InPrime: Indonesian Journal of Pure and Applied Mathematics
Vol. 7, No. 1 (2025), pp. 76 – 87, doi: 10.15408/inprime.v7i1.44234
p-ISSN 2686-5335, e-ISSN: 2716-2478

Fuzzy Unsupervised Artificial Learning Based on Credibilistic Fuzzy C-Means

Kangiama Lwangi Richard[1*], Kafunda Katalayi Pierre[2], Kabamba Baludikay Blaise[1], Mabela Makengo Matendo Rostin[2] and Munene Asidi Djonive[1]
[1]Department of Exploration-Production, Faculty of Oil, Gas and Renewable Energies, University of Kinshasa, D.R. Congo
[2]Department of Mathematics, Statistics and Computer Science, Faculty of Science and Technology, University of Kinshasa, D.R.Congo
Email: richard.kangiama@gmail.com

Abstract

This study proposes an unsupervised artificial learning approach based on the Credibilistic Fuzzy C-Means (CFCM) algorithm to enhance the governance and analysis of oil production data. The research focuses on supporting decision-making in managing oil output from the MOTOBA oil field, operated by PERENCO in Moanda, Democratic Republic of Congo, covering the period from 2018 to 2021. The methodology involves structuring and segmenting production data using the CFCM algorithm, which enables the identification of meaningful production patterns despite the presence of uncertainty and imprecision in the data. The analysis identified three distinct clusters: wells with low production, wells with moderate production, and wells with high production. These clusters offer valuable insights into the variability of well performance and provide a basis for optimizing operational strategies. The credibilistic enhancement of traditional fuzzy clustering allows for more effective handling of data uncertainty, resulting in a robust and interpretable model—particularly beneficial in complex and data-limited environments. This clustering framework supports more refined monitoring, resource allocation, and operational planning, making it well-suited for the dynamic nature of oil field management. Furthermore, the methodology demonstrates potential scalability and applicability to other industrial domains facing similar challenges in data quality and decision-making under uncertainty. Ultimately, this work contributes to the advancement of data-driven governance in natural resource management through a rigorous and adaptable analytical approach.
Keywords: Artificial learning; Clustering; Credibilist; Fuzzy C-means; Fuzzy logic.

Abstrak

Studi ini mengusulkan pendekatan pembelajaran buatan tanpa pengawasan berdasarkan algoritma Credibilistic Fuzzy C-Means (CFCM) untuk meningkatkan tata kelola dan analisis data produksi minyak. Penelitian ini berfokus pada dukungan pengambilan keputusan dalam mengelola produksi minyak dari ladang minyak MOTOBA, yang dioperasikan oleh PERENCO di Moanda, Republik Demokratik Kongo, yang mencakup periode 2018 hingga 2021. Metodologi ini melibatkan penataan dan segmentasi data produksi menggunakan algoritma CFCM, yang memungkinkan identifikasi pola produksi yang bermakna meskipun terdapat ketidakpastian dan ketidaktepatan dalam data. Analisis ini mengidentifikasi tiga klaster yang berbeda: sumur dengan produksi rendah, sumur dengan produksi sedang, dan sumur dengan produksi tinggi. Klaster ini menawarkan wawasan berharga tentang variabilitas kinerja sumur dan menyediakan dasar untuk mengoptimalkan strategi operasional. Peningkatan kredibilistik dari pengelompokan fuzzy tradisional memungkinkan penanganan ketidakpastian data yang lebih efektif, menghasilkan model yang kuat dan dapat ditafsirkan—terutama bermanfaat dalam lingkungan yang kompleks dan terbatas data. Kerangka pengelompokan ini mendukung pemantauan, alokasi sumber daya, dan perencanaan operasional yang lebih baik, sehingga sangat sesuai untuk sifat dinamis pengelolaan ladang minyak. Lebih jauh lagi, metodologi ini menunjukkan potensi skalabilitas dan penerapan pada domain industri lain yang menghadapi tantangan serupa dalam

*) Corresponding author
Submitted January 10th, 2025, Revised May 30th, 2025,
Accepted for publication May 31st, 2025, Published Online May 31st, 2025

kualitas data dan pengambilan keputusan dalam ketidakpastian. Pada akhirnya, karya ini berkontribusi pada kemajuan tata kelola berbasis data dalam pengelolaan sumber daya alam melalui pendekatan analitis yang ketat dan adaptif.
Kata Kunci: *Pembelajaran buatan; Pengelompokan; Kredibilitas; Fuzzy C-means; Logika Fuzzy.*

2020MSC: 68T05, 62H30, 90C90.

1. INTRODUCTION

The Democratic Republic of Congo has significant oil reserves, mainly located in the sedimentary basins of the Albertine Graben and the Central Kongo region. Extraction techniques include conventional production and, in some cases, more advanced methods of exploiting reserves [10]. National laws and regulations regulate the oil sector in the Democratic Republic of Congo. The government defines concession management, taxation and environmental standards [1]. The country is trying to balance the economic benefits of oil extraction with ecological and social concerns. However, challenges remain in oil production data management, where we carried out our research. The amount of data collected on oil fields is enormous and varied in the oil industry. The governance of this data is essential for optimal exploitation. It presents challenges in terms of data quality, integrity and value. Traditional data management methods show their limitations in the face of the complexity of data sources and the uncertainty inherent in oilfield operations [2].

With the rise in volumes of unstructured data and the variability of operational conditions, it is crucial to adopt robust methods for managing uncertainty in data. This raises the question of the applicability of fuzzy approaches, such as Fuzzy C-Means, in oil data governance. The variability and uncertainty of production data raise challenges for its effective management. The search for robust approaches is essential to improve data governance [3]. Other studies have shown that the credibilitic Fuzzy c-means algorithm represents a modified version of the Fuzzy c-means (FCM) algorithm, incorporating the principle of the degree of credibility to consider uncertainty and imprecision when categorising data. In this version of the Fuzzy c-means technique, data is assigned to one or more groupings with a certain level of reliability [4]. This offers flexibility in uncertainty management, unlike the traditional or standard FCM algorithm.

2. METHODS

The use of the Fuzzy C-Means in its credibility version is therefore particularly justified in our case, as it allows us to account for the uncertain nature of the oil production data, improve the quality of the clusters obtained, and reinforce the reliability of the decisions based on these clusters according to their production.

In this article, we will use the UML (Unified Modeling Language) approach combined with the analysis methodology in computer science and the data collection steps, implementing credibilistic fuzzy clustering algorithms [5]. The applied research methodology consists of the following steps: Artificial learning enabled us to familiarise ourselves with the basic concepts needed to address the research problem presented [5]. It aims to understand the concepts and stages of learning, the application of fuzzy logic, which has enabled us to integrate the idea of fuzzy logic into the work, and the application of oil data governance through decision-making on data quality and credibilistic fuzzy c-means. Having expressed the need for this study, we exploited the data and reconstituted a dataset on the production of the MOTOBA field from 2018 to 2021 [6]. We used Python to carry out the pre-processing and Matlab to train our credible Fuzzy C-Means model to obtain the results.

2.1. Fuzzy c-means credibility

In this study, we provide additional means for rejecting outliers by introducing a new variable, the credibility of a vector. Credibility measures the typicality of the vector concerning the entire dataset. An outlier is expected to have a lower credibility value than a non-outlier [7]. Using the new variable leads to the Credibilistic Fuzzy C Means (CFCM) algorithm [8]. Credibilistic Fuzzy c-means was introduced to deal with the noise sensitivity of fuzzy procedures, but it is unstable and fails to capture local non-linear interactions.

1. Formulating the problem as an algorithm

The credibilistic fuzzy c-means algorithm is an extension of the fuzzy c-means (FCM) algorithm, which uses the concept of degree of credibility to consider uncertainty and imprecision in data classification [9]. In this algorithm, each datum belongs to one or more clusters with a certain degree of credibility. This allows more flexible uncertainty management than the classic FCM algorithm[8]. The detailed description of the Fuzzy C-means algorithm is as follows:

Inputs:

1) Data: A set of data $X_1, X_2, \dots, X_n$, where each X_i is a point in a p-dimensional space.
2) Number of clusters c: number of clusters to be formed.
3) Blur parameter $m > 1$: controls the level of blur.
4) Credibility parameter β: controls the importance of the degree of credibility in the calculation of memberships.
5) Tolerance ε: convergence criterion for stopping the algorithm.

Outputs

Initialise the centroids of clusters $V_1, V_2, \dots, V_n$. Initialise a matrix of degrees of membership U, where u_{ij} represents the degree to which point X_i belongs to cluster *j*. Following several experiments in our research, we chose $m = 2$, which is in line with the recommendation often cited in the literature. This decision is due, in particular, to the fact that it presents an interesting compromise between the readability and stability of the groupings in the face of data with moderate overlaps. This is particularly relevant in the oil industry, where several wells may display close but not identical production profiles. The ideal choice turns out to be $\beta = 1$, as it maximises the internal density of the group while guaranteeing effective separation between the groups. This corresponds to a balance in data reliability representation, particularly in measurement fluctuations or noise, for example, pressure variations or sensor irregularities at oil sites.

2. Calculating the degrees of membership

Calculate the degree of membership for each point X_i and each cluster j using the distance between X_i and the centroid $\boldsymbol{V_j}$. For the credibility algorithm, membership is influenced by the credibility parameter β. The degree of membership u_{ij} is calculated as follows [10][11]:

$$\boldsymbol{u_{ij}} = \frac{\gamma_i}{\sum_{k=1}^{C}\left(\frac{\|x_i - v_j\|}{\|x_i - v_k\|}\right)^{\frac{2}{m-1}}}. \tag{1}$$

In the fuzzy credibilistic c-means model, the parameter γ_i symbolises a regularisation coefficient or an adjustment element of the membership function.

3. Calculating degrees of credibility

To integrate credibility, we define a degree of credibility c_{ij} for each cluster point, which modifies membership according to the consistency of the data in the cluster.

$$c_{ij} = exp(-\beta. \|x_i - v_j\|), \tag{2}$$

where β is a parameter regulating credibility. A low β gives more credibility to each cluster point.

4. The formula for updating centroids

Use the degrees of membership and credibility to update the centroids:

$$v_j = \frac{\sum_{i=1}^{n}(\boldsymbol{C_{ij}u_{ij}^{m}}X_i)}{\sum_{i=1}^{n}(\boldsymbol{C_{ij}u_{ij}^{m}})}. \tag{3}$$

5. Convergence criterion

Calculate the change in the membership degree matrix U. If this change is less than a tolerance threshold ϵ, stop the algorithm. Otherwise, return to step 2.

6. Output of results

The centroids $V_1, V_2, \ldots\ldots, V_c$ are the final centres of the clusters. Matrices U and C contain the degrees of membership and credibility, respectively, for each point in each cluster.

Algorithm :

Start

Inputs: Data X, Number of clusters fuzzy parameter m, credibility parameter β, tolerance ϵ

1) Initialise centroids $\{V_1, V_2, \ldots, V_c\}$,
2) Initialise the membership matrix *U* with random values

Repeat until convergence:

3) For each point X_i and each cluster *j*: Calculate the degree of membership $\boldsymbol{u_{ij}}$ and degree of credibility $\boldsymbol{C_{ij}}$,
4) Update the centroids V_i with the weights (u_{ij}*m * C_{ij}),
5) Calculate the change in the matrix *U*,
6) If the change is less than ϵ, stop.

Outputs:

Centroids V, membership matrix U, credibility matrix C.

2.2. Complexity of the credibilistic Fuzzy C-Means algorithm

The total time complexity of the Fuzzy C-Means algorithm is $O(T \times n \times c \times d)$, where n is the number of data points, c is the number of clusters, d is the dimension of data (number of features), and T is the number of iterations for convergence [12].

2.3. Working Environment

We utilized a combination of specific hardware and software environments to implement our system. The hardware used was a Lenovo ThinkPad computer with a 500 GB hard disk, 4 GB of RAM, an 11-inch screen, and a Dual-Core Intel processor running on the Windows 10 operating system. For the development environment, we employed Google Collaboratory Notebook, commonly known as "Colab," a platform provided by Google Research. Colab enables users to write and execute Python code directly from a web browser, making it particularly suitable for tasks involving machine learning, data analysis, and educational purposes.

In addition to Colab, we also used MATLAB (R2021a version), a mathematics software designed to handle matrix operations and numerical computations efficiently. MATLAB provides a robust environment for modeling, simulation, and algorithm development. Furthermore, we employed Python, a versatile, open-source, multi-paradigm programming language. Python allows developers to focus on problem-solving rather than syntactic constraints, making it an ideal tool for implementing complex algorithms and processing data effectively.

3. RESULTS

3.1. Description of the data

The initial dataset contains 1,644 entries from oil production reports collected between 2018 and 2021. This information includes daily or weekly records for various wells at an oil site, including volumes of oil, gas and water extracted and the date and well concerned. For the following reasons, a subset of 36 recordings was chosen. The primary purpose of the selection was to assess the robustness of the Fuzzy C-Means credibility algorithm on a small and representative sub-sample, thus facilitating manual analysis, visualisation of the clusters and validation of the results. In addition, the 36 recordings were selected to illustrate various sources and levels of production while preserving a certain temporal consistency (for example, over a specific period of one month), making it possible to study the algorithm's behaviour in typical practical cases. In addition, some preliminary experiments have been carried out in a hardware environment with limited computing resources, which justifies using a subset for the first iterations.

We use log and production data from the Motoba 13 ST and 15ST wells in our work. The statistics collected cover several years, but the production table varies from 2018 to 2021, with a total of 1644 records. We drew a sample of 36 records for efficiency of results depending on our material[12]. the producing wells of the MOTOBA field are: MOT_05,MOT_05D, MOT_06, MOT_06D, MOT_03, MOT_03D, MOT_O16, MOT_01X,MOT_13ST, MOT_08,MOT_09 in Excel format and structured as shown in Table 1 [13].

Table 1. Description of the attributes of our data source

Attribute	Description
Code	Code Production code
Well	Well Name of the motoba field well
Date	Date Date of production of the well
Production of oil	Oil production Quantity of oil production in barrels of oil
Production of gas	Gas production: Quantity of oil production in cubic feet of gas
Production of water	Water production Quantity of oil production in barrels of water

3.2. Application of fuzzy logic

Modelling variables with fuzzy logic is presented through the different windows of our memory from our well-cleaned data source product_oil.csv. Whose different well names begin with the prefix

MOT-XST. To carry out the fuzzification and defuzzification of a dataset on oil production from the MOTOBA oil field, we followed the fuzzification methodology as follows:

1. Choice of classes (fuzzy sets of the (RIS) output variable)
   For the system output (SIF), we have chosen a single output variable representing the well's production state. We have chosen three fuzzy sets for this variable corresponding to the different input parameters. The values represent the different fuzzy sets (classes) associated with the output variable [14]. The distribution of the database into classes is, therefore, as follows [15]:
   - From line 1 to line 5: class 1 or level (N1)
   - From the 5th line to line 20: class 2 or level (N2)
   - From line 20 to line 35: class 3 or level (N3)
2. Implementations of the System (SIF) via the graphical interface
3. Structure of the fuzzy inference system (FIS)
   We have chosen a Mamdani-type fuzzy inference system with three input variables and one output variable. The Figure 1 shows the block diagram of our system. We have chosen a fuzzy inference mechanism that works with the methods shown in the figure and must result in a single value for the output variable (state).

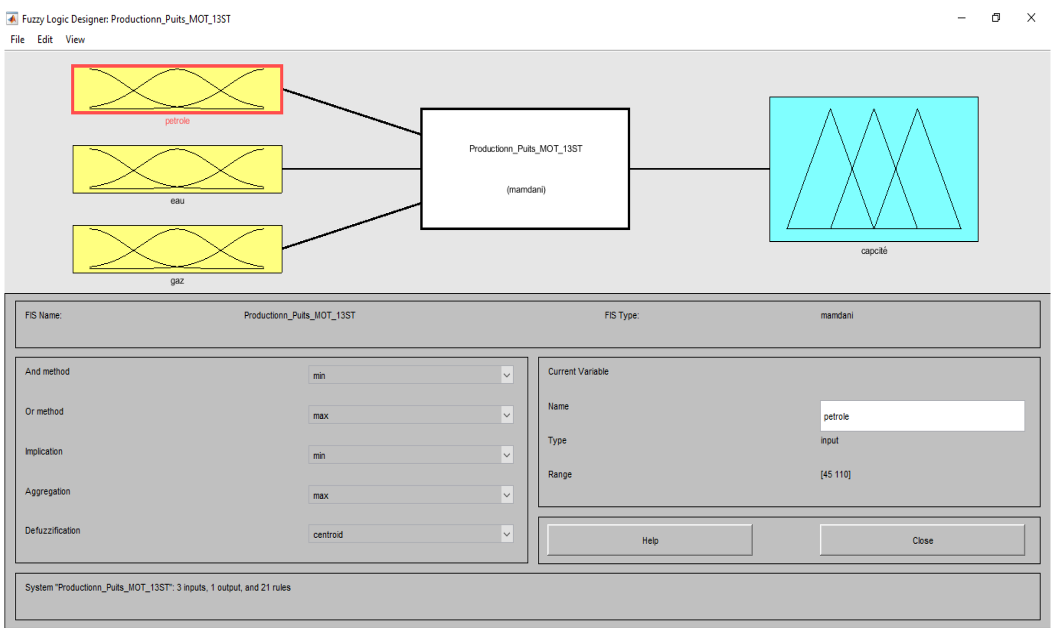


Figure 1. Fuzzy inference system block diagram

4. Fuzzification of input and output variables
5. Fuzzification of the input variable " qte_production_pétrole "
   This variable varies in the range (25,110). We, therefore, choose a range (25,45) [16] and three membership functions. The first function is trapezoidal (Oil) called very small(EP) with two parameters (25,39), the second function is trapezoidal(Moy) called <<PM>> with two parameters (39,45), the third function is trapezoidal(PE) called <<Moyen>> with two parameters (45,100). This process can be seen in detail in Figure 2.

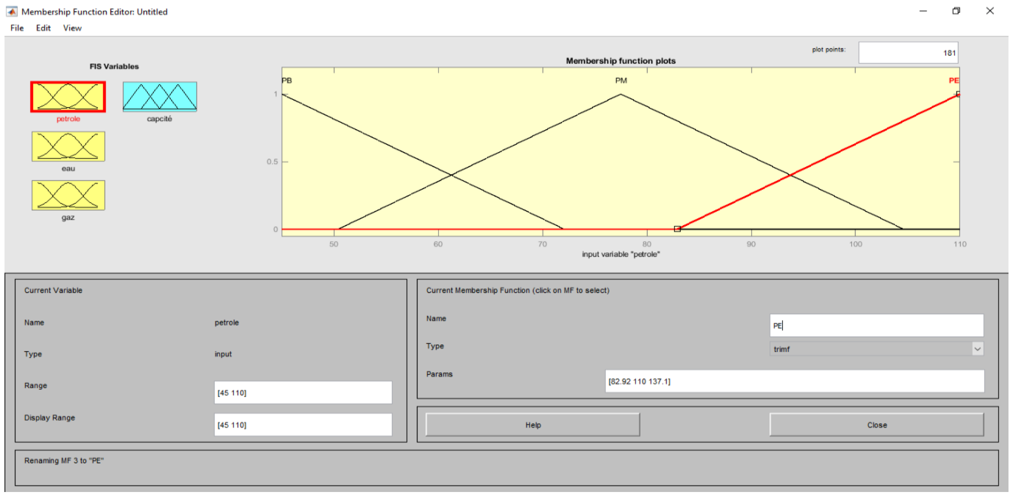


Figure 2. Fuzzification of the input variable production_oil

6. Defuzzification of the output variable
   We take an example of an application for an observation of the fourth class (N3): X= (175,526); Hypothesis: the production is 39,5 barrels of oil of 15/011/2020, the output of water is 243,6 barrels that of gas is 244.6 cubic feet of gas then the state of the studied device is considered as N4=7.We notice that our system classified this observation well [17]. These values are examples; the proposed system accepts all possible combinations. The Figure 3 shows the defuzzification for our practical case.

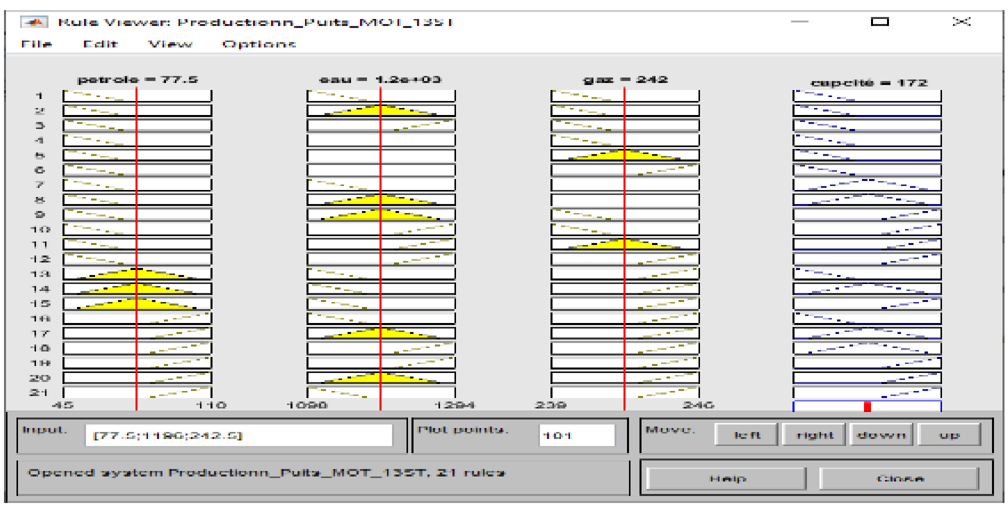


Figure 3. Defuzzification of the output variable (production state)

7. Definition of fuzzy sets and decision rules (Inference rules)
   The Adaptive Neuro-Fuzzy Inference System algorithm automatically produced the fuzzy rules designed to illustrate oil production from the Motoba field, merging fuzzy logic and learning. The different possible rules depending on the fuzzification of the various input and output variables chosen are as in Figure 4 (3 input variables and 3 functions for each of the variables, we obtain inference rules).

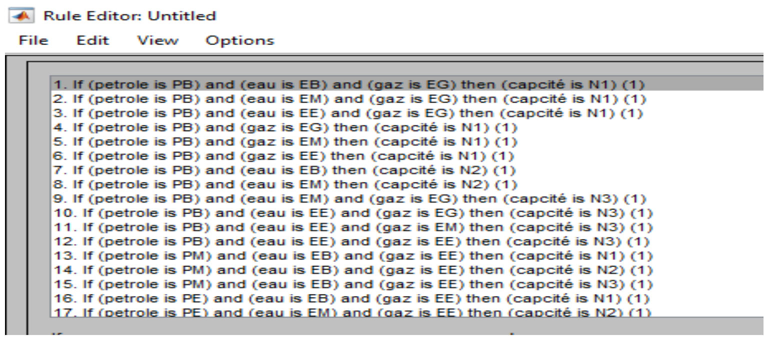
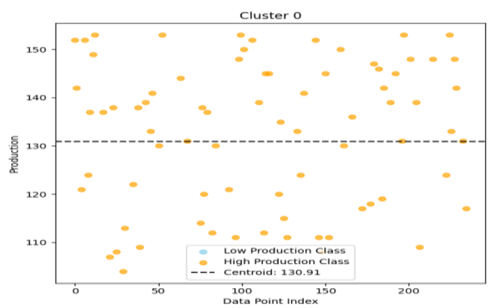


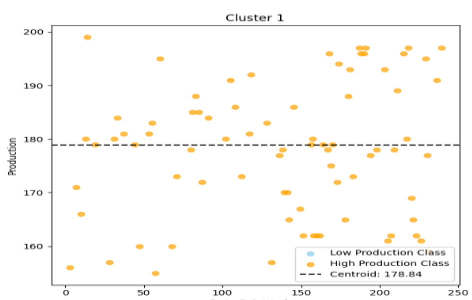
Figure 4. Inference rules.

3.3. The Cluster Centre

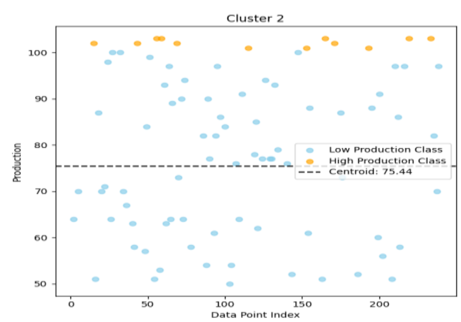
Cluster centres are found using fuzzy c-means clustering. Clustering stops when the improvement in the objective function is less than the specified minimum threshold [18][19].



(a)



(b)



(c)

Figure 5. Cluster description: (a) Cluster 1 representation, (b) Cluster 1 representation, and (b) Cluster 1 representation.

3.4. Visualisation of clusters with centres of gravity

A visualisation of the clusters has been carried out in Figure 6. The 2D graphs illustrate the differences between the clusters. The fuzzy C-Means credibilistic classification for oil well production data from the MOTOBA field in the Moanda territory [20][21]. The black dots in Figure 6 are the centres of gravity of our 3 classes, which group together the wells in the oil field's target sector. Some wells may belong to several classes with different degrees of membership [15].

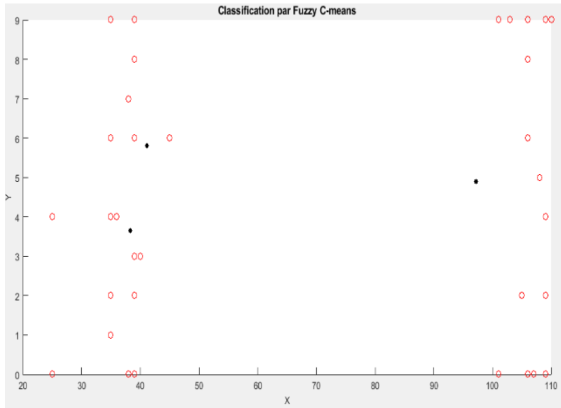


Figure 6. This is a figure. If there are multiple panels, they should be listed as (**a**) a Description of what is contained in the first panel and (**b**) a Description of what is contained in the second panel.

4. DISCUSSION

The conclusions drawn from the Fuzzy C-Means algorithm, which divides the wells in the Motoba field into three distinct categories (low, high and variable production wells), are fully consistent with those reported in the scientific literature. To illustrate, [22] identified a comparable set of low-producing wells due to unfavourable geological conditions. Similarly, the research of Gao et al. (2019) attests to the existence of a set of very efficient wells. Finally, [23] highlight the existence of wells with variable performance linked to technical variations or external elements, which correspond to our group 3.

4.1. Cluster analysis

After applying the Fuzzy C-Means algorithm, the data was successfully divided into three distinct clusters, each representing a group of wells with similar production characteristics. This clustering enables a more focused analysis of well performance. Cluster 1 represents wells with low to moderate production levels. The records in this group typically show values below the average, which may suggest less favorable geological conditions or sub-optimal operational management. Cluster 2 consists of wells with high production, characterized by above-average output. This optimal performance could be attributed to factors such as strategic well placement, efficient extraction techniques, or advantageous geological features. Meanwhile, Cluster 3 includes wells with highly variable production patterns, displaying both high and low values. Such fluctuations may be linked to seasonal influences, inconsistent reservoir behavior, or technical failures in the production process.

4.2. Evaluation of results/metrics

The silhouette index is a metric used to evaluate the algorithm's results. It is handy for determining the correct number of clusters to use in a clustering analysis. Silhouette indices (0.8730) confirm the quality of the clusters formed: A value of 0.8730 indicates significant separation between clusters. This means that the data points within each cluster are closer to each other than they are to points in different clusters [24][25]. This separation is crucial for the analysis, as it facilitates the identification of differentiated management strategies for each cluster.

4.3. Comparison of the two methods, FCM and FCM-C

4.3.1. Fuzzy C-Means (FCM)

Apply the FCM algorithm to this data using a fuzzy membership function to determine the production distribution of the wells. The model identifies three clusters: Cluster 1: Wells MOT_1ST and MOT_03ST (low production), Cluster 2: Well MOT_14 ST (medium production) and Cluster 3: Well MOT_06ST (high production). The comparison is detailed in Table 2.

Table 2. Clustering using FCM algorithm

Cluster	Well	Adjusted Average Production (barrels/day)
Cluster 1	MOT_1ST, MOT_03ST	462.5
Cluster 2	MOT_14ST	560
Cluster 3	MOT_06ST	715

4.3.2. Fuzzy C-Means Credibility (FCM-C)

Applying the FCM-C model allows the algorithm to incorporate credibility weights or uncertainties into the production data. For example, if we know that the production of Well MOT_1ST in 2020 was less reliable due to technical failures, we can adjust the weight of this data. Let's assume that incorporating uncertainty has led to a reassessment of good production as follows: Cluster 1: Wells MOT_1ST and MOT_03ST, but production is less influenced by the uncertain data. Cluster 2: Well MOT_06ST, which remains stable. The comparison is detailed in Table 3.

Table 3. Clustering using FCM-C algorithm

Cluster	Well	Adjusted Average Production (barrels/day)
Cluster 1	MOT_1ST, MOT_03ST	475
Cluster 2	MOT_06ST	710

Both models succeed in finding similar clusters. However, FCM can be influenced by outliers, whereas FCM-C produces a more robust assessment thanks to the inclusion of credibility. FCM shows a slightly lower average output for Cluster 1 due to a more substantial weighted influence of problematic data from Well MOT_1ST; FCM-C, by incorporating credibility, corrects this and rounds off the results. The comparison of results between Fuzzy C-Means and Credibility Fuzzy C-Means on oil well production data shows that FCM-C provides more reliable and interpretable results, taking into account the uncertainties in the data. This can help decision-makers make more informed strategic choices about oil production management.

5. CONCLUSION

In this study, we have summarised the main findings derived from the synthesis of key analytical results. Our contribution lies in both advancing theoretical understanding and offering practical value. Theoretically, the research introduces innovative methodologies that deepen insights into well-performance dynamics. Practically, it provides oil operators with tools to improve decision-making, optimise resource management, and foster more sustainable exploitation practices. These contributions highlight the significance of employing rigorous analytical methods that integrate technological advancements and best practices in data governance within the oil sector.

A central component of this research involved the application of Fuzzy C-Means (FCM) and Fuzzy C-Means Credibilistic (FCM-C) clustering algorithms to oil well production data. The analysis successfully identified distinct production clusters, each representing specific operational and geological characteristics. FCM proved effective in classifying wells flexibly, while FCM-C added robustness by handling uncertainty in noisy data. The results revealed that factors such as well depth, extraction technology, and environmental conditions significantly impact well performance. Based on these findings, we recommend targeted interventions for low-performing wells, strategic resource allocation, and the use of clustering outputs to guide investment planning. Furthermore, the clustering framework demonstrated strong potential for broader application in other areas of the oil industry, such as risk assessment and supply chain optimisation. These outcomes emphasize the importance of data-driven approaches in enhancing operational efficiency and long-term strategic planning.

REFERENCES

[1] W. S. Meddaugh, N. Champenoy, W. T. Osterloh, and H. Tang, "Reservoir Forecast Optimism – Impact of Geostatistics, Reservoir Modeling, Heterogeneity, and Uncertainty," in SPE Annual Technical Conference and Exhibition, vol. SPE Annual. 2011, p. SPE-145721-MS. doi: 10.2118/145721-MS.

[2] Frédéric SUR, *Apprentissage automatique Introduction*. 2025.

[3] M. Vilela, G. Oluyemi, and A. Petrovski, "Fuzzy Data Analysis Methodology for the Assessment of Value of Information in the Oil and Gas Industry," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2018, pp. 1–7. doi: 10.1109/FUZZ-IEEE.2018.8491628.

[4] E. Mehdizadeh, S. Sadi-Nezhad, and R. Tavakkoli-Moghaddam, "Optimization of fuzzy clustering criteria by a hybrid PSO and fuzzy c-means clustering algorithm," *Iran. J. Fuzzy Syst.*, vol. 5, no. 3, pp. 1–14, 2008.

[5] J. Yang, "Technology Focus: Data Analytics (October 2024)," *J. Pet. Technol.*, vol. 76, no. 10, pp. 88–89, 2024, doi: 10.2118/1024-0088-JPT.

[6] PENRECO, "Rapport sur Les puits producteurs au niveau de Pinda Supérieur du champ MOTOBA," 2021.

[7] P. E. Handbook, *General Engineering*, vol. I. 2007. doi: 10.1007/1-84628-080-x_16.

[8] J. Zhou, Q. Wang, C.-C. Hung, and X. Yi, "Credibilistic Clustering: The Model and Algorithms," *Int. J. Uncertainty, Fuzziness Knowledge-Based Syst.*, vol. 23, no. 04, pp. 545–564, 2015, doi: 10.1142/S0218488515500245.

[9] S. Liu and L. Xue, "The Application of Fuzzy Clustering to Oil and Gas Evaluation," in *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, 2008, pp. 644–647. doi: 10.1109/FSKD.2008.227.

[10] K. S. Al-Sultan and S. Z. Selim, "A global algorithm for the fuzzy clustering problem," *Pattern Recognit.*, vol. 26, no. 9, pp. 1357–1361, 1993, doi: https://doi.org/10.1016/0031-3203(93)90141-I.

[11] A. P. Dempster, "Upper and Lower Probabilities Induced by a Multivalued Mapping," *Ann.*

Math. Stat., vol. 38, no. 2, pp. 325–339, 1967.
[12] C. Paper, M. Sh, I. Technology, and A. N. Academy, "Opportunities and Challenges Big Data in Oil and Gas," no. August, 2016.
[13] B. Liu and Y. K. Liu, "Expected value of fuzzy variable and fuzzy expected value models," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 4, pp. 445–450, 2002, doi: 10.1109/TFUZZ.2002.800692.
[14] M. M. Khadidja, "Étude comparative des deux méthodes d'évaluation objective de la qualité d'expérience audiovisuelle par Logique Floue de type 1 ettype 2," Université Mohamed Sadik Benyahia de Jijel, 2021.
[15] Khezzan/abd arhmane and B. Mouhamed, "Modelisation flout basee sur les méthodes de clustering," Université 8 mai, 2015.
[16] P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, vol. 66, no. 2, pp. 191–234, 1994, doi: 10.1016/0004-3702(94)90026-4.
[17] John H. Holland, *Adaptation in Natural and Artificial Systems, An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. The MIT Press, 1992.
[18] H. Afghoul, F. Krim, D. Chikouche, and A. Beddar, "Design and real time implementation of fuzzy switched controller for single phase active power filter," *ISA Trans.*, vol. 58, pp. 614–621, 2015, doi: https://doi.org/10.1016/j.isatra.2015.07.008.
[19] A. C. Apelle, O. C. Olot, C. F. E. Aloigne, B. Spmi, and B. Marie, "Quantification des images couleur par un algorithme de classification crédibiliste Quantifi cation crédibiliste," pp. 1–4.
[20] Krim Amina, "Implémentation de l'algorithme Fuzzy C-Means sur des ensembles de données réelles," BADJI MOKHTAR – ANNABA UNIVERSITY, 2022.
[21] J. Su, S. Yao, and H. Liu, "Data Governance Facilitate Digital Transformation of Oil and Gas Industry," *Front. Earth Sci.*, vol. 10, no. May, pp. 1–11, 2022, doi: 10.3389/feart.2022.861091.
[22] F. Zhang, S. Deng, X. Wang, and L. Bai, "Why high porosity but low production wells occur in carbonate reservoirs: An explanation from the perspective of pore structure," *Interpretation*, vol. 11, no. 1, pp. 161–169, 2022.
[23] R. Zhang, M. Chen, H. Tang, H. Xiao, and D. Zhang, "Production performance simulation of a horizontal well in a shale gas reservoir considering the propagation of hydraulic fractures," *Geoenergy Sci. Eng.*, vol. 221, p. 111272, 2023, doi: https://doi.org/10.1016/j.petrol.2022.111272.
[24] L. Lovmar, A. Ahlford, M. Jonsson, and A.-C. Syvänen, "Silhouette scores for assessment of SNP genotype clusters," *BMC Genomics*, vol. 6, no. 1, p. 35, 2005, doi: 10.1186/1471-2164-6-35.
[25] G. Menardi, "Density-based Silhouette diagnostics for clustering methods," *Stat. Comput.*, vol. 21, no. 3, pp. 295–308, 2011, doi: 10.1007/s11222-010-9169-0.