



# A Review on “Management by Exception” Surveillance for Well Management: To maximize Oil Production

Azlinda Abdul Malik<sup>1</sup>, Mohd Hilmi Hasan<sup>2\*</sup>, Mazuin Jasamai<sup>3</sup>

<sup>1,3</sup>*Petroleum Engineering Department, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia*

<sup>2</sup>*Center for Research in Data Science, Computer and Information Sciences Department, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak, Malaysia*

*\*Corresponding author E-mail: mhilmi\_hasan@utp.edu.my*

## Abstract

The business processes and decisions of oil and gas operations generate large amounts of data, which causes surveillance engineers to spend more time gathering, and analyzing them. To do this manually is inefficient. Hence, this study is proposed to leverage on data driven surveillance by adopting the principle of management by exception (MBE). The study aims to minimize the manual interaction between data and engineers; hence will focus on monitoring well production performance through pre-determined parameters with set of rules. The outcome of this study is a model that can identify any deviations from the pre-set rules and the model will alert user for deviations that occur. The model will also be able to predict on when the well be offline if the problem keeps on persisting without immediate action from user. The objective of this paper is to present a literature review on the prediction and management by exception for the above mentioned well management. The results presented in this paper will help in the development of the proposed prediction and management model. The literature review was conducted based on structured literature review methodology, and a comparative study among the collected works is analyzed and presented in this paper.

**Keywords:** Anomaly detection; management by exception; prediction; well surveillance.

## 1. Introduction

Surveillance has become one of the main operational activities in oil and gas industry. Surveillance includes important activities such as monitoring of behavior, activities or other changing information. Not only that, from business point of view, surveillance is considered as an aggregation of multiple business practices configured for a particular asset. This coordination of practices intends to support and facilitate user decisions in ensuring that assets are managed to its potential in a safe and controlled manners. Meanwhile, from point of view of tools and processes, surveillance involves activities of observation, visualization, analytics, expert systems and process automation. These activities provide users with the insights of oil and gas asset performance, hence further actions can be taken for performance improvement (1).

In general, companies use production report containing production target, downtime and operating expenses; to be used for asset surveillance. All of these data are normally gathered manually. Unneland and Hauser argued that professionals could not put their focus on improving the quality of decisions because 60-80% of their time is spent on searching for and preparing data (2). Furthermore, in normal practice, wells are reviewed sequentially throughout the field on a calendar basis. This practice consumes a lot of time, with a potential of time lag between problem occurrence and diagnosis. On top of that, surveillance engineers normally in-charge a field comprising more than hundreds of wells; which requires monitoring on both of wells' surface and subsurface points. This scenario of huge amount of data handled by sur-

veillance engineers supports the argument saying that engineers occupy more of their time manually for monitoring the data instead of focusing on the highest value tasks. Consequently, opportunities for improved performance and risk reduction maybe missed (3).

Advancements in oil and gas surveillance has been in place, with the aims to solve the problems discussed above. Some companies have deployed monitoring centers that have the ability to provide real-time data visualization; enabling operation personnel to monitor key measurements on a real-time basis. Moreover, an integrated workflow automation model has been used to enable operational excellence in surveillance (4). These advancements have improved the way surveillance is conducted because they eliminate the passive and time consuming characteristics of the information delivery system. Through the real-time information as described above, surveillance has been able to leverage on analytics, expert systems, and process automation. As a result, a combined module of business or operational intelligence with automated technical calculations is generated. This means that, a new generation of hybrid solution in well surveillance is invented and utilized; which comprises data-driven methods such as management by exception (MBE), business intelligence (BI), and situational awareness (SA) (1).

In this study, we propose the adoption of MBE for monitoring well condition status. The main significance of this adoption is to provide engineers with best information before it is too late for remedial action. MBE is chosen due to the fact that it is the most fundamental and commonly used control and optimization technique. In its execution, MBE determines the occurrence of significant events, normally known as anomaly. An anomaly is an event or a value that deviates from a standard. In a normal practice, a

comparison between a set of target values (performance) that are anticipated a priori, and the actual measurements or values provided by surveillance is carried out. This comparison will identify a list of events that do not fall within the accepted set-point boundaries. This practice benefits the surveillance operation activities in a way that it reduce time consumption by removing the needs of manual data gathering and values comparisons (5). Overall, MBE carries out the monitoring duty upon oil wells by using data-driven well rate estimates (6). Despite this benefit, MBE, however, does not have the capability of explaining why a problem has occurred or the significance of the problem. Due to this fact, technical judgment by engineers is still required for them to act towards resolving the problem. Generally, a production well has a certain production trend; such as an oil well will have a certain production, water production and gas rates. Normally, this trend will not deviate significantly from these rates for another 1-3 months except that an anomalous event occurs at downhole or at reservoir level. For example, sudden increase in water or gas production will deter the production of oil. Significant deviation may cause a well to be closed (shut-in). During this shut-in period, an investigation and remedial action can be done to make the well be on-line again.

In nature, engineers need to deal with multiple workloads and handle large surveillance data, hence the above mentioned production anomalies of the wells are usually overlooked. In the existing technology, anomalies of well behavior are not detected earlier until the problem occurs. This causes the well to be shut-in, hence resulting in production loss. This will add up the time needed by engineers to manage the overall surveillance operations because they need to understand the problem and come out with remedial plan. By adopting MBE, any anomalies can be detected earlier, hence it allows for timely preventive and corrective actions. This will reduce the production lost, protect the hydrocarbon reserves and eventually increase the potential cost saving to the company.

Currently, our study is still at its infancy phase, which requires a comprehensive review on the literature. There are two research questions of this review; 1) What are the existing models of anomaly detection for well surveillance?, and 2) How do the identified existing models carry out the well surveillance tasks? Hence, the objective of this paper is to review the previous works related to MBE (anomaly detection) in well surveillance, specifically through unsupervised machine learning methods. The discussion on why unsupervised machine learning methods have become our attention is presented in section 1.1 of this paper. The focuses of this literature review are to identify existing unsupervised machine learning methods implementation in oil and gas well surveillance, and to discover how those methods carry out the surveillance tasks. This paper is organized as follows; section 2 contains the methodology to conduct this review, section 3 presents and discusses the results of the review, and section 4 concludes the findings and outlines the future works.

## 1.1 Anomaly Detection

As discussed above, MBE operates based anomaly detection procedure. Lazarevic et al. categorized anomaly detection methods into two approaches, namely supervised and unsupervised approaches (7). Supervised approach carries out anomaly detection based on labeled data set. For example, each instance of the data set must be labeled as either "normal" or "anomaly". In supervised approach, this labeled data set is required to train a learning algorithm. An advantage of supervised approach is that it is capable of detecting known anomalies with high accuracy. This is because the approach can simply detect any anomalies that have been known and learned through its labeled data set. A drawback of supervised approach is that it is unable to detect new types of anomalies that have not been learned before. On the other hand, unsupervised approach detects anomalies without having to have labeled data set. It carries out anomaly detection based on deviations from the normal model in data. An advantage of unsuper-

vised approach is it has the capability of detecting new anomalies that have not been learned before. This is done by detecting values that deviate from normal conditions. However, unsupervised approach suffers from a high degree of false alarm, hence unseen legitimate signatures may be treated as anomalies.

Chandola et al. (8) emphasized that data label is the most important information required for determining whether or not a single data point is anomalous. However, they argued that the job of data labeling is a daunting task, where in normal circumstances, getting anomalous labeled data points that cover all types of anomalous behavior is more difficult than getting labels for normal behaviors. Furthermore, anomalous behaviors are generally dynamic; new and unknown types of anomalies may arise so they do not have any specific label assigned to them (9). Labeling requires a lot of efforts and is done by human experts, normally. A number of previous works categorized anomaly detection approach into three, namely supervised, semi-supervised and unsupervised (8-10). As discussed above, the performance of supervised approach is greatly affected by the availability of normal and anomaly training data sets. In normal circumstances, the availability of anomalous data is much less than that of normal data in the training data set. This drawback will lead to a state of deceived high prediction accuracy, or generally known as imbalanced class distribution. Apart from that, the work also emphasized that obtaining accurate and representative labels, particularly for anomaly class is difficult, hence becomes another drawback for supervised anomaly detection approach (8, 10). They introduced another category of anomaly detection approach; semi-supervised. This approach comprises labeled data points only for normal class. Hence, semi-supervised approach is more applicable than supervised approach as it eliminate the difficult part of collecting anomalous data. Finally, the third approach, unsupervised, requires no labeled training data. Due to this characteristic, it is most widely applicable among the three approaches. To overcome the unsupervised approach's drawback of producing false alarm, it was proposed that the unlabeled data set is used in optimized learning algorithm so that the ability to detect anomalies can be made more robust.

Each of the three approaches discussed above can be implemented using some techniques. A number of previous studies categorized the techniques into three categories; statistical-based, data mining based and machine learning-based techniques (11, 12). The advantage of statistical technique is it can detect unknown anomalies because it does not require prior knowledge of anomalies. Statistical technique is however, prone to be manipulated by skilled attackers (prone to be hacked in the case of intrusion detection system). It is also difficult to identify the threshold of false positive and negative by using statistical technique, and it also requires accurate distribution in modelling the methods. In actual, accurate modeling cannot be done upon all behaviors; resulting performance reduction (11, 12). Meanwhile, machine learning based technique works based on learning from previous results. This means that it constructs a model with improved performance as a result of its learning process. This characteristic distinguishes machine learning technique with statistical technique that focuses more on understanding the data generation process. Machine learning technique also has high flexibility and adaptability (11). However, machine learning technique is resource expensive; it requires enormous amount of computational resources in order to produce good performance in anomaly detection. Furthermore, machine learning-based technique also depends highly on the assumption of the systems' behaviors. As for knowledge-based technique, its model is generally robust, flexible and scalable. Some methods of knowledge-based technique can keep the number of false positive at low level. This is because the models are trained by human experts rather than basing on previous results. However, getting and acquiring high quality knowledge is normally time consuming.

By practical, generating labeled data for oil well surveillance as in supervised approach is a challenging task. Engineers may not have

the knowledge of an anomalous state since it has never occurred before as the well is still in operation. On top of that, as discussed above, our study aims at helping engineers to reduce time spent on monitoring well production by the means of MBE. Data labeling such as required by supervised and semi-supervised approaches may require some amount of time. Hence, in this study, we will only be focusing on unsupervised approaches in detecting anomalies in surveillance data. As for the technique, we will be focusing on machine learning-based technique. This is because our aim is to develop predictive model for oil well surveillance by learning from previous data. Our study will not consider statistical technique because it requires accurate distribution model in order to generate effective performance. Our study will also not cover knowledge-based technique as our focus is to produce a model that detects anomalies through data-driven approach.

## 2. Methodology

To conduct this review, we applied a structure literature review method. It consisted of five phases of activities described as follows:

### 1. Framing questions for a review

The main objective of this phase is to firm up the problems to be addressed in this research. The activity involved is defining problems into set of questions in the form of clear, unambiguous and structured. The results of this phase have been presented in section 1 of this paper, which can be summarized as follows:

*Problem* – Well surveillance operation/management involves big data management, hence time-consuming. Also, the manual way of analyzing the data for discovering/predicting well problems based on anomaly detection is inefficient and not effective (may result in inconsistency of decisions, repeating doing similar tasks and poor knowledge management).

*Research question (RQ)* – the research questions of this review are as follows:

RQ1: What are the existing models of anomaly detection for well surveillance?

RQ2: How do the existing models identified in RQ1 carry out the well surveillance tasks?

*Objectives and scope* – as stated in the previous section, our review focuses on unsupervised machine learning technique in anomaly detection of well surveillance. Our scope is to review the three main unsupervised machine learning techniques, namely K-means, C-means and Self-organizing maps (SOM). The objectives of this review are; 1) to identify existing unsupervised machine learning methods implementation in oil and gas well surveillance, and 2) to discover how those methods carry out the surveillance tasks.

### 2. Identifying relevant works

This phase involved extensive work on searching for relevant works. The search criteria were relaxed, for example the relevant works were searched among multiple resources and no language restrictions. The search was based on the questions identified in previous phase. Any reasons for inclusion/exclusion of works was recorded. In this review, we searched for relevant works from the two main indexing databases, namely SCOPUS and ISI Web of Science. The keywords (KW) used for the searches are:

KW1: oil production AND c-means

KW2: oil production AND k-means

KW3: oil production AND self-organizing maps

KW4: well surveillance AND c-means

KW5: well surveillance AND k-means

KW6: well surveillance AND self-organizing maps

KW7: oil well surveillance AND self-organizing maps

KW8: oil well surveillance AND c-means

KW9: oil well surveillance AND k-means

The search was carried out for “all years”, and it collected a total of 156 manuscripts. Figure 1 shows the number of manuscripts collected per each of the used keywords.

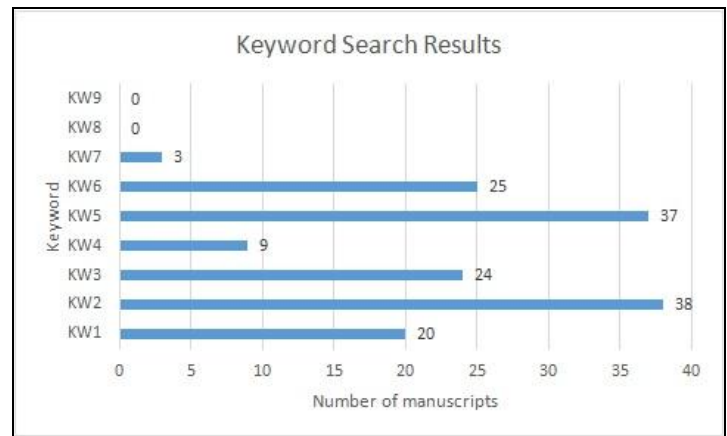


Fig. 1: Number of collected manuscripts from search

### 3. Assessing the quality of studies

The output of the second phase was a huge amount of information containing 156 manuscripts of relevant works. In this phase, those works were further refined and filtered based on stricter search criteria. This phase began with defining these new level of filtering criteria so that it can be used to determine whether or not any of the works will be reviewed. Firstly, the manuscripts were filtered to remove redundancy. Searches done in the previous phase were carried out using a number of keywords and taken from two different sources. Some of the manuscripts appeared in the search results of different keywords. This means that, redundancy occurred where a single manuscript may come out in the search results list of two or more keywords. Hence, these redundant manuscripts were removed from the list of collected manuscripts. Next, the manuscripts were filtered according to their title. Only relevant titles remained in the list. As a result, the number of manuscripts was further reduced to 37. Then, these manuscripts were filtered based on their abstract. The remaining number of manuscripts is 15.

### 4. Summarizing the evidence

The collected manuscripts of the related works were reviewed, synthesized and analyzed. The results of this activities are presented in this paper.

### 5. Interpreting the findings

In this phase all issues highlighted in the previous four phases will be consolidated and solved. The results of synthesis and analysis from previous phase will be converted into a guideline for effective implementation of well surveillance using unsupervised machine learning-based methods.

## 3. Literature Review

Bao and Guan proposed a model to predict crude oil output and main factors that affect those output (13). The model carried out the prediction task through classifying of output data according to related factors. A hybrid of data mining algorithms, namely Rough Set and C4.5 decision trees were applied to the model. This hybrid implementation is critical to add value to the traditional C4.5 decision tree algorithm. The addition of rough set into the model helped in attribute reduction; reducing the needed number of data by the model to generate a decision tree. This was where K-Means clustering algorithm came into the picture. K-Means was used for a data discretization, which was required to be performed before attribute reduction took place. K-Means involved in this data pre-processing stage by reducing the large number of data, and categorizing them into clusters.

Foroud et al. studied on production optimization under the state of geological uncertainty (14). The geological model was identified as the main source of uncertainty that diminished the viability of simulation results. In their study, Kernel K-Means (KKM) clustering method was used to select a representative subset of geological models. The clustering process was proposed due to the huge

computation time required to evaluate all geological realizations for oil production optimization. The clustering selected a number of representative reservoir models with distinct flow characteristic, from a very large set of models. The result produced by the KKM clustering was a set of clusters; each cluster containing similar realizations in terms of flow response characteristics. Their experimental results showed that KKM had successfully made the production optimization more efficient due to smaller number of simulation executions during evaluation. Their testing on Brugge field with 40 geological realizations had resulted in 9 realizations only by using KKM, with the representation of the whole set was still preserved.

Awoleke and Lane proposed a data analysis of Barnett Shale water production using conventional statistical techniques and neural networks (15). A data set of around 11,000 completions was used in their study, which firstly being clustered using SOM and K-Means algorithms. Then a feed-forward neural network was used to predict wells' water production. In their experiments, SOM and K-Means had resulted in the data set being clustered into two main clusters. SOM was used to see how data were clustered, while K-Means's purpose was to determine the number of clusters for the learning algorithm to carry out its tasks.

Hu et al. reported on their work to predict crude oil production (16). In their work, a new method applying transaction Boolean matrix and support count matrix (TB-SCM) was proposed. The method produced association rules of crude oil production based on minimum support threshold and minimum confidence threshold. In their proposed method, K-Means was also utilized in the data pre-processing part. This involved discretization process, which huge amount of data were categorized into clusters.

Jonge et al. reported on the need of an automation task to support automated data processing, event detection and user notification for reservoir surveillance (17). By applying data mining methods, this proposed expert system was able to analyze real-time data as well as optimize field production. Their work was mainly a conceptual work, where they discussed the elements required to be included in the expert system. Among the modules proposed were predictive functionalities using machine learning algorithms and also outlier detection using data mining algorithms. Another work reported in (18) also argued that conventional reservoir surveillance tools do not have the capability of providing high frequency real-time data to engineers. With gigantic amount of data at hand to handle, engineers and knowledge workers could not spend effective time for a higher level work to discover the great potential of those data. Hence, in this work, they proposed another conceptual approach, incorporating neural networks and simplified reservoir models to convert the data at hand into useful information.

Mota et al. emphasized that automatic surveillance system is critically important due to manual capturing of data from huge number of gas fields are challenging in nature (19). In the work, they introduced a new workflow for this surveillance management. The workflow comprises three modules; early detection of liquid loading in gas wells module, estimation and reconciliation of production rate module, and update of well decline curve module.

Popa et al. proposed a well data logs-based clustering method for optimal design perforation (20). The work proposed the use of fuzzy clustering methods to identify optimal perforation strategy. They used C-Means algorithm to cluster log data of wells from different areas of a reservoir. The clustering classified the perforation jobs as either "good" or "bad" regions. The experimental results showed that the proposed method of using log data for the wells, and applying C-Means clustering analysis can effectively identify rock intervals with similar properties.

Grieser et al. presented their investigation on Barnett completion strategy based on wells' completion (21), reservoir and production data sets. One of their investigations was to identify reservoir and stimulation characteristics of well productivity. For this purpose, they utilized a data set of 393 wells comprising information on perforation, frac, log and production. To carry out the identification task, they implemented data clustering with the aim to associ-

ate those data with similar behaviors. They implemented this task by using SOM algorithm. The results showed that the arrangement of data generated by SOM were consistent with that of generated by reservoir-engineering principles.

Cremaschi et al. and Shin and Cremaschi argued that estimating flow velocity in oil and gas pipelines is complex in nature (22, 23). This could be seen from the various outcomes from different prediction models though they were supplied with the same inputs. Their works proposed an approach to overcome this issue through data clustering, model selection and cluster identification techniques. The approach started with partitioning data into clusters based on similar transport characteristics. This was done by K-Means algorithms. The clustering results were then validated using clustering validation method, namely Bayesian Information Criterion (BIC). This validation was required to ensure that the clustering was performed in optimal way. The experiment results showed that the proposed approach performed significantly well by reducing the errors of the threshold velocity prediction by certain magnitudes.

Ding et al. proposed a collection of methods to determine the level and characteristics of high-permeability zones in water flooding reservoirs [24]. The obtained data of high-permeability zones is huge in size as there were many variables involved during the evaluation and data gathering activities. In the work, they divided the data into three types; 1) absence of high-permeability zone, 2) presence of light high-permeability zone and 3) presence of large high permeability zone. This clustering into three categories were done by using C-Means method.

Cui et al. proposed a method for improving oil recovery for high water-cut multilayered reservoirs (25). They applied layer regrouping which did the layering into several sets of production based on physical properties and percent of layer's recovery at high water-cut stage. This work proposed pseudo flow resistance contrast as an index to regroup the subjects. They applied K-Means clustering method to deliver this layer regrouping jobs.

Liu et al. studied on the measurement of water content in crude oil production (26). Water content is an important parameter, and measuring it accurately its level in crude oil can optimize production parameters and improve oil recovery rate. The work proposed General Regression Neural Network because it has advantages such as in approximation ability, capacity of classifying and speed of learning. They measured some parameters that affect the water content in crude oil by using a sensor-based technology. The gathered measurements were then processed using K-Means, before it was used for prediction procedure.

Singh et al. proposed uncertainty quantification measure for forecasted oil recovery (27). This was done using dynamic model ranking. The work aimed at quantifying uncertainty in the distribution of reservoir parameters governing reservoir connectivity. The work also aimed at predicting production under the state of uncertainty. One of the methods involved in their proposed model was K-Means clustering algorithm. K-Means was used to identify key reservoir models. K-Means managed to reduce the number of history-matched reservoir model from 40 to five models.

### 3.1 Discussion

As stated in section 1 and 2, this review focuses on anomaly detection as one type of MBE method in detecting the likelihood of well problems based on the behavior of production data. The production data that we are focusing on are oil rate, water rate and water cut, gas rate and gas-oil-ratio (GOR). Based on the summary of the reviewed manuscripts in Table 1, none of them presented on the surveillance production data as mentioned above. Hence, it is concluded that there is a gap in this area of research, which can provide significant contributions especially in overcoming the problems that we mentioned in section 1 and 2 of this paper.

Another two objectives of this paper are to identify existing unsupervised machine learning method deployed in well surveillance, and to discover how these methods perform their surveillance

tasks. The summary of this literature review regarding these two objectives is shown in Table 1. Only two of the manuscripts, i.e. Jonge and Stunder, and Jonge et al. had proposed the used of machine learning methods for surveillance (17, 18). They proposed neural networks for learning from historical data and the use of data mining algorithms for outlier detection. However, both of the manuscripts only present conceptual model of what automated reservoir surveillance must possess to ensure effective and efficient management.

Apart from that, the utilization of the three machine learning methods that we are proposing, namely K-Means, C-Means, and SOM, in the reviewed works is also presented in Table 1. It can be

seen that C-Means was specifically applied for data clustering. The outcomes of C-Means are clusters of data with fuzzy categorization using linguistic values such as “Good” and “Bad”. Meanwhile, SOM was also utilized in data clustering and seeing how data are clustered. K-Means was applied in majority of the collected manuscripts. The main purposes of applying K-Means were for data clustering and data discretion. The latter was mainly used with the intention to improve the time taken for each model to carry out their respective tasks by reducing the number of data through clustering.

**Table 1:** Summary of the reviewed manuscripts

| Source                            | Investigate well production data? | Purpose  | Use of machine learning method  |
|-----------------------------------|-----------------------------------|--|---|
| (Bao & Guan, 2016)                | No                                | Oil production prediction                              | K-Means – data discretion   |
| (Foroud, et al., 2016)            | No                                | Optimize production by looking at geological models    | K-Means – data discretion   |
| (Awoleke & Lane, 2011)            | No                                | Well water production prediction                       | SOM – see how data are clustered<br>K-Means – determine number of clusters<br>Neural Network – prediction |
| (Hu, et al., 2015)                | No                                | Oil production prediction                              | K-Means – data discretion   |
| (G. d. Jonge, et al., 2003)       | No                                | Conceptual model for reservoir surveillance            | Data mining - Outlier detection   |
| (G. J. d. Jonge & Stundner, 2002) | No                                | Conceptual model for reservoir surveillance            | Neural network - learning   |
| (Mota, et al., 2007)              | No                                | Workflow for surveillance management                   | None  |
| (Popa, et al., 2015)              | No                                | Perforation strategy optimization                      | C-Means – cluster log data  |
| (Grieser, et al., 2008)           | No                                | Overall well investigation                             | SOM – data clustering   |
| (Cremaschi, et al., 2015)         | No                                | Flow velocity estimation in pipelines                  | K-Means – data clustering   |
| (Shin & Cremaschi, 2014)          | No                                | Flow velocity estimation in pipelines                  | K-Means – data clustering   |
| (Ding, et al., 2015)              | No                                | Investigate high-permeability zone                     | C-Means – data clustering   |
| (Cui, et al., 2016)               | No                                | Oil recovery improvement for high water-cut reservoirs | K-Means – cluster/group the subjects  |
| (Liu, et al., 2009)               | No                                | Measurement of water content in crude oil              | K-Means – data preprocessing for prediction   |
| (Singh, et al., 2014)             | No                                | Measurement for forecasted oil recovery                | K-Means – data discretion   |

## 4. Conclusion

The conducted review has shown that our problem statements actually support the previous findings, hence the solutions to it, as proposed by our research will provide potential significant contribution to the field. The problems are well surveillance is time-consuming to be done manually, and manual analysis of surveillance production data degrades its efficiency and effectiveness. There is a gap exist in this area of research; applying machine learning methods for analyzing surveillance production data, as well as predicting the health of the wells based on the analysis. From the conducted review, we also can formulate a conceptualization that a comparative study among our three proposed methods; K-Means, C-Means and SOM, in detecting anomaly in surveillance production data is significant to be conducted. The study will imply a proactive measure, faster response time for well intervention, minimize well downtime, safeguard the production as well as contribute to cost saving. The study will also imply efficient surveillance operations through ensuring the engineers to focus on analysis rather than identification, eliminating of repetitive data gathering and reformatting tasks, ensuring consistency and repeatability of evaluation and better knowledge management. Our next direction is to conduct this comparative study, hence will produce the analysis and prediction model of surveillance production data as specified in the objectives of our study.

## References

- [1] Lochmann MJ. The Future of Surveillance. Paper presented at the The Intelligent Energy Conference. 2012.
- [2] Schipperijin P, Thavarajah R & Simonato A. Automated, “By Exception” Well Surveillance: A key to maximizing oil production.

- Paper presented at the SPE Digital Energy International Conference and Exhibition. 2009.
- [3] Nikolaou M, Cullick AS & Saputelli L. Production Optimization – A Moving-Horizon Approach. Paper presented at the SPE Intelligent Energy Conference. 2006.
- [4] Sankaran S, Olise M, Meinert D & Awasthi A. Realizing Value from Implementing i-field TM in Agbami - A Deppwater Greenfield in an Offshore Nigeria Development. SPE Economics and Management. 2011; 31-44.
- [5] Yero J & Moroney TA. Exception Based Surveillance in Shell. Paper presented at the Intelligent Energy Conference and Exhibition. 2010.
- [6] Agrawal A, Leeuwen PV, Demirbas B & Hajj J. Surface Management of Gas Breakthrough in Thin Oil Rim Waxy Reservoir. Paper presented at the Intelligent Energy International Conference and Exhibition. 2016.
- [7] Agrawal A, Lie J, Leeuwen PV, Adun D, Briers J, Pei-Huey P, Snieker A & Enns R. Leveraging Data-driven Exception Based Surveillance to Maximize Returns in Times of Low Oil Prices. Paper presented at the ntelligent Energy International Conference and Exhibition. 2016.
- [8] Chandola V, Banerjee A & Kumar V. Anomaly Detection: A Survey. ACM Computing Surveys. 2009.
- [9] Lazarevic A, Ertoz L, Kumar V, Ozgur A & Srivastav J. A comparative study of anomaly detection schemes in network intrusion detection. Paper presented at the SIAM Conference on Data Mining. 2013.
- [10] Kaur R & Singh S. A survey of data mining and social network analysis based anomaly detection techniques. Egyptian Informatics Journal. 2016; 17:199-216.
- [11] Garcia-Teodoro P, Diaz-Verdejo J, Maria-Fernandez G & Vazquez E. Anomaly-based network intrusion detection: techniques, systems and challenges. Computer and Security. 2009; 28(18-28).
- [12] Patcha A & Park JM. An Overview of anomaly detection techniques: existing solutions and latest technological trends. Computer Networks. 2007; 51: 3448-3470.

- [13] Bao X & Guan X. A Method of Predicting Crude Oil Output Based on RS-C4.5 Algorithm. Paper presented at the 3rd International Conference on Information Science and Control Engineering. 2016.
- [14] Foroud T, Seifi A & AminShahidy B. An Efficient Optimization Process for Hydrocarbon Production in Presence of geological uncertainty using a clustering method: A case study on Brugge field. *Journal of Natural Gas Science and Engineering*. 2016; 32:476-490.
- [15] Awoleke O & Lane R. Analysis of Data From the Barnett Shale Using Conventional Statistical and Virtual Intelligence Techniques. *SPE Reservoir Evaluation & Engineering*. 2011; 14(5).
- [16] Hu H, Zhang R & Guan X. Application on crude oil output forecasting based on TB-SCM algorithm. Paper presented at the 5th International Conference on Electronics Information and Emergency Communication (ICEIEC). 2015.
- [17] Jonge Gd, Stundner M & Zangl G. Automated Reservoir Surveillance through Data Mining Software. Paper presented at the Offshore Europe, Society of Petroleum Engineers. 2003.
- [18] Jonge GJd & Stundner M. How Routine Reservoir Surveillance with Neural Networks and Simplified Reservoir Models can Convert Data into Information. Paper presented at the European Petroleum Conference. 2002.
- [19] Mota MF, Campos OM & Gonzalez HME. Automatic Surveillance System for Large Gas Fields With Multifrequency Measurements. Paper presented at the SPE Annual Technical Conference and Exhibition. 2007.
- [20] Popa A, Hao M, Cassidy SD, Cover A, Shepherd D & Jikich S. Clustering-Based Optimal Perforation Design Using Well Logs. Paper presented at the SPE Eastern Regional Meeting. 2015.
- [21] Grieser WV, Shelley RF, Johnson BJ, Fielder EO, Heinze JR & Werline JR. Data Analysis of Barnett Shale Completions. *SPE Journal*, 2008; 13(3).
- [22] Cremaschi S, Shin J & Subramani HJ. Data clustering for model-prediction discrepancy reduction – A case study of solids transport in oil/gas pipelines. *Computers & Chemical Engineering*. 2015; 81: 355-363.
- [23] Shin J & Cremaschi S. Data Clustering for Uncertainty Reduction – A Case Study of Solids Transport in Oil/Gas Pipelines. *Computer Aided Chemical Engineering*. 2014; 34: 735-740.
- [24] Ding S, Jiang H, Wang L, Liu G, Li N & Liang B. Identification and Characterization of High-permeability Zones in Waterflooding Reservoirs With an Ensemble of Methodologies. Paper presented at the SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Bali, Indonesia. 2015.
- [25] Cui C-Z, Xu J-P, Wang D-P, Liu Z-H, Huang Y-S & Geng Z-L. Layer regrouping for water-flooded commingled reservoirs at a high water-cut stage. *Petroleum Science*. 2016; 13(2): 272-279.
- [26] Liu C-L, Niu H-F, Wang J-Q & Sun X-W. Research into prediction model of water content in crude oil of wellhead metering based on General Regression Neural Network. Paper presented at the IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2009. 2009.
- [27] Singh AP, Maucec M, Carvajal GA, Mirzadeh S, Knabe SP, Al-Jasmi AK & Din IHHE. Uncertainty Quantification of Forecasted Oil Recovery using Dynamic Model Ranking with Application to a ME Carbonate Reservoir. Paper presented at the International Petroleum Technology Conference. 2014.