Data based sensing of Shale Oil yield in Oil Shale Retorting process

Hasan Qayyum Chohan¹ and Iftikhar Ahmad¹,*

¹School of Chemical and Materials Engineering, National University of Sciences and Technology, Islamabad, Pakistan
*Email: iftikhar.salarzai@scme.nust.edu.pk

Abstract. Oil shale is sedimentary organic rocks that are being converted into useful shale oil and shale gas. North American regions, Canada and China are exploring the oil shale reserves to accommodate the depletion of natural oil and gas resources. Oil shale retorting technology is being utilized to convert the shale rocks into shale oil and shale gas. The major product is oil that is further treated to convert it into gaseous form. In this study, machine learning techniques like ensemble learning (least square boosting and bagging) and artificial neural network (ANN) are employed for data sensing of oil shale retorting process and being compared. Data is generated for ensemble models through MATLAB-Excel-Aspen interfacing. The proposed framework shows that ANN provides higher accuracy as compare to other models for oil shale retorting process for efficient oil recovery.

Keywords: Oil Shale, Retorting; Machine learning; Artificial neural network; Ensemble learning; MATLAB; Excel; Aspen.

1. Introduction

Energy is an integral part of world. Energy resources are depleting day by day due to significant advancements and developments in economics and population growth. Therefore, world is focusing to resolve the energy challenges and looking for alternative solutions. Oil and gas reservoirs / petroleum reservoirs are based on hydrocarbons present in porous and fractured rocks. They are classified into conventional as well as unconventional petroleum reservoirs. Conventional reservoirs of oil and gas are trapped in low permeable rocks while unconventional rocks have high porosity factor. The naturally occurring hydrocarbons are trapped in rocks due to high porosity factor. Extraction of oil and gas requires extensive exploration methods for efficient recovery of oil and gas present in underlying rocks. Pyrolysis of organic component ‘kerogen’ present in sedimentary rocks converts it into oil and gas. The sedimentary rock is called oil shale and product is known as ‘shale oil’ and ‘shale gas’ [1, 2]. In-situ and ex-situ techniques are employed to convert the oil shale into desired shale oil and other industrial fuels.

According to worldwide production reports of oil and gas 2018 presented in oil and gas journal (OGJ), 165.1 trillion bbl and 69,550 tcf oil and gas are produced per year. World is dependent on naturally occurring oil resources. Canada, USA and China are the key players in exploration of oil and gas through ‘oil shale’ [3]. Institute of energy research provides an overview of countries having major oil shale resources shown in Figure 1 [4].
Modeling and simulation is playing an important role in the industrial processes for optimization and reliability. Oil shale has become an interesting research area in many countries. In literature, various studies have been reported. Luo et al. examined the performance of oil shale processing of hydrogenation technology and analyzed the process of hydrogen through exergy analysis and found 69.20% exergy efficiency of simulated model [5]. Zhou et al. addressed the problems related to low utilization rate of oil shale raw material, low yield and heavy components presence. Techno-model economic was performed with addition of an indirectly heating moving bed in simulation model. The simulation modeling resulted a decrease of 12% in cost of oil shale production as compared to conventional processing [6]. Mu et al. studied parametric study of oil shale retorting for exergy analysis on China comprehensive utilization system [7]. Wang et al. determined the environmental sustainability and resource utilization efficiency factors through energetic life cycle assessment by Aspen Plus of Fushun type oil shale retorting. [8].

Yang et al. utilized the MATLAB interfacing with Aspen Plus to evaluate technical and economic aspects of solid heat carrier technology in oil shale retorting process [9]. Raja et al. investigated the presence of high valued hydrocarbon fuels and hydrogen in oil shale retorting process using Aspen Plus V8.4 modeling. [10]. Li et al. compare two technologies of oil shale retorting Dagong and Fushun. Exergy analysis through Aspen Plus [11]. Tamm et al. also observed thermodynamic parameters of three different samples of oil shale using Aspen Plus V8.6 [12]. Lv et al. studied the conversion of oil shale into bio fuel. The exergy analysis was performed using Aspen Plus [13]. Li et al. studied the effect of deposition of carbon on the overall performance of oil shale retorting process [14].

In this study, soft sensor is developed using machine learning techniques like ensemble learning and artificial neural network for oil shale retorting technology. Data based soft sensor is developed to predict the product composition as well as yield of shale oil. Aspen Plus V10.0 based model is used to generate the data using interface of Aspen Plus MS Excel 2010 and MATLAB R2018a. This work predicts our quantities of interest i.e. oil yield through retorting process and compare different machine learning techniques.

Section II of this study explains the process description for oil shale retorting process followed by modeling techniques in section III. Section IV describes the results and discussion while section V concludes the work.

2. Process description
Oil Shale is converted into shale oil (constituent of crude oil) through retorting process. Ammer et al. developed Aspen Plus V10.0 model for retorting as given in Figure 2 [15]. The major component of oil shale ‘Kerogen’ is crushed into small pieces to increase surface area of particles for efficient reaction.
Oil Shale (FD-SHALE) is heated with in a pre-heater (SHA-HEAT) and moisture is removed through (MOIS-Sep). Pre-heated oil shale is entered in fluidized bed equipment i.e. called “Retort”. Oil shale is converted to gas, oil and char in retort by providing temperature up to 950°F. The temperature requirement is accomplished with help of burned shale in combustor (ORG-COMB) using hot combustible gas (COMB-GAS) coming from MINR-DEC. Air (FD-AIR) is introduced to through heater (AIR-HEAT) for combustion reactions. The spent shale is expelled to a combustor from retort in which hot air coming from pre-heater decomposes spent shale minerals like dolomite and calcite etc. The hot gas of combustor is used for pre-heating of oil shale while some part of burned shale is recycled to the retort through splitter (SSPLIT) and other is separated for disposal (SLD-OUT). The other components are recycled to retort and further processed for ultimate oil recovery. Recycle streams are also utilized for heat transfer mechanism to increase the exergy of process.

3. Machine learning techniques

3.1. Ensemble learning
Ensemble learning models are developed using least square boosting and bagging techniques. Boosting and bagging develop efficient models using combination of different weak models. The concept of ensemble learning models is demonstrated in Figure 3 [16]. The models are trained on stochastic data in several rounds and then models predict data by increasing their respective aggregated weights.

3.2. Artificial neural network
Artificial neural network (ANN) is biological brain type neural network that develops machine learning models to predict the accuracy of stochastic data. Neurons of network receive inputs and predict outputs using transfer function as given in Figure 4 [17]. The accuracy of predicted model is achieved by optimizing number of hidden layers. The increase of hidden layers in network and presence of more stochastic data provides more accuracy in predicted data.
4. Results and discussions

This section deals with the results obtained for production of shale oil yield through retorting process. The results are also provided details of carbon dioxide extracted in flue gases. The results are derived using machine learning techniques i.e. least square boosting, bagging and artificial neural network. The model is developed using 13 input variables given in Table 1. Oil shale and air temperature, elemental composition of kerogen, retort and fractionator specifications are varied in Aspen Plus model to predict the machine learning models. The composition of illite is also altered to determine the impact on production of oil through retorting. Shale oil flow rate and carbon dioxide flow rate in flue gases are output variables enlisted in Table 2.

Table 1. Process input variables

| No. | Process variable            | No. | Process variable            |
|-----|------------------------------|-----|------------------------------|
| 1   | Air flowrate                 | 11  | GL-Separator Temperature     |
| 2   | Illite mass fraction         | 12  | GL-Separator Pressure        |
| 3   | Silica mass fraction         | 13  | G-Split Fraction             |
| 4   | Carbon Composition           |     |                              |
| 5   | Hydrogen Composition         |     |                              |
| 6   | Shale Heater Temperature     |     |                              |
| 7   | Air Heater Temperature       |     |                              |
| 8   | Retort Pressure              |     |                              |
| 9   | Retort Volume                |     |                              |
| 10  | Retort vol. fraction         |     |                              |
Table 2. Process output variables

| No. | Process Variable               |
|-----|--------------------------------|
| 1   | Oil flowrate (Product)         |
| 2   | CO₂ flow rate (Flue Gases)     |

Ensemble models are developed for oil recovery and carbon dioxide in flue gases. The results are being compared for three different ensemble learning techniques. Data is generated for 1350 cases using MATLAB-Excel-Aspen interfacing in which input variables are varied at some extent. Data is divided into training and testing data. Training data is utilized for model development while 20% testing data to validate ensemble model. The modeling framework is given in Figure 5.

![Modeling Framework](image)

Figure 5. Modeling Framework

Data sensing accuracy in prediction of shale oil in final product is 97.8%, 99.8% and 99.6% respectively for boosting, bagging and ANN models. Similarly, for carbon dioxide, boosting, bagging and ANN models provide accuracy up to 98.9%, 99.7% and 99.9% respectively. The results for three ensemble models are given in Figure 6-11. Predicted data is optimized by increasing or decreasing of ensemble layers. For better accuracy of artificial neural network predicted data, hidden layers of neurons are increased. The presence of maximum stochastic data of industrial operations helps more in data sensing and much accurate prediction.

Root mean square error values also exhibit that bagging model is given much accuracy with lowest value of RMSE for oil production (0.0597) as compare to LSBoost (0.2467) and ANN (0.1033). Table 3 shows the exhibiting values of accuracy and RMSE for three models. Table 4 shows the accuracy of validated data of oil shale retorting process.
Figure 6. LS Boosting Regression Performance of Oil

Figure 7. Bagging Regression Performance of Oil

Figure 8. ANN Regression Performance of Oil
Figure 9. LS Boosting Regression Performance of CO₂

Figure 10. Bagging Regression Performance of CO₂

Figure 11. ANN Regression Performance of CO₂
Table 3. RMSE values for output variables

| No. | Process variable              | LSBoost  | Bag     | ANN    |
|-----|-------------------------------|----------|---------|--------|
| 1   | Oil flowrate (Product)        | 2.467e-1 | 5.97e-2 | 1.03e-1|
| 2   | CO₂ flow rate (Flue Gases)    | 6.143e-1 | 3.509e-1| 2.169e-1|

Table 4. Regression values for output variables

| No. | Process variable              | LSBoost | Bag     | ANN    |
|-----|-------------------------------|---------|---------|--------|
| 1   | Oil flowrate (Product)        | 97.8%   | 99.8%   | 99.6%  |
| 2   | CO₂ flow rate (Flue Gases)    | 98.9%   | 99.7%   | 99.8%  |

5. Conclusions
In this study, ensemble learning models of least square boosting and bagging are compared with artificial neural network model. The results predicted that ANN provides more accuracy as compare to least square boosting while bagging is providing comparable behavior to ANN. The predicted data show that machine learning techniques are efficient methods to evaluate the accuracy for real time applications in industries.

6. Nomenclature
ANN  Artificial Neural Network
OGJ  Oil and Gas Journal
FD-SHAL  Oil Shale Feed
SHA-HEAT  Shale Pre-Heater
MOIS-Sep  Moisture Separator
ORG-COMB  Combustion Chamber
COMB-GAS  Combustible Gas
FD-AIR  Air Inlet Stream
AIR-HEAT  Air Pre-Heater
SSPLIT  Burned-Shale Splitter
SLD-OUT  Spent Shale
GL-SEP  Gas-Oil Fractionator
OIL-OUT  Product (Oil)

7. References
[1] Zhi Y, Caineng Z, Songtao W, Senhu L, Songqi P, Xiaobing N, Guangtian M, Zhenxing T, Guohui L and Jiahong Z 2019 Formation, distribution and resource potential of the” sweet areas (sections)” of continental shale oil in China Marine and Petroleum Geology 102 48-60
[2] Seewald J S 2003 Organic–inorganic interactions in petroleum-producing sedimentary basins Nature 426 327
[3] Guo H, Dong J, Wang Z, Liu H, Ma R, Kong D, Wang F, Xin X, Li Y and She H 2018 2018 EOR Survey in China-Part 1. In: SPE Improved Oil Recovery Conference: Society of Petroleum Engineers)
[4] Golpour H and Smith J 2017 Oil Shale Ex-Situ Process-Leaching Study of Spent Shale International Journal of Engineering and Science Invention 6 45-53
[5] Luo X, Guo Q, Zhang D, Zhou H and Yang Q 2018 Simulation, exergy analysis and optimization of a shale oil hydrogenation process for clean fuels production Applied Thermal Engineering 140 102-11
[6] Zhou H, Zeng S, Yang S, Xu G and Qian Y 2018 Modeling and analysis of oil shale refinery process with the indirectly heated moving bed Carbon Resources Conversion
[7] Mu M, Han X and Jiang X 2018 Combined fluidized bed retorting and circulating fluidized bed combustion system of oil shale: 3. Exergy analysis Energy 151 930-9

[8] Wang Q, Ma Y, Li S, Hou J and Shi J 2018 Exergetic life cycle assessment of Fushun-type shale oil production process Energy Conversion and Management 164 508-17

[9] Yang Q, Li X, Qian Y and Zhang D 2018 Technical and economic analysis of an oil shale comprehensive utilization process with solid heat carrier technology Carbon Resources Conversion

[10] Raja M A, Chen H, Zhao Y, Zhang X and Zhang S 2017 Process simulation and assessment of hydrogen and high valued hydrocarbon fuels products from oil shale International Journal of Hydrogen Energy 42 4922-34

[11] Li X, Zhou H, Wang Y, Qian Y and Yang S 2015 Thermoeconomic analysis of oil shale retorting processes with gas or solid heat carrier Energy 87 605-14

[12] Tamm K, Kallaste P, Uibu M, Kallas J, Velts-Jaenes O and Kuusik R 2016 LEACHING THERMODYNAMICS AND KINETICS OF OIL SHALE WASTE KEY COMPONENTS Oil Shale 33

[13] Lv Q, Yue H, Xu Q, Zhang C and Zhang R 2018 Quantifying the exergetic performance of bio-fuel production process including fast pyrolysis and bio-oil hydrodeoxygenation Journal of Renewable and Sustainable Energy 10 043107

[14] Li H, Guo C and Yuan D 2017 Study on Carbon Deposition Characteristics of Oil Shale Dry Distillation Gas and Influencing Factors of Carbon Deposition Reaction Journal of Nanoelectronics and Optoelectronics 12 1339-42

[15] Ammer J 1986 Identification of data gaps found during the development of a zero-order model for a fluidized-bed retort/combustion process. USDOE Morgantown Energy Technology Center, WV)

[16] Ahmad I, Ayub A, Mohammad N and Kano M 2019 Data-Based Prediction and Stochastic Analysis of Entrained Flow Coal Gasification under Uncertainty Sensors 19 1626

[17] Ahmad I, Mabuchi H, Kano M, Hasebe S, Inoue Y and Uegaki H 2011 Data-Based Fault Diagnosis of Power Cable System: Comparative Study of k-NN, ANN, Random Forest, and CART IFAC Proceedings Volumes 44 12880-5