Developing multilevel statistical models of determining bottom-hole flowing pressure in commercial oil well operations

Inna N. Ponomareva * , Vladislav Ign. Galkin , Alexander V. Rastegaev ,
Sergey Vl. Galkin

Department of Oil and Gas Technologies, Perm National Research Polytechnical University, Russia

https://doi.org/10.54139/revinguc.v28i1.3

Abstract.- One of the major tasks of monitoring oil production well operations is to determine bottom-hole flowing pressure (BHFP). The overwhelming majority of wells in the Perm krai are serviced using borehole pumps, which makes it difficult to take direct bottom-hole flowing pressure measurements. The BHFP in these wells is very often determined by recalculating the parameters measured at the well mouth (annulus pressure, dynamic fluid level depth). The recalculation is done by procedures based on analytically determining the characteristics of the gas-liquid mixture in the wellbore, which is very inconsistent to perform due to the mixture's complex behavior. This article proposes an essentially different approach to BHFP measurements, that relies on the mathematical processing of the findings of more than 4000 parallel mouth and deep investigations of the oil production wells of a large oil-production region. As a result, multivariate mathematical models are elaborated that allow reliably determining the BHFP of oil-production wells in operation.

Keywords: oil production well; bottom-hole flowing pressure; BHFP determination technique; multivariate statistical model; regression analysis; multilevel modeling.

Desarrollo de modelos estadísticos multinivel para determinar la presión de flujo de fondo en operación de pozos petroleros

Resumen.- Una de las principales tareas del monitoreo de las operaciones de los pozos de producción es determinar la presión de flujo del fondo del pozo. La inmensa mayoría de los pozos en Perm krai reciben servicio mediante bombas de pozo, lo que dificulta la toma de mediciones directas de la presión de flujo del fondo del pozo. La presión de flujo de fondo en estos pozos se determina muy a menudo recalculando los parámetros medidos en la boca del pozo (presión anular, profundidad del nivel de fluido dinámico). El recálculo se realiza mediante procedimientos basados en la determinación analítica de las características de la mezcla gas-líquido en el pozo, que es muy inconsistente de realizar debido al complejo comportamiento de la mezcla. Este artículo propone un enfoque esencialmente diferente para las mediciones de la presión de flujo del fondo del pozo, que se basa en el procesamiento matemático de los hallazgos de más de 4000 investigaciones paralelas de boca y profundidad de los pozos de producción de petróleo de una gran región de producción de petróleo. Como resultado, se elaboran modelos matemáticos multivariados que permiten determinar de manera confiable la presión de flujo de fondo de los pozos productores de petróleo en operación.

Palabras clave: pozo de producción; presión de flujo del fondo del pozo; técnicas de determinación de BHFP; modelo estadístico multivariable; modelado multinivel.

Received: December 20, 2020.
Accepted: February 24, 2021.

1. Introduction

Well operation monitoring is an integral part of oil and gas production optimization [1, 2, 3]. One of the topical tasks of this monitoring is to determine the BHFP [4, 5, 6]. In addition, the BHFP is the parameter linking the work of the elements of the reservoir-well system. The BHFP level can be used to control and manage the operation of downhole equipment [7]. As a matter of practice, if a well is operated using deep-pumping equipment, the BHFP is determined by calculation [8]. If the suction manifold is

* Correspondence author:
*e-mail:*ponomarevaIN@pstu.ru (I. Ponomareva)

equipped with a special instrument (sensor), the pressure measured with its help and referred to as the suction pressure is recalculated to the BHFP. As a rule, these calculations are not impaired by any particular difficulties and are made fairly stably. If the deep-pumping equipment configuration has no room for installing a measuring instrument at the suction manifold, the BHFP is determined by recalculating the parameters that are measured at the well mouth and include dynamic fluid level depth and annulus pressure [9]. This recalculation is mathematically based on the hydrostatic equation. In this case, the quantity to determine is the wellbore fluid density [10]. This fluid is a gas-liquid mixture the parameters of which are very difficult to describe by analytical equations [11, 12, 13]. In this article, all of the BHFP determination procedures based on recalculating the parameters measured at the well mouth and recalculated considering the gas-liquid mixture density are called density-based techniques.

In the Perm krai oil is produced at more than 5.000 wells only 30 % of which are equipped with downhole instruments. The BHFP at other wells is determined by recalculating the mouth parameters by various density-based procedures [14, 15, 16]. The accuracy of these techniques has been evaluated by several specialized investigations [17, 18, 19], which have revealed their low consistency and significant errors in the BHFP measurements, especially at deposits with reservoir oil highly saturated with gas.

The current investigations are aimed at choosing optimal BHFP control techniques [20, 21]. When the analytical description of physical processes is inapplicable for whatever reason, it is relevant to apply statistical techniques based on processing collected facts mathematically [22]. Statistical techniques are often easier to apply and more accurate than the attempt to elucidate analytical regularities in the behavior of complex physical systems. This approach does not require any significant reductions and suppositions, applies for any distribution laws, to systems of any complexity and multiplicity of states, and is restricted only by the actual parameters of the original sample. Statistical techniques are successfully applied for solving various engineering tasks [23, 24, 25].

In this vein, the article describes an original multilevel approach elaborated by the authors. This technique allows constructing multivariate statistical models of determining the BHFP of wells in operation. These models are generalized for vast territorial entities.

2. Materials and methods

2.1. *Characteristics of original data*

To solve the formulated task, we used the data from n = 4,145 investigations in which the measurements were taken synchronously at the mouth and bottom of the flowing wells of the oil deposits confined to the Solikamsk depression, a major oil-production area in the Perm krai. The deposits were Un'vinskoye, Gagarinskoye, Siberian, Magovskoye, Ozernoye, Chashkinskoye, Shershnevskoye, and the one named after Sukharev. To create multivariate multilevel models, we used the data of the downhole BHFP measurements in the wells equipped with measuring devices at the suction manifold; in addition, such well performance characteristics were used as liquid rate (Q_l, m^3/day), oil rate (Q_o, t/day), water cut (w, %), fluid pressure in the annulus between tubing and casing (P_{an}, MPa), dynamic fluid level depth (H_{dfl}, m), pump measured depth (H_{pump}, m), distance from the pump hinge point to the dynamic fluid level depth (H_{pd}); oil-water contact (OWC) measured depth, and reservoir pressure (P_r). The BHFP recalculated from the suction manifold pressure was used as actual (P_{bh}, MPa).

2.2. *Research features*

A distinct feature of our approach is the different degree of differentiating the facilities for which the models are constructed; that is, the approach is multilevel. The levels for which the models are constructed are exposed below:

- Level one involves using the entire sample, without separating deposits and production targets (occurrences).
- Level two consists in considering each deposit without separating occurrences.

- Level three consists in considering each occurrence in a generalized sense, without taking the deposits into account.

- Level four implies constructing the model specifically for each occurrence within a deposit.

We also constructed for practical application the multivariate model covering the BHFP determination results for all of the levels.

3. Literature review

The investigation was conducted using available mathematical statistics tools [26].

In the preliminary phase, we analyzed the correlations between the BHFP and the well operation properties that might affect the BHFP level. To achieve this purpose, we calculated not only coefficient r of the correlation between the input parameters and the BHFP but also coefficient r of the mutual correlation among the input parameters. The calculated results are exposed as correlation matrices, correlation fields, and equations of regression between the BHFP and the well performance indicators. We conducted the investigations in question for the models from all of the levels.

The next phase of the investigation is the construction of multivariate multilevel models using the original cumulative sample technique. According to this approach, the initial data are all tentatively graded against the range of BHFP levels from minimal to maximal. The first model is constructed according to the first three graded data (the amount of data per sample is $n = 3$). Then the model is constructed for $n = 4$. Thus multivariate models are consecutively constructed until all of the available data are used. In these variants, the multivariate models are constructed by step-by-step regression analysis. The dependent feature is P_{bh}, whereas the values of the rest of the above specified well performance indicators are used as independent factors. Step-by-step regression analysis was used not only to derive the equation of regression, defined as a multivariate statistical model, but also to identify the existence and kind of influence of the independent factors on the dependent variable.

The regression coefficients in the elaborated models were calculated by the least square technique. The functionality of each model was assessed in several ways. First of all, such statistical characteristics of the models were calculated in each case, as multiple correlation coefficient R, significance level p, and standard error S_0. The applicability limits of each model are defined. To assess the model of each level for functionality, it was used to calculate the BHFP also compared with the actual levels. The comparison was made while visually analyzing the correlation fields, analyzing the equations of regression between the actual and the calculated BHFP levels, and with the help of such other mathematical statistics means as Student's t-test and Pearson's chi-squared test [27, 28].

Thus the models were constructed for all of the four distinguished levels. Four domes are distinguished within the confines of the Un'vinskoe deposit. To take this peculiarity into account, we additionally constructed the models of each of the domes as part of modeling level four.

The analysis of constructed models should be considered a major part of any kind of modeling, multivariate modeling included. In this case, the analysis of all of the models involved examining the succession and frequency of including each of the input parameters in the equations of regression. It is considered that, the more often is a particular parameter found in multivariate models and the earlier it is used in model construction, the larger will be its effect on the BHFP. Thus this analysis allowed distinguishing the factors that had the biggest effect on the BHFP level registered when the commercial wells of the oil deposits in the considered region were in operation.

The multivariate equations of regression derived at all of the modeling levels are supposed to be used together, for which purpose the generalized multivariate model is constructed. It is proposed for use as the mathematical basis of the technique of measuring the BHFP in well operation.

The final phase was the investigation of the accuracy of determining the BHFP using the elaborated multivariate models. For that purpose, we calculated the model BHFP levels (PbhM) and then compared them with the actual BHFP levels (Pbh). This phase also involves comparing the elaborated technique, based on using multivariate models, with the currently applied BHFP calculation technique based on using the density of the gas-liquid mixture in the wellbore. For that purpose, we also calculated PbhM using the conventional technique and then compared the results with the actual BHFP levels. The elaborated technique based on multivariate multilevel models and its currently applied counterpart based on calculating the density of the gas-liquid mixture in the wellbore were compared by drawing the correlation fields when investigating the correlation (equations of regression and their characteristics) between the actual and the calculated BHFP for the whole sample and the separate deposit development targets.

It is supposed that the joint consideration of the regression equation parameters and their statistical characteristics will allow evaluating not only the tightness but also the kind of relations between the actual BHFP and the BHFP calculated by the two techniques.

4. Results

4.1. Analyzing correlations between BHFP and well performance indicators

The correlation matrix drawn for the whole examined sample (level one) and characterizing the influence of the well performance indicators on the BHFP level is shown in Table 1.

The correlation matrix of level two is represented for the Un'vinskoye deposit as the largest one in the region and exposed in Table 2.

Similar correlation matrices were drawn for all of the deposits in the considered region. The example of a correlation matrix for investigation level three is presented for the occurrences confined to Bobrikovian sedimentations (Table 3).

The correlation matrices at level four were made up for each occurrence within the deposits. The example exposed in Table 4 is the correlation matrix for the Bobrikovian occurrence of the Un'vinskoye deposit. The making up of the matrix involved calculating coefficients r using all the data (upper string) and, separately, for the domes (lines two, three, and four, and the bottom line are for the Un'vinskii, Palasherskii, Southeastern, and Bystrovskii dome, respectively).

Similar correlation matrices were made up for all of the occurrences of all the deposits. All in all, 29 correlation matrices were made up for different modeling levels; the matrices included 1305 values of r.

4.2. Multivariate BHFP determination models

Then, multivariate models were constructed for all of the levels. The model for level one is recorded as is shown in equation (1)

$$\begin{aligned} P_{bh}^{M1} =& 1,163 + 0,0042 \cdot H_{bound} \\ &+ 0,009 \cdot H_{owc} - 0,009 \cdot H_{pump} \\ &+ 0,022 \cdot w + 0,601 \cdot P_{an} \\ &+ 0,097 \cdot Q_l + 0,040 \cdot P_r . \end{aligned} \tag{1}$$

The statistical accuracy characteristics determined for the model were multiple correlation coefficient $R = 0,763$, significance level $p < 0,0000$, and standard calculation error $S_0 = 1,76$ MPa. The model forming sequence is provided in the equation of regression. The coefficients describing the reliability of the statistical relations changed as follows: $r = 0,505$; $R = 0,612$; 0,703; 0,736; 0,760; 0,762; 0,763.

The multivariate models at level two were constructed separately for the deposits, without distinguishing the occurrences. The model provided as an example was made up for Un'vinskoye deposit as the largest deposit in the region and is recorded in equation (2)

$$\begin{aligned} P_{bh}^{M2-U} =& -6,637 - 0,00718 \cdot H_{dfl} \\ &+ 0,0086 \cdot H_{owc} + 0,021 \cdot w \\ &+ 0,684 \cdot P_{an} + 0,3054 \cdot P_r \\ &- 0,0021 \cdot H_{bound}, \end{aligned} \tag{2}$$

at $R = 0,822$, $p < 0,0000$, $S_0 = 1,52$ MPa. The model was formed in the sequence presented in the

Table 1: Matrix of the correlation between bottom-hole flowing pressures and well performance indicators (level one)

	P_{bh}	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
P_{bh}	1	-0,45*	0,15*	0,19*	0,38*	0,28*	0,48*	0,2*	0,51*	0,32*
H_{dfl}		1	0,13*	0,05*	-0,26*	-0,29*	0,04*	0,07*	-0,87*	-0,02
P_{an}			1	-0,03	0,21*	0,19*	0,07*	0,06*	-0,09*	0,09*
W				1	0,17*	-0,25*	0,02	0,04*	-0,05*	0,33*
Q_l					1	0,83*	0,3*	0,27*	0,34*	0,36*
Q_o						1	0,27*	0,27*	0,38*	0,26*
H_owc							1	0,75*	0,3*	0,54*
H_{pump}								1	0,39*	0,53*
H_{pd}									1	0,2*
P_r										1

Note: *-significant correlations

Table 2: Matrix of the correlation between bottom-hole flowing pressures and well performance indicators (level two, Un'vinskoye deposit)

	P_{bh}	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
P_{bh}	1	-0,72*	0,09	-0,02	0,3*	0,26*	0,42*	-0,07	0,65*	0,17*
H_{dfl}		1	0,14*	0,21*	-0,24*	-0,31*	-0,24*	0,02	-0,92*	-0,13*
P_{an}			1	-0,1*	0,31*	0,33*	0,13	-0,03	-0,15	-0,07
W				1	0,03	-0,43*	-0,28*	-0,12*	-0,24*	0,22*
Q_l					1	0,85*	0,2*	0	0,22*	0,09*
Q_o						1	0,24*	0,05	0,31*	0,01
H_{owc}							1	0,31*	0,34*	-0,25*
H_{pump}								1	0,36*	-0,14*
H_{pd}									1	0,07
P_r										1

Note: *-significant correlations

Table 3: Matrix of the correlation between bottom-hole flowing pressures and well performance indicators (level three, Bobrikovian sedimentations)

	P_{bh}	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
P_{bh}	1	-0,51*	0,02	0,25*	0,21*	0,09*	0,13*	-0,26*	0,37*	0,26*
H_{dfl}		1	0,22	0,15	-0,18*	-0,29*	-0,01	0,02	-0,92*	-0,04
P_{an}			1	-0,05	0,28*	0,28*	0,18*	-0,03	-0,21*	0,08
W				1	0,19*	-0,37*	0,02	-0,11*	-0,18	0,22*
Q_l					1	0,77*	0,2*	0,01	0,17*	0,16*
Q_o						1	0,11*	0,07*	0,3*	0
H_{owc}							1	0,2*	0,09	0,59*
H_{pump}								1	0,38*	0,04
H_{pd}									1	0,05
P_r										1

Note: *-significant correlations

equation of regression. The coefficients describing the reliability of statistical relations varied as follows: $R = 0,717$; 0,762; 0,790; 0,808; 0,815; 0,822. Similar models were also constructed for the Siberian, Shershnevskoye, Gagarinskoye, Ozernoye, Magovskoye, and Chashkinskoye deposit. Level three implies constructing multivariate models for the main occurrences distinguished in particular deposits within the considered region. The model for the Bobrikovian occurrence developed at almost all of the deposits is recorded in equation (3)

$$P_{bh}^{M3-bb} = 19,684 - 0,00404 \cdot H_{dfl} + 0,025 \cdot w - 0,0037 \cdot H_{pump}, \quad (3)$$

at $R = 0,644$, $p < 0,0000$, $S_0 = 1,9$ MPa. The model was formed in the sequence exposed in the

Table 4: Correlation matrix of the influence of performance indicators (level four, Un'vinskoye deposit and Bobrikovian sedimentations)

	P_{bh}	Q_l	Q_o	B	P_{an}	H_{dfl}	H_{pump}	H_{owc}	P_r
	1	0,22*	0,16*	0,05	0,21*	-0,44*	-0,19*	0,23*	0,07
	1	0,11*	0,03	0,17*	0,08	-0,46*	-0,11*	-0,01	-0,05
P_{bh}	1	0,28*	0,28*	-0,06	0,17*	-0,65*	-0,33*	0,45*	0,27*
	1	-0,15	0,57*	-0,39	-0,23	-0,48*	-0,12	-0,33	0,23
	1	-0,71	-0,72*	0,79*	0,81*	0,52	-0,72*	-	-
		1	0,77*	0,05	0,33*	-0,25*	-0,04	-0,17*	-0,02
		1	0,76*	0,08	0,14*	-0,36*	0,09	-0,39*	-0,06
Q_l		1	0,76*	0,12*	0,37*	-0,28*	-0,22*	-0,11	-0,12
		1	-0,15	0,54*	-0,09	-0,26	0,88*	-0,58*	-0,9*
		1	0,97*	-0,78*	-0,59	-0,08	0,75*	-	-
			1	-0,5*	0,31*	-0,23*	-0,05	-0,1	-0,04
			1	-0,48*	0,22*	-0,33*	0,09	-0,29*	-0,04
Q_o			1	-0,51*	0,48*	-0,27*	-0,23*	-0,04	0,01
			1	-0,72*	-0,11	-0,30	-0,4	-0,37	0,07
			1	-0,88*	0,53	-0,01	0,86*	-	-
				1	0,14*	0	0	-0,07	-0,18*
				1	0,02	0,05	-0,04	0,01	-0,26*
B				1	-0,22*	0,08	0,08	-0,17*	-0,27*
				1	-0,06	0,07	0,69*	0,23	-0,32*
				1	0,48	0,05	-0,99*	-	-
					1	0,18*	-0,01	0,19*	0,04
					1	0,12*	0,01	0,05	-0,37*
P_{an}					1	0	-0,03	0,01	-0,14
					1	0,5*	-0,22	0,23	-0,06
					1	0,73*	-0,38	-	-
						1	0,09*	0,27*	0,13*
						1	0,03	0,4*	-0,09
H_{dlf}						1	0,25*	-0,16*	0,33*
						1	-0,3	0,5*	0,18
						1	0,07	-	-
							1	-0,29*	0,08
							1	-0,2*	0,05
H_{pump}							1	-0,5*	0,17*
							1	-0,54*	-0,64*
							1		
								1	0,14*
								1	-0,09
H_{owc}								1	-0,31*
								1	0,52*
								1	-
									1
									1
P_r									1
									1
									1

Note: *-significant correlations

equation of regression. The values of coefficient R changed as follows: 0.512; 0.608; 0.644.

The model for all of the occurrences was derived similarly.

Then we built multivariate level-four models specifically for the occurrences within the deposits.

For example, the presented level-four model is recorded for the Bobrikovian occurrence of the Un'vinskoye deposit is shown in equation (4)

$$P_{bh}^{M4-U-bb} = 17,902 - 0,0053 \cdot H_{dfl} + 0,0257 \cdot w + 0,5906 \cdot P_{an} - 0,0026 \cdot H_{pump}, \quad (4)$$

at $R = 0,763$, $p < 0,0000$, $S_0 = 1,62$ MPa. The model was formed in the sequence exposed in the equation of regression. The values of coefficient

R changed as follows: 0,683; 0,723; 0,748; 0,763. For the ranges, in which all of the above recorded models can be used, see Table 5.

The multivariate models for each of the domes of the Un'vinskoye deposit are recorded as follows:

a) Un'vinskiy dome:

$$P_{bh}^{M4-1} = 4,430 - 0,0024 \cdot H_{dfl} + 0,003 \cdot H_{owc} + 0,011 \cdot w + 0,449 \cdot P_{an} - 0,0055 \cdot Q_l, \quad (5)$$

at $R = 0{,}552$, $p < 0{,}0000$, $S_0 = 1{,}44$ MPa. The model was formed in the sequence exposed in the equation of regression (5). The values of coefficient R changed as follows: 0,460; 0,497; 0,532; 0,549; 0,552.

b) Palasherskiy dome:

$$P_{bh}^{M4-2} = -57,862 - 0,0045 \cdot H_{dfl} + 0,0075 \cdot H_{owc} + 0,7014 \cdot P_{an} + 0,0172 \cdot Q_l - 0,0177 \cdot Q_o, \quad (6)$$

at $R = 0{,}770$, $p < 0{,}0000$, $S_0 = 1{,}37$ MPa. The model was formed in the sequence exposed in the equation of regression (6). The values of coefficient R changed as follows: 0,648; 0,734; 0,573; 0,759; 0,764; 0,770.

c) Southeastern dome:

$$P_{bh}^{M4-3} = -201,705 + 0,375 \cdot Q_o - 0,001 \cdot H_{dfl} + 0,104 \cdot H_{owc} + 0,004 \cdot H_{pump}, \quad (7)$$

at $R = 0{,}729$, $p < 0{,}0121$, $S_0 = 0{,}91$ MPa. The model was formed in the sequence exposed in the equation of regression (7). The values of coefficient R changed as follows: 0,566; 0,652; 0,706; 0,729.

d) Bystrovskiy dome:

$$P_{bh}^{M4-4} = -6,775 + 5,321 \cdot P_{an} + 0,22617 \cdot w + 0,0041 \cdot H_{dfl} \quad (8)$$

at $R = 0{,}943$, $p < 0{,}02155$, $S_0 = 0{,}94$ MPa. The model was formed in the sequence exposed in the equation of regression (8). The values of coefficient R changed as follows: 0,805; 0,927; 0,943.

The value ranges of the indices, in which the models derived for each elevation can be used, are exposed in Table 6.

The correlation of the calculated and the actual BHFP levels was examined by making up equations of regression and exposed in Table 7.

The frequency of inclusion (incidence) of the indices in the models of all the levels was calculated during their analysis and is shown in Table 8.

The generalized model is recorded as shown in equation (9)

$$P_{bh}^{MM} = -0,089 + 0,689 \cdot P_{bh}^{M-3} + 0,361 \cdot P_{bh}^{M-2} - 0,039 \cdot P_{bh}^{M-1}, \quad (9)$$

at $R = 0{,}941$, $p < 0{,}0000$ and the mean standard error is $0{,}45$ MPa. The model was formed in the sequence exposed in the equation of regression. The values of coefficient R describing the strength of the statistical relations changed as follows: 0,848; 0,851; 0,941.

The data of the results of comparing which are exposed below were obtained by two techniques, the new one based on applying multivariate models and the conventional one based on calculating the gas-liquid mixture density.

The correlation fields for the sample in general are presented in Figure 1.

The equations of regression between the actual BHFP and the BHFP calculated by the two techniques are shown in Table 9 as well as the statistical characteristics of these equations, including the correlation coefficient, the significance level, and the standard calculation error. The equations are compared at two levels, i.e., using all the data (level one) and, separately, for the occurrences of all of the deposits (level two).

In addition, to comparatively evaluate the functionality of the techniques as applied to specific occurrences, the respective correlation fields were drawn by the example of Un'vinskoye as the region's most representative deposit (Figure 2).

Table 5: Applicability ranges of the models of four levels

Model's index	Applicability range of model of levels			
	1	2	3	4
H_{dfl}, m	-	93,9-1946,5	206,9-1946,5	304,9-1946,5
P_{an}, 10^6 Pa	0-8,5	0,06-6,05	-	0,06-6,05
w, %	0,0-99,9	0,0-98	0-99,9	0,0-97,6
Q_l, m^3/day	-	-	-	-
Q_o, t/day	0-119,3	-	-	-
H_{owc}, m	1426,0-2332,7	1837,7-2246,7	-	-
H_{pump}, m	1201,4-2287,6		1439,3-2177,2	1445,3-2083,1
H_{pd}, m	3,8 -1720,5	17,1-1554	-	-
P_r, 10^6 Pa	6,4-23,3	13,9-20,4	-	-

Table 6: Applicability ranges of the models of the domes of the Un'vinskoye deposit (Bobrikovian occurrence)

Model's index	Applicability range of the models derived for domes			
	Un'vinskiy	Palasherskiy	Southeastern	Bystrovskiy
H_{dfl}, m	210 – 1935	513 – 2076	916 – 1416	1 194 – 1744
P_{an}, MPa	0,14 – 4,79	0,37 – 3,71		0,94 – 1,3
w, %	1 – 98,4			5-20
Q_l, m^3/day	0,8 – 98,7	0,5 – 107,2		
Q_o, t/day		0,2 – 71,2	0,11 – 6,36	
H_{owc}, m	2088 – 2519,5	2152 – 2553,2		
H_{pump}, m			1715 – 1970	
H_{pd}, m				
P_r, MPa				

Table 7: Equations of regression between the actual BHFP levels and the BHFP values calculated according to the level-four model, considering the discrimination of domes within the Un'vinskoye deposit

Dome	Equation of regression	r	p
Un'vinskiy	$P_{bh}^{M4-1} = 5,416 + 0,303 \cdot P_{bh}$	0,549	0,0000
Palasherskiy	$P_{bh}^{M4-2} = 3,484 + 0,598 \cdot P_{bh}$	0,766	0,0000
Southeastern	$P_{bh}^{M4-3} = 3,769 + 0,532 \cdot P_{bh}$	0,729	0,0002
Bystrovskiy	$P_{bh}^{M4-4} = 0,818 + 0,890 \cdot P_{bh}$	0,943	0,0004

Table 8: Incidence rate of the indices in the models of all the levels

Modeling level	Models input parameters								
	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
One		0,142	0,142		0,142	0,142	0,142	0,142	0,142
Two	0,136	0,136	0,136	0,090	0,113	0,136	0,136	0,045	0,068
Three	0,121	0,121	0,090	0,121	0,090	0,121	0,090	0,121	0,121
Four	0,147	0,132	0,102	0,132	0,102	0,088	0,132	0,102	0,068
Total	0,131	0,131	0,111	0,111	0,105	0,111	0,125	0,095	0,079

5. Discussion

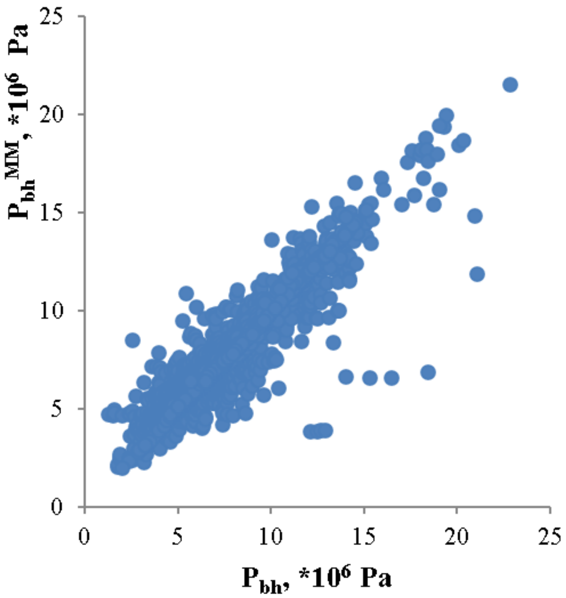
The exposed investigations should be considered the justification of the expediency of applying probabilistic statistical techniques to determining the bottom-hole flowing pressure treated as a major task in the oil extraction industry.

As shown by all of the conducted investigations, the creation of a stable BHFP determination technique is a complex task to solve.

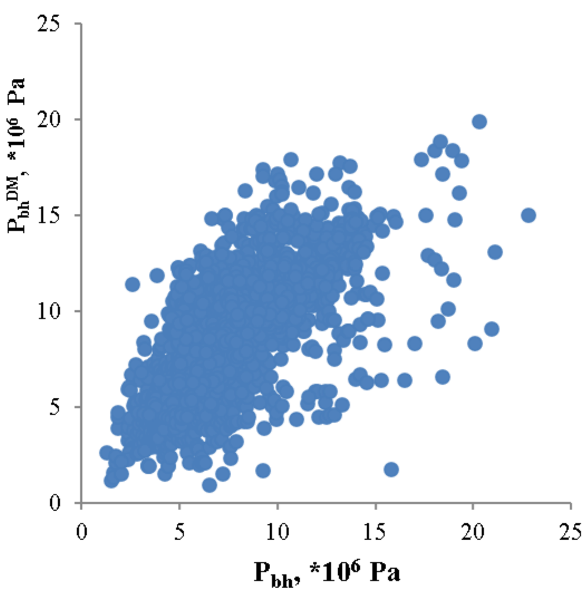
For example, as evidenced by the correlation analysis, in which 29 correlation matrices with 1305 correlation coefficient values were built for the four levels of investigation, the well

Table 9: Equations of regression between the actual BHFP and the BHFP calculated by the two techniques

First Level of comparison			
	Type and amount of data, sedimentation index	Equations of regression for technique	
	Whole sample	$P_{bh}^{MM} = 0,840 + 0,893 \cdot P_{bh}$ $r = 0,945; p < 0,0000$ $S_0 = 0,893$ MPa	$P_{bh}^{DM} = 2,309 + 0,791 \cdot P_{bh}$ $r = 0,671; p < 0,0000$ $S_0 = 2,068$ MPa
Second Level of comparison			
Un'vinskoye deposit	$C_1t - D3fm$ $n = 174$	$P_{bh}^{MM} = 0,088 + 0.993 \cdot P_{bh}$ $r = 0,998; p < 0.0000$ $S_0 = 0,158$ MPa	$P_{bh}^{DM} = 1,633 + 0,856 \cdot P_{bh}$ $r = 0,847; p < 0.0000$ $S_0 = 1,491$ MPa
	C_1v $n = 889$	$P_{bh}^{MM} = 0,032 + 0,996 \cdot P_{bh}$ $r = 0,997; p < 0,0000$ $S_0 = 0,189$ MPa	$P_{bh}^{DM} = 2,561 + 0,823 \cdot P_{bh}$ $r = 0,770; p < 0,0000$ $S_0 = 1,670$ MPa
	C_2b $n = 143$	$P_{bh}^{MM} = 0,144 + 0,988 \cdot P_{bh}$ $r = 0,997; p < 0.0000$ $S_0 = 0,216$ MPa	$P_{bh}^{DM} = 0,739 + 1,042 \cdot P_{bh}$ $r = 0.862; p < 0.0000$ $S_0 = 1,660$ MPa
	C_2vr $n = 31$	$P_{bh}^{MM} = -0,013 + 0,988 \cdot P_{bh}$ $r = 0.998; p < 0.0000$ $S_0 = 0,045$ MPa	$P_{bh}^{DM} = 14,369 - 0,424 \cdot P_{bh}$ $r = 0.374; p < 0.0378$ $S_0 = 0,968$ MPa
Chashkinskoye deposit	$C_1t - D_3fm$ $n = 89$	$P_{bh}^{MM} = 0,513 + 0,946 \cdot P_{bh}$ $r = 0,972; p < 0.0000$ $S_0 = 0,713$ MPa	$P_{bh}^{DM} = 2,704 + 0,710 \cdot P_{bh}$ $r = 0,876; p < 0,0000$ $S_0 = 1,225$ MPa
	C_1v $n = 161$	$P_{bh}^{MM} = 2,476 + 0,750 \cdot P_{bh}$ $r = 0,865; p < 0,0000$ $S_0 = 0,965$ MPa	$P_{bh}^{DM} = 5,955 + 0,467 \cdot P_{bh}$ $r = 0,497; p < 0,0000$ $S_0 = 2,272$ MPa
Ozernoye deposit	$C_1t - D_3fm$ $n = 579$	$P_{bh}^{MM} = 1.746 + 0.727P_{bh}$ $r = 0,852; p < 0,0000$ $S_0 = 0,655MPa$	$P_{bh}^{DM} = 0.612 + 1.124P_{bh}$ $r = 0,617; p < 0.0000$ $S0 = 2,107$ MPa
	C_2b $n = 99$	$P_{bh}^{MM} = 0.241 + 0.958P_{bh}$ $r = 0,979; p < 0.0000$ $S_0 = 0,268$ MPa	$P_{bh}^{DM} = -0.226 + 0.932P_{bh}$ $r = 0,964; p < 0,0000$ $S_0 = 0,347$ MPa
Magovskoye deposit	$C_1t - D_3fm$ $n = 123$	$P_{bh}^{MM} = 1.063 + 0.890P_{bh}$ $r = 0.943; p < 0.0000$ $S_0 = 1,304$ MPa	$P_{bh}^{DM} = 5.136 + 0.495P_{bh}$ $r = 0.689; p < 0.0000$ $S_0 = 2,172$ MPa
	C_2b $n = 33$	$P_{bh}^{MM} = 1.893 + 0.640P_{bh}$ $r = 0.799; p < 0,0000$ $S_0 = 0,681$ MPa	$P_{bh}^{DM} = 2.488 + 0.901P_{bh}$ $r = 0,640; p < 0,0000$ $S_0 = 1,532$ MPa
Gagarinskoye deposit	$C_1t - D_3fm$ $n = 288$	$P_{bh}^{MM} = 3.182 + 0.570P_{bh}$ $r = 0,754; p < 0,0000$ $S_0 = 1,568$ MPa	$P_{bh}^{DM} = 3.635 + 0.598P_{bh}$ $r = 0,724; p < 0,0000$ $S_0 = 1,791$ MPa
	C_2b $n = 193$	$P_{bh}^{MM} = 0.928 + 0.837P_{bh}$ $r = 0,901; p < 0.0000$ $S_0 = 1,053$ MPa	$P_{bh}^{DM} = 2.691 + 0.534P_{bh}$ $r = 0,653; p < 0.0000$ $S_0 = 1,680$ MPa
Shershnevskoye deposit	$C_1t - D_3fm$ $n = 152$	$P_{bh}^{MM} = 2.847 + 0.532P_{bh}$ $r = 0.716; p < 0.0000$ $S_0 = 0,694$ MPa	$P_{bh}^{DM} = 1.618 + 0.445P_{bh}$ $r = 0,389; p < 0.0000$ $S0 = 1,429$ MPa
	C_1v $n = 112$	$P_{bh}^{MM} = 1.852 + 0.798P_{bh}$ $r = 0,893; p < 0.0000$ $S_0 = 1,344$ MPa	$P_{bh}^{DM} = 3.079 + 0.599P_{bh}$ $r = 0,594; p < 0.0000$ $S_0 = 2,716$ MPa
Sukharev deposit	$C_1t - D_3fm$ $n = 50$	$P_{bh}^{MM} = 0.535 + 0.948P_{bh}$ $r = 0,976; p < 0,0000$ $S0 = 0,886$ MPa	$P_{bh}^{DM} = 0.004 + 0.925P_{bh}$ $r = 0,947; p < 0.0000$ $S_0 = 1,240$ MPa
	C_1v $n = 61$	$P_{bh}^{MM} = 0.917 + 0.917P_{bh}$ $r = 0,957; p < 0.0000$ $S_0 = 0,651$ MPa	$P_{bh}^{DM} = 1.495 + 0.800P_{bh}$ $r = 0,737; p < 0.0000$ $S_0 = 1,728$ MPa
	C_2b $n = 15$	$P_{bh}^{MM} = 2.013 + 0.802P_{bh}$ $r = 0,895; p < 0.0000$ $S_0 = 1,048$ MPa	$P_{bh}^{DM} = 6.005 + 0.305P_{bh}$ $r = 0,358; p < 0.0000$ $S_0 = 2,088$ MPa



(a) BHFP calculated by the technique based on multivariate models



(b) BHFP calculated by model based on calculating the gas-liquid mixture density

Figure 1: Field of correlation (whole sample) of the actual BHFP

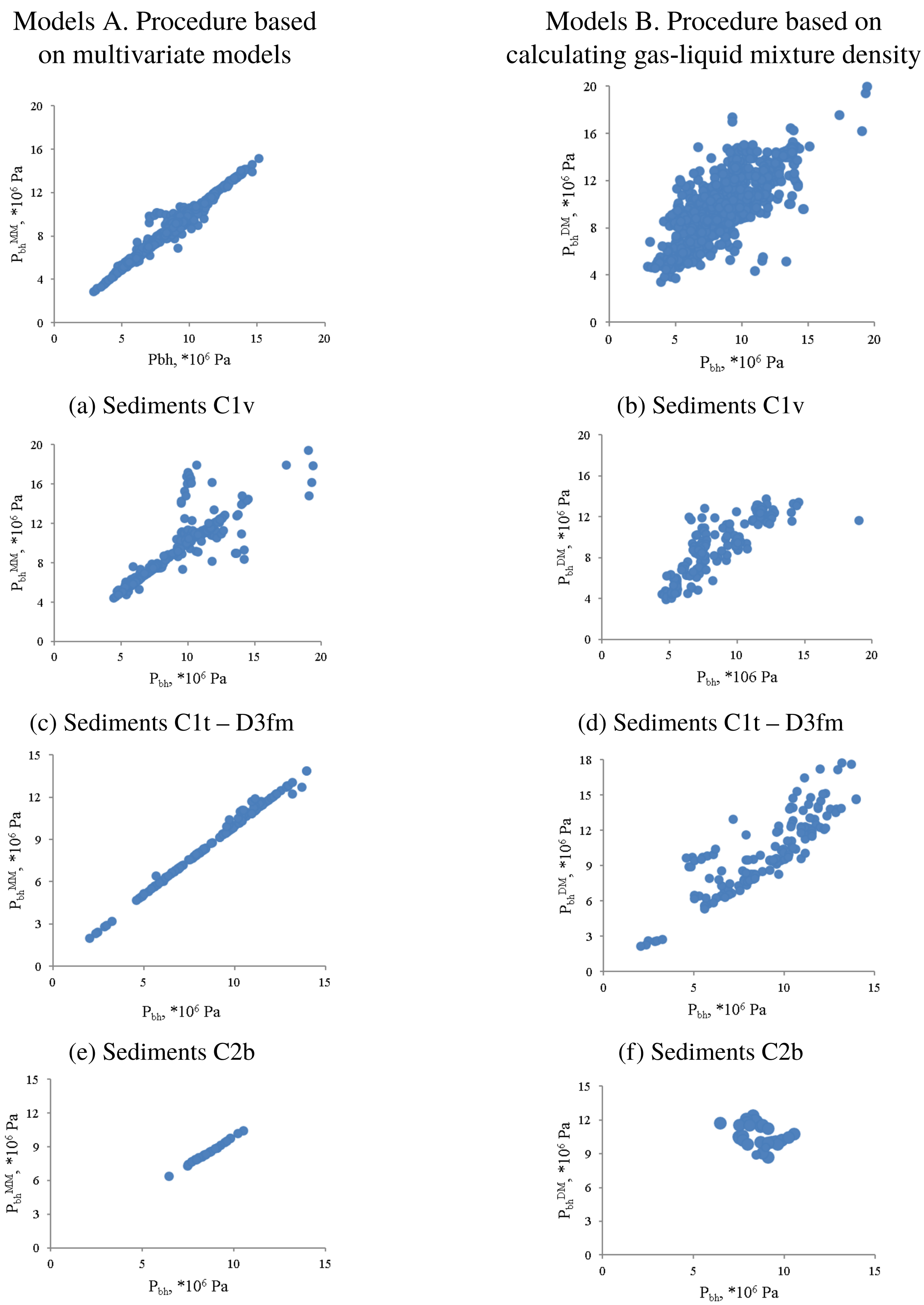
performance indicators have a complex effect on the BHFP level. The mutual correlations among the indicators vary in a broad range from −0,87 to 0,83. As found out by the correlation analysis one and the same performance indicator may have a different effect on the BHFP level in various conditions, i.e., the input parameters have a differentiated effect on the calculated quantity. For example, coefficient r between Pbh and Pan changed from -0.08 to 0.51 at various levels of investigation; sometimes, the relation among these parameters is negative, sometimes – positive and statistically significant. On the whole, the identical direction of affecting Pbh is followed by such parameters, as H_{dfl}, H_{owc}, H_{pd}, and P_r, whereas the others produce various effects in terms of both, direction and force.

All of this shows that the BHFP is affected by the performance indicators both, together and individually. Thus the correlation analysis has allowed finding out that none of the performance indicators makes it possible to reliably predict BHFP levels. The BHFP formation during well operation follows very complicated and non-stationary laws, which is why the analytical solution making it possible to determine the BHFP in a firm and reliable manner should be considered an extremely complex task to solve.

As mentioned earlier, when analytical techniques do not apply due to the low accuracy of their results, it seems expedient to use statistical (probabilistic statistical) techniques). For this reason, it is not analytical solutions but multivariate statistical models that are proposed in this paper as the techniques for determining the BHFP level during well operation. Therefore, the rest of the investigation is dedicated to constructing multivariate statistical models.

It should be noted that these models were built using the original approach, i.e., by a pre-ordered distribution. Not only does this approach allow deriving the model, which allows reliably determining the BHFP as the target parameter, but it also allows distinguishing all the regularities of its formation in various conditions.

The conclusions derived from analyzing the resulting models are exposed below.



(a) Sediments C1v
(b) Sediments C1v
(c) Sediments C1t – D3fm
(d) Sediments C1t – D3fm
(e) Sediments C2b
(f) Sediments C2b
(g) Sediments C2vr
(h) Sediments C2vr

Figure 2: Actual-to-calculated BHFP correlation fields by the example of the Un'vinskoye deposit occurrences

Specifically, the original approach consisting in using a pre-ordered distribution allowed making a detailed analysis of the frequency with which the input parameters were included in the resulting equations of regression. As found out by the analysis, the sole indicator with the prevalent effect on the BHFP level could not be distinguished for any of the modeling levels. This conclusion reveals the complexity of the law, according to which the BHFP is formed during well operation, the existence of an integrated effect of the input parameters on the BHFP, and the expediency of using these parameters by applying the built multivariate models of all the levels.

Another characteristic feature of this paper is that the models were constructed for several levels, with different degrees of differentiating modeled objects. The chosen approach is evaluated for correctness by analyzing the values of multiple correlation coefficient R, one of the major indicators characterizing the functionality of a derived model. The model for level one has $R = 0{,}763$. The average R for the models of levels two, three, and four is 0,790, 0,801, and 0,822, respectively. That is, the model of each subsequent level has a higher degree of functionality.

To consider the established complex effect of the input parameters on the predicted BHFP level, it is proposed to use all of the modeling levels in an integrated manner, by constructing a general multivariate model. This model has very high performance capabilities, for example, multiple correlation coefficient R of 0,941. Thus using all of the distinguished levels of multivariate mathematical modeling has allowed deriving a functional multivariate model the application of which as the mathematical basis of the respective technique will allow determining the BHFP during well operation with a high degree of certainty.

The certainty of determining the BHFP by the technique based on applying the elaborated multilevel multivariate mathematical models was evaluated by a special detailed analysis.

It included comparing the actual (measured) BHFP with the BHFP calculated by the elaborated technique, and the BHFP calculated by the technique currently applied in the region and based on determining the density of the gas-liquid mixture in the wellbore. To derive the most valid conclusions, we made the comparison at different differentiation levels and by means of various tools, including visual analysis of the correlation fields, and derivation and analysis of the equations of regression among the actual BHFP and the BHFP calculated by the two techniques.

The comparative visual analysis of the fields of correlation between the actual and the calculated BHFP allows concluding that the elaborated technique based on applying multilevel multivariate models is characterized by essentially higher accuracy of calculation: the correlation field for the elaborated technique has a much tighter and even shape, its points group around the line with a slope close to one. The derived regularity is typical of both, the correlation fields drawn for the sample in general and for the fields drawn individually for the occurrences from separate deposits.

The visual analysis of the correlation fields has allowed evaluating, by comparison, the certainty of determining the BHFP by the two techniques at the level of quality. The quantitative comparison of these fields has made it necessary to derive the equations of regression between the actual and the calculated BHFP for the two levels of investigation as well, i.e., for the sample in general (level one) and, individually, for occurrences in separate deposits (level two). The sign of the higher certainty of a particular technique is that the value of the free term in the equation of regression is close to zero and the angular ratio close to one. In addition to the regression equation parameters, their statistical characteristics have been calculated as well, including the correlation coefficient, significance level, and standard calculation error. As shown by analyzing the statistical characteristics of the exploitation targets from all of the deposits, not only the values of correlation coefficient r differ in all of the four cases but the equation of regression themselves. In all of the cases, the coefficients at PBHFP found by the multilevel technique are higher than those found by the density-based technique. The standard errors calculated according to the multivariate models for all of the exploitation targets are much lower than the standard errors calculated by the density-based technique. As shown by comparing the average reduced characteristics values against the t criterion, statistical differences are found in each case. The enumerated facts convincingly prove that the BHFP determination by the elaborated technique based on applying multivariate multilevel (statistical) models has a higher degree of certainty.

6. Conclusion

Thus this article validates the technique of determining the BHFP during oil-production well operation by means of the constructed multivariate multilevel models.

The models were constructed proceeding from the significant accumulated experience in parallel depth and estuarine measurements conducted when servicing the commercial wells in the Perm krai.

The constructed models have high statistical capability characteristics. A distinct feature of the models is that the parameters they use as the sole original data are easy to determine in practice. This fact should be considered the main strength of the developed technique as compared with its multiple density-based counterparts.

The high capability of the constructed multivariate BHFP models is mainly stipulated by the original approach to making them that consists in creating a model according to a sequence of pre-ordered original data.

Not only has the construction of multivariate mathematical models allowed determining the BHFP in practice but it has also allowed identifying the regularities of its formation and in-operation behavior individually for each considered exploitation target.

The BHFP determination technique based on the developed multivariate models is much more functional than its conventional density-based counterparts.

It is worth noting individually that the new technique should not be considered an alternative to density-based methods. The joint application of these techniques is supposed to ensure a reliable BHFP control during production well operation.

Acknowledgments

We thank OOO LUKOIL-PERM and Sergey Cherepanov and Irina Chernykh in person for assistance in the pursuance of the research

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

7. References

[1] L. Nazarova and E. Nechaeva, “The analysis of influence of bottomhole pressure decrease under saturation pressure on the oil recovery,” *Oil Industry*, no. 1, pp. 83–85, 2014.

[2] Y. Kashnikov and S. Yakimov, “Geomechanical and hydrodynamic estimation of the bottom-hole pressure influence on the well performance,” *Oil and Gas Business*, vol. 11, pp. 111–115, 2019.

[3] T. Bikmukhametov and J. Jäschke, “Oil Production Monitoring using Gradient Boosting Machine Learning Algorithm,” *IFAC-Papers online*, vol. 52, no. 1, pp. 514–519, 2019.

[4] S. Natarajan, K. Ghosh, and R. Srinivasan, “Collaborative Multi - Agent based Process Monitoring System for Offshore Oil and Gas Production. Comput,” *Computed Aided Chemical Engineering*, vol. 27, pp. 1227–1232, 2009.

[5] D. Martyushev, “Determination of the rational bottomhole pressure of producing wells in the development of carbonate reservoirs,” *Drilling and oil*, no. 11, pp. 22–24, 2014.

[6] D. Martyushev and V. Mordvinov, “Change in the flow rate of wells in an oil and gas condensate field with a decrease in reservoir and bottomhole pressures,” *Oil Industry*, no. 167–69, 2014.

[7] A. Codas, B. Ordoñez, and U. Moreno, “Sucker-Rod Pumping System Fault Detection and Isolation Method Using Bottom Hole Pressure Measurement,” *IFAC Proccedings Volumes*, vol. 42, no. 8, pp. 1031–1036, 2009.

[8] V. Mordvinov, A. Lekomtsev, and D. Martyushev, “Determination of pressure at the intake of electric centrifugal pumps when pumping out low-foam carbonated oil,” *Oil industry*, no. 61–63, 2014.

[9] M. Ali Ahmadi, M. Galedarzadeh, and S. Reza-Shadizadeh, “Low parameter model to monitor bottom hole pressure in vertical multiphase flow in oil production wells,” *Petroleum*, vol. 2, no. 3, pp. 258–266, 2016.

[10] E. Levitina, “Influence of density gas liquid phase on parameters of well pressure,” *Oil and Gas Stud*, no. 1, pp. 35–40, 2010.

[11] H. Yang, J. Li, G. Liu, X. Xing, H. Jiang, and C. Wang, “The effect of interfacial mass transfer of slip-rising gas bubbles on two-phase flow in the vertical

wellbore/pipeline," *International Journal of heat and mass transfer*, vol. 150, no. 119326, 2020.

[12] A. Hasan and C. Kabir, "A Study of Multiphase Flow Behavior in Vertical Wells," *SPE Production Engineering*, vol. 3, no. 2, pp. 263–272, 1988.

[13] J. Brill and H. Mukherjee, *Multiphase Flow in Wells*. Richardson: SPE, 1999.

[14] A. Lekomtsev, E. Zhelanov, and I. Chernykh, "Statistical approach to the evaluation of bottomhole pressure in producing wells," *Oil and Gas Business*, vol. 10, pp. 98–101, 2016.

[15] A. Lekomtsev and D. Martyushev, "Comparative analysis of methods for determining BHP during well test," *Oil Industry*, vol. 6, 2014.

[16] S. Bikbulatov and A. Pashali, "Analisys and selection of methods for calculating the pressure gradient in the wellbore," *Oil and Gas Business*, vol. 21, no. 2, pp. 1–12, 2005.

[17] I. Chernykh, V. Galkin, and I. Ponomareva, "Comparative analysis of the methods for defining bottom hole pressure at well operation of Shershnevsky field," *A Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, vol. 328, no. 8, pp. 41–47, 2017.

[18] V. Galkin, I. Ponomareva, I. Chernykh, and E. Filippov, "Methodology for estimating downhole pressure using multivariate model," *Oil Industry*, vol. 1, pp. 40–43, 2019.

[19] D. Martyushev and Y. Slushkina A, "Assessment of informative value in determination of reservoir filtration parameters based on interpretation of pressure stabilization curves," *Bulletin of the Tomsk Politechnic University*, vol. 330, no. 9, pp. 26–32, 2019.

[20] M. Nait Amar, N. Zeraibi, and K. Redouane, "Bottom hole pressure estimation using hybridization neural networks and grey wolves optimization," *Petroleum*, vol. 4, no. 4, pp. 419–429, 2018.

[21] W. Chen, Q. Di, F. Ye, J. Zhang, and W. Wang, "Flowing bottomhole pressure prediction for gas wells based on support vector machine and random samples selection," *International Journal of Hydrogen Energy*, vol. 42, no. 29, pp. 18 333–18 342, 2017.

[22] A. Sánchez-Fernández, F. Baldán, G. Sainz-Palmero, J. Benítez, and M. Fuente, "Fault detection based on time series modeling and multivariate statistical process control," *Chemometrics and Intelligent Laboratory Systems*, vol. 118, pp. 57–69, 2018.

[23] C. Kumar Singha, A. Kumar, S. Shashtrib, A. Kumar, P. Kumar, and J. Mallick, "Multivariate statistical analysis and geochemical modeling for geochemical assessment of groundwater of Delhi, India," *Journal of Geochemical Exploration*, vol. 175, pp. 59–71, 2017.

[24] D. Cecconet, S. Bolognesi, S. Daneshgar, A. Callegari, and A. Capodaglio, "Improved process understanding and optimization by multivariate statistical analysis of Microbial Fuel Cells operation," *International Journal of Hydrogen Energy*, vol. 43, no. 34, pp. 16 719–16 727, 2018.

[25] J. Hua, J. Li, M. Ouyang, L. Lu, and L. Xu, "Proton exchange membrane fuel cell system diagnosis based on the multivariate statistical method," *International Journal of Hydrogen Energy*, vol. 36, no. 16, pp. 9896–9905, 2011.

[26] S. Barkovskiy, V. Zakharov, A. Lukashov, and A. Nurutdinova, *Multidimensional Data Analysis by Applied Statistics Techniques*. Kazan: KSTU, 2010.

[27] R. Kissell and J. Poserina, *Optimal Sports Math, Statistics, and Fantasy*. Academic Press, 2017.

[28] E. Wentzel, *Operations Research*. Moscow: Mir Publishers, 1983.