The results of applying the method of mathematical statistics to predict technological effectiveness of hydraulic fracturing

Al F Shakurova and Ay F Shakurova

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

E-mail: alsu0017@mail.ru

Abstract. The article discusses the apparatus of mathematical statistics for predicting the technological effectiveness of hydraulic fracturing. The authors made an attempt to improve the methods for processing experimental data in the conditions of a small sampling. The main method used for solving this problem was the second-order regression equation in the form of a polynomial. The comparison of calculated and actual data on technological results for hydraulic fracturing based on the solution of second-order polynomial regression equations for different fields showed a high level of convergence, which allows us to conclude that the proposed methodology for constructing a regression equation in the form of a power polynomial is acceptable.

1. Introduction
The most common way of processing experimental data is the regression analysis, which allows us 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, reliability of the mathematical description of the technological process improves [1]. The practice of processing experimental data has shown that in most cases experimental results in the form of a tabular function are reflected with a sufficient approximation by a complete cubic polynomial. The third degree is sufficient and excessive, i.e. the number of terms of the polynomial can be reduced without significant loss of accuracy [2].

2. Materials and methods
As a research tool, we use the standard Microsoft Office software package (MS Excel). Functions that implement statistical methods of data processing and analysis are implemented in MS Excel in the form of special software - Analysis Package; for solving optimization problems, we used "Solution Search" [2].

3. Results and discussion
The apparatus of mathematical statistics has been widely used to process measurements of field data when evaluating the effectiveness of various methods of increasing oil recovery, as well as to predict the technological effect when planning hydraulic fracturing. The hydraulic fracturing
technology is sufficiently developed in relation to various geological, physical and geological and technical conditions [3].

However, the widespread practice of using this apparatus shows that in order to converge calculation results the actual data, a large sampling of measurements is required, which is not always possible in real production conditions.

According to [4, 5], the technological effect of hydraulic fracturing is influenced by geological, physical, and geological and technological parameters of the objects. Based on paired correlation dependencies, parameters that have a greater effect on the effectiveness of hydraulic fracturing was used. These parameters were used to forecast hydraulic fracturing results.

An attempt was made to improve the methods for processing experimental data in the conditions of a small sampling. As the main method for solving this problem was the second-order regression equation in the form of a polynomial.

For this purpose, the geological production results of hydraulic fracturing in the Bavlin field were analyzed for wells of Bobrikov and Devonian deposits in the Pavlovsk and East Leninogorsk areas of the Romashkin field. Based on the paired correlation dependencies, parameters that have a greater effect on the effectiveness of hydraulic fracturing were used. These parameters were used to predict the results of hydraulic fracturing.

The most common way of processing experimental data is the method of regression analysis, which allows us 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, reliability of the mathematical description of the technological process improves.

The practice of processing experimental data has shown that in most cases the experimental results in the form of a tabular function are reflected with a sufficient approximation by a complete cubic polynomial. The third degree is not only sufficient, but also excessive, i.e. the number of members of the polynomial can be reduced without significant loss of accuracy.

Tabular functions were constructed for each hydraulic fracturing object.

Based on the tabular functions and using the MS Excel software, for each hydraulic fracturing object, second-order regression equations were constructed in the form of a polynomial:

\[
Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 + b_6 x_6 + b_7 x_7 + b_8 x_8 + b_9 x_9 + b_{10} x_{10} + b_{11} x_{11} + b_{12} x_{12} + \ldots + b_{66} x_{66} + b_{67} x_{67} + b_{68} x_{68} + b_{69} x_{69} + b_{70} x_{70} + \ldots + b_{27} x_{27} + b_{28} x_{28} + b_{29} x_{29} + \ldots + b_{78} x_{78},
\]

where \( Y = \Delta Q_0 \) (additional oil production), \( x_1 = K_p \) (porosity coefficient), \( x_2 = K_c \) (clay factor), \( x_3 = h_{\text{perf}} \) (perforation power), \( x_4 = h_{\text{eff}} \) (effective reservoir power), \( x_5 = K_{\text{os}} \) (oil saturation coefficient), \( x_6 = K_{\text{pr}} \) (permeability coefficient), \( x_7 = V_f / V_p \) (ratio of fracture fluid volume to proppant volume), \( x_8 = P_{\text{hp}} - P_p \) (difference between pressure of the hydraulic fracturing and layer pressure).

Let us study the components of the polynomial by means of a correlation analysis. Correlation is used to quantify the relationship of two data sets presented in a dimensionless form. The correlation coefficient of the sample is the covariance of two data sets divided by the product of their standard deviations [6].

The correlation analysis makes it possible to establish whether large values from one data set are associated with large values of another set (positive correlation), or, conversely, whether small values of one set are associated with large values of another one (negative correlation), or the data of two data sets are not connected in any way (the correlation is close to zero).

Then, using the “Analysis Package”, the terms included in equation (1) was studied to assess the degree of their relationship and exclude identical ones from this equation.

To search for the coefficients of the regression equation, we used the “Solution Search” in “MS Excel”.

“Solution Search” is a powerful data analysis tool in “MS Excel”. It can determine at what values of the cells the formula in the target cell takes the desired value (minimum, maximum, or equal to any value). You can set restrictions for the solution search procedure, and it is not necessary that the same
influencing cells are used. To calculate the set value, various mathematical search methods are used. You can set the mode in which the obtained values of the variables are automatically entered in the table. In addition, the results of the program can be presented in the form of a report.

The program “Solution Search” (in the original “Excel Solver”) is an additional add-in of the table processor “MS Excel”, which is designed to solve systems of equations, as well as linear and nonlinear optimization problems. The size of the problem that can be solved using the basic version of this program is limited by the following indicators:

- the number of unknowns is 200;
- the number of formula restrictions on the unknown is 100;
- the number of limit conditions for the unknown is 400.

Solution Search Options:

- Maximum time serves to limit the time allowed to find a solution. In this field you can enter the time in seconds, not exceeding 32,767 (approximately nine hours); a default value of 100 is acceptable for most simple tasks.

- The limiting number of iterations controls the time of solving the problem by limiting the number of computational cycles (iterations).

- Relative error determines the accuracy of calculations. The lower the value of this parameter, the higher the accuracy of the calculations.

- Tolerance is designed to specify the tolerance for deviation from the optimal solution, if the set of values of the influencing cell is limited to the set of integers. The larger the tolerance value, the less time it takes to find a solution.

In conclusion, using the methodology for finding the minimum standard deviation between the actual and calculated values (Table 1), the numerical values of the coefficients included in the regression equation (1) were determined with respect to each object of impact using a hydraulic fracturing technology.

As a result, the following second degree regression equations were obtained [7]:

- for Bobrikov deposits:
  \[ Y = 1.943 + 19.499x_1 + 1.929x_2 - 12.92x_3 + 78.277x_5 - 73.183x_6 + 6.6x_7 + 54.217x_8 - 88.92x_9 - 282.227x_{10} - 20.788x_{11} - 3.572x_{12} + 9.045x_{13} + 18.903x_{14} + 5.608x_{15} + 26.116x_{16} - 80.932x_{17} - 4.025x_{18} + 123.273x_{19} + 76.774x_{20} - 24.542x_{21} - 44.803x_{22} + 6.606x_{23} + 0.087x_{24} + 1.214x_{25} - 7.075x_{26}^2, \]
  \( (2) \)

where \( Y = \Delta Q_{os}, x_1 = K_p, x_2 = K_s, x_3 = h_{eff}, x_4 = x_5 = K_{os}, x_6 = K_{per}, x_7 = V_0/V_{pr}, x_8 = P_{hfp} - P_p. \)

This equation (2) ensured that the calculated and actual values coincided with an accuracy of 1.94%.

- for Devonian sediments of the Baivlin deposit:
  \[ Y = -0.0104 - 0.157x_1 - 0.042x_2 - 0.017x_3 - 0.253x_4 - 0.451x_5 + 0.048x_6 - 0.773x_7 - 0.045x_{10} - 0.348x_{12} - 3.179x_{13} - 0.342x_{14} + 1.25x_{15} - 7.18x_{16} - 7.697x_{17} - 3.96x_{18} - 0.95x_{19} - 0.466x_{20} + 12.518x_{21} - 2.578x_{22} + 0.954x_{23} - 0.017x_{24} + 5.774x_{25} - 0.062x_{26} + 0.881x_{27} - 0.354x_{28}, \]
  \( (3) \)

where \( Y = \Delta Q_{os}, x_1 = K_p, x_2 = K_s, x_3 = h_{eff}, x_5 = K_{os}, x_6 = K_{per}, x_7 = V_0/V_{pr}, x_8 = P_{hfp} - P_p. \)

This equation (3) ensured that the calculated and actual values coincided with an accuracy of 0.95%.

- for Devonian deposits of Pavlovsk and East Leninogorsk areas:
  \[ Y = -4.482 - 42.359x_1 + 43.28x_2 + 36.319x_3 - 190.331x_4 - 304.513x_5 + 51.451x_6 - 968.722x_7 - 513.079x_{10} + 6.361x_{12} - 138.067x_{13} + 341.51x_{14} + 84.749x_{15} + 22.675x_{16} + 22.674x_{17} + 108.98x_{18} + 846.583x_{19} - 121.647x_{20} + 24.626x_{21} + 0.019x_{22} - 48.525x_{23} - 2.435x_{24}, \]
  \( (4) \)

where \( Y = \Delta Q_{os}, x_1 = K_p, x_2 = K_s, x_3 = h_{eff}, x_5 = K_{os}, x_6 = K_{per}, x_7 = V_0/V_{pr}, x_8 = P_{hfp} - P_p. \)

This equation (4) ensured that the calculated and actual values coincided with an accuracy of 1.1%.
Table 1. Comparison of calculated values of additional oil production with actual field values obtained for Bobrikov deposits of the Bavlin field

| No | Calculation | Fact  | Deviation | Square deviation | Error, % |
|----|-------------|-------|-----------|-----------------|----------|
| 1  | 2956.2696   | 2956.3| -0.0303515| 0.000921214     | 0.01     |
| 2  | 6402.1293   | 6402  | 0.12928633| 0.016714955     | 0.02     |
| 3  | 4993.3307   | 4994  | -0.66933333| 0.448007102     | 0.13     |
| 4  | 4857.6724   | 4857.9| -0.22761772| 0.051809824     | 0.05     |
| 5  | 1732.6263   | 1733  | -0.37365975| 0.139621607     | 0.22     |
| 6  | 3402.8769   | 3402  | 0.87693764 | 0.769019619     | 0.26     |
| 7  | 2371.8132   | 2373  | -1.18675585| 1.408389451     | 0.50     |
| 8  | 907.48939   | 907.6 | -0.11060918| 0.012234391     | 0.12     |
| 9  | 1603.2146   | 1602.6| 0.61458128 | 0.377710147     | 0.38     |
| 10 | 5004.9362   | 5005.3| -0.36378511| 0.132339604     | 0.07     |
| 11 | 9813.8538   | 9812.92| 0.93383358 | 0.872045153     | 0.10     |
| 12 | 2924.4344   | 2924.3| 0.13437208 | 0.018055856     | 0.05     |
| 13 | 3826.3922   | 3826.5| -0.10782055| 0.011625272     | 0.03     |
|    | Total       |       | ∑4.258494195| ∑1.94%          |          |

4. Conclusion

Thus, it was proved that even in the conditions of a small representative data sampling the power regression equation makes it possible to accurately predict the technological effect of hydraulic fracturing for the Bobrikov and Devonian deposits, provided that they are constructed taking into account specific geological, production and technological data corresponding to these deposits within each of these productive areas.

The comparison of calculated and actual technological results in hydraulic fracturing based on the solution of second-order polynomial regression equations for different fields showed a high level of convergence, which allows us to conclude that the proposed method for constructing a regression equation in the form of a power polynomial is acceptable.

References

[1] Mirzadzhanzade AKh and Stepanova G S 1977 *The mathematical theory of an experiment in the oil and gas industry* (Moscow: Nedra)

[2] Shashkov V B 2003 *Applied Regression Analysis. Multivariate Regression: Manual.* (Orenburg: OSU)

[3] Economides M J and Nolte K G 1989 Reservoir Stimulation *Prentice Hall, Eglewood Cliffs, New Jersey 07632*

[4] Shakurova Al F and Shakurova Ay F 2018 The influence of hydraulic fracturing on the estimated ultimate recovery *IOP Conference Series: Earth and Environmental Science* 194(8) 082039 1-4.

[5] Shakurova A F 2010 The study of parameters affecting the efficiency of hydraulic fracturing in the conditions of the Romashkinskoye field *Scientific and technical journal "Oil field business"* 1 P. 17-19.

[6] Kuleshova L S, Mukhametshin V V and Safiullina A R 2019 Applying information technologies in identifying the features of deposit identification under conditions of different oil-and gas provinces *Journal of Physics: Conference Series (ITBI 2019 – International Conference "Information Technologies in Business and Industry")* 1333(7) 1-5 DOI: 10.1088/1742-6596/1333/7/072012

[7] Shakurov Al F and Shakurova Ay F 2013 Determining the economic feasibility of applying hydraulic fracturing technology in the low-permeable formation zone *Electronic scientific journal "Oil and Gas Business"* 2 254-261, available at: http://www.ogbus.ru/authors/Shakurova/Shakurova_3.pdf