Geosciences and Engineering, Vol. 12, No. 1 (2024), pp. 75–87.
https://doi.org/10.33030/geosciences.2024.01.005

EXPLORING UNCERTAINTY IN FLOW UNIT IDENTIFICATION AND PERMEABILITY PREDICTION

FRANKLIN GÓMEZ[1*], MARIANNA VADÁSZI[2]
[1*]*Institute of Mining and Energy, University of Miskolc, Hungary;*
gomez.soto.franklin.vinicio@student.uni-miskolc.hu
[1*]*Petroleum Department, Faculty of Geology and Petroleum Engineering,*
Escuela Politécnica Nacional, Ecuador
[2]*Institute of Mining and Energy, University of Miskolc, Hungary;*
marianna.vadaszi@uni-miskolc.hu
[1]https://orcid.org/0000-0003-4367-4972
[2]https://orcid.org/0000-0003-0649-311X

Abstract: The study proposes a comparative uncertainty analysis of the main methods for permeability prediction or estimation, including the Cluster analysis (K-means), the Kozeny-Carman (KyC) equation for flow unit identification, and the K-nearest neighbor Density Estimate (KNN) algorithm, Kozeny-Carman equation, and One Flow Unit (OFU) for permeability prediction or estimation. The proposed analysis is applied to 13 wells in the Sacha field located in the Amazon region of Ecuador, targeting the Hollin and Napo formations, which mainly consist of sandstone, limestone, and shale. The selected wells have a sufficient number of laboratory measurements of permeability and electrical logs of porosity, permeability, natural gamma ray, medium, and deep resistivity. Initially, the K-means clustering and KyC methods are applied to identify the flow units, followed by a regression process to calculate the permeability using the KNN, KyC, and OFU methods. During the clustering process, the KyC method yielded better results, with the experimental data exhibiting uncertainties of less than ±35 mD, except in the outlier flow unit with an average porosity of 16.86% ±3.87% (Flow Unit D) whose average permeability is 407.52 mD and uncertainty of ±504.10 mD. For software simulation purposes, it is recommended to utilize the KyC method, as it employs basic concepts and equations in accordance with hydraulic principles.

Keywords: *uncertainty, permeability prediction, Sacha, flow units, K-nearest neighbor density estimate algorithm, cluster analysis*

1. INTRODUCTION

The process of modeling an underground reservoir requires the permeability property determination, but this property can be measured by pressure test or directly by core samples (Remeczki et al., 2020, Fanchi, 2018). This determination of the permeability is only developed in small portions of the reservoir. Therefore, the prediction process of the permeability is essential. The results obtained from permeability values vary widely between the laboratory data and the predictions of the different methods, however these predictions are used for modeling. Why are these values used in modeling despite their relatively wide difference? What is the method that provides the best results? It is very important to know the uncertainties involved in each of the values of the methods used to predict permeability, since in this way the most appropriate method can be used.

The uncertainty in any measurement process has three components (Berg et al., 2021): the variation between the sample and the same rock type with heterogeneity, experimental uncertainty during the measurement, and the interpretation process related to the model application. Few research evaluates and gives directly the values of permeability (Johnston and Beeson, 1944) or others evaluate the uncertainty of the relative permeability (Berg et al., 2021, Mathias et al., 2013) but not the sample data permeability.

Normally, laboratory measurements are not carried out in all wells and at all depths, so it is decided to use well logs to obtain permeability values and identify flow units through different processes. In some cases, the Kozeny-Carman equation is used to discretize the porosity values and based on a certain number of laboratory measurements of permeability, identify the flow units, and obtain the permeability by regression for the other zones and wells. For instance, (Belhouchet and Benzagouta, 2019) use the Kozeny–Carman equation with the DRT (Discrete Rock Typing method) for predict the permeability in Algerian B-H oil field reservoir. This reservoir is composed of sandstone, limestone and dolomites and the results are based in the Correlation Coefficient R to determine if the estimation is right or wrong. However, they never expo the data of the permeability regression. Abbaszadeh et al. (1996), Amaefule et al. (1993) and Perez et al. (2005) use the same equations with different variations but with the main similar idea.

On the other side, Aminian et al. (2003) proposes the use of neural networks to base on training data predict the permeability and identify the flow units. It is similar to using the K-nearest-neighbor Density Estimate (KNN) which needs training data. In the case of study of the Norcan East Field (Bhattacharya et al., 2008) the methodology offers a complete process which includes validation based on the water saturation and capillary pressure which appear in the transition's zones. This methodology is more complete but never calculates the uncertainty of the permeability values. Data from 13 wells were used in this research, which are in total 428 samples measured its porosity and permeability. Three methods are proposed to identify the flow units, one based in the size of porous that we call Kozeny–Carman (KyC) method, other is the cluster analysis algorithm called K-means which identified centroids and classified the data according to the nearest points [Szabó et al. (2019), Ali, Sheng-Chang (2020)]. In addition to, on this research the called One Flow Unit (OFU) method which as its name implies all the data belongs to one flow unit, it is because in the process of modeling it is always used and certainly sometimes It gets the best history match with the real fluid production (Krause et al., 2009).

Permeability was estimated using the Kozeny–Carman equations with DRT (Discrete Rock Typing) (Belhouchet and Benzagouta, 2019; Amaefule et al., 1993) but also uses the K-nearest-neighbor Density Estimate (KNN) method (Gómez et al., 2022) which is based on a test data for learning and estimate the permeability. The third method uses exponential matching in One Flow Unit (OFU).

In addition, the uncertainty for the real sample data is determined with *Equations (14), (15)* and *(16)*. And for the estimation of the permeability the equation used will be determined with exponential matching (Papadopoulos and Yeung, 2001). However, the calculation of the error develops using the *Equation (17)*. The results show that the error and the uncertainty do not have any similarity and the uncertainty is always present.

2. METHODOLOGY

Permeability and porosity data are measured on 428 samples obtained from 13 wells in the Sacha field located in northeastern Ecuador (Baby et al., 2014). This data was matched with the well logs of natural gamma ray, superficial, intermediate, and deep resistivity, neutron porosity, gamma-gamma density and spontaneous potential. The methodology is divided into three stages:

- Flow unit identification,
- Permeability estimation,
- Uncertainty and error of permeability determination.

2.1. Flow Unit Identification

Using the Kozeny-Carman methodology four flow units were identified: A, B, C and D. The equations used are (Amaefule et al. 1993):

$$\theta_z = \frac{\theta}{1-\theta} \tag{1}$$

$$RQI = 0.0314\frac{K}{\theta} \tag{2}$$

$$FZI = \frac{RQI}{\theta_z} \tag{3}$$

$$DRT = 2Log_e FZI + C \tag{4}$$

where Θ_z is normalized porosity in volume fraction, RQI is Rock quality index, K is permeability in mD, Θ is porosity in volume fraction, FZI is Flow zone indicator, DRT is Discrete rock typing. According to the theory in a graphic Log RQI vs Log Θ_z must draw parallels lines to differentiate the flow units (see *Figure 1*).

Discrete rock typing (DRT) gives integer numbers, in this case ranging from 6 to 27, and they are organized into larger groups called flow units. In this case, using MATLAB, the grouping of the values was easier because the parallel lines in *Figure 1* must have the same slope and the numbers are in order. In Table 1 are the numbers that are related to the flow units A, B, C, D.

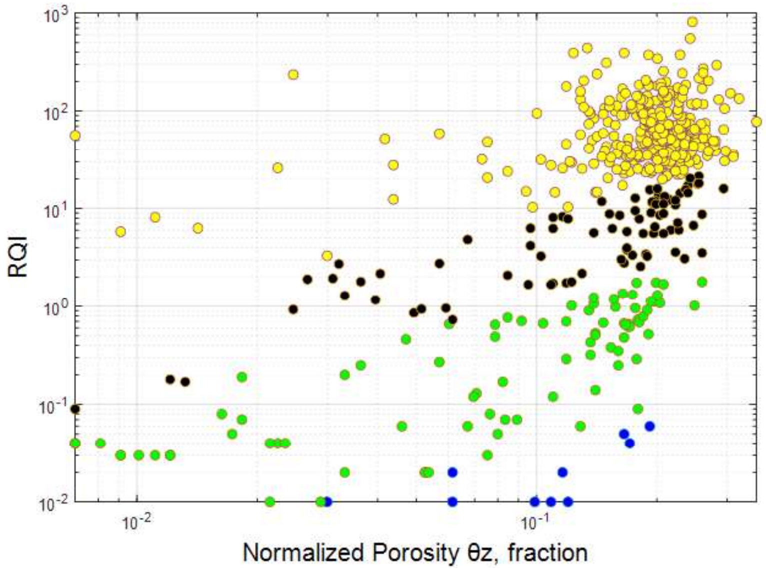

Figure 1

Log RQI vs log Θz of 428 samples of the Sacha Field. Blue, green, black and yellow represent the corresponding flow units A, B, C, D identified

The other method used to identify the flow units was using the Cluster Analysis methodology. K-means cluster analysis is a simple unsupervised statistical method that orders the objects of a multivariate dataset into groups (flow units) using the information of similarities given by metric distance (see *Figure 2*). The method is sensitive to the scale difference between variables, so normalization of the data set is required. Then, each object is designated into a non-overlapping group of great homogeneity and large differences from other groups. For further explanation see Szabó et al. (2019) and Ali and Sheng-Chang (2020). Mathematically, the method is expressed by the following equation:

$$J = \sum_{j=1}^{k} \sum_{i=i}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \tag{5}$$

where J is objective function, x_i is the i-th analyzed object, i= 1,…, n, c_j is the j-th cluster centroid, j= 1,.., k, K is optimal number of clusters. The objective function converges at the minimum sums of square deviation of objects x_i, from the cluster centroid c_j.

The "City Block" distance metric is incorporated for this study. Here, each cluster centroid is the component-wise median of the points in the cluster

$$D_1 = \sum_{k=1}^{N} \left| x_k^{(i)} - x_k^{(j)} \right| \tag{6}$$

where D is sum of lengths between points, $x_k^{(i)}$ is distance of x_k from the centroid in component i, $x_k^{(j)}$ is distance of x_k from the centroid in component j. In this case, the four flow units identified were named 1, 2, 3 and 4. The data classification can be seen on *Figure 2*:

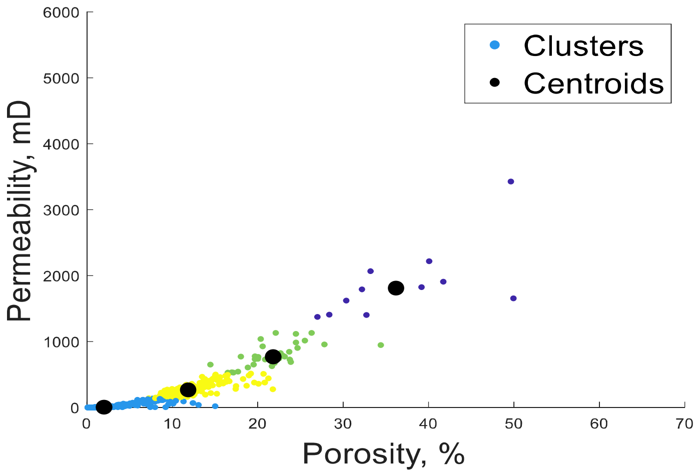

Figure 2

Permeability and porosity of 428 samples classified by K-means clustering algorithm. Light blue, yellow, green, and purple are the flow units 1, 2, 3, and 4 respectively. Black points are the centroids

Finally, the One Flow Unit (OFU) method to complete permeability data is presented in *Figure 3*.

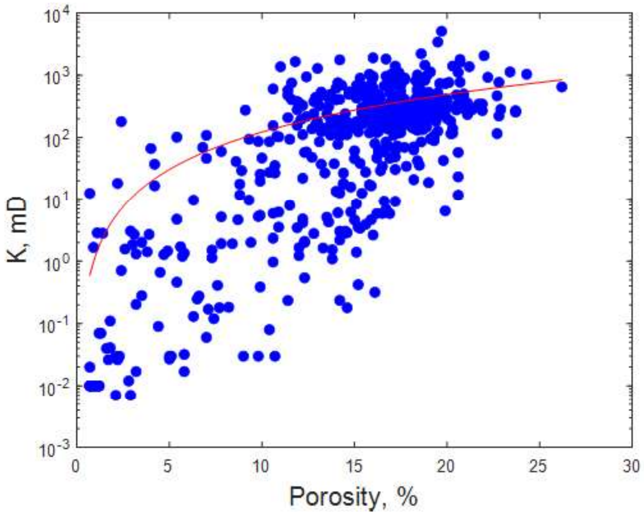

Figure 3

Logarithm of permeability vs. porosity (%) of the 428 samples.
Red line represents the exponential regression to predict the permeability

2.2. Permeability estimation

The permeability was estimated first using an exponential regression for each flow unit A, B, C and D. The empirical equations are given below

$$K_A = 4.804 * 10^{-6}\theta^{3.994} \tag{7}$$

$$K_B = 0.00179\theta^{2.862} \tag{8}$$

$$K_C = 0.01604\theta^{2.849} \tag{9}$$

$$K_D = 1.521\theta^{1.959} \tag{10}$$

The second method used for estimating the permeability was the K-nearest neighbor density estimate (KNN) which is a non-parametric method of estimating a probability density function. The algorithm estimates a function that predicts the rock type(z) according to the log-well registered values. Every interpreted rock category x is a p-dimensional random variable X. This means the interpretation of the rock permeability will depend on the pre-established rock type z. The d(x,z) represents the Euclidean distance between x and z. X is an example of z, consequently x is the permeability measured and matched with the logs before the prediction (Hu et al., 2008, Mitra et al., 2002). The hypersphere of radius r about z is designated by *Equation 11*

$$A_{r,z} = \{X \setminus d(x,z) \leq r\} \tag{11}$$

where $A_{r,z}$ is volume of the hypersphere, r is radius of the hypersphere, x isa categorical class, and X is a variable with p dimensions. Then *Equation 12* defines the density function:

$$f_N(z) = \frac{k(N)}{N} * \frac{1}{A_{rk(N),z}} \tag{12}$$

where $f_N(z)$ is a function f to estimate z with N, k(N) is a sequence of positive integers from x_1 to x_N, x is a rock type interpreted in a set of data. By the other hand, the OFU method determine the exponential equation. See *Figure 3*:

$$K_T = 0.97330\theta^{2.825} \tag{13}$$

2.3. Uncertainty and error of permeability determination

Here, the real sample data was measured in poropermeameter equipment whose precision is 0.1% for the porosity and 0.01 mD for the permeability. The process of divide the permeability in flow units suppose that in the modeling process the same equation will be used for all the data that belongs to this flow unit. It is similar to the up-scaling where a group of data of wells is assigned to a cell of a grid. For this reason, the *Equation 14* is used for determining the uncertainty in the experimental data. The *Equation 15* is used to determine the uncertainty when a regression process is developed. The *Equation 16* is the porosity uncertainty necessary for the calculation of the permeability uncertainty (JCGM member organizations, 2008)

$$u(k) = \sqrt[2]{P_K^2 + S_K^2} \tag{14}$$

$$u(k) = \frac{\partial K}{\partial \theta} u(\theta) \tag{15}$$

$$u(\theta) = \sqrt[2]{P_\theta^2 + S_\theta^2} \tag{16}$$

where u is the uncertainty of the permeability and porosity, P is a precision of the poropermeameter equipment for permeability and porosity, S is the Standard Deviation for porosity and permeability. The calculation of the error was also developed using the equation:

$$\%E = \frac{RV - EV}{RV} * 100 \tag{17}$$

where %E is the data estimation Error with respect to the sample measured, RV is the real value with refer to the sample measured, and EV is the estimated value with the different proposed methods.

3. RESULTS

The results of the flow units and the permeability estimation according the KyC methodology is presented on the *Figure 4, Table 1* and *Table 2*. The θ_{mean} is the mean porosity in the flow unit, k_corr is the measured permeability, K_NN is the permeability estimated using the KNN method, E_KNN is the error between the measured and the estimated permeability, K_KyC is the permeability in the flow unit and E_KyC is the error. In the calculation of the errors the equation used was *Equation 17*.

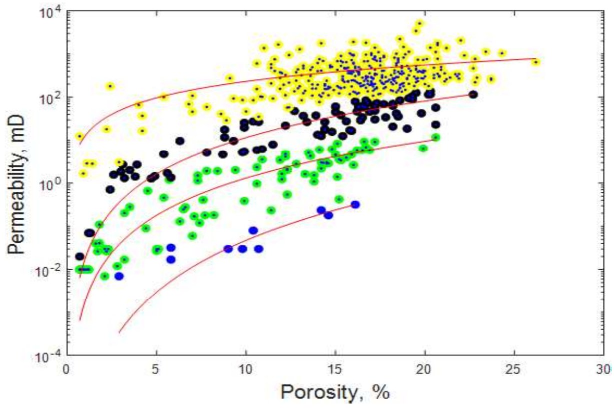

Figure 4

Logarithm of permeability vs. porosity (%) of the 428 samples divided in four flow units. Red lines represent the exponential regression to predict the permeability

Table 1

Flow units A, B, C, and D determined using the KyC method with the DRT (discrete rock typing) with porosity (Θ_{mean}), permeability (k_corr) measured in laboratory, and permeability calculated with KNN(K_NN) and KyC (K_KyC)

DRT	FLOW UNIT	NUMBER ELEMENTS	Θ_{mean} %	k_corr mD	K_NN mD	E_KNN	K_KyC mD	E_KyC
6	A	3.00	9.83	0.03	0.03	0.00%	0.07	131.88%
7		2.00	8.10	0.05	0.05	0.00%	0.05	4.28%
8		5.00	10.72	0.15	0.15	0.00%	0.28	81.82%
9	B	8.00	6.70	0.10	0.10	0.00%	1.36	1220.58%
10		5.00	6.56	0.13	0.13	0.00%	0.58	356.38%
11		5.00	9.50	0.46	0.46	0.00%	2.18	377.81%
12		13.00	7.27	0.60	0.60	0.00%	1.79	198.15%
13		19.00	10.35	2.04	2.04	0.00%	4.44	117.34%
14		21.00	9.50	2.88	2.88	0.00%	4.15	43.97%
15		15.00	10.81	4.28	4.28	0.00%	3.67	14.29%
16	C	21.00	10.41	8.42	8.42	0.03%	38.74	360.08%
17		11.00	14.59	23.46	19.83	15.46%	69.42	195.93%
18		16.00	13.09	30.12	30.12	0.00%	58.20	93.23%
19		33.00	14.78	62.62	62.62	0.00%	79.05	26.25%
20	D	41.00	16.95	142.72	141.69	0.72%	616.22	331.77%
21		79.00	17.69	230.72	227.98	1.19%	647.59	180.68%
22		66.00	16.91	327.78	327.78	0.00%	585.22	78.54%
23		29.00	16.39	551.21	545.49	1.04%	588.57	6.78%
24		26.00	15.07	770.47	770.47	0.00%	514.63	33.21%
25		6.00	14.85	1412.19	1412.19	0.00%	525.87	62.76%
26		2.00	17.75	2665.86	2665.86	0.00%	636.04	76.14%
27		2.00	15.75	3377.07	3377.07	0.00%	523.76	84.49%

In *Table 2* $u_{R\theta}$ is the porosity uncertainty in % determined using *Equation 16*, u_{RK} is the uncertainty in the permeability determined using *Equation 14* with the S_k of each flow unit of KyC, u_{KNN} is the uncertainty in the KNN permeability determined with *Equation 14* and the $u_{R\theta}$ for each KyC flow unit, U_{CK} is the uncertainty in the KyC permeability determined with *Equation 15* and $u_{R\theta}$ for each KyC flow unit.

Table 2

Uncertainty and Error values in the KyC method for flow units ($u_{R\Theta}$ Equation 16, u_{RK} Equation 14, U_{KNN} Equation 14, u_{CK} Equation 15 and 16) in KNN and KyC methods for permeability estimation

FLOW UNIT	# SAMPLES	Θ_{mean} %	$u_{R\Theta}$ %	K_R mD	u_{RK} mD	KNN _mD	U_{KNN} mD	E_KNN	K_kyC_mD	u_{CK} mD	E_KC
A	10	9.93	±4.26	0.10	±0.11	0.10	±0.15	0.00%	0.17	±0.08	77.81%
B	86	9.15	±5.65	2.04	±2.56	2.04	±2.56	0.00%	3.19	±1.78	56.77%
C	81	13.29	±5.66	36.83	±34.74	36.34	±34.94	1.34%	63.17	±30.90	71.54%
D	251	16.86	±3.87	407.52	±504.10	405.83	±504.68	0.42%	601.49	±173.21	47.60%

Table 3 presents the flow units according the K-means clustering method: $u_{R\theta}$ is the porosity uncertainty in % determined using *Equation 16*, u_{RK} is the uncertainty in the permeability determined using *Equation 14* with the Sk of each flow unit of K-means, u_{KNN} is the uncertainty in the KNN permeability determined with *Equation 14* and the $u_{R\theta}$ for each K-means flow unit, U_{KyC} is the uncertainty in the KyC permeability determined with *Equation 15* and $u_{R\theta}$ for each K-means flow unit.

Table 3

Uncertainty and error values in the K-means clustering method for flow units ($u_{R\Theta}$ Equation 16, U_{RK} Equation 14, U_{KNN} Equation 14, U_{KyC} Equation 15 and 16) in KNN and KyC methods for permeability estimation

FLOW UNIT	# SAMPLES	Θ_{mean} %	$u_{R\Theta}$ %	K_R mD	U_{RK} mD	KNN_mD	U_{KNN} mD	EKNN	KkyC Md	U_{KyC} mD	EKyC
	173	14.98	±4.98	233.03	±514.99	231.55	±515.26	0.64%	368.43	±38.31	58.10%
	79	13.00	±6.57	199.37	±270.59	198.83	±270.69	0.27%	287.29	±31.07	44.10%
	134	15.20	±4.93	254.43	±341.91	254.43	±341.91	0.00%	396.75	±59.86	55.94%
	42	12.80	±7.16	364.01	±535.80	360.05	±535.90	0.01	399.22	±435.10	9.67%

In *Table 4* we present the uncertainty and error with One Flow Unit method.

Table 4

Uncertainty and Error values in the OFU method for permeability estimation

K_OFU mD	u_{UFO} mD	E_KOFU
438.70	±349.59	78.73%

4. DISCUSSION

Unquestionably in this case, the KNN method gives the best result taking account the %Error (E_KNN) (*Table 1* and *2*) for both grouping methods, it attributes because KNN uses test data for estimate the permeability. If the permeability property is needed the KNN is the best option. However, it recommends using KyC method in modeling process. For modeling the up scaling Fanchi (2018) means grouping permeability for assign values or equations to a complete cell or cells. In his publication, Wang (2018) arrives at a similar conclusion regarding the identification of flow units for lithological purposes. In this work, Wang writes, “KNN (K-Nearest Neighbors) clustering in machine learning is a very efficient clustering method applied in lithology identification, reservoir type recognition, flow unit classification, and so on. However, it is not consistently effective due to its inherent limitations, such as its initial center selection and clustering center shift caused by outliers.” Wang states that while the KNN method is generally efficient, it encounters problems with clustering, which he attributes to issues with the initial selection center (Wang et al., 2018). In the publication by Silva (2019) titled Petrofacies Classification using Machine Learning Algorithms, the precision for predicting petrofacies using KNN is 92% for homogeneous layers, while for heterogeneous layers, it is reduced to 85%. In this research, they utilize an 80/20 split, meaning 20% of the data is used for testing, and 80% is for evaluation. The challenge in petrofacies identification lies in classifying heterogeneous layers. Permeability is a crucial aspect of petrofacies. This confirms the assertion because the data grouping directly addresses the heterogeneity of layers (Silva et al., 2020).

In this investigation % Error is calculated by comparing the measured data with the estimated data, and it is an individual comparison between two values, while the uncertainty first depends on the precision of measuring equipment and after of the clustering process. The standard deviation is an important part of the determination of uncertainty.

In the application of the KyC for identifying flow units, the suitable value of porosity varies from 9.15 ±5.65% to 16.86 ±3.87%. While the experimental data of permeability has a suitable value from 0.1 ±0.11 mD to 407.52 ±504.10 mD. Like, the estimated permeability varies from 0.10 ±0.15 mD to 405.83 ±504.68 mD with the KNN method. Comparable, in the KyC method the permeability varies from 0.17 ±0.08 mD to 601.49 ±173.21 mD. Certainly, the permeability has a huge uncertainty, and it increases when is tried to estimate with different methods.

By the other hand, in the application of the K-means method for identify flow units the suitable value of porosity varies from 12.80 ±7.16% to 15.20 ±4.93%. While the experimental data of permeability has a suitable value from 199.37 ±270.59 mD to 364.01 ±535.80 mD. Like, the estimated permeability varies from 198.83 ±270.69 mD to 360.05 ±535.90 mD with the KNN method. Comparable, in the KyC method the permeability varies from 287.29 ±31.07 mD to 399.22 ±435.10 mD. Here, the uncertainty is less in the KyC method than the KNN method. However, the % Error is higher for every mean permeability. Consequently, the KNN is better in individual analysis but in grouped data the KNN has higher uncertainty.

In the application of the OFU method for identifying flow units the suitable value of permeability is 438.70 ±349.59 mD with an error of 78.73%. If we compare this value with the other methods this value still inside the limits of the permeability measured or estimated for them. It means that the use of one flow unit in modeling is not a bad option considering that the uncertainty of permeability with the other methods has a huge variation.

5. CONCLUSIONS

According to our study, the best method to use in modeling process is KyC for flow units identification and permeability estimation. The uncertainty in *Table 2* and *3*, u_{KNN} (uncertainty in KNN method) and u_{CK} (uncertainty in method KyC) show smaller values in KyC method. The best method for determining the permeability is the KNN because the error is almost zero in all the flow unit classifications.

The flow unit D has uncertainty higher than the estimated value except in the KyC method, it is because the D flow unit is the one with bigger porous size and permeability measures. The KyC method always uses the size of the porous to grouping the data.

The KyC method uses the size of the porous-like base for dividing the data while K-means uses only mathematics for dividing. However, using K-means is disquieting to clustering the data of natural gamma ray or resistivity logs to identify some match with the groups of permeability with porosity. If a good correlation is found, is possible to develop equations for predicting permeability with log data. Certainly, it is recommended to use K-means in K vs. natural gamma ray or K vs. Resistivity or K vs. density and compare with K-means of K vs. porosity.

ACKNOWLEDGMENTS

The author, Franklin Gómez, would like to express his deepest gratitude to National Polytechnic School of Ecuador for their sponsorship and support as he embarks on his doctoral studies at the University of Miskolc. Furthermore, Franklin would like to extend his appreciation to the University of Miskolc and the Hungarian people who contribute their taxes to make the Hungarian Stipendium scholarship possible. Lastly, Franklin would like to thank Pierre Kuhmmert and E.P. Petroecuador for their collaboration in sharing valuable data. The methodology of the research was made in the frame of Project no. RRF-2.3.1-21-2022-00009, titled National Laboratory for Renewable Energy has been implemented with the support provided by the Recovery and Resilience Facility of the European Union within the framework of Program Széchenyi Plan Plus. The simulation process and the results were funded by the Sustainable Development and Technologies National Program of the Hungarian Academy of Sciences.

REFERENCES

Abbaszadeh, M., Fujii, H., Fujimoto, F. (1996). Permeability prediction by hydraulic flow units theory and applications. *Society of Petroleum Engineers*, 11 (4), pp. 263–271. Retrieved from https://doi.org/10.2118/30158-PA

Ali, A., Sheng-Chang, C. (2020). Characterization of well logs using k-mean cluster analysis. *Journal of Petroleum Exploration and Production Technology*, 10 (6), pp. 2245–2256. Retrieved from https://doi.org/10.1007/s13202-020-00895-4

Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., Keelan, D. K. (1993). Enhaced reservoir description: using core and log data to identify hydraulic(flow) units and predict permeability in uncored intervals/wells. *Software: Practice and Experience.* pp. 205–220. Retrieved from https://api.semanticscholar.org/CorpusID:54908443.

Aminian, K., Ameri, S., Oyerokun, A., Thomas, B. (2003). Prediction of flow units and permeability using an artificial neural networks. *SPE West. Reg. Pacific Sect.*, pp. 299–305. Retrieved from https://api.semanticscholar.org/CorpusID:140543065.

Baby, P., Rivadeneira, M., Roberto, B. (2014). *La cuenca oriente: geologia y petroleo*. Institut français d'études andines.
Retrieved from https://doi.org/10.4000/books.ifea.2984

Belhouchet, H. E., Benzagouta, M. E. (2019). Rock typing: reservoir permeability calculations using discrete rock typing methods (DRT): case study from the Algerian B-H oil field reservoir. *Advances in Petroleum Engineering and Petroleum Geochemistry*. Retrieved from https://api.semanticscholar.org/CorpusID:134624668.

Berg, S., Unsal, E., Dijk, H. (2021). Non-uniqueness and uncertainty quatification of relative permeability measurements by inverse modelling. *Computers and Geotechnics*, 132. Retrieved from https://doi.org/10.1016/j.compgeo.2020.103964

Bhattacharya, S., Byrnes, A. P., Watney, W. L., Doveton, J. H. (2008). Flow unit modeling and fine-scale predicted permeability validation in Atokan sandstones: Norcan East Kansas. *AAPG Bulletin*, 92 (6), pp. 709–732. Retrieved from https://doi.org/10.1306/01140807081

Fanchi, J. R. (2018). *Principles of Simulation*. Fourth ed. Texas, Elsevier.

Gómez, F. V., Flores, Y., Vadászi, M. (2022). Comparative analysis of the k-nearest-neighbour method and k-means cluster analysis for lithological interpretation of well logs of the shushufindi oilfield, Ecuador. *Rudarsko-geolosko-naftni zbornik*, 37 (4), pp. 155–165. Retrieved from http://dx.doi.org/10.17794/rgn.2022.4.13

Hu, Q., Yu, D., Xie, Z. (2008). Neighborhood classifiers. *Expert Syst. Appl*., 34 (2), pp. 866–876. Retrieved from https://api.semanticscholar.org/CorpusID:11609344.

JCGM member organizations (2008). Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" Propagation of distributions using a Monte Carlo method. In: JCGM organizations: *Evaluation of Measurement data*. pp. 1–80.

Johnston, B. N., Beeson, C. M. (1944). Water permeability of reservoir sands. *Society of Petroleum Engineer Journal*, 160, pp. 43–55. Retrieved from https://doi.org/10.2118/945043-G

Krause, M., Perrin, J. C., Benson, S. M. (2009). Modeling permeability distributions in a sandstone core for history matching coreflood experiments. *SPE International Conference CO2 Capture Storage, and Utilization*, pp. 25–38. Retrieved from http://dx.doi.org/10.2118/126340-PA

Mathias, S. A., Gluyas, J. G., Martinez, G. J., Bryant, S. L., Wilson, D. (2013). On relative permeability data uncertainty and CO2 injectivity estimation for brine aquifers. *International Journal of Greenhouse Gas Control*, 12, pp. 200–212. Retrieved from https://doi.org/10.1016/j.ijggc.2012.09.017

Mitra, P., Murthy, C. A., Pal, S. K. (2002). Density-based multiscale data condensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (6), pp. 734–747. Retrieved from https://doi.org/10.1109/TPAMI.2002.1008381

Papadopoulos, C. E., Yeung, H. (2001). Uncertainty estimation and monte carlo simulation method. *Flow Measurement and Instrumentation*, 12 (4), pp. 291–298. Retrieved from https://doi.org/10.1016/S0955-5986(01)00015-2

Perez, H. H., Datta-Gupta, A., Mishra, S. (2005). The role of electrofacies, lithofacies, and hydraulic flow units in permeability prediction from well logs: a comparative analysis using classification trees. *SPE Reservoir Evaluation & Engineering*, 8 (2), pp. 143–155. Retrieved from https://doi.org/10.2118/84301-PA

Remeczki, F., Szabó, P. N., Dobróka, M. (2020). Flow rate and permeability determination in rock samples from unconventional reservoirs to support the geophysical inversion model. *Geosciences and Engineering*, 8, pp. 154–166.

Silva, A. A., Tavares, M. W., Carrasquilla, A., Misságia, R., Ceia, M. (2020). Petrofacies classification using machine learning algorithm. *Geophysics*, 85, pp. 101–113. Retrieved from https://doi.org/10.1190/geo2019-0439.1

Szabó, N. P., Nehez, K., Hornyák, O., Piller, I., Deák, C., Hanzelik, P. P., Ott, K. (2019). Cluster analysis of core measurements using heterogeneous data sources: an application to complex miocene reservoirs. *Journal of Petroleum Science and Engineering*, 178, July 2019, pp. 575–585.
Retrieved from https://doi.org/10.1016/j.petrol.2019.03.067

Wang, X., Yang, S., Zhao, Y., Wang, Y. (2018). Lithology identification using an optimized KNN clustering method based on entropy-weighed cosine distance in Mesozoic strata of Gaoqing field, Jiyang depression. *Journal of Petroleum Science and Engineering*, 166, pp. 157–174.
Retrieved from https://doi.org/10.1016/j.petrol.2018.03.034