ICCPGE 2016, 1, 25 - 30

Recovery Predictions for Polymer Flooding Method

Laila D. Saleh[1,*], Muna M. Hammuda[2], Abdussalam M. Mohamed[1], Haiat K. Alhaj[1], Mahmoud O. Elsharafi[3], Baojun Bai[4]
[1]Department of Petroleum Engineering, University of Tripoli, Tripoli, Libya
[2]Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, Missouri, USA
[3]McCoy School of Engineering, Midwestern State University, Wichita Falls, Texas, USA
[4]Department of Geosciences and Geological and Petroleum Engineering, Missouri University of Science and Technology, Rolla, Missouri, USA
*Corresponding Author: l.saleh@uot.edu.ly

Abstract

Future performance is very important to evaluate whether an enhanced oil recovery (EOR) is successful or not. The most important step in prediction is to determine the amount of oil that can be recovered after applying an EOR process. Polymer flooding is one of the most important EOR techniques used to improve the mobility ratio and, therefore, sweep efficiency. Multiple linear regression techniques were used to develop the equation that can be used to predict the oil recovery based on rock and fluid properties in field data set. A dataset was created by collecting information from EOR surveys of Oil and Gas Journal (1974 - 2012). A total of 481 field projects was considered to construct the dataset. Unfortunately, this data contained a number of problems (duplicate, missing, and inconsistent data) that affected the dataset's quality. To ensure the quality of the dataset before running any analyses, box plots and cross plots were used to identify data problems. Graphical and statistical methods were used to analyze and describe the results of the dataset. After enhancing the data quality, only 82 fields were used for the predictions.75 fields were used to build the model. The remaining fields were selected to validate the models. Parameters were chosen for the models: area, oil gravity, oil viscosity, porosity, saturation start, permeability, depth, and temperature. The stepwise technique was used to establish the independent model that affects the response variable significantly. Two models were constructed; one to predict oil recovery and another to predict oil saturation (So(end)) after polymer flooding. Equations for both models were presented in this paper. The equation for So(end) appears to represent the best model based on R2 values.

Keywords: Polymer flooding; predication; regression; oil saturation model; recovery model.

1. Introduction

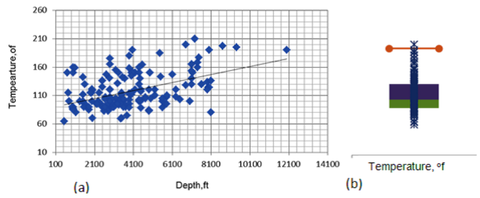
One important Enhanced oil recovery (EOR) techniques is polymer flooding (Water – soluble polymers). Polymer flooding has been used as an EOR technique since the early 1960s [1]. To improve efficiency of water flooding, some additives added to water. Detling (1944) suggested raising water viscosity to improve efficiency of water flooding [2]. Water-soluble polymers have also been used to improve sweep efficiency by increasing the viscosity of water. Barnes (1962) proposed using a viscous water slug to improve water sweep efficiency in reservoirs partially invaded by bottom water [3]. The result of his study indicated that the use of viscous water slugs does increase ultimate oil recovery. In the Daqing oil field (in China), the oil recovery is 12% higher when using polymer flooding than water flooding and the oil increment is 120 tons for every ton of polymer injection [4]. Future performance is very important to evaluate whether an EOR is successful or not. The most important step in prediction is to determine the amount of oil that can be recovered after applying an EOR process. Statistical method is used to determine the predicted performance of polymer flood. The prediction was based on the actual available data from fields application. This study explored regression analysis as tools to reveal the complicated relationships among oil recovery and other reservoir and fluid properties, and correlations generated can be used as prediction models. A dataset was created by collecting information from enhanced oil recovery (EOR) surveys of Oil & Gas journal. The dataset reviewed in this study included 481 polymer flooding projects. The reservoir and geology data collected contained information on well spacing (area, acres), oil gravity (^{o}API), oil viscosity (cp), porosity (%), permeability (md), depth (ft), temperature (^{o}F), initial oil saturation (%), end oil saturation (%), and formation type (sandstone or carbonate). These data have quality problems that previous research has not addressed . These problems (duplicate, missing, and inconsistent data), will describe briefly in this paper. literature research did not find applications of regression analysis in oil recovery prediction using the same or similar set of reservoir and fluid parameters. This study explored regression analysis as tools to reveal the complicated relationships among oil recovery and other reservoir and fluid properties, and correlations generated can be used as prediction models.

2. DESCRIPTION OF THE MODEL DATASET

Table 2.1 includes a portion of the worksheet taken from the field dataset that contains both duplicate fields and missing values. Many fields are listed numerous times with the same values and either in the same year or over different years of publications. Moreover, several fields within the dataset were missing one or more pieces of information, including oil saturation (start and end), permeability, depth, and temperature.

Data quality is critical in ensuring the quality of the analysis results [5]. A box plot helps both to summarize a dataset and to detect outliers within the data . A cross plot was used to plot a pair of variables from the dataset. The plot helps to uncover the relationships between these variables and to detect outliers [5]. The box and cross plots were combined to yield additional clarity. As Example for combination between the box and cross plot see Figure 2.1 [5]. Figure 2.1 (a) shows the cross plot between temperature and depth, while Figure 2.1

Figure 2.1: Temperature versus depth (a) cross plot, (b) box plot



(b) shows the box plot for temperature. The box plot in Figure 2.1 (b) shows that data from only one field exceeded the upper limit of the dataset (orange line). However, this data cannot be considered an outlier because the field temperature value is consistence with other field characteristics, such as depth and porosity; and the polymer flooding project was successful at this temperature.

After removing the duplicate records and those having missing values, only 82 projects were used to construct the predication models, among 75 records were randomly selected to build the model, and the remaining records were used to validate the model. Among reported reservoir and fluid properties, the following parameters were chosen for the model: area, oil gravity, oil viscosity, oil saturation before polymer flooding, formation porosity, permeability, depth, and temperature. These parameters are determined based on the availability in the EOR survey data and the polymer flooding EOR mechanism.

3. REGRESSION ANALYSIS

Regression analysis is a statistical technique that is used to investigate and model the relationships between one or more independent variables(X) which are also known as either predictors or explanatory variables, and a single dependent variable (Y) or the response. Thus, the regression explains how the response (Y) changes as the predictors (X) change [6]. Among varied regression analysis methods, linear regression is one of the first considerations. The linear regression can be classified into simple and multiple linear regression.

Based on available variables ,multiple linear regressions are used to evaluate the relationship between a single response (Y) and more than one predictor variable ($x1, x2, x3, \ldots, xp$). The general form of the multiple linear regression equation is given by

$$yi = \beta o + \beta 1x1 + \beta 2x2 + \cdots + \beta pxp + \varepsilon i \quad (3.1)$$

Table 2.1: Polymer Flooding Projects in *EOR* Survey that Contain Both Duplicate and Missing Data

	A	B	C	D	E	F	G	H	I
1	Year Reported	Country	Field	Area(Acres)	Oil Gravity (API)	Oil Viscosity (cp)	Porosity(%)	Oil Saturation Start	Oil Saturation End
2	1998	China	.	3803	33	.	.	.	.
3	1998	China	.	4740	33.5	.	.	.	.
4	1998	China	.	2989	15.5	.	.	.	.
5	1998	China	.	1536	13.6	.	.	.	.
6	1998	China	.	2717	32.08	.	.	.	.
7	1996	USA	ALba North Central Unit	512	15	75	21	62	56
8	1998	USA	ALba North Central Unit	512	15	75	21	62	56
9	2000	.	Alba North Central Unit	512	15	75	21.4	62	56
10	2000	.	ALba Notrth Centeral Unit	512	15	75	21	62	56
11	2000	.	Alba SEFB Unit	731	15.5	75	23.3	69.5	51.8
12	1986	USA	Alba Southeast	731	40	0.44	23.3	69.5	.
13	1996	USA	Alba West Unit	567	15	75	21	65	56
14	1998	USA	Alba West Unit	567	15	75	21	65	56
15	2000	.	Alba West Unit	567	15	75	21	65	56
16	1996	USA	Albh SEFB Unit	731	15.5	75	23.3	69.5	51.8
17	1998	USA	Albh SEFB Unit	731	15.5	75	23.3	69.5	51.8
18	1986	USA	Apache Pool	1060	39	1.8	13.6	36	25
19	1986	USA	Atlantic	1920	39	0.3	25	34	32
20	1986	USA	Beverly Hills/West Pico	221	32	1	28	37.5	.
21	1986	USA	Black Diamond	200	48	0.4	25	68	41
22	1988	Germany	Bockstedt	57	34	11	23	39	33
23	1996	USA	bracken unit	120	20.2	16.6	16.5	76.1	59.4
24	1998	USA	bracken unit	120	20.2	16.6	16.5	76.1	59.4
25	1980	USA	Brelum	427	23	12	29	88	75
26	1984	USA	Brelum	427	23	12	29	80	65
27	1986	USA	Byon	1500	23	15	13.9	.	.
28	1986	USA	Carthage	1640	39	2.8	19	51	.
29	1986	USA	Carthage, N.E.	1024	40	2	18.6	45	44.5
30	1986	USA	C-Bar	2600	36	5	10	39	.
31	1986	USA	C-Bar	2600	33	4.8	10	76	.
32	1984	USA	Cement	1510	35	5	19.2	.	.
33	1986	USA	Cement	1510	35	4	19.2	34	.
34	1986	USA	Cement	1120	28	20	36.1	57	.
35	1984	USA	Cement (W. Cement Unit)	1120	35	6	19.1	77	.

The βs are the regression coefficients (unknown parameters). When $\beta 0$ is equal to zero (without an intercept). Equation 3.1 can be written as

$$yi = \beta 1x1 + \beta 2x2 + \cdots + \beta pxp + \varepsilon i \qquad (3.2)$$

A number of criteria are used to evaluate the fitness of a regression model (e.g., the coefficient of determination '$R2$', Mallows' Cp, and the mean square error [MSE]

4. Results and Discussion

4.1. Full Linear Regression Model for Recovery

In the full multiple linear regression model for oil recovery defined in the Equation 4.1, the variables include reservoir areal size, oil gravity, oil viscosity, formation porosity, oil saturation before polymer flooding, formation permeability, formation depth, and reservoir temperature. These variables can be selected as independent variables or predictors in this study. The dependent, or response variable (Y), is the oil recovery defined as the difference of oil saturation before and after the polymer flooding, as expressed in the Equation 4.1.

$$Recovery = (So(start) - So(end)) \qquad (4.1)$$

Where $So(start)$ is average oil saturation before polymer flooding, and $So(end)$ is average oil saturation after polymer flooding.
Figure 4.1 illustrates the regression diagnostic or residual (the difference between observed and predicted value) plots of the predictor variables. Most of these plots exhibit a null plot and exhibit no obvious pattern, which indicates
In this regression analysis, the $R2$ and the adjusted $R2$ were 79 % and 77%, respectively, as shown in Table 4.1.

Table 4.1: Regression Analysis for Full Recovery Model

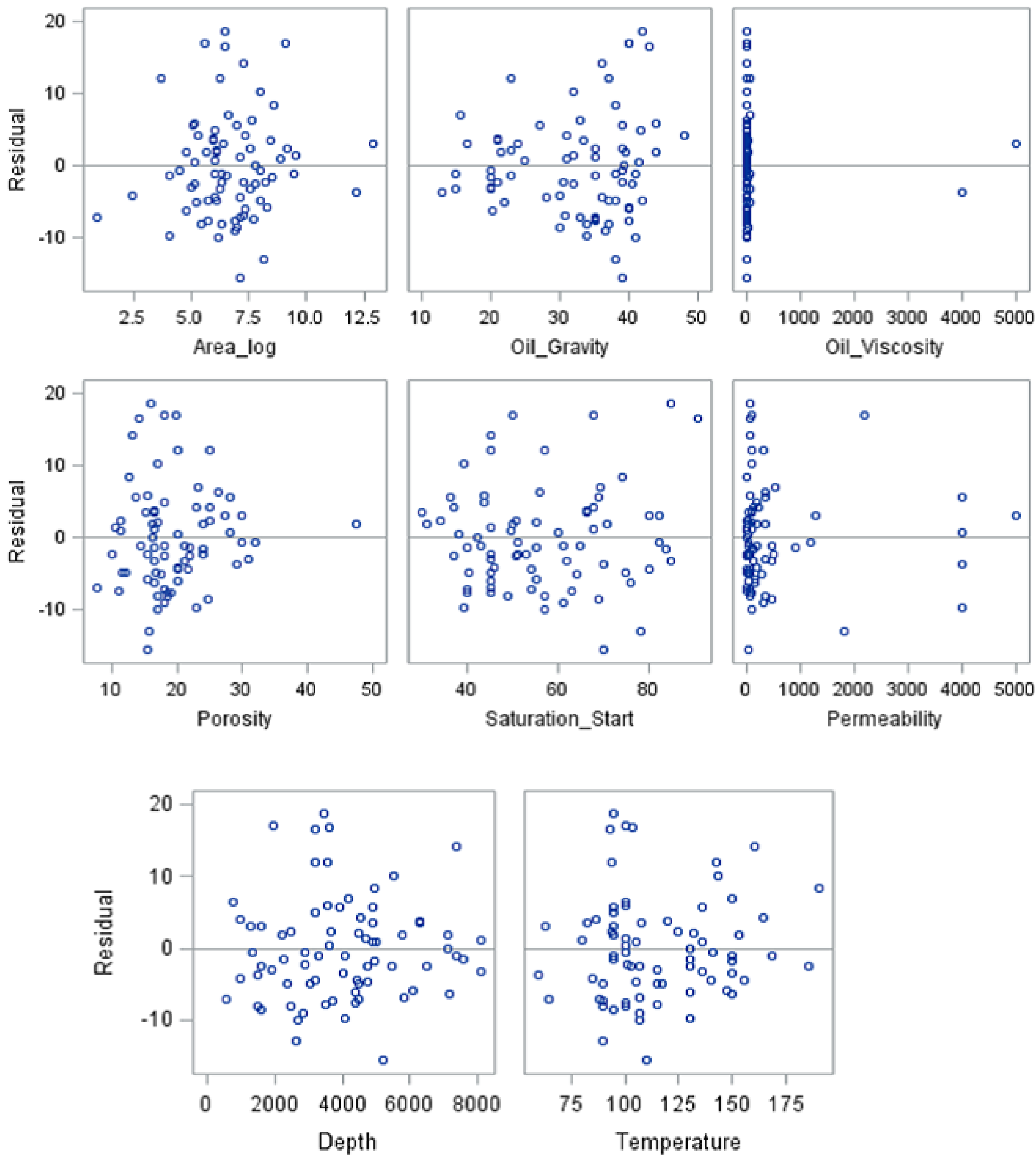
Root MSE	7.21366	R-Square	0.7512
Dependent Mean	10.28406	Adj R-Sq	0.7318
Coeff Var	70.14412		

Another suggestion was made to improve fitting the model. This suggestion was to predict the saturation end (after flooding) rather than the recovery. After the saturation end was predicted correctly, the recovery would be calculated with Equation 4.1.

4.2. Full Linear Reduced Model for $So(end)$.

The dataset used in the saturation end model was the same data used in the recovery model. The full model used to predict the saturation end included

Figure 4.1: Residual plots of selected variables



a single response ($So(end)$) and eight independent variables as discussed in the Section 4.1.The most regression diagnostic plots exhibit a null plot and not obvious pattern, which indicates the correctness of linearity and normality assumptions, as shown in Figure 4.2.
The results gathered from this model were quite similar to the recovery model with two exceptions:-

- $R2$improved from 79 % in the recovery model to 97.7% in the $So(end)$ model.
- And adj $R2$ improved from 77% in the recovery model to 97.4 % in $So(end)$ model, as shown in Table 4.2.

Table 4.2: Regression Analysis for Full$So(end)$ Model

Root MSE	7.51039	R-Square	0.9769
Dependent Mean	44.81467	Adj R-Sq	0.9741
Coeff Var	16.75877		

The insignificant variables in the recovery model were the same variables in the $So(end)$ model (Table 4.3). The stepwise method was also used to remove the insignificant parameters.

Table 4.3: Parameters Estimate from the $So(end)$ Model

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Area_log	1	3.01422	0.62271	4.84	<.0001
Oil_Gravity	1	-0.28940	0.09562	-3.03	0.0035
Oil_Viscosity	1	-0.01081	0.00209	-5.16	<.0001
Porosity	1	0.16273	0.15609	1.04	0.3009
Saturation_Start	1	0.68034	0.05631	12.08	<.0001
Permeability	1	-0.00082219	0.00113	-0.73	0.4690
Depth	1	-0.00109	0.00062425	-1.75	0.0845
Temperature	1	-0.01248	0.03959	-0.32	0.7536

4.3. Reduced Model for $So(end)$

In order to construct the reduced model, the stepwise technique was first used to establish the independent model that affects the response variable significantly. Both response variable and the independent variables were the same as Full $So(end)$ model. The model selection criteria assessed the selected model. The model fit statistics (criteria), Adj $R2$, CP, were included in the model. The selected model is the model listed at step 5 in Figure 4.3.

All of the independent variables selected for the best model were significantly better than those used in Full $So(end)$ model as shown in Table 4.4.

Table 4.4: Fit Statistic for Selected Model

Step	Variable	F Value	Pr > F
1	Saturation_Start	1761.05	<0.0001
2	Oil_Viscosity	17.32	<0.0001
3	Area_log	11.73	0.0010
4	Oil_Gravity	6.45	0.0133
5	Depth	6.24	0.0149

The final saturation end equation is

$$So(end) = 3.044629\, A - 0.252698\, \gamma_o - 0.011081\, \mu_o + 0.692322\, So(start) - 0.001286\, D \tag{4.2}$$

where A is the logarithmic of the reservoir areal size in acres, γ_o is the oil gravity in ^oAPI, μ_o is the oil viscosity in cp, and D is the formation depth in ft. In order to validate the model constructed, a validation set was selected randomly. The predicted results and reported values are compared, as listed in Table 4.5.

Figure 4.2: Residual plots for full So (end) model

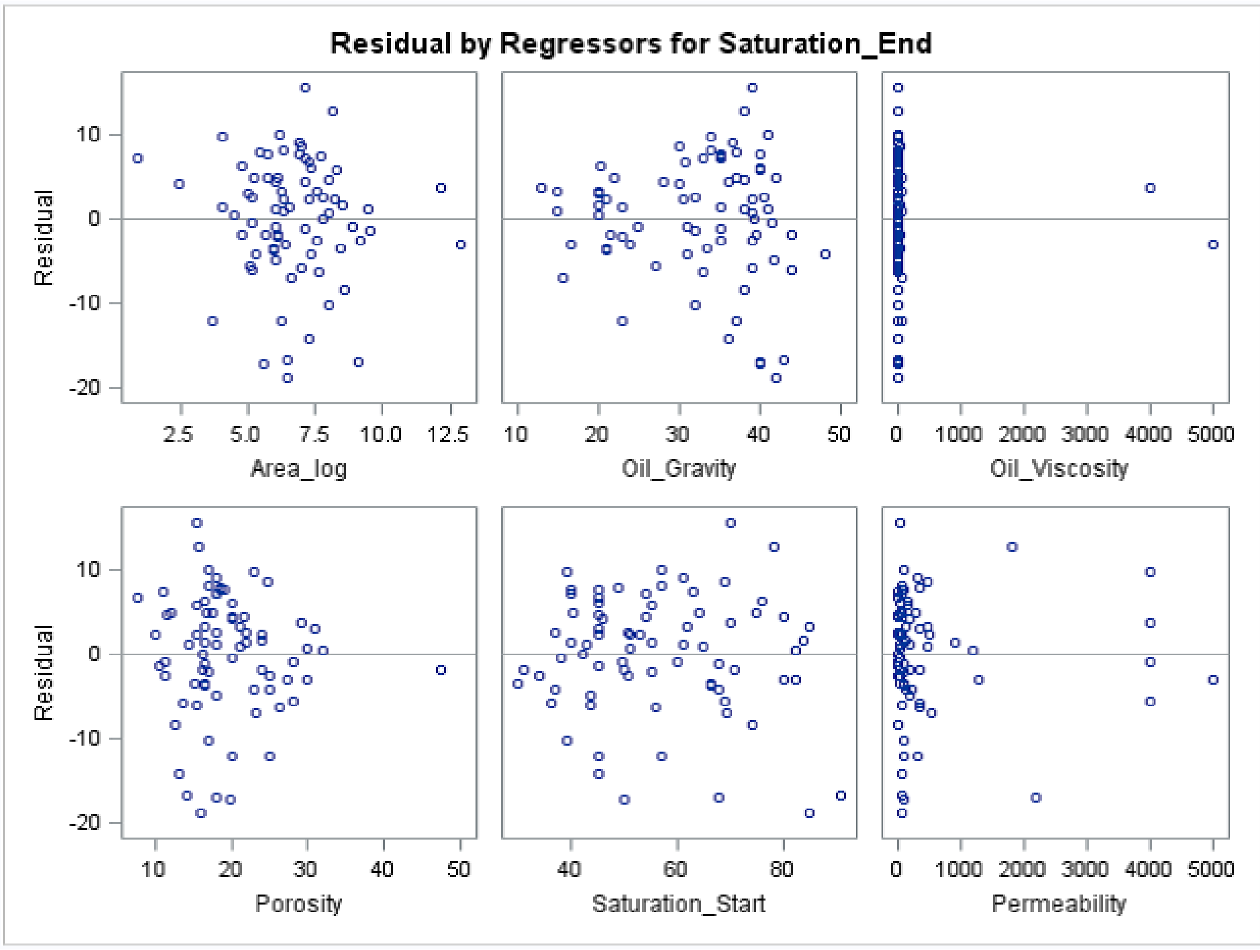


Figure 4.3: Fit criteria for saturation end model

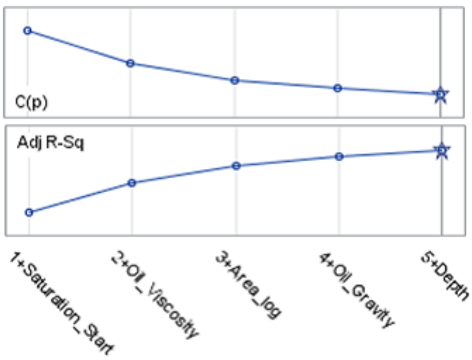


Table 4.5: Validation Data

Area, acres	Oil Gravity, ^{o}API	Oil Viscosity,cp	Saturation Start,%	Depth, ft	Actual values	Predicted values
427	23	12	80	1950	65	65
3800	30.8	1.47	45	5818	44	41
145	20	10	83	1900	68	65
1440	39	3	48	2900	44	42
2448	40.6	0.92	37	6500	34	31
2368	39.3	0.885	42	7100	34	34
9360	35	2.8	50.7	3650	46	49

5. Conclusion

- Multiple linear regression techniques were used to develop the equation that can be used to predict the oil recovery based on rock and fluid properties in field dataset.
- Multiple linear regressions provide a useful tool to evaluate the effect of several parameters on oil recovery.
- The stepwise technique was used to select the best model based on significant parameters. A coefficient of determination of 97.6 % was achieved for saturation end model.

Acknowledgment

The authors gratefully acknowledge the financial support given for this work by the Al-Mergib University. Also the authors would like to thank department of Petroleum Engineering, University of Tripoli, Tripoli, Libya ; McCoy School of Engineering, Midwestern State University, Wichita Falls, Texas, USA and Department of Geosciences and Geological and Petroleum Engineering, Missouri University of Science and Technology, Rolla, Missouri, USA for their support.

References

[1] Sandiford, B.B. Laboratory and Field Studies of Water Floods Using Polymer Solutions to Increase Oil Recoveries. Journal of Petroleum Technology, 1964, Volume 16, Number 8, 917-922, DOI 10.2118/844-PA

[2] Delting, K. Process of reservoir oil from oil sands. U.S. 1944 patent No. 2341500, Feb.

[3] Barnes, A.L. The Use of a Viscous Slug to Improve Waterflood Efficiency in a Reservoir Partially Invaded by Bottom Water. Journal of Petroleum Technology, 1962 Volume 14, Number 10, 1147-1153, doi. 10.2118/334-PA

[4] Yabin, N.; jian, O.; zhuoyan, Z.; Guijiang , W.; Guanghua, S.; and lijun, S. Research on Hydrophobically Associating water-soluble polymer used for EOR. 2001, Proceedings of SPE International Symposium on Oilfield Chemistry, 13-16 February, Houston, Texas, doi. 10.2118/65378-MS

[5] Saleh, L. D.; Wei, M.; and Bai, B.. Data Analysis and Updated Screening Criteria for Polymer Flooding Based on Oilfield Data. Journal Reservoir Evaluation & Engineering. 2014, doi:10.2118/168220-PA

[6] Weisberg, S. Applied Linear Regression. 2005, Third Edition Book. Published by John Wiley & Sons, Inc., Hoboken, New Jersey