International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-1, May 2020

Applications of Data Mining in Hydrocarbon Exploration, Constraints on Geology and Petroleum Reservoir

Rajesh Kanna, Sivasankar P, Kalpana S

Abstract*: *In the present scenario, low oil prices and explorations shave triggered energy industries to look into cost depression of supply chains vibrantly. Advanced and new technology had been identified and experimentally considered for new hydrocarbon exploration prevailing data mining. Geoscience and geophysical problems are dominated as data obtained for mining process is enforced for geology and reservoir issues and properties. Our present research deals about understanding the geological problems clearly analysed in the literature and with few experiments, a variety of data mining methods for the outcome has been concluded, which provides solution for a betterment understanding about gas and oil exploration with that of data miners and the geoscientists. Collection of data for various hydrocarbon wells has been addressed with seismic surveys, for identification of source, segregating and forecasting using iteration methods and neural networks had been discussed for betterment exploration of new wells without any constraints.

Keywords*: *Hydrocarbon, CBM, Recovery, Shale, Geoscience

I. INTRODUCTION

The present two techniques utilizing stabilization and profit derived from corporate and finance perspective; in the form of revenues are improved or the costs are deceased. It is to be noted that there was steep decline in oil prices since few decades and more specifically during mid of 2010, till recently, reduction in cost is and more attention are focussed in the energy industry. Oil and gas industries, particularly the logistics and supply chain have four main phases: exploration, refining, transportation and distribution are the major phases in oil industries [2]. The process of exploration and exploitation is the technique of reducing the expenses so that the commercial oil and gas deposits which are due to exploratory activities carried out using infrared satellite, radar and microwave survey. Aerial imaging, geochemical predictions and geobotany are the methods for identifying source of hydrocarbon. Moreover, surveys such as aerial-magnetic and electro-magnetic gravity surveying, seismic surveying along with exploratory wells (generating well logs) are recent advancements in prediction. The ultimate idea is to improve the accuracy with zero or negligible errors, so that results generated are reliable. It has to be noted that most valuable and crucial results might be incomparable with activities in very limited time so that reduction in cost and fast exploration leads to recovery of hydrocarbon [5]. In present new technology termed as data mining applied in oil and gas exploration has created minimal finance and brought about considerable gains. Data mining has the potential of permitting analysis for more reliable results so that an accurate decision on high volumes of data gathered can be utilized [6]. On the other hand, data mining has the most critical potential which lies in the ability to extract most value data in the form of larger volumes that can be used for a wide variety of data by utilizing new discoveries. Most of the academic institutions, industries related to hydrocarbon industries and business analytics service providers, are contributed on real time applications. The purpose of this article is to provide a detailed update on applications and production of hydrocarbon utilizing data mining, which tries to create a common understanding between the geological experts and data miners. Perhaps considering two major categories of geoscientific problems in which mining of data is applied and, moreover in brief, reviews a few consolidated and interesting works in each of them. The predominant and promising area that includes properties and issues on geology and reservoir so that former, data mining assists in creating a authenticated images so that betterment of most heterogeneous geological structures, in the reservoir. At a given point of time and depth by exploiting attributes of seismic surveying data (seismic attributes); while in the latter, this technology helps in predicting about accuracy of the reservoir along with the characteristics like lithology using both seismic well logging data. The second part of the paper is organized so that it will cover and update about the structural geology and reservoir issues prevailing to their properties. Part of the section that reviews the most relevant works in chronological order. The earlier research was addressed to initiate since a long, whereas none has come up with a fruitful suggestion so that geophysical and reservoir problem were analyzed and addressed. Hence, we concluded the article by lending crisp suggestions for enhancement in the particular area.

Manuscript received on April 02, 2020.
Revised Manuscript received on April 21, 2020.
Manuscript published on May 30, 2020.
* Correspondence Author
Rajesh Kanna*, Department of Petroleum Engineering, Academy of Maritime Education and Training (AMET), Chennai, India. Email: rajeshkanna.a@ametuniv.ac.in
Sivasankar P, Department of Petroleum Engineering, Indian Institute of Petroleum and Energy, Visakhapatnam, India. Email: sivasankar.petro@iipe.ac.in
Kalpana S, Department of Physics, Academy of Maritime Education and Training (AMET), Chennai, India. Email: kalpanaphy@ametuniv.ac.in

II. STRUCTURAL GEOLOGY - ISSUES

Typical hydrocarbon process, requires consistent and accuracy in top, bottom and lateral mapping of the specific structures of the rock matrix which may potentially yield hydro carbons. Hence it should be ensured that these positioning, for the possible availability of hydrocarbon, assists in determining the depth and thickness of the crude layers considered as economic factor.

Eventually, consideration for accurate mapping, utilizing the most optimized formation of the well is the priority. In other means, the multi-layered reservoirs are targeted for enhanced production, where sand and shale layers are covered, may be out from a combined single well that will reduce the cost. To the best of authors knowledge [8] studies shows that initial and proper application of various analysis in crude oil exploration. Usage of non- linear mapping and a non-parametric scheme to estimate the density function with that of a function of distribution for data in clusters. [8] Advancement for mapping such as laminated sand, shale sequences were analyzed with the help of seismic data at Gulf of Mexico's reservoir. It is correlated with 10 seismic studies, using Hilbert -Wavelet transformation for determining the thicknesses from the exploratory wells [6]. Some of the back propagation (BP) and neural network regressor had been deployed for predicting the effectiveness in the form of thickness in Chinese oil field. They are introduced to discussion on various seismic attributes to propose a dynamically key parameters and neural network regressor for predicting the thickness of the deposit in determining a nonlinear genetic algorithms (GAs), annealing stimulation (SA) and BP neural network [10]. One of the disadvantages of fusion deposit with similar technique is eventually more robust neural network. Usage of GA designed the network structure and a hybrid SA-BP is employed for prediction of weights optimally based on the network [10]. Since the two steps has to be repeated continuously for collective seismic analysis, training set with available data combination is done until arriving optimal network. This particular method had been successfully applied for an oil reservoir situated in Central China. Further, it is optimized that network must be made up of attributes with eight seismic surveys, out of 60 available ones that were not deviated to the known reservoir depth and layer of 18 wells. In contradiction, a comparison study is made with other methods in the literature such as algorithm of GA-SA, proves the superiority of this data, which is significant with a smaller mean squared error (MSE) that evaluated with the available test data [1]. Compared to traps, the most vibrant for storage and transportation of hydrocarbons are the faults. Hence, precision in imaging of faults which may encounter displacements will determine their corner side of the fault in which rocks moves towards upward direction and downthrown sides is less important [9]. In current scenario, detection of fault is mainly comprised of an iterative workflow for seismic analysis so that methodology with a BP neural network as an image processing technique is fairly adapted. This network may be transformed most of the attributes for a new attribute with the probability of the fault for the studies for example fault and non-fault positions. Some of well performing and trained network is employed for seismic section to reproduce a new section that has high probability for the presence of hydrocarbon sources. Adequate idea on fault orientation with shape can be processed for improving the resolution of the targeted faults for the declaration of hydrocarbon. It is appreciated for few networks that gave accurate idea 3D seismic section that has been characterized as fault and non-fault. Followed by application of image zooming and removal of filters with predeveloped algorithms were used to extract fault planes in a viable controlled method, that demonstrated the added value, for fault continuity and contrast, of the combination of seismic attributes rather than a single attribute. Hence, some of the reservoir which mainly comprised of complex fault system were subjected to test in parts of Texas, resulted in par with the confirmation of the method so that potential to bifurcate noise to have broad idea of the faults. This method is considered to be instrumental in both exploration and evaluations for hydrocarbon in economic quantity and research projects. In hydrocarbon industries, it is used as one of the most potential risk reduction tools.

III. PROPERTY OF HYDROCARBON RESERVOIR & ISSUES

Previous discoveries on hydrocarbon deposits for finding structural traps were based on simple seismic surveys, whereas now different type of trap come in to picture and deemed to be important, for stratigraphic mapping. Marr has given a clear information and difference between the two as: Structural trap and permeability fold. The properties of rock are that the ability for fluids to flow and percentage of void space in the rock. These are termed as stratigraphy. The analysis of the stratigraphic structure will deduce possible lithologies and probable presence of crude hydro carbons. It is denoted as litho-facies analysis that provides data about horizontal trap which provides human interaction with access can be received using pattern recognition that are automated [6]. The variation in clusters of the rock matrix and seismic section center point, are considering attribute of seismic analysis which eventually results in groups of lithofaciel ideology. The most important and conventional method called ITERATE, is a technique which provides data using redistribution operator in making clusters by a common iterative analysis for clustering lithofacies in seismic techniques. More authentic results, can be achieved by quantifying the micro quality clusters which are generated, utilizing the two probabilistic measures such as cohesion and distinctness i.e., intra and inter class dissimilarity respectively. In order to determine the incremental predictability of value of the objects in the data set cohesion is used whereas distinctness is used to analyze a probability match measure only. The methods for evaluation measure like MSE are found to be more probabilistic as ITERATE method of cluster definitions are certainly meant for probability analysis. It has been approved and accepted that the application of unsupervised self-organizing maps neural networks in lithofacies analysis is slightly compatible with that of the ITERATE method. For determining the petroleum reservoir in India, clustering of the lithofacies is employed as two stage process. The first stage involves the tracing of shape into horizontal interval of the 3D seismic section of the particular fields that are further analyzed by SOM where a series of model traces are represented by difference in the observed shapes.

The second stage is the implementation of each trace in a particular and given interval of time is compared to all traces that had been assigned to the one the accurate correlation. A colored seismic section with similar lithofacies alike area is obtained. Further, the usage of selective Generative Topographic Mapping technique for a given lithofacies of clustering formation is reported in Texas Panhandle, district in USA. GTM is a method widely for non-linear projection so that probabable representation of the data vectors were drawn interest prevailing lower dimensional latent space that features and describe the data of the 3D seismic section of geometry and texture inversion seismic which involves the coherence, gray-level co-occurrence matrices (GLCM) entropy coherent energy, peak frequency, heterogeneity, p-impedance, and convergence reflectance. Brazil reservoir followed two clustering techniques were used for horizontal and vertical geological lithologic analysis for prediction and presence of hydrocarbon materials. Moreover, SOM and c-means are other techniques that were being imposed. It is further reported that three attributes such as cosine phase Hilbert Transformation and amplitude were used in order to prevailing 3D seismic section of the fault and trap. Algorithms for 2-10 partitions were run to screen the optimized seismic section and arrived with three attributes for optimal value. Recursive feature elimination (RFE) which is an algorithm has been developed for the linear SVM and computes the weights for variable attributes. Since this feature for smaller weights process is continued until no useless attribute remains. Also, a linear SVM is trained to predict the accurate presence of hydrocarbon deposits. Radial basis function (RBF), spline and polynomial considered for further expressions. Expanded features are made to cover more related classification results. SVM is used for used for the selection of unpaid attributes. The problem of optimization with respect to regression of a convex formulation by applying Lagrange multipliers found to be more ineffective. SVM regressor found to perform better with that of a lower MSE in evaluation process. In exploratory wells one of the most important issue in determining permeability and production in the identified deposit in more economic feasibility method. A method is proposed for assessing permeability by analyzing three well logs that are not limited to bulk volume of induction log response and gamma ray. With the aid of permeability in core samples derived from several exploratory wells of a hydrocarbon deposit in US, a BP neural network regressor is applied. In initial stage regression neural network is generally applied for identifying network design and to optimize. A modelled estimation of permeability based on well logs such as shallow resistivity, flushed zone resistivity, bulk density, neutron porosity, photoelectric factor, gamma ray, deep resistivity and SVM regression problem; permeability and the relative well logs were more used for an exploratory well in a reservoir offshore Russian province [8].

Shear wave velocity estimation in wells is a challenging task are important in determining the petrophysical studies crude availability. Most of the identified probable wells found out that determined velocity is appropriately by broken instruments, worst well conditions and limited equipment only. It is further suggested that these methods for identifying BP neural network and shear velocity only. It has been reported that access can be compiled to core samples velocity for a sand stone formation at Iran using the well logs of the sand stone oil field with neutron porosity; and; neutron porosity, bulk density, transit time with X and Y coordinates respectively. Post evaluation showed that the neural network predicted that shear wave velocity slightly better, to that of the multi regression method. It has been reported previously that few of the companies who used few correlations mainly the government of US and Russia performed certain data mining in-order to determine the effectiveness of water injection in enhancing hydrocarbon recovery Nebraska's petroleum field. Attributes on Eight locations were identified as common. Important data for crude sample such as oil water ratio been extracted from the well and the average porosity in the deposit beneath. Although the average permeability, and the initial well pressure were continuously monitored for real time data, a neural network regressor is induced to develop and correlation to that of previous water injection data [4].

Few neural network authorities namely Bayesian regulation follows stratigraphic constraints which were recommended to reconfirm for the prediction of shear velocity. Shear wave velocity with vicinity data points from nearly 30 exploratory wells in an oil prospect onshore India and Kuwait were analyzed by data mining. The effectiveness was termed as primary cum secondary oil extraction (P/S) ratio [7]. The usage of fuzzy logic for ranking for determining the important data attributes, out of 30 available ones from existing water flood data. Compared to the BP neural networks a developed neural network has provided smoother output which further imposes stratigraphic constraints by segregating heavy and bulk data into geologic formations [9]. Successful results in faster training timing had evolved NNBR with multi regression method for effective output followed in various parts of the universe.

IV. RESULTS

Surveys and Analysis are conducted in order to predict the hydrocarbon source utilizing the available raw dataset collected. REF along with SVM computed differential attributes in prediction of geological and reservoir nature and source of hydrocarbon (HC). Our results obtained gave an authentic HC source with the help of shear wave velocity analysed by data mining. BP neural network and iteration methods included pre-processing the dataset, the obtained result is improved to 85% accuracy in prediction of hydro carbon source with reservoir heterogeneity.

V. CONCLUSIONS

Present paper provides an overview update on data mining applications in petroleum exploration sector. We addressed two major categories of geoscientific heterogeneities and with data mining efforts under structural reservoir and geological issues. Based on the surveys and analysis we had elaborated surveys using seismic, 3D, Iteration and neural network for prediction of source of a well so that exploration is performed with correctness.

REFERENCES

1. Zhaojun L, Rong L, Oil shale resource state and evaluating system. Earth Science Frontiers, vol.5, no.10, pp. 349 – 354, 2005.
2. Shaopeng L, Shixin Z, Baozhong W, Relationship between industrial index and source rock evaluation parameters of oil shale. Natural Gas Geoscience, vol. 3, no. 7, pp. 129 – 132, 2012.

3. Lucia M. P , Shale Oils. The science and technology of unconventional oil, vol . 4 no. 5, pp. 063 – 071, 2001.
4. Shen C, Reservoir simulation study of an in-situ conversion pilot of green-river oil shale, vol . 4 no. 5, pp. 063 – 071, 2007.
5. Ryan RC, Fowler TD, Beer GL, Nair V, Oil shale: A solution to the liquid fuel dilemma, chapter shell's in situ conversion process from laboratory to field pilots. American Chemical Society, vol . 11 no. 6, pp. 463 – 470, 2010.
6. Burnham AK, Braun RL, Global kinetic analysis of complex materials. Energy and Fuels. vol . 4 no. 5, pp. 053 – 065, 2009.
7. Panbarasan M, A substantial study on environmental legal framework for exploration and exploitation of unconventional hydrocarbon resources (Uhr) Around The Globe, IJSTR, vol. 8, no.11, pp. 3339 – 3342, 2019.
8. Ahmed Abou-Sayed, Data mining applications in the oil and gas industry, JPT, vol. 64, no. 10, pp. 88 – 95, 2012.
9. Alistair R Brown, Seismic attributes and their classification. Leading Edge, vol. 15, no. 10, pp. 1090 – 1090, 1996.
10. Gautam Biswas, Jerry B Weinberg, Fisher DH, Iterate: A conceptual clustering algorithm for data mining. IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 28, no. 2, pp. 219 – 230, 1998.
11. Oluwatobi Olorunsola, Jie Qi, Lennon Infante, Bryce Hutchinson, Kurt Marfurt. Multiattribute seismic facies expressions of a complex granite wash formation: A buffalo wallow field illustration. SEG pp. 1884–1888, 2016.
12. Zhi Hong Zheng, Payam Kavousi, and Hai Bin Di. Multi-attributes and neural network-based fault detection in 3D seismic interpretation. Advanced Materials Research, vol. 838, pp. 1497 – 1502, 2014.

AUTHORS PROFILE

Dr. A. Rajesh Kanna is currently working as Associate Professor, Dept. of Petroleum Engineering, AMET, Chennai. He has more than 8 years of experience both in India and abroad. Current area of research includes Data Mining, IoT, Automation, Enhanced Oil Recovery. Published more than 12 research articles indexed in Scopus and 16 articles in UGC Journal. Life member in SPE, SPG and AMI. Graduated with B.Tech, M.Tech and Ph.D from Anna University Trichy Campus, Anna University Chennai Campus and IIT Madras respectively.
E-mail:rajeshkanna.a@ametuniv.ac.in; petrorajesh@gmail.com

Dr. P. Sivasankar is Assistant Professor in Dept. of Petroleum Engineering at Indian Institute of Petroleum and Energy, Visakhapatnam. His area of research is Numerical Reservoir Simulation and Data Mining. He is a member in SPE. E- mail: sivasankar.petro@iipe.ac.in

Dr. S. Kalpana is Associate professor in Physics Department at AMET, Chennai. Has more than 20 years of teaching experience. Research Interest includes: Biomaterials, Nanoscience and IoT.
E-mail: Kalpanaphy@ametuniv.ac.in