Mathematical model as a means of predicting effectiveness of well intervention for near-wellbore space

K F Gabdrakhmanova, G R Izmaylova, M O Mikhaylov

Ufa State Petroleum Technological University, Branch of the University in the City of Oktyabrsky, 54a, Devonskaya St., Oktyabrsky, Republic of Bashkortostan, 452607, Russia

E-mail: gulena-86@mail.ru

Abstract. The conclusion about the appropriateness of using this or that type of well intervention equipment should be based not only on the technological and economic efficiency but also on a mathematical model. The existing geological and field restrictions for the selection of impact objects for each type of well intervention should not also be overlooked due to the fact that in this case the use of a mathematical model is more appropriate. A fundamentally new method of multifunctional regressive mathematical modeling for evaluating the technological efficiency of processing the bottom-hole zone with an increase in the oil recovery coefficient is described. A new algorithm for determining the dependence of the technology efficiency is proposed, the obtained multivariate regression model is checked for statistical reliability and significance, the forecast for the increase in oil recovery coefficient is calculated.

1. Introduction

The last stage of an oil and gas field development is characterized by the importance of searching for the most effective technologies aimed to increase oil recovery and well injection capacity with simultaneous energy savings in the pressure maintenance system. The effective application of one technology in a particular area does not always guarantee its success in another one due to the heterogeneity of reservoir layers. Therefore, it is necessary to treat each site of the field individually, taking into account the analysis of the results concerning well intervention on the basis of the methods of correlation and statistical analysis of the geological and physical as well as geological-field characteristics of the development targets, taking into account the level of economic profitability of the well stock. In this connection, there arose the need to develop new criteria for the efficiency of technologies aimed to increase the oil recovery coefficient on the basis of a mathematical apparatus and stochastic approach. In the works [1, 2], mathematical methods for defining oil recovery coefficient were first proposed. A formula denoting a simple oil recovery coefficient by factor multipliers has been presented:

\[ \eta = \eta_1\eta_2 \]  

(1)

where \( \eta_1 \) is displacement efficiency, \( \eta_2 \) is volumetric displacement efficiency. It is known that the main criterion for changing the oil recovery coefficient occurs under the influence of three main geological and physical factors being macro- and microinhomogeneities of a layer, viscous forces, and surface tension forces.
This engineering approach was further developed [3] in assessing a deposits productivity factor at the stage of the initial design documents development and modeling oil recovery processes in the fields that have been in operation for a long time.

A methodology for assessing the potential effectiveness of new technologies in the development of the oil zone is proposed in [4]. In [5, 6], the role of a multi-parameter analysis of the technologies effectiveness was considered. In studies [6], the use of the Nernst – Planck – Poisson equation is proposed. It is an applied equation in terms of its applicability to the description of phenomena in various media. Algorithms for data mining and machine data processing that provide the effectiveness of a modeling alternative under the conditions when the basic physical relationships between system variables are very complex and non-linear have been proposed [7].

Researchers [8] propose a mathematical model that enables to accurately assess the effectiveness of hydraulic fracturing. Linear regression analysis is suggested for unifying the obtained research results [9–12].

2. The purpose and objectives of the study

The aim of the study was to determine a unified regresional dependence of well intervention effectiveness aimed to increase the oil recovery coefficient on the parameters and indicators of the development status.

To achieve this goal, the following tasks were set:

1. To analyse the experience of using statistical methods for determining the effectiveness of the recovery factor and the technologies used.
2. To derive the regression equation.
3. To study the effectiveness of the derived regression equation for small samples.

3. Materials and methods

A sufficiently large sample of measurements is required for obtaining reliable results, which is not always possible in real production conditions. The path to constructing a regressive equation in the form of a polynomial of the second degree was chosen as the main method for solving this problem.

For this purpose, the geological and production results of the use of thermal oil and gas treatment (thermogas barometric treatment) at 13 production wells of the Pashensky field were systematized. The results are presented in the form of a summary table 1.

It is known that the most common way of processing experimental data is the method of regression analysis, which enables to obtain a mathematical description of the technological process on the basis of experimental data in the form of an algebraic power polynomial. It is known that with an increase in the number of its members, the reliability of the mathematical description of the technological process increases.

Processing experimental data has shown that in most cases the results of an experiment in the form of a tabular function are reflected with a sufficient approximation by a full cubic polynomial, so that the number of the polynomial members can be reduced without significant loss of calculation accuracy.

Of course, in this case, there arises the question of the sample presentability and the results reliability, which in turn leads to the formal use of statistical methods of analysis in solving the problem under study. Despite the low reliability of the formally applied probabilistic-statistical methods, it can be stated that they are widely used to predict the oil recovery coefficient in the absence of multidimensional filtration models of liquid hydrocarbon deposits.
Table 1. Initial data of geological and field studies of the Pashenskoye field

| Well number | \(y\) | \(x_1\) | \(x_2\) | \(x_3\) | \(x_4\) | \(x_5\) | \(x_6\) | \(x_7\) | \(x_8\) | \(x_9\) | \(x_{10}\) | \(x_{11}\) |
|-------------|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1304        | 0.7  | 13.9   | 0.88   | 0.108  | 6.5    | 0.15   | 8.3    | 57     | 6.2    | 0.3    | 0.01   | 800.6  |
| 1273        | 0.8  | 16.79  | 0.90   | 0.510  | 7.0    | 0.19   | 14     | 60     | 19     | 0.41   | 0.34   | 1100   |
| 1230        | 1.1  | 14.49  | 0.90   | 0.510  | 7.2    | 0.19   | 6.8    | 43.1   | 8      | 0.44   | 0.34   | 500    |
| 4126        | 4.9  | 15.55  | 0.86   | 0.184  | 5.4    | 0.16   | 25.6   | 67.9   | 18.3   | 0.41   | 0.36   | 1000   |
| 1353        | 5.5  | 12.27  | 0.90   | 0.510  | 8.8    | 0.19   | 11     | 66     | 15     | 0.41   | 0.34   | 1100   |
| 1660        | 5.5  | 19.02  | 0.86   | 0.184  | 3.6    | 0.19   | 9.3    | 27     | 22.8   | 0.41   | 0.48   | 900    |
| 1657        | 7.7  | 18.62  | 0.90   | 0.510  | 6.0    | 0.19   | 9.7    | 68.5   | 6.2    | 0.41   | 0.34   | 900    |
| 2457        | 8.6  | 11.67  | 0.90   | 0.510  | 3.2    | 0.19   | 12.5   | 66.7   | 15     | 0.41   | 0.34   | 970    |
| 2406        | 9.9  | 15.03  | 0.90   | 0.510  | 6.0    | 0.16   | 5.5    | 65     | 22.6   | 0.44   | 0.34   | 1000   |
| 1153        | 10.7 | 15.38  | 0.90   | 0.510  | 6.4    | 0.16   | 15.6   | 66.2   | 2      | 0.44   | 0.34   | 800    |
| 1764        | 11.3 | 24.21  | 0.90   | 0.184  | 2.6    | 0.147  | 8.8    | 73.4   | 14.8   | 0.35   | 0.36   | 800    |
| 1460        | 13   | 19.24  | 0.86   | 0.184  | 7.0    | 0.19   | 8.2    | 67.4   | 14     | 0.41   | 0.36   | 1100   |
| 2497        | 14.6 | 16.64  | 0.86   | 0.07   | 4.4    | 0.14   | 15.8   | 50.4   | 5      | 0.17   | 0.12   | 800    |

In table 1, the following notation is used: \(y\) is an increase in flow rate (t per day), \(x_1\) is a reservoir pressure (MPa), \(x_2\) is an oil saturation coefficient (unit), \(x_3\) is a permeability coefficient (\(\mu m^2\)), \(x_4\) is oil-bearing thickness (m), \(x_5\) is a porosity coefficient (d. ed), \(x_6\) is proppant mass in the layer (t), \(x_7\) is a final pressure (MPa), \(x_8\) is a water cut coefficient (\%), \(x_9\) is an initial oil recovery coefficient (d. ed), \(x_{10}\) is a current oil recovery coefficient (unit fraction), \(x_{11}\) is proppant concentration (kg/m³). For a table function (table 1), we compose a polynomial of degree 2 having the following form:

\[
y_0 = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_4 \cdot x_4 + b_5 \cdot x_5 + b_6 \cdot x_6 + ... + b_{11} \cdot x_1 x_1 + b_3 \cdot x_3 x_3 + ... + b_{10} \cdot x_{10} x_{10} + b_{11} \cdot x_{11} x_{11},
\]

where \(x_1 - x_{11}\) are table values 1.

Let us study the correlation between \(y\) and \(x_1, x_2, x_3, ..., x_6\). We obtain the following results (table 2). It follows from table 2 that a close relationship between these components was not detected, with the exception of \(x_1, x_3, x_6, x_{10}\).

Table 2. Source data correlation

| \(y\) | \(x_1\) | \(x_2\) | \(x_3\) | \(x_4\) | \(x_5\) | \(x_6\) | \(x_7\) | \(x_8\) | \(x_9\) | \(x_{10}\) | \(x_{11}\) |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1    | –      | –      | –      | –      | –      | –      | –      | –      | –      | –      | –      |
| 0.362368 | 1      | –      | –      | –      | –      | –      | –      | –      | –      | –      | –      |
| 0.21639 | -0.19663 | 1      | –      | –      | –      | –      | –      | –      | –      | –      | –      |
| -0.23358 | -0.42532 | 0.818062 | 1      | –      | –      | –      | –      | –      | –      | –      | –      |
| -0.38609 | -0.44314 | 0.196188 | 0.3993763488 | 1      | –      | –      | –      | –      | –      | –      | –      |
| -0.35047 | -0.19566 | 0.200137 | 0.552918087 | 0.33673027 | 1      | –      | –      | –      | –      | –      | –      |
| 0.019696 | -0.14985 | -0.35728 | -0.1892059 | -0.0965889 | -0.24307 | 1      | –      | –      | –      | –      | –      |
| 0.315304 | 0.037097 | 0.386946 | 0.226750393 | 0.10820026 | -0.18506 | 0.189767 | 1      | –      | –      | –      | –      |
| 0.14406 | 0.109873 | -0.12117 | 0.025844413 | -0.1814912 | 0.287751 | -0.04256 | -0.11387 | 1      | –      | –      | –      |
| -0.27952 | -0.11848 | 0.360131 | 0.626746969 | 0.26578869 | 0.595272 | -0.02354 | 0.203481 | 0.43033 | 1      | –      | –      |
| 0.068136 | 0.278573 | 0.124420 | 0.3972i1074 | -0.1230575 | 0.58673 | 0.00094 | -0.05368 | 0.56600 | 0.745226 | 1      | –      |
| 0.133594 | -0.06204 | -0.13070 | 0.07936459 | 0.17977783 | 0.281831 | 0.208175 | 0.401373 | 0.56006 | 0.298175 | 0.231855 | 1      |
Let us analyze the pairwise components, which are summarized in Table 3.

### Table 3. Source data containing products of variables

| Well number | $x_1/x_2$ | $x_1/x_3$ | $x_1/x_4$ | $x_1/x_5$ | $x_1/x_6$ | $x_1/x_7$ | $x_1/x_8$ | $x_1/x_9$ | $x_1/x_{10}$ | $x_1/x_{11}$ |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-------------|-------------|
| 1           | 12.232    | 90.35     | 2.085     | 115.37    | 792.3     | 86.18     | 4.17      | 11128.34  |             |             |
| 2           | 15.111    | 117.53    | 3.1901    | 235.06    | 1007.4    | 319.01    | 6.8839    | 18469     |             |             |
| 3           | 13.041    | 104.328   | 2.7531    | 98.532    | 624.519   | 115.92    | 5.9409    | 7245      |             |             |
| 4           | 13.373    | 83.97     | 2.488     | 398.08    | 1055.845  | 284.565   | 6.842     | 15550     |             |             |
| 5           | 11.043    | 107.976   | 2.3313    | 134.97    | 809.82    | 184.05    | 5.0307    | 13497     |             |             |
| 6           | 16.3572   | 68.472    | 3.6138    | 176.886   | 513.54    | 433.656   | 7.7982    | 17118     |             |             |
| 7           | 16.758    | 111.72    | 3.5378    | 180.614   | 1275.47   | 115.444   | 7.6342    | 16758     |             |             |
| 8           | 10.503    | 37.344    | 2.2173    | 145.875   | 778.389   | 175.05    | 4.7847    | 11319.9   |             |             |
| 9           | 13.527    | 90.18     | 2.4048    | 82.665    | 976.95    | 339.678   | 6.6132    | 15030     |             |             |
| 10          | 13.842    | 98.432    | 2.4608    | 239.928   | 1018.156  | 30.76     | 6.7672    | 12304     |             |             |
| 11          | 21.789    | 62.946    | 3.55887   | 213.048   | 1777.014  | 358.308   | 8.4735    | 19368     |             |             |
| 12          | 16.5464   | 134.68    | 3.6556    | 157.768   | 1296.776  | 269.36    | 7.8884    | 21164     |             |             |
| 13          | 14.3104   | 73.216    | 2.3296    | 262.912   | 838.656   | 83.2      | 2.8288    | 13312     |             |             |

Using the Data Analysis function of the MS Excel program, we perform a correlation analysis of the data in Table 3 for the presence of indirect relationships. Since there is no close relationship between the components, all of them can be included in the equation. We carry out a correlation analysis for the presence of indirect connections between $x_2 \cdot x_3$, ..., $x_2 \cdot x_{11}$ based on the obtained results, shown in Table 4.

### Table 4. Product variations correlation $x_2 \cdot x_3$, ..., $x_2 \cdot x_{11}$, $x_3 \cdot x_6$, ..., $x_3 \cdot x_{11}$

| $x_2 \cdot x_3$ | $x_2 \cdot x_4$ | $x_2 \cdot x_5$ | $x_2 \cdot x_6$ | $x_2 \cdot x_7$ | $x_2 \cdot x_8$ | $x_2 \cdot x_9$ | $x_2 \cdot x_{10}$ | $x_2 \cdot x_{11}$ |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 0.39918        | 0.83900        | 0.53853        | 0.50740        | 0.93102        | 0.80115        | 0.58709932     | 0.838499       | 0.517801      |
| 0.21923        | 0.434624       | 1              | 0.394102       | 0.46132        | 0.860115       | 0.55709932     | 0.838499       | 0.517801      |
| 0.684567       | 0.695567       | 0.81169169     | 1              | 0.860115       | 0.55709932     | 0.838499       | 0.517801      | 0.517801      |
| 0.02131        | 0.84260        | 0.101567       | 0.7093721      | 0.0585135      | 0.14155        | 1              | 0.860115       | 0.517801      |
| 0.36434        | 0.24418        | 0.640023       | 0.23411030     | 0.4356033      | 0.50818        | 0.1842         | 1              | 0.860115       |
| 0.02844        | 0.010515       | 0.125967       | 0.79583967     | 0.1026033      | 0.21178        | 0.06801        | 0.011231      | 1              |
| 0.514429       | 0.104687       | 0.3207678      | 0.493352851    | 0.64709105     | 0.379546       | 0.111766       | 0.432319       | 0.518986      |
| 0.425489       | 0.232446       | 0.4347509      | 0.684482183    | 0.41800242     | 0.690178       | 0.250228       | 0.561565       | 0.639215      |

Observing the tight connection between the variables, excluding those with a weak connection, we come to the equation in the following primitive form:

$$y_0 = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + ... + b_6 \cdot x_6 + b_{12} \cdot x_1 x_2 + b_{14} \cdot x_1 x_4 + b_{14} \cdot x_1^2 + b_{22} \cdot x_2^2 + ... + b_{1111} \cdot x_{11}^2$$  \(3\)

To search for the coefficients of this regression equation, we use the MS Excel’s Solution search function. Let us make up a comparison table of calculated and actual values (Table 5) by sequential substitution of the initial data from Table 1 for all wells into the obtained equation. The obtained results show a high level of convergence of the calculated and actual data (the relative error in this case is not more than 0.001%).
Table 5. Comparison of calculated and actual values $y$

| Well number | Estimated Values | Actual Values | Deviation | Quadratic deviation | Error, % |
|-------------|------------------|---------------|-----------|---------------------|---------|
| 1           | 0.70             | 0.7           | 0.001111  | 0.000001234         | 0.00126 |
| 2           | 0.80             | 0.8           | 0.000176  | 0.00000031          | 0.00011 |
| 3           | 1.10             | 1.1           | 0.001240  | 0.00001538          | 0.01120 |
| 4           | 4.90             | 4.9           | -0.001146 | 0.00001313          | 0.00250 |
| 5           | 5.50             | 5.5           | -0.000245 | 0.0000060           | 0.00063 |
| 6           | 5.50             | 5.5           | -0.000575 | 0.00000331          | 0.02212 |
| 7           | 7.70             | 7.7           | -0.000709 | 0.00000503          | 0.00043 |
| 8           | 8.60             | 8.6           | 0.001910  | 0.00003648          | 0.00344 |
| 9           | 9.90             | 9.9           | -0.000419 | 0.00000176          | 0.00055 |
| 10          | 10.70            | 10.7          | 0.001138  | 0.00001295          | 0.00194 |
| 11          | 11.30            | 11.3          | -0.000730 | 0.00000533          | 0.00619 |
| 12          | 13.00            | 13.0          | -0.000247 | 0.00000061          | 0.00048 |
| 13          | 14.60            | 14.6          | -0.001867 | 0.00003486          | 0.00275 |

According to the graph (figure), the difference between the actual and calculated results is minimal.

If we turn to the graph, we will determine that the deviations from the actual and the forecast results are very small.

4. Conclusion
The authors have proposed a methodology for assessing the forecast of technological efficiency of measures aimed to increase oil recovery from productive reservoirs of oil fields. This methodology is based on the analysis of the results of experimental and industrial testing of a specific EOR technique (increasing oil recovery of productive reservoirs). According to the results obtained, a correlation analysis was carried out in order to assess the magnitude and range of diagnostic criteria that ensure the expected technological effect. The authors have been working for many years on the use of methods of
statistical analysis of the applied technologies. In the study, it is necessary to continue work on determining the range of the effectiveness criterion and the choice of the applied technologies parameter.

1. The analysis of the statistical methods enables to state that each enterprise has its own methods for assessing effectiveness. The considered estimates have their advantages and disadvantages. The main disadvantages include the following:
   - inability to establish the exact reasons for the increase or decrease in the effect of exposure;
   - no accurate prediction of the effectiveness of technology used by enterprises
2. The regression equation contributing to determining the efficiency of technologies used by enterprises and tight links between indicators, as well as the one which can serve as a criterion for predicting the applied technologies effectiveness has been derived.
3. The technique for forecasting EOR based on the construction of a regression equation of the second degree proposed by the authors contributes to obtaining predicted results in a limited sample with a high degree of convergence with actual data in the range 0.5–1.5%.
4. The technique enables to evaluate the effectiveness of the results obtained not only in the oil industry but also in other areas of activity that introduce various technologies.

References
[1] Gabdrakhmanova K F, Izmailova G R, Larin P A, Vasilyeva E R, Madjidov M A, Marupov S R 2018 Nomogram method as means for resource potential efficiency predicative aid of petrothermal energy Journal of Physics: Conference Series 1015(3) 032036 1-5
[2] Andreev V E, Chizhov A P, Chibisov A V and Mukhometshin V Sh 2019 Forecasting the use of enhanced oil recovery methods in oilfields of Bashkortostan IOP Conference Series: Earth and Environmental Science (International Symposium «Earth sciences: history, contemporary issues and prospects») 350(1) 1–6 DOI: 10.1088/1755-1315/350/1/012025
[3] Gabdrakhmanova K F, Izmaylova G R, Larin P A 2018 The way of using geothermal resources for generating electric energy in wells at a late stage of operation IOP Conference Series: Earth and Environmental Science 194(8) 082012 1-9
[4] Akhmetov R T, Mukhametshin V V and Kuleshova L S 2019 Simulation of the absolute permeability based on the capillary pressure curves using the dumbbell model Journal of Physics: Conference Series (ITBI 2019 – International Conference "Information Technologies in Business and Industry") 1333(3) 1-8 DOI: 10.1088/1742-6596/1333/3/032001
[5] Tyncherov K T, Mukhametshin V Sh, Paderin M G, Selivanova M V, Shokurov I V, Almukhametova E M 2018 Thermoacoustic inductor for heavy oil extraction IOP Conference Series: Materials Science and Engineering 327(4) № 042111 DOI: doi:10.1088/1757-899X/327/4/042111
[6] Rogachev M K, Mukhometshin V V and Kuleshova L S 2019 Improving the efficiency of using resource base of liquid hydrocarbons in Jurassic deposits of Western Siberia Journal of Mining Institute 240 711-715 DOI: 10.31897/PMI.2019.6.711
[7] Chakma A and Jha K N 1992 Heavy-Oil Recovery From Thin Pay Zones by Electromagnetic Heating SPE 24817 p 525 – 534
[8] Davletbaev A, Kireev V, Kovaleva L, Zainullin A, Minnigalimov R 2016 Cold heavy oil production and production by radio-frequency electromagnetic radiation: Comparative numerical study AIP Conference Proceedings: International Conference of Computational Methods in Sciences and Engineering ICCMSE 1790 p 150021
[9] Effati S, Janfada M, Esmaeili M 2008 Solving the optimal control problem of the parabolic PDEs in exploitation of oil Journal of Mathematical Analysis and Applications 340(1) p 606-620
[10] Zhangxin Chen 2007 Homogenization and simulation for compositional flow in naturally fractured reservoirs Journal of Mathematical Analysis and Applications 326(1) p 12-32
[11] Tiago M. Magalhaes, Denise A. Botter, Monica C. Sandoval 2013 Asymptotic skewness for the beta regression model Statistics & Probability Letters 839(10) p 2236-2241
[12] Theodore S. Glickman and Feng Xu 2008 The distribution of the product of two triangular random variables *Statistics & Probability Letters* 78(16) p 2821-2826