Implementation of Data Science and Decision Analysis to Determine Shale Gas Sweet Spot Depth Interval

Muhamad Aditya, Andri D. Setiawan, Armand O. Moeis
Industrial Engineering Department, Faculty of Engineering, Universitas Indonesia
Kampus UI Depok, Depok 16424, Indonesia
muhamad.aditya01@ui.ac.id, a.d.setiawan@ui.ac.id, armand.omar@ui.ac.id

Abstract

Shale gas has been regarded as one of the most promising energy sources to sustain the world's energy demand. However, its exploration is still underdeveloped in several countries due to a lack of methods and technology implementation compared to conventional hydrocarbon exploration. In addition, the technology, methods, and data available in various oil and gas companies are currently still concentrated on conventional hydrocarbon exploration. The purpose of this study is to propose a new comprehensive method in shale gas exploration by utilizing the existing conventional hydrocarbon exploration data using data science and decision analysis approaches. The methods used in this study are K-Mean Clustering to cluster the similar rock characters (TOC, Porosity, Water Saturation, and Poisson Ratio) then continued by Multi-Criteria Decision Analysis to determine the best rock cluster for shale gas exploration. The study takes Banuwati Shale Formation in Asri Basin as a case which is well known as one of the promising source rocks in Indonesia. Based on this study, the rocks in the study area can be classified into three clusters. Cluster 1 is determined as "High Fractability Cluster", Cluster 2 is determined as "Water Saturated Cluster" and Cluster 3 is determined as "High Organic Content Cluster" based on its physical and chemical properties. Meanwhile, Cluster 3 is determined as the best cluster with 10212 ft – 10412 ft (3113 m – 3174 m) depth interval preferred as the sweet spot for Shale Gas exploration based on Multi-Criteria Decision Analysis result.

Keywords
Shale Gas, Data Science, K-Mean Clustering, Decision Analysis, Asri Basin

1. Introduction
Fossil energy is predicted to remain as the main commodity to fulfill energy needs in Indonesia until 2050 (Malik, 2021). The same thing happened globally where fossil energy sources, especially natural gas, dominate a quarter of the world's energy consumption (Tan et al. 2018). To respond to these conditions, several countries such as the United States, Russia, and China have started to increase shale gas exploration to increase the reserves of hydrocarbon energy. Meanwhile, shale gas exploration in is considered still not much developed in several countries around the globe. Indonesia is one of the finest examples among oil and gas producing countries (PWC, 2019). With 574 TCF (trillion cubic feet) of shale gas potential (LEMIGAS, 2018), Indonesia has a great opportunity to increase the reserves of domestic hydrocarbon energy sources through shale gas commodities. The lack of shale gas exploration is predicted caused by low studies, methods, and technology implementation compared to conventional oil and gas exploration. In addition, the technology, methods, and data available in various oil and gas companies are currently still concentrated on conventional hydrocarbon exploration. A new and comprehensive method to utilize conventional hydrocarbon exploration data for shale gas exploration seems to be one of the promising moves not only to speed up shale gas exploration but also minimize effort and operational cost. The purpose of this study is to determine shale gas sweet spot depth interval in one well using data science and decision analysis approaches by utilizing existing conventional hydrocarbon exploration data. This method is expected to be an option to perform an optimum operation to do shale gas exploration both in the side of effort and cost so the unconventional hydrocarbon exploration can be accelerated.

2. Literature Review
An in-depth analysis and literature review were carried out before building a research framework to find gaps from previous research. The literature review was done to make sure that the study conducted does not replicate any

existing research so that it can become a novelty in science. Based on a literature study of similar previous studies, there are several advantages and disadvantages in each study that can be used as guidelines in this study.
The research entitled K-Mean Cluster Analysis for Better Determining the Sweet Spot Intervals of the Unconventional Organic-Rich Shale: A Case Study by Akbar and Nugraha (2018) used the K-Mean Clustering method in classifying data from hydrocarbon exploration wells. This method is considered to be able to determine the desired number of clusters based on the number of parameters that depend on the distance between the data or the centroid. Even so, the determination of the best cluster is considered to have the potential to generate bias since there was no quantitative analysis to measure each cluster quality to determine the cluster's ranking.

Passey et al. (2010), through the research entitled From Oil-Prone Source Rock to Gas-Producing Shale Reservoir – Geologic and Petrophysical Characterization of Unconventional Shale-Gas Reservoirs explain the derived equations from several petrophysical parameters to obtain total organic carbon or TOC values in one well for exploration of shale gas reservoirs. Similar research that is Total Organic Carbon Prediction of Well Logs Data: Case Study Banuwati Shale Member Fm., Asri Basin, Indonesia by Basyir, et al. (2020) also explains the use of the multilinear regression method to get the TOC value. The parameters used are gamma-ray, density log, neutron porosity log, and P wave log. The results show the value of R2 = 0.9054, which shows a strong correlation. These two research are used as the guide to determine TOC value in this study.

The research entitled A GIS-Based Multi-Criteria Decision Analysis Approach For Exploring Geothermal Resources: Akarcay Basin in 2017 by Yalcin & Gul (2017) used the Analytical Hierarchy Process (AHP) method to determine the best alternative location for geothermal exploration. The results of the study show consistent results so that it can be concluded that the GIS-based AHP method can be accepted in the feasibility study stage of geothermal exploration. A similar outcome was also obtained from research entitled Geothermal Resource Potential Assessment Utilizing GIS-Based Multi-Criteria Decision Analysis Method by Meng, et al. (2021) and Assessment of Groundwater Potential Zones in Chittar Basin, Southern India Using GIS-Based AHP Technique by Nithya, et al. (2019). The studies showed that the use of AHP is considered efficient in selecting alternatives based on qualitative and quantitative criteria, especially in natural resources.

As to conclude, the results of the literature study show that the K-Mean Clustering method shows optimal results to cluster petrophysical data from exploration well. Even so, the end results are considered not fully maximized and the selection has the potential to generate bias because each cluster has its own advantages and disadvantages. On the other hand, Multi-Criteria Decision Methods such as Analytical Hierarchy Process (AHP) and Multi-Attribute Utility Theory (MAUT) could be the solution to determine the best options based on several parameters. The combination of those methods is considered to be the answer to the current gap and can be a novelty in the field of science especially unconventional hydrocarbon exploration.

3. Methods

The framework of this study is shown in the system diagram in Figure 1. McKeon (2013) stated that there are several parameters that could determine if shale reservoir could produce hydrocarbon commercially or not, those are permeability, porosity, pressure, total organic carbon (TOC), water saturation, shale thickness, moderate clay content and britlleness index. Based on the availability of data in this study, several petrophysical properties such as Total Organic Carbon (TOC), Porosity (POR), Water Saturation (SAT), and Poisson Ratio (PRT) are selected as parameters. K-Mean Clustering method was selected to cluster the data based on its value similarity to map the general data distribution. The number of the cluster was defined by observing the biggest difference of the Sum of Square Error (SSE) in each possible cluster number.

After clusters are defined, decision analysis is then performed to determine the best cluster as a shale gas exploration target. The selected method to do this operation were Analytical Hierarchy Process (AHP) and Multi-Attribute Utility Theory (MAUT).

A questionnaire method was performed to collect data from experts regarding the weight of the importance of each parameter (TOC value, porosity, poisson ratio, and water saturation) to sustain the decision analysis. The outcome of the questionnaire was the pairwise comparison of each parameter that will be calculated with some matrix calculations. The experts who participated in this study came from exploration geologists and petrophysics backgrounds with more than eight years of experience in oil and gas exploration to make sure the judgment credibility.

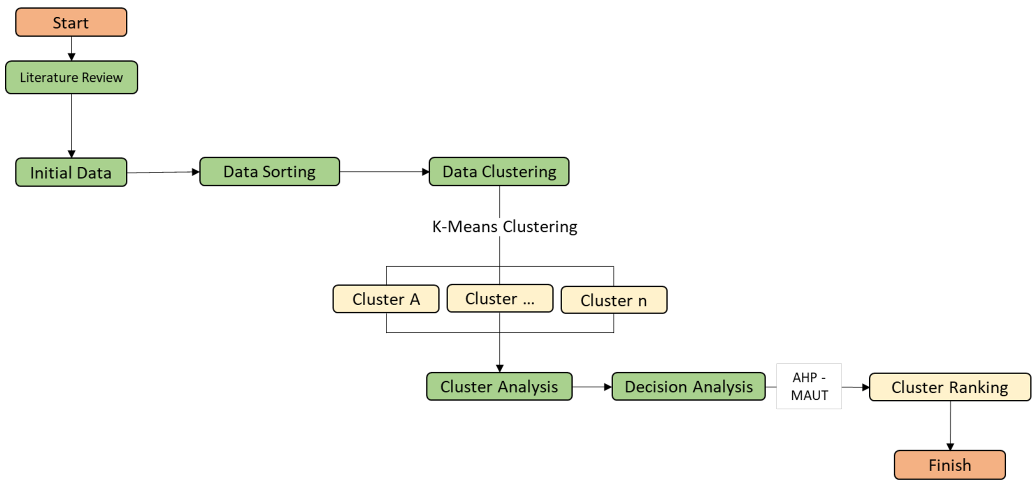

Figure 1. System Diagram of the Study

4. Data Collection

A total of 2123 data was obtained from 10174 ft to 11297 ft depth interval from one exploration well located in Asri Basin, Indonesia. The target of study is Banuwati Shale Formation that is classified as lacustrine-originated source rock (Satyana, 2017) that is well known as one of the promising source rock in Indonesia. Petrophysical properties such as gamma-ray, bulk density, porosity, poisson ratio, delta-T shear velocity, and water saturation were collected for the clustering process. Porosity, water saturation, and poisson ratio from the wells data are directly used as clustering parameters while another petrophysical property is used to determine the TOC value through a derivative equation based on Basyir et al. (2020). Questionnaire results in pairwise comparisons from five experts were also obtained as data to perform decision analysis.

5. Results and Discussion

After performing several operation as mentioned previously, the study has shown a promising result. The clustering result showed a clear data distribution and the decision analysis also showed a consistent result that produced decent decision support for shale gas exploration. A detailed explanation will be explained below.

5.1 TOC Value Calculation

Total organic carbon (TOC) is a measure of the dry weight percent of organic carbon within hydrocarbon source rocks. It is one of the most important parameter to assess the quality of source rock in one petroleum system also has a big part to decide hydrocarbon exploration success. Waples (1985) stated that rocks that have a TOC value of less than 0.5% are not considered as potential rocks. The amount of hydrocarbon produced in the rock is so small that expulsion is almost not going to occur. The current conditions, the availability of TOC data on shale as source rock are very limited. Most of conventional hydrocarbon exploration dataset has a very limited TOC value data due to expansive cost to define the total organic carbon contained in one rock sample. Usually this analysis only performed to desired point in several depth to observe the general value total organic carbon in one depth interval. In unconventional hydrocarbon or shale gas exploration, TOC value is having a crucial role to determine shale gas reservoir. Because of that the complete TOC value from top to bottom of the target interval is required. To overcome the cost issue, the value of TOC can be calculated using formula derived from several petrophysical properties. Basyir et al. (2020) explained one of the equation to calculate TOC value using Multi Linear Regression (MLR) method. This method uses the general equation of MRL (formula 1).

$$yi = (a0) + (a1 \times x1) + (a2 \times x2) + (a3 \times x3) + (a4 \times x4) + \cdots + (an \times xn) \quad (1)$$

Basyir (2020) suggested that the parameter used in this MRL equation are gamma-ray (GR), bulk density (RHOB), neutron porosity (NPHI), and delta-T shear velocity (DT). The result of the MRL equation to define the TOC value in this study is shown in formula 2.

$$TOC = 9.639 + (-0.018 \times GR) + (-2.801 \times RHOB) + (16.071 \times NPHI) + (-0.038 \times DT) \tag{2}$$

The regression result between real TOC value with calculated TOC showed a promising result with 0.8816 R^2 result quantitatively and shown a positive correlation qualitatively in the scatter plot diagram (Figure 2).

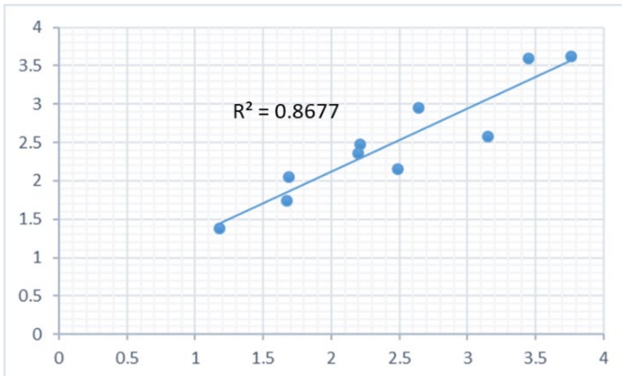

Figure 2. Scatter Plot between Real TOC Value and Calculated TOC Value

5.2 K-Mean Clustering

The number of clusters should be defined before performing the cluster analysis. It will determine the centroids of each cluster so the nearest value will be clustered. The number of cluster actually can be decided based on requirement but to generate a natural result deciding the number of the cluster through calculation is recommended like what has been done in this study.

The number of cluster in this study is determined using Elbow Method. The idea is to see the biggest difference of Sum Square of Error (SSE) in each possible cluster number. Calculation of SSE in this study is shown in Figure 3.

Number of Clusters	Akaike's Information Criterion (AIC)	AIC Change[a]	Ratio of AIC Changes[b]	Ratio of Distance Measures[c]
1	5900.206			
2	3582.176	-2318.030	1.000	1.976
3	2416.809	-1165.367	.503	3.599
4	2104.524	-312.284	.135	1.030
5	1801.862	-302.662	.131	1.516
6	1607.714	-194.149	.084	2.125
7	1524.817	-82.897	.036	1.030
8	1444.847	-79.970	.034	1.065
9	1370.703	-74.144	.032	1.131
10	1306.995	-63.708	.027	1.184
11	1255.646	-51.349	.022	1.038
12	1206.774	-48.872	.021	1.433
13	1177.514	-29.260	.013	1.074
14	1151.387	-26.128	.011	1.113
15	1129.520	-21.867	.009	1.117

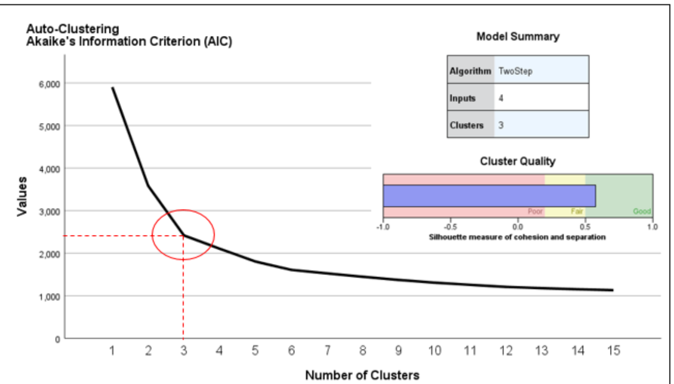

Figure 3. The Sum Square of Error Calculation (left) and its Cross Plot Diagram (right)

The SSE calculation results in this study show that the optimum number of clusters is 3 clusters. This number was determined based on SSE change from the number of cluster 3 to number cluster 4, which is dramatically changed with an 853.083 deficit. This result is reflected in the diagram where the correlation line experiences the biggest curve before it slowly goes flat.

After deciding the number of cluster that would be created, the next step is to determine the cluster center or centroid. The centroid determination needs several times of iteration before the centroid of each cluster is steady or in the convergent condition. The centroid of each cluster in this study reaches its convergent condition after 14 times of iteration (Table 1). The final centroid of each cluster is shown in Table 2.

Table 1. Iteration History of Clusters' Centroid Determination

Iteration	Change in Cluster Centers		
	1	2	3
1	1.878	0.2883	2.527
2	0.234	0.491	0.310
3	0.036	0.312	0.323
4	0.007	0.209	0.289
5	0.200	0.144	0.171
6	0.299	0.131	0.090
7	0.236	0.112	0.063
8	0.134	0.059	0.028
9	0.058	0.029	0.015
10	0.039	0.016	0.005
11	0.019	0.007	0.002
12	0.009	0.006	0.004
13	0.006	0.003	0.002
14	0.000	0.000	0.000

Table 2. Final Centroid of Each Cluster

Cluster	Zscore(TOC)		Zscore(POR)		Zscore(SAT)		Zscore(PRT)	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
1	-0.58046	0.42788	-0.34962	0.47699	0.66987	0.26526	0.01668	0.49069
2	-0.63863	0.55249	-1.19980	0.34214	-1.30100	1.30100	-1.43226	0.77514
3	1.16535	0.68078	1.18193	0.39837	0.69271	0.69271	0.82075	0.62100
Combined	0.00000	1.00000	0.00000	1.00000	1.00000	1.00000	0.00000	1.00000

The centroid of each cluster is affected by the parameters used in this study. ANOVA result is also provided to make sure that all parameters (TOC, porosity, water saturation, and poisson ratio) have a significant influence on centroid determination. The significance of all parameters showed a low even zero value (Table 3) that indicates all parameters have an influence in calculating the centroid and determining the cluster.

Table 3. ANOVA Table of Parameters

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore(TOC)	741.649	2	0.301	2120	2461.702	0.000
Zscore(POR)	846.655	2	0.202	2120	4186.597	0.000
Zscore(SAT)	590.558	2	0.444	2120	1330.646	0.000
Zscore(PRT)	617.135	2	0.419	2120	1473.790	0.000

After having the ideal number of clusters and determining each centroid, the next step is grouping all data to the nearest cluster. This process is what we called as data clustering. All 2123 data are well clustered without any single data missing. Cluster 2 is the biggest cluster with 1047 data, followed by cluster 3 with 714 data and cluster 1 with 362 data clustered. All clusters were analyzed to see the properties of all clusters by calculating the average value of each parameter (Table 4). Cluster 1 is determined as High Fractability Cluster due to its lowest poisson ratio that is 0.25. The lower poisson ratio value, the easier the rock to be cracked in shale gas production method for example hydraulic fracturing. The crack produced from the hydraulic fracturing process will act as artificial porosity that can sustain the hydrocarbon flow. Akintorinwa (2020) classified poisson ratio of 0.25 as low poisson ratio and the implication to surface structure considered as weak material. Cluster 2 is determined as Water Saturated Cluster due to its highest water saturation percentage. The high percentage of water saturation is actually not a good sign since it is suggested that the rock is already full of water so there is no space for hydrocarbon storage. Cluster 3 is determined as a High Organic Content Cluster due to its highest TOC value. The higher the TOC value the better its potential to generate hydrocarbon.

Table 4. Petrophysical Properties Summary of Each Cluster

Cluster	Average				Name
	TOC	Porosity	SW	Poisson Ratio	
1	0.57	9%	55%	0.25	High Fractability Cluster
2	0.61	16%	98%	0.3	Water Saturated Cluster
3	2.16	30%	78%	0.33	High Organic Content Cluster

The summary of data clustering is shown in Figure 4 where all parameters are delivered in log data to better visualize its distribution in the exploration well. Cluster 3 tends to be located on the top of the formation, cluster 2 tends to be located at bottom of the formation and cluster 1 is well distributed from top to bottom of the formation. The geological hypothesis is quite hard to explain from this condition since additional data still need to be included in the study.

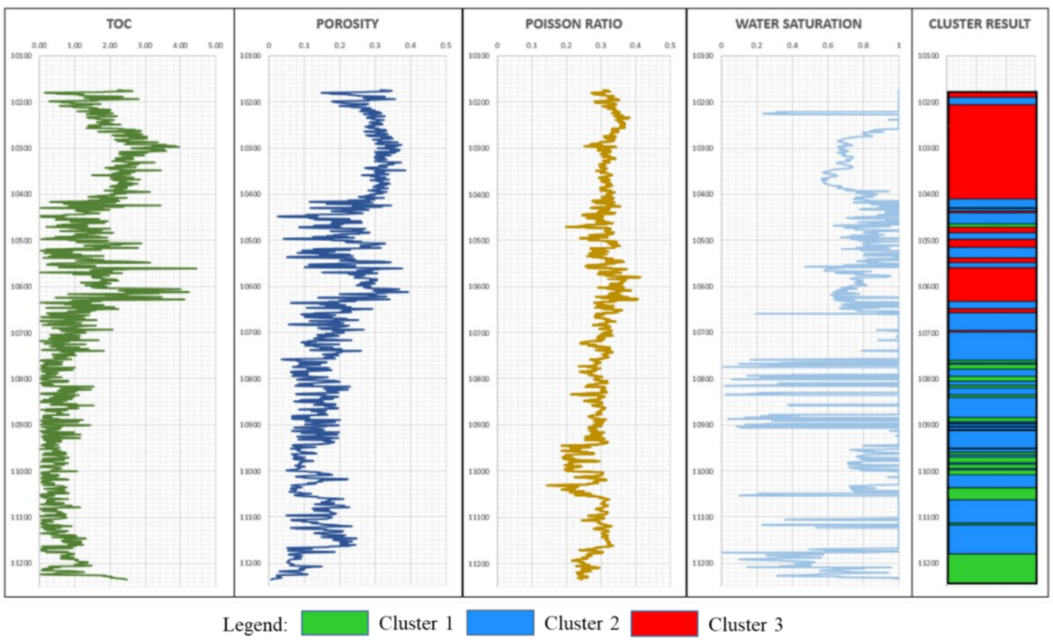

Figure 4. Clustering Log Result

5.3 Multi-Criteria Decision Analysis

The methods used for decision analysis in this study are Analytical Hierarchy Process (AHP) and Multi-Attribute Utility Theory (MAUT). AHP is an approach in decision making that structures multiple choices or criteria into a hierarchy, assessing the relative importance of the criteria, comparing alternatives, and determining the overall ranking of alternatives (Wisianto and Andriansyah 2007).

Saaty (1994) also stated that AHP is a mathematical technique for multi-criteria decision-making. Aside from AHP, MAUT is also used in this study to determine the overall ranking of the alternatives. The use of an additional method is required since not all parameters used in this study are proportional to its value. It means some parameters value are considered in better condition in the lower value, such as water saturation and poisson ratio.

MAUT has utility function calculations that enable users to determine if the parameters are expected high or low. To describe the decision analysis framework, the parameters and alternatives used, authors created an AHP tree / structure shown in Figure 5.

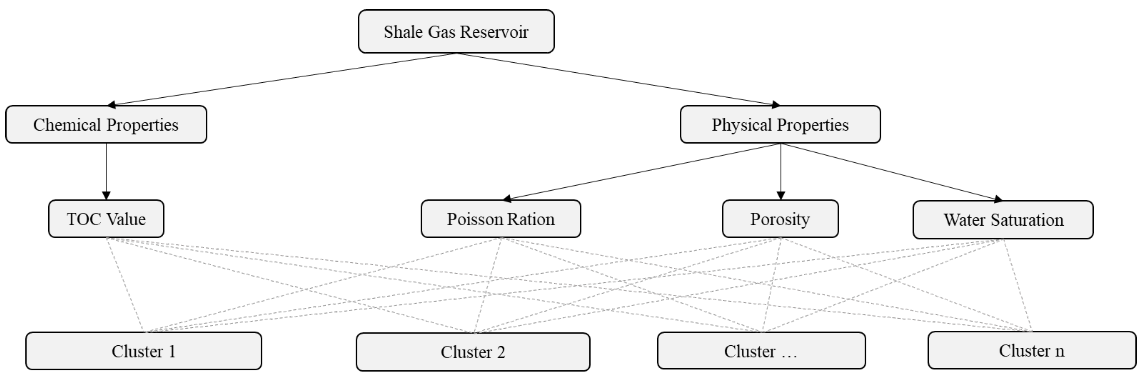

Figure 5. Decision Analysis Framework

AHP process started with collecting the data obtained from experts' questionnaires. Experts were asked to give their judgment regarding the relative importance of each parameter based on the Pairwise Comparison Scale from Saaty (1994). The Saaty Pairwise Comparison Scale is shown in Table 5. The geomean of the data needs to be first calculated to calculate the average of the set of values (Table 6).

Table 5. Saaty Pairwise Comparison Value

Numerical Value	Description
1	Equal Importance
3	Slight importance of one over another
5	Moderate importance of one over another
7	Very strong importance
9	Extreme importance of ne over another
2, 4, 6, 8	Intermediate value between two adjacent values

Table 6. Geomean Calculation of Overall Data

Respondent	TOC vs POR	TOC vs PRT	TOC vs SAT	POR vs PRT	POR vs SAT	PRT vs SAT
1	4.00	5.00	7.00	6.00	7.00	7.00
2	8.00	5.00	8.00	0.25	0.25	1.00
3	9.00	5.00	9.00	0.11	3.00	7.00
4	1.00	5.00	3.00	1.00	0.50	2.00

5	6.00	6.00	5.00	6.00	5.00	5.00
Geomean	4.44	5.19	5.97	1.00	1.67	3.45

Pairwise comparison is the next crucial step in AHP calculation. After collecting all data and calculating geomean, the data is then normalized and arranged in one comparison matrix to see the weight of the importance of each parameter (Table 7). The result of the pairwise comparison is the priority of each parameter in the study. The result showed that the TOC value is considered as the most important parameter in shale gas exploration with the 0.61 priority value, followed by poisson ratio of 0.17, porosity with 0.14, and water saturation with a 0.08 priority value.

Table 7. Geomean Calculation of Collected Data

Criteria	TOC	POR	PRT	SAT	Priority
TOC	0.63	0.63	0.69	0.49	0.61
POR	0.14	0.14	0.13	0.14	0.14
PRT	0.12	0.14	0.13	0.29	0.17
SAT	0.11	0.08	0.04	0.08	0.08
S.O.R	1.00	1.00	1.00	1.00	1.00

Several important AHP parameters should be determined to make a decent and consistent AHP calculation. Those are Eigen Value, Consistency Index, and Consistency Ratio. Eigen Value is obtained from the summation of products between each element of the Eigen vector and the sum of columns of the pairwise matrix. Eigen Value is then used to calculate the consistency index of the pairwise comparison. The formula of Eigen Value is shown in formula 3.

$$CI = (\lambda \max - n) \,/\, (n - 1) \tag{3}$$

λmax is the maximum Eigen Value of the judgement matrix and n is the order number of the matrix. The next important AHP parameter is Consistency Ratio (CR) where this value will determine whether the matrix used is good enough to be advanced or not. Saaty (1980) suggests that if that ratio exceeds 0.1, the set of judgments may be too inconsistent to be reliable. CR is obtained from the division of Consistency Index (CI) and Random Consistency Index (RI) (formula 4). Saaty (1980) suggested the Random Consistency Index value based on the matrix size (Table 8).

$$CR = \frac{CI}{RI} \tag{4}$$

Table 8. Random Consistency Index

Matrix Size	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.1	1.45	1.49

The element consistency limit measured by the RI value which has been classified by Saaty (1994) through an experiment of 500 samples where if numerical considerations are carried out randomly from a scale of 1/9, 1/8, …, 1, 2, …, 9, the average consistency value will be obtained as shown in Table 8. RI of 0.90 used since there are four parameters that produced a 4 x 4-matrix calculation in this study. From the explanation above, it can be calculated that those AHP parameters in this study are shown in Table 9. The value of CR is 0.06 so it can be stated that the matrix calculation is consistent.

Table 9. AHP Parameters Calculation in the Study

AHP Parameter	Value
Eigen Value	4.17
CI	0.06
RI	0.9
CR	0.06

After having a consistent comparison matrix, the next step is calculating the matrix square to determine the Eigen Vector (EV). The matrix square calculation can be iterated until reaching a stable state of the matrix, in which the difference of each Eigen Vector is equal to zero. The matrix square operation of this study is shown in Table 10, Table 11, and Table 12.

Table 10. Matrix Square Calculation: First Iteration

Criteria	TOC	POR	PRT	SAT	S.O.C	EV
TOC	4.00	17.63	16.54	37.26	75.44	0.62453
POR	0.92	4.00	3.65	8.14	16.72	0.13840
PRT	1.19	4.92	4.00	9.73	19.84	0.16421
SAT	0.53	2.23	2.05	4.00	8.80	0.07286
TOTAL					**120.79**	**1**

Table 11. Matrix Square Calculation: Second Iteration

Criteria	TOC	POR	PRT	SAT	S.O.C	EV	1st EV - 2nd EV	
TOC	71.55	305.51	272.98	602.59	1252.63	0.61922	TOC	0.005
POR	16.01	68.40	61.16	135.09	280.66	0.13874	POR	0.000
PRT	19.17	82.01	73.54	162.19	336.92	0.16655	PRT	-0.002
SAT	8.70	37.17	33.21	73.64	152.71	0.07549	SAT	-0.003
TOTAL					**2022.92**	**1**		

Table 12. Matrix Square Calculation: Third Iteration

Criteria	TOC	POR	PRT	SAT	S.O.C	Eigen Vector	2nd EV - 3rd EV	
TOC	20485.91	87540.60	78301.22	173035.88	359363.60	0.61939	TOC	0.000
POR	4588.47	19607.51	17538.06	38756.95	80490.99	0.13873	POR	0.000
PRT	5505.59	23526.56	21043.52	46503.59	96579.26	0.16646	PRT	0.000
SAT	2494.59	10659.90	9534.80	21070.80	43760.08	0.07542	SAT	0.000
TOTAL					**580193.92**	**1**		

The matrix square reaches a stable state after the third iteration which the EV difference with the previous iteration is equal to zero. The latest EV obtained from the latest matrix square calculation reflected the parameters' priority or importance in shale gas exploration. Parameter priority for shale gas exploration is shown in Table 13.

Table 13. Parameter Priority

Criteria	Rank	Priority Percentage (%)
TOC Value (TOC)	1	61.9
Poisson Ratio (PRT)	2	16.6
Porosity (POR)	3	13.9
Water Saturation	4	7.5

After obtaining the parameter priority through the AHP method, the next operation will be done with the MAUT method to determine the alternative preference, or in this study is cluster preference. The first step of MAUT is to set the utility function to determine if each parameter is expected to be higher or lower value to be in its optimum condition. TOC value is expected to be higher since the higher the value, the higher the possibility of hydrocarbon expulsion from the rock. Porosity is also expected to be in higher value since the higher porosity will help the hydrocarbon flow in the production phase. Poisson ratio is expected to be in the lower value since the lower poisson ratio, the easier the rock is to be cracked during the hydraulic fracturing process. The last, water saturation is expected to be in the lower value since the higher the water saturation value, the lower possibility of hydrocarbon contained in the rocks. The utility function operation in this study is shown in Table 14.

Table 14. Utility Function of MAUT Calculation

Option/Criteria	TOC	POR	PRT	SAT
Expectation	max	max	min	min
Cluster 1	0.57	0.09	0.25	0.55
Cluster 2	0.61	0.16	0.3	0.98
Cluster 3	2.16	0.30	0.33	0.78
Option -	0.57	0.09	0.33	0.98
Option +	2.16	0.3	0.25	0.55

The result of the utility function is then normalized using formula 5 below, the result is shown in Table 15.

$$ui(x) = \frac{xi - xi^{-}}{xi^{+} - xi^{-}} \quad (5)$$

Table 15. Normalized Utility Function of MAUT Calculation

Option/Criteria	TOC	POR	PRT	SAT
Expectation	max	max	min	min
Cluster 1	0.00	0.00	1.00	1.00
Cluster 2	0.03	0.33	0.38	0.00
Cluster 3	1.00	1.00	0.00	0.47

The result of the utility function is then multiplied with the parameters weight as explained previously in Table 13 to generate the final cluster preference in this study. The formula used in this operation is shown in formula 6, and the result is shown in Table 16.

$$u(x) = \sum_{i=1}^{n} wiui(x) \quad (6)$$

Table 16. The Result of Cluster Preference

Ranking	Preference	Cluster
1	0.793197056	Cluster 3
2	0.241883502	Cluster 1
3	0.124248361	Cluster 2

6. Conclusion

This study analyzes the new method to determine shale gas weet spot depth interval using data science and decision analysis approaches. The data obtained from one exploration well data located in Asri Basin, Indonesia with Banuwati Shale Formation as the main target of the study. The result of sweet spot determination is shown in Figure 6.

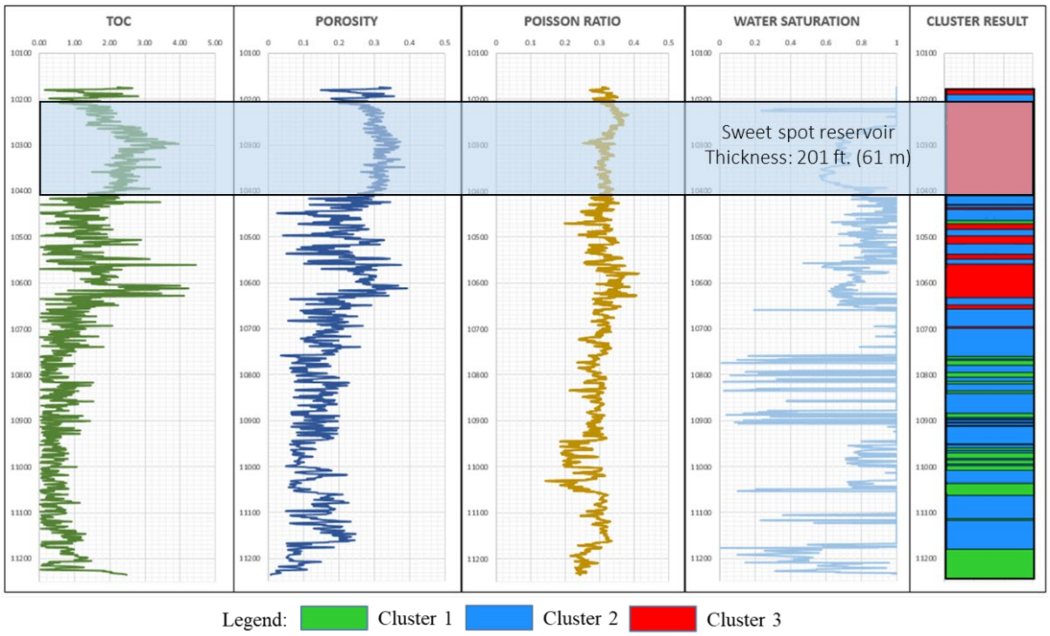

Figure 6. Sweet Spot Determination Result for Shale Gas Exploration

The methods used in this study are K-Mean Clustering, Analytical Hierarchy Process (AHP), and Multi-Attribute Utility Theory (MAUT). The result of data clustering shows that the data in this study can be clustered into three clusters. Cluster 1 is determined as High Fractability Cluster, cluster 2 is determined as Water Saturated Cluster and cluster 3 is determined as High Organic Content Cluster. The preferred cluster was obtained from AHP and MAUT methods using experts judgment questionnaires as the data. The result suggested that Cluster 3 is the most preferred cluster with a 0.793 preference score followed by Cluster 1 with 0.242 and Cluster 2 with a 0.124 preference score.

Mckeon (2013) stated that the minimum thickness of shale layer to be able to produce hydrocarbon commercially is 100 ft. or equal to 30.5 m. Referring to the actual condition in the exploration well, there is one horizon of Cluster 3 rocks which exceed 100 ft. thick and can be determined as sweet spot zone that is the 10212 ft. – 10413 ft. (3113 m – 3174 m) interval or equal to 201 ft. (61 m) thick. This interval determined as the most preferred spot for shale gas exploration.

6.1 Proposed Improvement

Future research is highly recommended for the more accurate and more comprehensive result of the study. Future research suggested utilizing more petrophysical properties as parameters to minimize uncertainty. Aside from that, more data utilization of the well is also highly recommended so the result of each well can be correlated and produce more comprehensive sweet spot mapping in 2D and 3D results.

References

Akbar, M. N., & Nugraha, S. T., K-Mean Cluster Analysis for Better Determining the Sweet Spot Intervals for Unconventional Organic-rich Shale: A Case Study, 2018.

Akintorinwa, O. J., Ajayi, C., & Ademilua, L., Empirical Relationship between Gravimetric and Mechanical Properties of Basement Rocks in Ado-Ekiti, Southwestern Nigeria. *Journal of Geology and Geophysics*, 486 - 499, 2020.

Basyir, A., Haris, A., & Bachtiar, A., Total Organic Carbon Prediction of Wells Log Data: Case Study Banuwati Shale Member Fm., Banuwati Basin, Indonesia. *Proceedings of the 5th International Symposium on Frontier of Applied Physics.* AIP Publishing, 2020.

LEMIGAS. , *Update Indonesia Shale Hydrocarbon Basin.* Bangkok: LEMIGAS., 2018.

Malik, C. L., Indonesia Country Report. *Energy Outlook and Energy Saving Potential in East Asia 2020*, 102-121, 2021.

McKeon, M., *Horizontal Fracturing in Shale Plays.* Halliburton, 2013.

Meng, F., Liang, S., Xiao, C., & Wang, G., Geothermal Resource Potential Assessment Utilizing GIS - Based Multi Criteria Decision Analysis Method. 101969, 2021.

Nithya, C. N., Srinivasa, Y., Magesh, N., & S., K., Assessment of Groundwater Potential Zones in Chittar Basin, Southern India Using GIS Based AHP Technique. *Remote Sensing Applications: Society and Environment, 15*, 100248, 2019.

Passey, Q. R., Bohacs, K. M., Esch, W. L., Klimentidis, R., & Sinha, S., From Oil-Prone Source Rock to Gas Producing Shale Reservoir - Geologic and Petrophysical Characterizaion of Unconventional Shale-Gas Reservoir. *SPE International Oil & Gas Conference and Exhibition.* Beijing, 2010.

PwC.,*Oil and Gas in Indonesia, Investment and Taxation Guide* (Vol. 10), 2019.

Saaty, T. , *Fundamentals of Decision Making and Priority Theory.* Pittsburgh: RWS Publications. 1994.

Saaty., T. L. , The Analytic Hierarchy Process. New York: McGraw-Hill.1980.

Satyana, A. H. , Regional Petroleum Geochemistry of Indonesian Basins; Updated, and Implications for Future Exploration. *41st Indonesian Petroleum Association Convention and Exhibition.* Jakarta.2017.

Tan, L., Zuo, L., & Wang, B. ,. Methods of Decline Curve Analysis for Shale Gas Reservoir. *Energies, 11*, 552, 2018.

Waples, D. W. (1985). *Goechemistry in Petroleum Exploration.* Boston: D. Reidel Publishing.1985

Wisianto, A., & Andriansyah, A. ,. Application of the Analytical Hierarchy Process (AHP) to Determine Gas Allocation Under Shortage Condition. *31st IPA Annual Conference Proceeding.* Jakarta: Indonesian Petroleum Association., 2007.

Yalcin, M., & Gul, F. K.,A GIS-Based Multi Criteria Decision Analysis Approach for Exploring Geothermal Resources: Arkacay Basin (Afyonkarahisar) . *Geothermics, 67*, 18-28. , 2017.

Biographies

Muhamad Aditya graduated from Geological Engineering program from Universitas Trisakti, Indonesia in 2019. His research was about geochemistry analysis and oil to source rock correlation study in one petroleum basin in Indonesia. He is currently pursuing his master degree in Industrial Engineering program at Universitas Indonesia focusing on Design and Modelling System. His research interest includes energy, system dynamics, and policy analysis. He currently works as Junior Consultant in one multinational consulting company in Indonesia.

Dr. Andri D. Setiawan is currently an assistant professor and member of the Systems Engineering Modeling and Simulation Laboratory at the Industrial Engineering Department, Universitas Indonesia. He is also a visiting research fellow at the School of Innovation Sciences, Eindhoven University of Technology (TU Eindhoven), The Netherlands. He earned a bachelor's degree in Aeronautics and Astronautics Engineering from Bandung Institute of Technology (ITB) in 2001. Since then he worked with Microsoft business partner as a consultant and system analyst in Jakarta until 2008. His interest in technology and policy led him to study master in the Netherlands and obtained a master's degree in Engineering and Policy Analysis from Delft University of Technology (TU Delft) in 2010. He then worked as a junior researcher, conducting research on the assessment of technological options for gas transition pathways in the Netherlands, under the Energy Delta Gas Research (EDGaR) program at the Faculty of Technology and Management, TU Delft for almost three years. He earned a doctorate in Industrial Engineering and Innovation Sciences from TU Eindhoven in 2018. His doctoral research focused on developing approaches for applying responsible innovation in different contextual conditions, combining the use of quantitative and qualitative approaches. He has specialties in the analysis of the multi-actor systems, decisions under risk and uncertainty, modeling and simulation, policy analysis, Q-methodology, system dynamics, and technology assessment. His disciplinary research interests include energy, renewable energy, renewable and sustainable energy systems, responsible innovation, and technology adoption.

Dr. Armand O. Moeis is an assistant professor in the Industrial Engineering Department, Faculty of Engineering, Universitas Indonesia. His research focuses mainly on Maritime Logistics, while his minor is in Renewable Energy and techno-economic policy. He especially deploys methods based on System Dynamics, Simulation Gaming, and Policy Analysis.