WORLD HEAVY OIL CONGRESS 2015
Edmonton, Alberta, Canada

WHOC15-358

Correlation to predict hot oil rate in cyclic steam stimulation processes (CSS). Bare field, Orinoco Oil Belt

R. Mago, PDVSA; E. Morales, PDVSA; J. Rodríguez, PDVSA; F. Armas; R. Vásquez PDVSA.

This paper has been selected for presentation and/or publication in the proceedings for the 2015 World Heavy Oil Congress. The authors of this material have been cleared by all interested companies/employers/clients to authorize dmg::events (Canada) inc., the congress producer, to make this material available to the attendees of WHOC2015 and other relevant industry personnel.

ABSTRACT

A regression analysis is a statistical process that estimates the relationships between variables. In order to quickly, simple and efficiently estimate oil production response in wells submitted to Ciclic Steam Stimulation in Bare field, a correlation was developed to predict hot oil volume in wells, based in historic data of Downhole Monitoring (Pressure and Temperature) and evaluation of the stratigraphic sequences with CSS application. To achieve this, a correlation matrix was built, determining the relationship or impact between the selected independent variables: oil viscosity, sand thickness and permeability being the most involved, aiming to discard possible multicollinearity between variables and obtain a representative regression of the studied phenomenon.

The most impacting variables of the variance in the studied variable were viscosity with 49%, formation thickness with 41% and Permeability 10% of impact. The resulting regression presented a logarithmic linear adjustment, according to the presence of more than two independent variables in the study. Distribution Analysis theory was applied revealing that the obtained regression follows a normal behavior with a distribution average of 135 BPD for CSS wells in Bare field. This regression presents the following range of application: Permeability from 4 to 10 Darcy, Thickness between 20 and 90 ft and oil viscosities reductions from 70 cPs to 530 cPs.

The present correlation coefficient in the adjustment resulted 89%, which indicates that the variables taken into account explain the variance of the dependent variable. The F-statistic value or Fisher test resulted in 28, being the critical value 4.77, which concludes that the studied variables are significant and does not depend on randomness. Finally, the determination coefficient generated was 94% with an arithmetic error of 15%, which indicates that the logarithmic regression presents an excellent goodness of fit and can be used as a predictive model. According to all these results, the obtained regression is an excellent tool to easily and efficiently estimate hot oil production rate of wells submitted to cyclic steam stimulation in the Orinoco Oil Belt.

KEY WORDS

Cyclic Steam Stimulation, regression analysis, logarithmic regression, hot oil production rate.

INTRODUCTION

For multiple reasons, in the cases of viscous oils, thermal methods are applied instead of other recovery methods. When heat is applied, oil viscosity reduces as product of the raise of temperature, which allows oil to flow easily. The principal variable to monitor in thermal recovery projects is temperature, this one being a relative measure of how hot or cold is an object. In thermal processes, steam is the most commonly used medium for transmitting heat into the formation, with the weakness that only a part of the applied heat reaches the objective, as heat losses to surface, to the wellbore and to adjacent formations are generated.

Bare field has been subject to several steam injection processes for more than 40 years, showing excellent results. To the date more than 400 wells with cyclic steam stimulation delivered an extra potential average of 150 BPD/well, with an injection cycle of approximately 1.5 years. During this period field dynamic data was obtained in all the CSS projects which led us to understand and optimize the Cyclic Steam Stimulation processes.

In the present investigation the already stated data was combined with the petrophysical properties of Bare field's reservoirs, aiming to quickly, simply and efficiently estimate the hot oil production rate for CSS wells using an statistical regression model.

MONTE CARLO SIMULATION

Monte Carlo simulation is a quantitative technique that makes use of statistics and computers to imitate, through mathematical models, the non-dynamic real systems randomness behavior. Monte Carlo simulation creates a mathematical model of the system, process or activity to analyze, identifying those variables which random behavior determines the global behavior of the system.

Once these random variables are identified, a following experiment generates, with the help of a computer, random samples and analyzes the system behavior in front of those generated values. After repeating n times this experiment, we count with n observations regarding the system behavior. Useful to understand the system functioning, the analysis will be more precise as more n experiments are available.

REGRESSION ANALYSIS

Nonlinear regression is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. There are cases when the contribution of one independent variable is not sufficient: in many problems more than two variables are inherently involved and it becomes necessary to explore the nature of their relation.

The linear regression equation has the following expression:

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad \text{Ec. 1}$$

Where:

y = Dependent variable.

β_0 = Ordinate origin.

β_0 = Slope.

x = Independent variable.

ε = Random mistake.

Initially an analysis, recollection and validating process of the data involved in the CSS wells in Bare field was done. Injection records were studied to select the wells that were injected according to the optimal injection conditions established for the field: a steam injection rate of 500 tons per day for a period of ten days and applying a soak time of 7 days. 50 wells were studied obtaining a final population of 21 wells.

Then, over the well population, the most theoretically impacting (static and dynamic) variables on the hot production rate during a steam injection were analyzed: viscosity (μ), permeability (K) and sand thickness (H). These variables were declared as independent variables, being the increment of the rate of oil the dependent variable or studied variable.

A correlation matrix was built aiming to verify if independent variables weren't a linear combination between them, in consequence discarding the existence of the mullticollinearity phenomenon that could originate a linear regression not appropriate to predict the studied variable. The following table confirms that no linear combination exists given that the correlation coefficients differ from 1 and -1.

Correlations	μh (cPs)	Permeability (D)	Thickness (ft)
μh (cPs)	1	-0.7349	-0.9115
Permeability (D)		1	0.7648
Thickness (ft)			1

Table 1: Correlation Matrix

According to Ec. 1, the multiple linear regression used to calculate the hot oil rate increment is given by the following mathematical expression:

$$\Delta Q_o = \beta_1 . \mu + \beta_2 . h + \beta_3 . K + \beta_0 \qquad \text{Ec. 2}$$

Where β_1, β_2 & β_3 are the regression coefficients and β_0 is the origin ordinate.

It is possible to find the values of the coefficients β_1, β_2 & β_3 using the Least Squares method, same as in simple linear regression methods.

The method, in this case wants to solve three linear equations with three unknowns, these equations, known as normal equations, are:

$$\sum y = n.\beta_0 + \beta_1 \cdot \left(\sum x_1\right) + \beta_2 \cdot \left(\sum x_2\right) + \beta_3 \cdot \left(\sum x_3\right) \ldots \beta_n \cdot \left(\sum x_n\right)$$

$$\sum x_1 y = \beta_0 \cdot \left(\sum x_1\right) + \beta_1 \cdot \left(\sum x_1^2\right) + \beta_2 \cdot \left(\sum x_2^2\right) + \beta_3 \cdot \left(\sum x_3^2\right) \ldots \beta_n \cdot \left(\sum x_n^2\right)$$

$$\sum x_2 y = \beta_0 \cdot \left(\sum x_2\right) + \beta_1 \cdot \left(\sum x_1 \cdot x_2\right) + \beta_2 \cdot \left(\sum x_2^2\right) + \beta_3 \cdot \left(\sum x_3^2\right) \ldots \beta_n \cdot \left(\sum x_{4n}^2\right) \qquad \text{Ec. 3}$$

Solving the previous system of equations, based on the established dependent and independent variables, generates the following multiple logarithmic regression applying a reliability level of 95%:

$$\Delta Q_o = -40.2 \ln \mu + 33.5 \ln K + 48 \ln H + 118$$

Ec. 4

To determine the correlation coefficient of the obtained regression, the **Pearson** method can be used which allows calculating the correlation between a determined group of variables. The relation intensity between each one of the independent variables and the dependent variable previously declared was calculated. The Pearson's correlation coefficient can be measured following this formula:

$$R = \frac{SSR}{SST}$$

Ec. 5

Where:

SSR = sum of regression squares.

SST = sum of total squares.

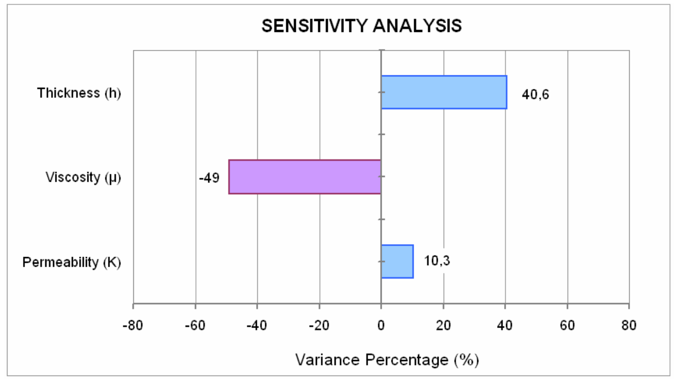
Pearson's correlation coefficient presented a value of 0.89, which indicates that independent variables explain the dependent variable variance in an 89%, on following table show linear terms and their interactions, values "p" for each interactions are less than Significance level "α", indicating that Viscosity, permeability and thickness explain behavior to dependent variable, if reliability level is 95% then significant level is 0.05, so "p" values is less than significant level. The variance inflation factor shows that there does not exist strong mullticollinearity between independent variables.

Term.	Coef	P Value	VIF
Constant	118	0.29	
LN µh	-40.2	0	3.45
LN µh	-40.2	0	3.45
L N H	48	0.034	3.85

Table 2: ANOVA Analysis

Finished the resulting logarithmic linear regression elaboration, a sensitivity analysis was done to verify the impact of each one of the independent variables established in the hot oil rate increment, or dependent variable.

Results indicated that the most impacting variable in the equation is viscosity (µ) with 49%, sand thickness (h) with 40% of statistical variance, followed by permeability (K) 11%.



Graph 1: Sensitivity analysis

In order to use the regression equation obtained with predictive purposes, it is necessary to apply hypothesis tests or statistical significance to verify that the obtained linear regression doesn't depend on randomness and is representative for the studied variable.

Statistical significance test F or Fisher distribution expresses the hypothesis that the average of multiple populations normally distributed with the same standard deviation are the same:

$$F = \frac{MSR}{MSE}$$

Ec. 5

Where:

MSR = Regression square average.

MSE = Error square average.

Applying Ec.5 resulted in an F value of 28.7.

To calculate the F statistical value it is necessary to determine the F critical value, given by:

$$F_\alpha = (1.n-2)$$

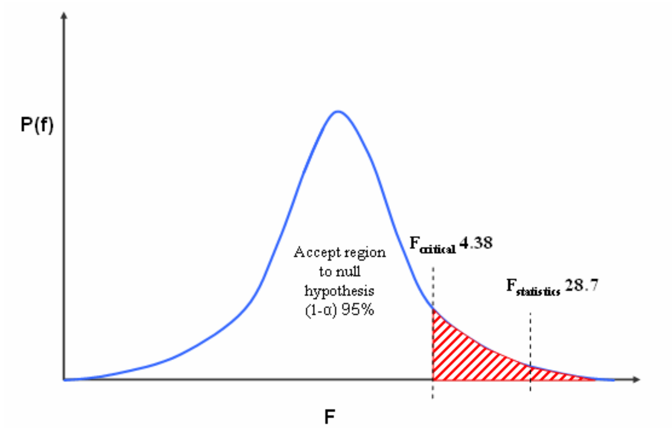
Ec. 6

Where:

α = Significance level of the test.

n= Number of observations.

When calculating the F critical value a level of reliability of 95% was set and the observation number was 21, according to the number of tests studied. Having all this, $F_{critical}$ resulted in the following expression: F0.05 = (1.19). Using Fisher's F-Distribution table, $F_{critical}$ value was determined to be 4.38. Given that $F_{statistical} > F_{critical}$ it is concluded that null hypothesis must be reject and the obtained regression does not depend on randomness and can be used to predict the hot oil rate increment behavior for wells submitted to cyclic steam stimulation.



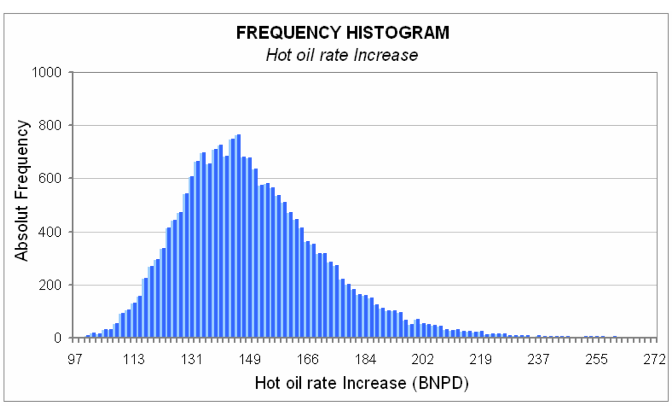
Graph 2: Fisher's F-Distribution graph

Once the resulting logarithmic regression is confirmed not to be originated by randomness and the behavior is really represented, a detailed analysis to this regression was done via statistical simulation programs with the Monte Carlo method, in order to determine the resulting distribution in the studied population, statistic percentiles, each variable's impact in the resulting logarithmic regression and the corresponding sensibility studies. To achieve this study, minimum and maximum standards were established for each one of the independent variables declared for Bare field, seen in Table 2.

Variable	Min	Max
µh (cPs)	70	530
Permeability (D)	4	10
Thickness (ft)	20	90

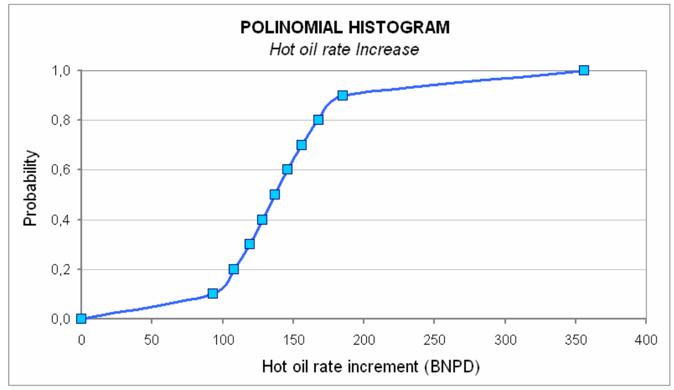
Table 2: Minimum and Maximum standards

Throughout the simulation study, the frequency histogram analysis was obtained showing that the given correlation adjusts to a normal-type distribution with an average of 135 STBPD and a standard deviation of 20 STBPD.



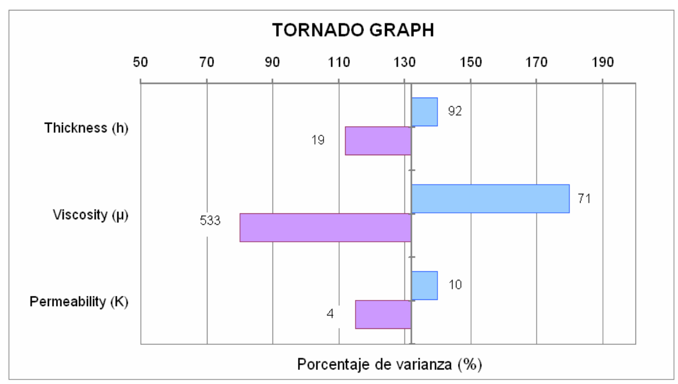
Graph 3: Frequency histogram

The percentiles of the frequency distribution calculated showed that there exists a 10% probability that CSS wells generate increments of 100 STBPD, a 50% probability in the sample that the increments round 135 STBPD and, at last, a 90% probability that the increments reach 250 STBPD.



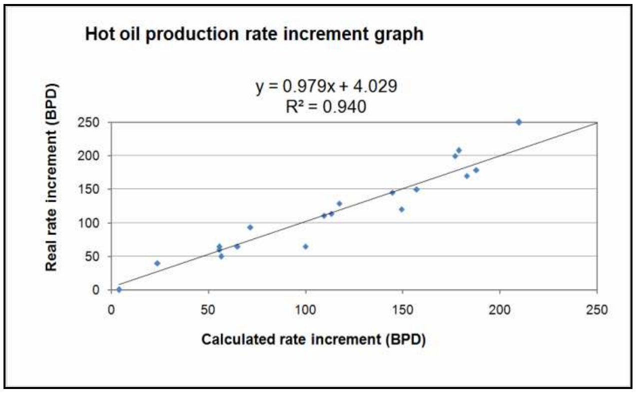
Graph 4: Frequency polynomial

Later, a Tornado diagram was built to evaluate the impact in the dependent variable as consequence of the variation of the independent variables. Maximum values of sand thickness allows the generation of additional oil rate greater than 140 STBPD, equally oil viscosity values near 71 cPs could generate increasing close to 180 STBPD.



Graph 5: Tornado graph

The cross plot comparison presented in the following picture, shows the adjustment of the real values vs. the calculated values of hot oil rate increment by the multiple linear regression model. The given points distribution shows a determination coefficient (R2) of 94%, an arithmetic average error (AE) of 15% and a standard deviation (SD) of 20 STBPD.



Graph N°6 Cross comparative graph

CONCLUSION

According to the obtained results it is concluded that the present static and dynamic parameters: sand thickness, permeability and viscosity affect the hot oil rate increment in wells submitted to cyclic steam stimulation. Pearson's correlation coefficient indicates that 89% of the studied population impacts the dependent variable and the significance F test showed that the resultant logarithmic regression can be used with a 95% reliability level to predict the hot oil rate increment.

ACKNOWLEDGEMENTS

The authors want to express their appreciation to PDVSA E&P for the pertinent permissions to publish the data included in this work.

REFERENCES

1. Rojo Abuin J.M., *Regresión lineal multiple*, Universidad de Madrid (2007).

2. Baran S. A., *Consistent estimator for nonlinear regression models.* Metrika 2005: 62: 1-15.

3. Canavos, George C. *Probabilidad y Estadística. Aplicaciones y Métodos.* (1988). Ed.: Mc Graw Hill.

4. Rodríguez P., *Generación de Vapor para inyección a Yacimientos Petrolíferos.* Venezuela (1986).