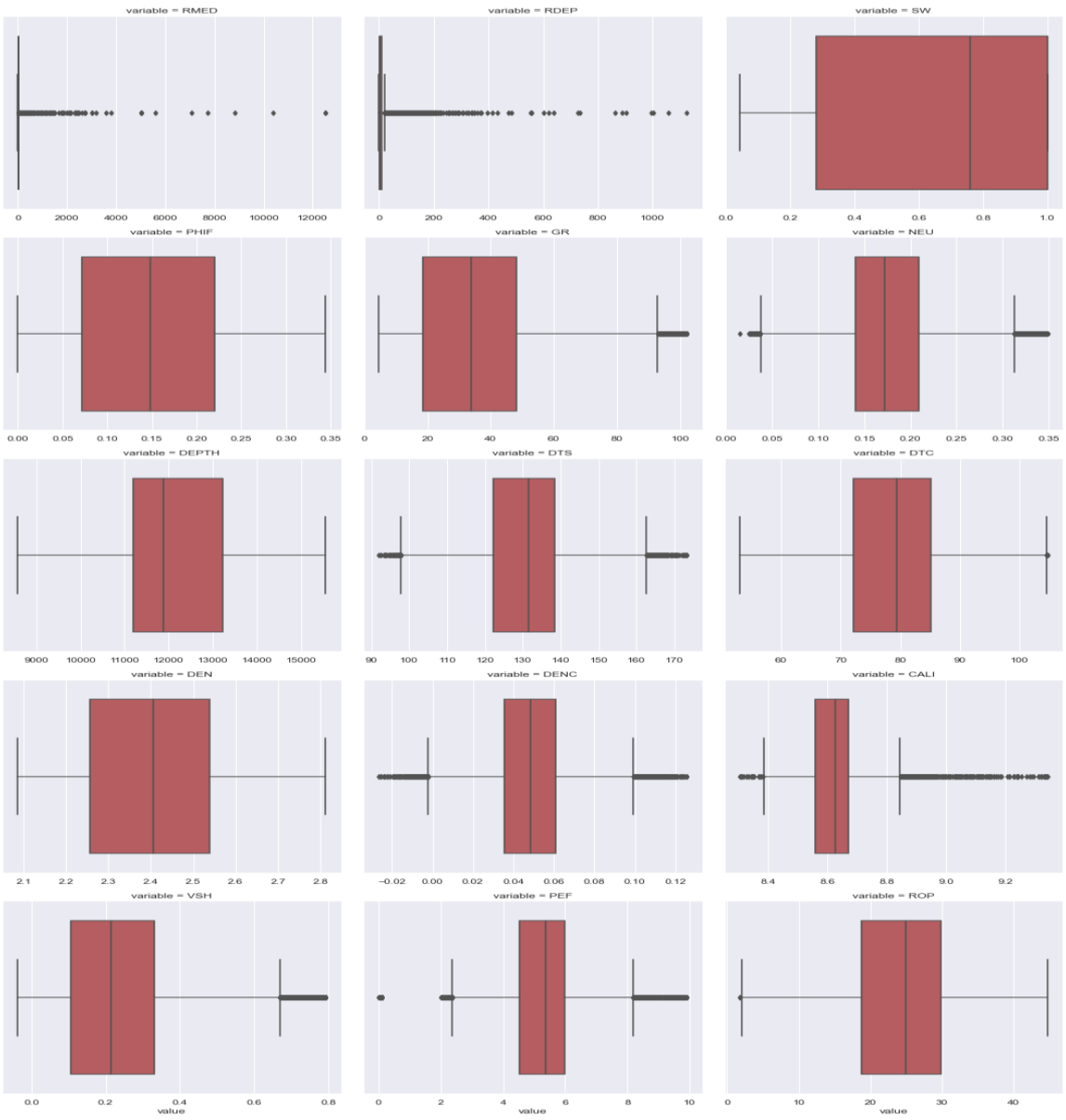


Fig 10: Boxplot after outlier removal

7. Machine Learning Techniques

Both Supervised and unsupervised learning techniques will be employed in this work.

7.1. Unsupervised Learning

K-means and agglomerative clustering would be explored in clustering well logs into electrofacies (zonation). The clusters would be compared with lithological maps from the acquired dataset to ascertain distinct formations and validate the performance of the clustering algorithm.

7.2. Supervised Learning

Regression models would be developed using the following machine learning algorithms; Linear Regression, K-Nearest Neighbors, and Random Forest. Gradient boosting and other optimization techniques would be employed where necessary. A train-test split would be used to divide the training data set for model training and validation. Stratification would be ensured in splitting to avoid bias in data sampling for training and validation. Grid-search cross-validation would be conducted to prevent over-fitting models and ensure optimal model performance. The target variables would be predicted using the test data.

8. Process Validation

As an extra step to validate the performance of the best model above, the predicted target variables would be represented on a log plot and compared with log plots obtained from the equinor data set. The plots will be compared on a depth basis.

9. References

1. Archie, G. E. (1952). Classification of carbonate reservoir rocks and petrophysical considerations. Aapg Bulletin, 36(2), 278-298.
2. Alberty, M. (1992). Standard interpretation: Part 4. wireline methods.
3. Brown, A. A., & Bowers, B. (1958). Porosity determinations from neutron logs. The Petroleum Engineer, 30.
4. Tenchov, G. G. (2016). Porosity evaluation from acoustic log using the theory of mixtures.
5. Moradi, S., Moeini, M., Al-Askari, M. K. G., & Mahvelati, E. H. (2016, October). Determination of shale volume and distribution patterns and effective porosity from well log data based on cross-plot approach for a shaly carbonate gas reservoir. In IOP Conference Series: Earth and Environmental Science (Vol. 44, No. 4, p. 042002). IOP Publishing.
6. Hearst, Joseph R., and Philip H. Nelson. "Well logging for physical properties." (1985).
7. Mondol, Nazmul Haque. "Well logging: Principles, applications, and uncertainties." In Petroleum Geoscience, pp. 385-425. Springer, Berlin, Heidelberg, 2015.
8. Moore, William R., Y. Zee Ma, Jim Urdea, and Tom Bratton. "Uncertainty analysis in well-log and petrophysical interpretations." (2011): 17-28.
9. Wang, Jun, Junxing Cao, Jiachun You, Ming Cheng, and Peng Zhou. "A method for well log data generation based on a spatio-temporal neural network." Journal of Geophysics and Engineering 18, no. 5 (2021): 700-711.
10. Equinor Volve Data set; https://www.equinor.com/energy/volve-data-sharing

11. American Association of Petroleum Geologists.(2022). Density-Neutron Log Porosity. https://wiki.aapg.org/Density-neutron_log_porosity
12. Sen, S., & Ganguli, S. S. (2019, April). Estimation of pore pressure and fracture gradient in Volve field, Norwegian North Sea. In SPE oil and gas india conference and exhibition. OnePetro.
13. Gupta, I., Tran, N., Devegowda, D., Jayaram, V., Rai, C., Sondergeld, C., & Karami, H. (2020). Looking ahead of the bit using surface drilling and petrophysical data: Machine-learning-based real-time geosteering in Volve field. SPE Journal, 25(02), 990-1006.