Updated Screening Criteria for Steam Injection Projects Based on Oil World Wide Survey

Laila D. Saleh [1*], Usama M. Al-Ahmadi [2], Mohamed A. Mgeg [3], Qais M. Ruqyah [4]

[1] ladasa08@gmail.com, [2] usama.etman@gmail.com, [3] mohamed.mgeg@yahoo.com, [4] kmuhammed96@gmail.com

[1,2,3,4] Department of Petroleum Engineering, University of Tripoli, Libya

*Corresponding author email: ladasa08@gmail.com

ABSTRACT

Enhanced oil recovery (EOR) screening criteria are considered the first step in evaluating potential EOR techniques for candidate reservoirs. Therefore, as new technologies are developed, it is important to update the screening criteria. Many of the screening criteria for steam flooding that have been described in the literature were based on data collected from EOR surveys biennially published in the Oil & Gas Journal. However, these datasets contain some problems, including outliers, missing data, inconsistent data and duplicate data, which could affect the accuracy of the results. Despite the importance of ensuring the quality of a dataset before running analyses, data quality has not been addressed in previous research related to EOR screening criteria. The objective of this current work was to update the screening criteria for steam flooding by using a database that had been cleaned. The original dataset included 1078 steam flooding field projects from around the world (Brazil, Canada, China, Colombia, Congo, France, Germany, Indonesia, Trinidad, U.S. and Venezuela). These projects had been reported in the Oil and Gas Journal from 1996 to 2014. After detecting and deleting the duplicate projects, only 221 field projects remained. To analyze and describe the results of the dataset, both graphical and statistical methods were used. A box plot and cross plots were used to detect and identify data problems, allowing for the removal of outliers and inconsistent data. Histogram distributions and box plots were used to show the distribution of each parameter and present the range of the dataset. New screening criteria for steam flooding were developed based on these statistics and the defined data parameters. After removing outlier data, a model to estimate the recovery factor were built by using the Minitab software. Regression analysis is a statistical technic that is used to investigate and model the relationships between one or more independent variable and single dependent variable in order to construct a better model.

Keywords: steam, statistics, criteria, regression, model, recovery factor

1. Introduction

Enhanced oil recovery (EOR) means the processes to recover oil by the injection of materials not normally present in the reservoir which increases the areal and volumetric sweep efficiency and make the oil become less viscous and more mobile to get toward the production well . Enhanced oil recovery (EOR) processes include all methods that use external sources of energy and/or materials to recover oil that cann't be produced, economically by conventional means [1].

In general, EOR methods can be classified into two major groups: thermal and non-thermal processes. Thermal methods include: cyclic steam stimulation, steam drive, hot water drive, and in-situ combustion, Non-Thermal method includes: first Chemical methods: polymer flooding, surfactant flooding, enhanced alkaline (ASP), biological/MEOR, second Miscible/Non-Miscible methods including: hydrocarbon gases, CO_2, air/nitrogen [2].

EOR screening are usually the first step in determining if further investigation into the application of EOR is warranted by its nature, screening can only be considered a coarse judgment on the suitability of an EOR process [3].

This paper is a part of our previous researches to update screening criteria for different EOR methods [3, 4, and 5]. This study aims to update screening criteria for steam injection to include reported EOR projects and new reservoir and rock properties from 2014 in Oil & Gas Journal that not collected before. To achieve this objective, it is essential to ensure that the data is of high quality and produces reliable results. The methodology to clean and analysis data have been documented previously [3], so this paper provides a brief description of data analysis.

2. Data Collection

A dataset was created by collecting thermal flooding field project data from the Worldwide EOR Survey biennially published in the Oil & Gas Journal from 1996 to **2014**.

The original dataset included a total of 1,206 thermal flooding field projects from all over the world including 1,078 steam flooding field projects. Figure 1 shows data percentage for every thermal process.

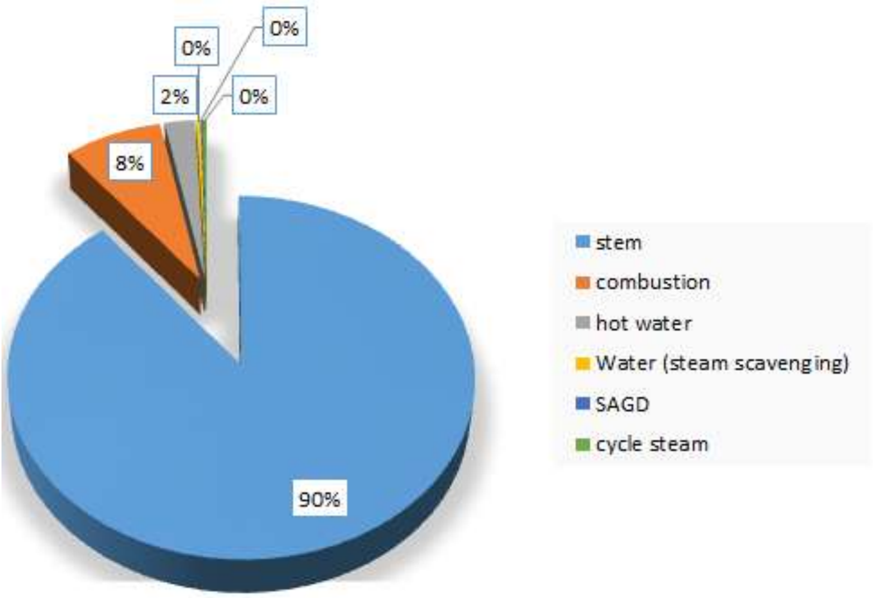

Figure 1. Data percentage for thermal process

The U.S shared more than the half of these projects, with 520 Steam flooding project, 67 in-situ combustion project and 22 hot water flooding project. Figure 2 shows percentage of projects for every country.

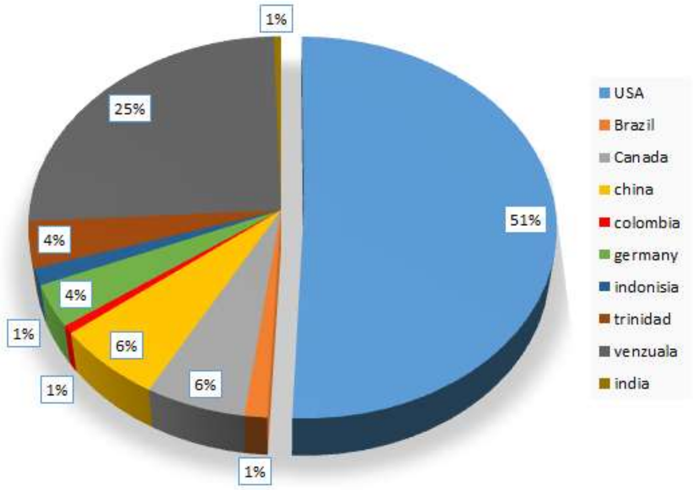

Figure 2. Percentage of projects for countries.

Figure 3 shows the number of thermal projects (steam, in-situ combustion & hot water) and steam projects that was reported in Oil & Gas Journal from 1996 to 2014.The number of thermal projects increased from 1996 to 1998 then between 1998 and 2000 the thermal projects remained constant while a slightly decrease happened to thermal projects before they decrease in a high rate from 2000 to 2002 and then to 2006 then number of thermal projects got slightly increasing and decreasing between 2006 and 2012, from 2012 to 2014 number of projects decreased in a high rate again.

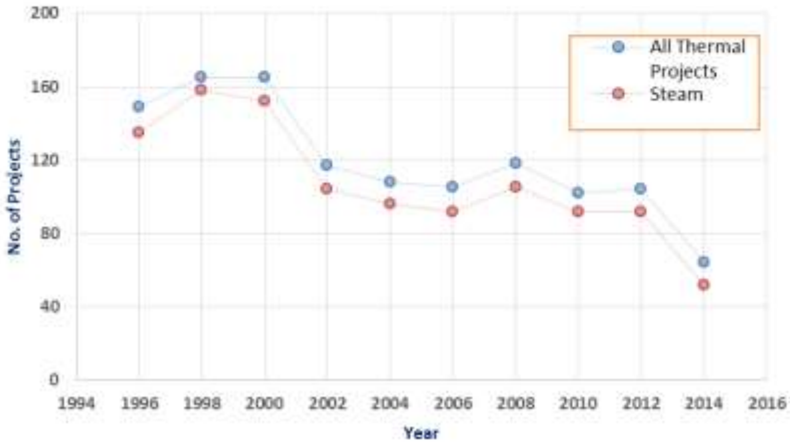

Figure 3. Number of projects reported per year.

3. Data Cleaning

Data quality is essential to ensuring the validity of screening criteria results. EOR survey data contain many types of problems that can affect the quality of the dataset, in particular, duplicate projects, missing data, inconsistent data and outliers [3, 5].

3.1 Duplicate Data

The duplicate data problem was observed while collecting data from the worldwide EOR surveys. Many fields were listed more than once with the same values in different years of publication.

This duplication may have occurred because some countries may not have updated their EOR information for several years.

To avoid the problems caused by the duplicate data, a series actions has been conducted, as shown in figure 4 Firstly, the duplicate projects were deleted if projects are with the exact same records. Then for the data with little different records, two different cases were under consideration. If the difference is caused by the missing values in the project and the other values remain the same, project with missing values were deleted. If projects that have just one feature different, the comparison of irrelevant information between projects were conducted (like report year, project locations). If the irrelevant information is the same, the projects were considered as repeating projects; if they have different irrelevant information, the reservoir properties and fluid properties might coincidentally be the same.

Therefore, projects with different irrelevant information were kept in the data sets.

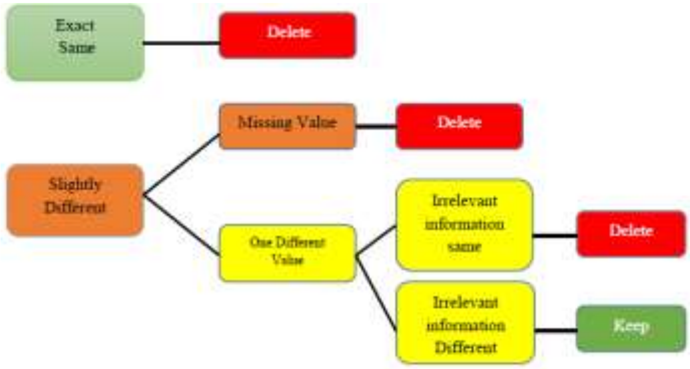

Figure 4. Solving duplicated data

3.2 Missing Data

Some fields within the dataset were missing one or more pieces of information, including oil saturation (start and end), viscosity, permeability, depth, API, gravity, porosity, formation type and temperature. These missing values were ignored during the analysis.

3.3 Inconsistent Data

Data are considered inconsistent if they contain either discrepancies or impossible values. Several pieces of information in the dataset were inconsistent, such as (porosity >50%). This problem was solved by using cross plots and boxplots.

Figure 5 A shows the cross plot of the temperature vs. the depth, and Figure 5 B shows the box plot for the reservoir temperature of the dataset. The box plot illustrates that the temperature data from several fields exceeded the upper limit (Orange line, 139˚F) of the dataset. The reservoir temperatures in these fields exceeded 200˚F and also were inconsistent with the corresponding reservoir depths, as shown in the cross plot. These projects have been circled and marked as outliers.

Figure 5 C shows the box plot for the reservoir depth of the dataset, which indicates that the depth data from several fields exceeded the upper limit (Orange line, 3,400 ft).

Therefore, the box plot indicates that only three fields in China (Gaoshen 2-3, Xiao Wa, Shu I) considered as outliers because these value lies far from the majority of the data and the trend.

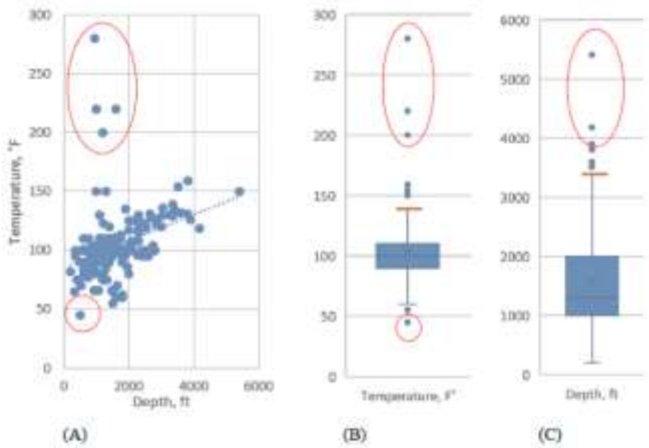

Figure 5. (A) Cross plot of temperature vs. depth, (B) Box plot of temperature, (C) Box plot of depth.

The cross plot in figure 6 A shows the relationship between the reservoir porosity and depth. The box plot in figure 6 B shows the reservoir porosity ranges for the dataset. The box plot illustrates that porosity of two fields in USA (Cymric, Midway-Sunset Diatomite) exceeded the upper limit in the box plot. The upper limit is represented by an orange line and equals approximately 40%. These fields also appear on the cross plot, where they have been outlined with a circle. The porosities of these outlier fields is approximately about 65%. One of these two projects was applied in shale formations, while the other project contained no formation type data. It is known that the shale formations have high porosity and low permeability but cannot be considering these values as special case because they are only two fields for this reason they were considered as outliers

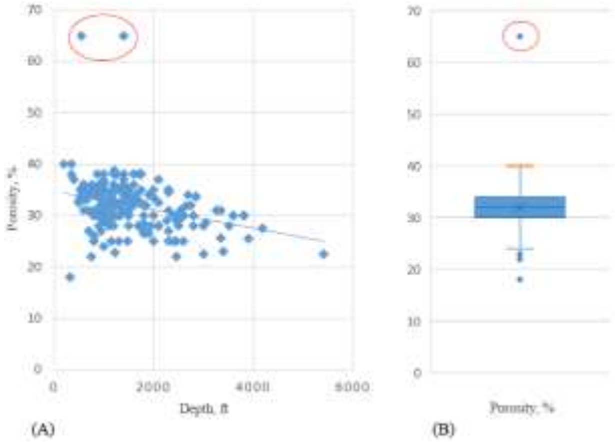

Figure 6. (A) Cross plot of porosity vs depth, (B) Box plot of porosity

The cross plot in figure 7 A shows the relationship between the oil gravity and the viscosity, and the box plot in figure 7 B shows the oil gravity ranges of the dataset. The cross plot shows a few fields lying far from the majority of the data and exhibiting different behavioral trends. The four projects outlined in a red circle, all implemented in China, and they had high oil gravity with high oil viscosity and were considered to be special cases.

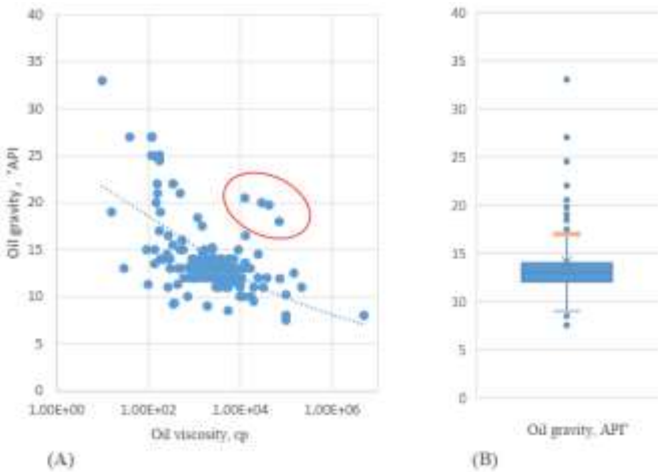

Figure 7. (A) Cross plot of oil gravity vs oil viscosity, (B) Box plot of oil gravity

Figure 8 A shows the cross plot of the oil viscosity vs. the reservoir temperature, and figure 9 B shows the box plot for the oil viscosity of the dataset. The oil viscosity values of several fields exceeded the upper limit in the box plot. The upper limit is represented by an orange line and equals approximately 100,000 cp. Therefore, both of the box plot and cross plot indicates that only one fields in Canada (Athabasca Oil Sands) considered as outliers because these value lies far from the majority of the data and the trend. In addition, the boxplot and cross plot shows that three projects that fell far from the majority of the data and the trend. These data have been circled and marked as outliers. These data were considered outliers because the other field parameters for the projects, such as the reservoir temperature, oil viscosity and oil gravity, were inconsistent.

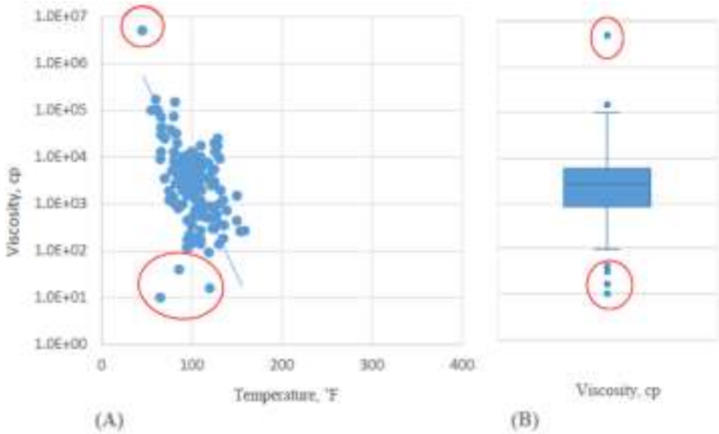

Figure 8. (A) Cross plot of viscosity vs temperature, (B) Box plot of oil viscosity

After cleaning and removing the duplicated projects and inconsistent data we just got 195 steam flooding project ,17 in-situ combustion project and 24 hot water flooding project. So this study was made on steam flooding projects.

Figure 9 shows the formation type for successful steam flooding projects in this study, Sandstone considering the most common and suitable formation type for steam flooding as 171 steam projects was succeeded in sandstone with 88% percentage of all projects.

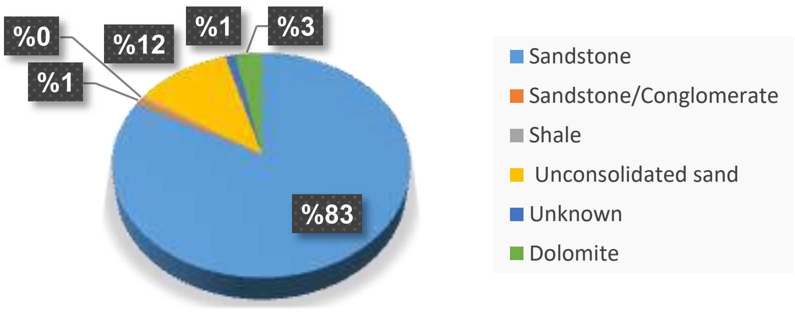

Figure 9. The percentage of formation type for steam flooding.

4. Methods for Displaying the Data Set

In this project to display and analyses the data set of 221 project, two methods were used, the first method is the graphical analysis which include Histograms and boxplots. The second method is the statistical values such as (min, max, mean, median and standard deviation).

Figure 10 shows the reservoir permeability distribution of the dataset across 208 reservoir permeability data points. The highest permeability frequency in the distribution is between 1000 and 2000 md. Approximately 82.2% of the data points fall between 100 and 4000 md.

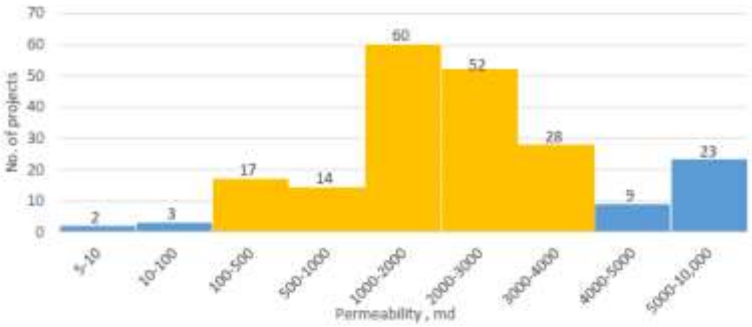

Figure 10. Reservoir permeability distribution of the dataset.

Figure 11 illustrates the oil viscosity distribution of the dataset a cross 214 oil viscosity data points. The distribution is skewed to the right. The most frequent viscosity is in the 1,000 to 5,000 cp range. Approximately 81.3% of the data points fall between 100 and 10,000 cp.

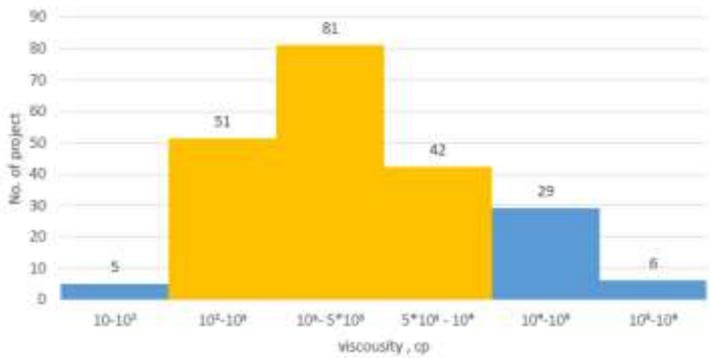

Figure 11. Oil viscosity distribution of data set.

Figure 12 shows the reservoir temperature distribution of the dataset across 215 reservoir temperature data points. The distribution is normal or symmetrical shape. The distribution shows that steam flooding projects applied in different reservoirs temperature. The lowest between 55 and 65 ˚F, and the highest is between 155 and 165 ˚F. Approximately 66.04% of the data points fall between 85 and 115 ˚F.

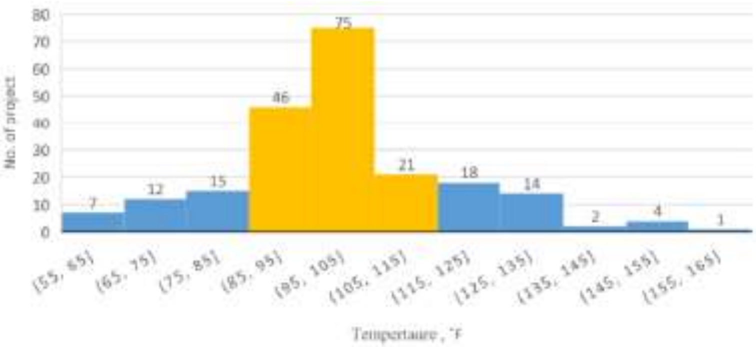

Figure 12. Reservoir temperature distribution of the dataset.

Figure 13 shows the reservoir depth distribution of the dataset across 218 reservoir depth data points. The distribution is skewed to the right and the figure shows two reservoir depth value peak, one between 800 and 1,100 ft and the other between 1,100 and 1400 ft. The majority of the data fall between 500 and 2,000 ft, representing 73.8% of the field projects. There is a tail on the right side of the distribution.

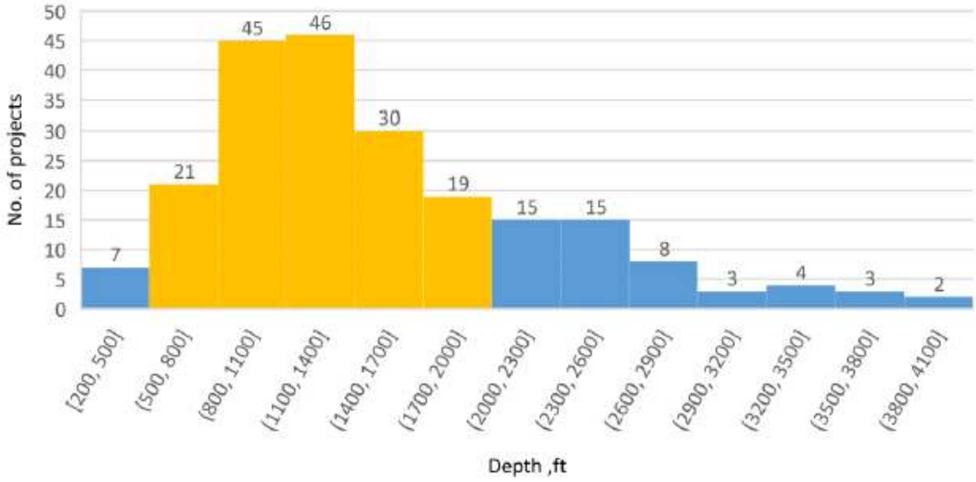

Figure 13. Reservoir depth distribution of the dataset.

Figure 14 shows the oil saturation (start) distribution of the dataset across 213 oil saturation (start) data points. The distribution is random and shows two oil saturation (start) value peak, one between 55 and 60% and the other between 60 and 65 %. Approximately 91.5 % of the data points fall between 45 and 85%.

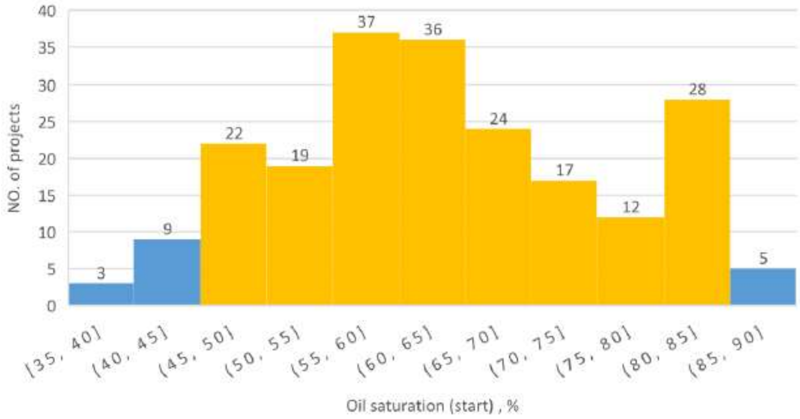

Figure 14. Oil saturation (start) distribution of the dataset.

The Box plots were used to detect outliers, also to display the ranges and summarize the dataset for each variable, as shown in figure 15 Data value ranges were provided for each parameter (minimum and maximum value) after removing outliers. These ranges are illustrated by the distance between the opposite ends of the whiskers. Also, the box plot displays additional information, such as the mean and median of the dataset. The mean value is indicated by orange circles, and the median value by horizontal line between each set of boxes.

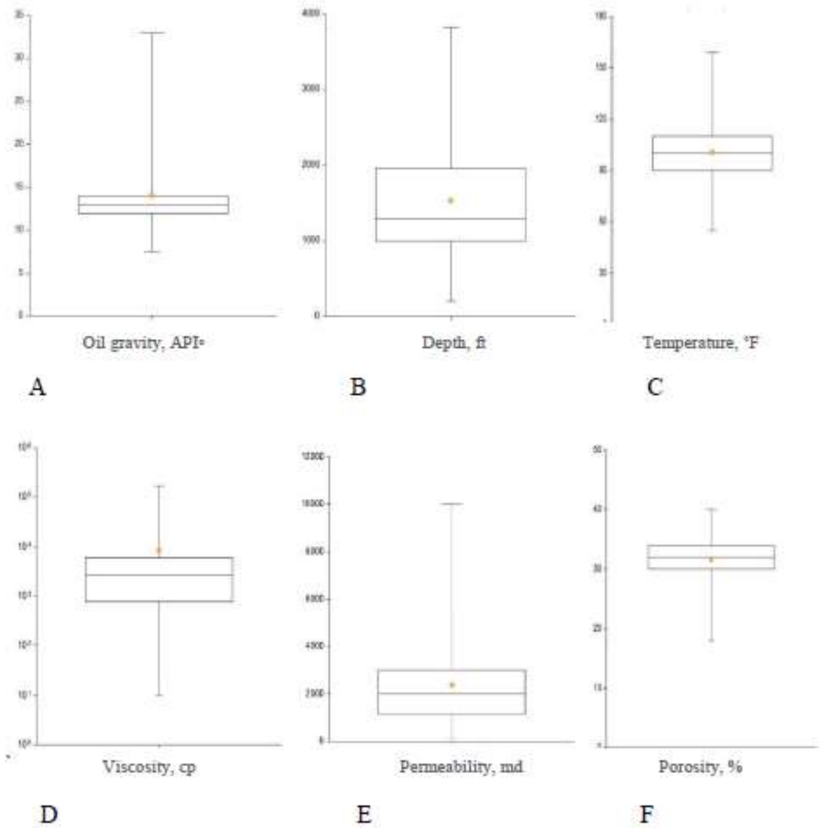

Figure 15. (A) Box plot of oil gravity, (B) Reservoir depth, (C) Reservoir temperature, (D) Oil viscosity, (E) Reservoir permeability, (F) Reservoir porosity.

5. Summarizing Screening Data

Table 1 provides a summary of the updated steam flooding criteria derived from the preceding statistical analysis of the cleaned dataset. This summary includes the screening parameters that have led to the success or failure of steam flooding projects. These parameters include the oil gravity, oil viscosity, reservoir porosity, oil saturation start and end, reservoir permeability, reservoir depth and reservoir temperature. The standard statistics used to describe the criteria are the mean, median, standard deviation, and minimum and maximum values.

Table 1. *Screening Criteria for steam flooding project in data set*

Statistic	Oil Gravity, °API	Oil Viscosity, cp	Porosity, %	Oil saturation, Start	Oil saturation, End	Average Permeability, md	Depth, ft	Temperature, °F
Mean	14.09	8429.98	31.62	65.15	30.54	2367.56	1535.76	100.59
Median	13	2,725	32	65	20	2000	1300	100
Standard Deviation	4.11	21,607.34	3.91	12.59	20.02	1769.19	750.24	18.08
Minimum	7.5	10	18	35	10	5	200	55
Maximum	33	170,000.0	40	90	82	10000	3820	159

6. Recovery Model

Recovery factor is the percentage of the hydrocarbon in place that can be produced with each production plan: primary, secondary and tertiary. In this part, there is a model of recovery for thermal steam injection however, this model was just based on reservoir parameters.
Minitab software was used to estimate those models by using stepwise method.

Table 2 summarizes the results from the multiple linear regression analysis some variables were significant were based on P-value including K, API, So start, Log(K/μ).

Table 2. *The results from the multiple linear regression analysis*.

Term.	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
permeability	0.0000733263339551	0.00002	(0.00003; 0.00011)	3.63	0.000	6.93
API	0.0547749262118239	0.00659	(0.04168; 0.06787)	8.31	0.000	18.84
So start	-0.4841173635166619	0.152	(-0.78652; -0.18171)	-3.18	0.002	23.92
Log K/Visco	-0.1486141900137918	0.0601	(-0.26800; -0.02923)	-2.47	0.015	2.64

In this regression analysis, the R^2 and the adjusted R^2 were 92.3% and 91.78%

The Estimated Model for Recovery factor:

$$RF = 0.000073 \text{ permeability} + 0.05477 \text{ API} - 0.484 \text{ So start} - 0.1486 \text{ Log(K/Viscosity)} \quad (1)$$

7. Conclusions

This paper illustrates the steps and techniques that are used to obtain results which can be applied as criteria for implementing successful steam flooding projects. As a first step, the data were collected from oil and gas journal EOR survey, and then cleaned by removing the missing and duplicate data. After that a statistical analysis techniques were used, which represented in (cross plots and box plots) to show the relationships between different parameters and to detect outlier values, also to determine inconsistent data.

After data cleaning, a powerful statistical tool represented in histogram plot is applied to show the distribution for each parameter of successful steam flooding projects. The box plot was used again to show and clarify the screening criteria of the data set.

Finally, new screening-criteria for steam flooding projects have been introduced and contrasted to previously published criteria. The steam flooding screening criteria can be summarized as follows: oil gravity >33˚API; oil viscosity <170,000 cp; temperature <159˚ F; porosity >18%; permeability >5 md; start oil saturation >35%; . Steam projects were applied mostly in sandstone and unconsolidated sand formations.

Finally, The stepwise technique was used to select the best model based on significant parameters. The estimated model for recovery factor can be applied when recovery factor is greater than 50%.

References

[1] Taber, J. J., et ai, 1996. "EOR Screening Criteria Revisited," SPE/DOE 35385

[2] Laura R. Z., 2012. Advances in Enhanced Oil Recovery Processes, University of New Brunswick, Chemical Engineering Department, Canada

[3] **Saleh, L**., Wei, M., and Bai, B., (203) . Data analysis and updated screening criteria for polymer flooding based on oil field data. SPE reservoir evaluation & reservoir engineering journal. SPE 168220-PA.

[4] **Salah, L**.. Wei, M., and Bai, B., (2016). Data Analysis for polymer flooding that is based on a comprehensive database, Missouri University of Science and Technology.

[5] **Saleh, L**., M.Q. Hama, B. Bai, M. Wei (2014). Updated Screening Criteria for Steam Flooding Based on Oil Field Projects Data. Accepted to present at SPE Heavy Oil Conference-Canada 10-12 June.