# Factorization of flow profile data in production and injection wells based on Bayesian modeling

Konstantin Sidelnikov[1] and Rinat Faizullin[2]

[1] Department of Reservoir Energy State Control, Izhevsk Petroleum Scientific Center, Izhevsk, Russia

[2] Department of Management, Kalashnikov Izhevsk State Technical University, Izhevsk, Russia

sidelkin@yandex.ru

**Abstract.** This problem is ill-posed, and the process of adding information in order to solve it is needed (regularization). Several methods of deterministic regularization based on $\ell_2$ -norm minimization of the unknown vector are considered. It is shown that such approaches do not take into account the petrophysical properties of the reservoir and cannot cover all possible factorization combinations. Bayesian regularization is proposed to factorize the flow profile data. According to this method, all relative factors are defined by the corresponding probability distribution functions. Core studies are used to determine the joint probability distribution of rock permeability and porosity. Layer productivity ratio distributions are calculated separately for each well based on its log interpretation data. Bayesian statistical inference is used to obtain the general drawdown ratio distribution for the entire field. This approach was tested on real data obtained from three fields.

**Keywords:** Regularization, Bayes' theorem, statistical inference, well test, flow profile, productivity index, injectivity index

## 1 Introduction

In the process of analyzing the oil field development, it is required to solve the problems of determining the injected water front, assessing the oil recovery factor separately for the producing formations, etc. In order to distribute the cumulative oil production and the volume of injected water between the producing formations, it is necessary to have the values of the oil and water flow rate for each formation separately [1-7].

Investigation of the flow profiles in formations allows obtaining the distribution of the produced and injected fluid over the entire cross-section of the pay zone. As a result, the dependence of the amount of produced or injected fluids on the depth of the completed interval is established. To obtain flow profiles over the formation thickness, the results of measurements by flow meters are usually used, as well as thermometry data in the wellbore [8-15].

Bayesian methods are widely used in the practice of developing hydrocarbon fields: processing the results of well tests [5-6; 11] and well logging [9; 12-13], in reservoir simulation [7-8; 10], in statistical prediction of oil-field performance [3-4].

## 2 Materials and methods

**Formulation of the problem.**

The paper discusses the results of determining the flow profile for a two-layered reservoir (Figure 1). It is also assumed that the inflow data from the investigated well interval contain the value of the total oil and water production rate without exact information about fractional flow of each fluid.

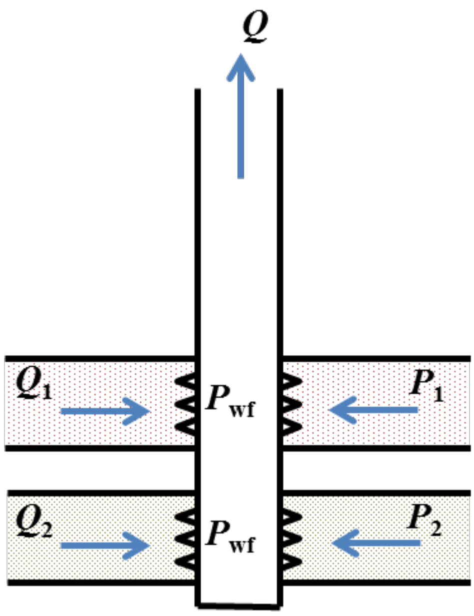



**Fig. 1.** Well inflow model.

Thus, as a result of the production log, the following flow rate ratio for 2 production layers becomes known:

$$q = \frac{Q_1}{Q_2}, \tag{1}$$

Where:

$$Q_i = \pm J_i \left( P_i - P_{\text{wf}} \right) = J_i \cdot \Delta P_i,\ i = \overline{1,\ 2} \tag{2}$$

$J_i$ – productivity of $i$ -th layer; $\Delta P_i$ – drawdown for the $i$-th layer in the case of production ("+" sign) or injection ("–" sign), respectively.

The problem is to factorize (1) resulted in decomposition of the known $q$ flow rate ratio into the product of the following unknown ratios:

$$q = j \cdot p \tag{3}$$

Where $j = J_1 / J_2$ – productivity / injectivity ratio; $p = \Delta P_1 / \Delta P_2$ – drawdown ratio.

Possible interpretations of nonuniform recovery of reserves based on factorization (3) are:

- For relatively homogeneous reservoirs in terms of productivity, it can be approximately assumed that $j \approx 1$; therefore, the observed flow profile will be determined mainly by the ratio $p$, which will depend on the initial reserve and dynamics of the reservoir energy consumption in relation to each reservoir;
- In the case of hydrodynamic equilibrium between the reservoirs and sufficiently fast process of repressuring between them, it can be assumed that $p \approx 1$; therefore the main factor of the flow profile non-uniformity will be the heterogeneity of the reservoir properties, expressed as a ratio $j$.

There may also be other possible values of the productivity and drawdown ratios, reflecting more complex hydrodynamic processes of multilayer reservoir development.

Obviously, under such conditions, factorization problem (3) has no solution without additional information. In particular, if it is possible to directly calculate the productivity $J_i$ of each interval, then with the known $q$, it is possible to obtain the drawdown for each reservoir. However, direct calculation of $J_i$ requires knowing the transmissibility of each layer, the current inflow regime of the well, the reservoir geometry, etc. [15]. The absence or incompleteness of such information generally precludes this approach.

It is known that, to control the development of oil and gas fields, in addition to the production log, well testing is carried out as well. The values of the productivity index and reservoir pressure obtained from the results of these studies, as a rule, are integral in nature and determine the properties of the multilayered system only as a whole. Thus, on the basis of well testing, only the total fluid flow rate (injection) can be factored:

$$Q = J \cdot \Delta P \tag{4}$$

Where $Q = Q_1 + Q_2$, $J = J_1 + J_2$ and weighted average drawdown:

$$\Delta P = \frac{J_1}{J} \Delta P_1 + \frac{J_2}{J} \Delta P_2 \tag{5}$$

Obviously, however, for an unambiguous factorization (3) one well-known expression in the form (4) is not enough. Moreover, this problem can be attributed to an ill-posed problem, the solution to which requires some additional constraints to its conditions (regularization) [16].

**Deterministic regularization.**

When solving factorization problem (3) on the basis of (4), in fact, there is a vector of four unknowns:

$$\mathbf{u} = \begin{bmatrix} J_1 & J_2 & \Delta P_1 & \Delta P_2 \end{bmatrix}^{\mathrm{T}} \tag{6}$$

And a system of three equations:

$$\begin{cases} Q_1 = J_1 \cdot \Delta P_1 \\ Q_2 = J_2 \cdot \Delta P_2 \\ J = J_1 + J_2 \end{cases} \tag{7}$$

The system of equations (7) has many solutions; therefore, one of the methods for obtaining a unique solution is to impose an additional constraint on the norm of the vector space [16]. For example, consider the minimization of the $\ell_2$-norm of the vector (6):

$$\|\mathbf{u}\|_2^2 = \mathbf{u}^{\mathrm{T}}\mathbf{u} \to \min \tag{8}$$

Subject to the fulfillment of (7).

Problem (8) formally belongs to the class of nonlinear programming problems [17] due to additional conditions (7), some of which are nonlinear functions with respect to variables. Therefore, a special solution method is required, suitable for constrained optimization problems [17].

Conditions (7) can be linearized by replacing the variables for drawdown as follows:

$$\mathbf{x} = \begin{bmatrix} J_1 & J_2 & 1/\Delta P_1 & 1/\Delta P_2 \end{bmatrix}^{\mathrm{T}} \tag{9}$$

As a result, system (7) can be represented as:

$$\mathbf{Ax} = \mathbf{b} \tag{10}$$

Where:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & -Q_1 & 0 \\ 0 & 1 & 0 & -Q_2 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} J \\ 0 \\ 0 \end{bmatrix} \tag{11}$$

Further, there are two essentially equivalent ways of solving the problem:

$$\|\mathbf{x}\|_2^2 = \mathbf{x}^{\mathrm{T}}\mathbf{x} \to \min \tag{12}$$

*Method 1 (optimization methods).* On the one hand, problem (12) with constraints (11) is a quadratic programming problem [17]. Due to the fact that all constraints are equalities (there are no inequalities), it can be reduced to solving a system of linear equations:

$$\begin{bmatrix} \mathbf{I} & \mathbf{A}^{\mathrm{T}} \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix} \tag{13}$$

Where **I** is the identity matrix, the elements of the main diagonal of which are equal to one; **λ** is the vector of Lagrange multipliers that appear along with the solution **x**.

*Method 2 (methods of linear algebra).* On the other hand, system (10) itself belongs to the class of underdetermined systems of linear equations; therefore, one of the ways to solve it is to obtain a system of normal equations, in which the number of equations will already be equal to the number of unknowns:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b} \tag{14}$$

Unfortunately, for a given matrix **A** in the form (11), system (14) has no solution, since the determinant of a normal matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is 0.

Nevertheless, solution (10) remains possible on the basis of special algorithms for decomposition of rectangular matrices: QR decomposition, singular value decomposition (SVD), etc.

Thus, factorization problem (3) can be solved in the formulation of minimizing $\ell_2$-norm of either vector **u** (8) or vector **x** (12). Table 1 shows the results of calculations in two ways. Here $Q = 50$ m$^3$ / day and $J = 1$ m$^3$ / day / bar.

As follows from the results in table 1, the solutions obtained in two ways differ significantly from each other, with the exception of the case $q = 1$. This is due, of course, to the fact that in one case the values of the drawdown are minimized, and in the other case, their reciprocal values.

**Table 1.** Solutions based on minimizing the $\ell_2$-norm.

| $q$ | Solution in the form **u** | | | | Solution in the form **x** | | | |
|---|---|---|---|---|---|---|---|---|
| | $J_1$, m$^3$/day/bar | $J_2$, m$^3$/day/bar | $\Delta P_1$, bar | $\Delta P_2$, bar | $J_1$, m$^3$/day/bar | $J_2$, m$^3$/day/bar | $\Delta P_1$, bar | $\Delta P_2$, bar |
| 50/50 | 0.500 | 0.500 | 50.0 | 50.0 | 0.500 | 0.500 | 50.0 | 50.0 |
| 70/30 | 0.637 | 0.363 | 54.9 | 41.4 | 0.501 | 0.499 | 69.9 | 30.1 |
| 90/10 | 0.812 | 0.188 | 55.4 | 26.6 | 0.510 | 0.490 | 88.3 | 10.2 |
| 99/1 | 0.955 | 0.045 | 51.8 | 11.2 | 0.833 | 0.167 | 59.4 | 3.0 |

It is also interesting to note that for the most frequently encountered in practice range of $1 \le q \le 9$ values, the solution in the form of **x** corresponds to formations that are practically homogeneous in terms of productivity. It can also be noted that for both cases, when $q > 1$ the values of *j* and *p* obtained in two ways are always greater than one.

The main problem of factorization (3) based on regularization (8) or (12) is that it is, in fact, an artificial mathematical technique that is used to mechanically solve a problem without taking into account its specifics. In addition, with this approach, there is an incomplete coverage of possible combinations of relations $j$ and $p$ that may take place in reality. In fact, for cases $q > 1$, either it will always be that $j > 1$ or $j \approx 1$, which may contradict the petrophysical concept of the distribution of reservoir properties for different formations.

**Bayesian regularization.**

Consider the factorization problem (3) from a probabilistic point of view. We will assume that the observed values of $q$ for different wells are some finite sample from the generally unknown true distribution of a random variable $Q$. Accordingly, the variables $j$ and $p$ are also instances of random variables $J$ and $P$. From the point of Bayesian statistics, regularization corresponds to the addition of some prior distributions on the required parameters, i.e. a method for calculating the distribution density functions $f_J(j)$ and $f_P(p)$ is required.

Neglecting the influence of different mobility of fluids flowing in or out of formations, different degrees of wellbore damage (skin factor), etc., we will assume that:

$$j = \frac{\sum_n (k_n h_n)_1}{\sum_m (k_m h_m)_2} \tag{15}$$

Where $(k_n h_n)_l$ is the product of the absolute permeability and the thickness of the $n$-th interlayer for the $l$-th layer.

The randomness of $J$ is due to the uncertainty related to absolute permeability values. As it is known, according to the core study data, a positive correlation is observed between the permeability and porosity of the samples, which can be represented as the density of the two-dimensional distribution of a random vector $X = [K, F]^{\mathrm{T}}$:

$$X \sim f_{K,F}(k, \varphi) \tag{16}$$

According to well log interpretation data, the values of the porosity are known for each interlayer. As a result, the distribution for the random values of the permeability of the $n$-th interlayer $K_n$ can be interpreted as the conditional distribution density $K$ at $F = \varphi_n$ obtained on the basis of (16):

$$K_n \sim f_{K|F}(k|\varphi_n) = \frac{f_{K,F}(k, \varphi_n)}{f_F(\varphi_n)} \tag{17}$$

Thus, for each $w$-th well, using (15) and (17), the distribution density $f_J(j|\boldsymbol{\varphi}_w)$ is calculated, where $\boldsymbol{\varphi}_w$ is the vector of interlayer porosity values.

The use of different distribution functions $f_J(j)$ for different wells naturally takes into account the chaotic change in the reservoir properties of oil formations from one zone (well) to another [14]. However, assuming the continuity of the spatial change in reservoir pressure, its distribution across reservoirs is mainly determined by production mechanism of a reservoir and well operation conditions that have been established at the current stage of development. Thus, the main task is to statistically derive a single distribution function $f_P(p)$ for the entire field.

The calculation is attended with certain difficulties. On the one hand, it is possible to use well test data for wells that completed in only one layer. However, as a rule, the amount of such information is not enough to construct a sample distribution function. Instead, we use Bayesian inference to derive the posterior distribution function for *P* by combining all available observations $q_w$ and known functions $f_J\left(j|\boldsymbol{\varphi}_w\right)$, where *w* is the well number.

According to Bayes' theorem, the posterior distribution of *P* with respect to the available data *q* can be calculated as:

$$f_{P|Q}\left(p|q\right)=\frac{f_{Q|P}\left(q|p\right)f_P\left(p\right)}{f_Q\left(q\right)} \tag{18}$$

Where $f_{Q|P}\left(q|p\right)$ is the likelihood of the data *q* at $P=p$; $f_P\left(p\right)$ is prior distribution P; $f_Q\left(q\right)$ is the marginal likelihood of the data *Q*.

The formula for a function $f_Q\left(q\right)$ is not so important, since it plays the role of a simple normalization factor [1; 2]. An uninformative distribution of a random variable *P* whose values belong to an interval of finite length is used as a prior distribution. In this case the probability density of $f_P\left(p\right)$ will be constant throughout this interval. The likelihood of observations $q_w$ can be calculated as:

$$f_{Q|P}\left(q_w|p\right)=f_J\left(\frac{q_w}{p}\middle|\boldsymbol{\varphi}_w\right) \tag{19}$$

Where *w* is the well number. Thus, we get:

$$f_{P|Q}\left(p|\mathbf{q}\right)\propto\prod_w f_J\left(\frac{q_w}{p}\middle|\boldsymbol{\varphi}_w\right) \tag{20}$$

Where $\mathbf{q}$ is the vector of observations $q_w$.

Using a different likelihood function for each observation *q* differs from classical Bayesian statistical inference. This approach is typical for hierarchical models in which some of the prior distributions (so-called hyperdistributions) are shared as parameters of lower-level distributions [1]. In this case, the role of such a prior hyperdistribution plays $f_P\left(p\right)$, through which information is exchanged between different groups of observations (flow profile and reservoir properties of the wells) to obtain a more stable (reduced) estimate of the posterior distribution [1].

In addition, let us explain some details of the practical implementation of the proposed scheme. As it turned out, more stable estimates of the distribution parameters are obtained if we carry out a logarithmic transformation of problem (3):

$$\log q = \log j + \log p. \tag{21}$$

Note also that, as a rule, the results of studies on the core are presented in the form of pairs of sample $(\log k,\ \varphi)$ values.

**Methodology of factorizing the flow profile data.**
As a result of Bayesian regularization, posterior distributions $j$ and $p$ are calculated, which can be used as additional constraints when solving the factorization problem (3). For example, one can solve this problem in the following form:

$$(j - j_{\mathrm{MAP}})^2 + \alpha (p - p_{\mathrm{MAP}})^2 \to \min \tag{22}$$

Subject to fulfillment of (3), where $j_{\mathrm{MAP}}$, $p_{\mathrm{MAP}}$ are the maximum aposteriori estimates; α is weight coefficient.

In fact, under condition (3), problem (22) is reduced to finding a real positive root of the quartic equation:

$$j^4 - j_{\mathrm{MAP}} j^3 + \alpha q p_{\mathrm{MAP}} j - \alpha q^2 = 0 \tag{23}$$

In some cases, equation (23) can have two real positive roots. This situation arises with an equivalent contribution of the first and second terms to the total sum (22), i.e. the problem has two equivalent solutions.

Further, on the basis of (4), one can calculate the individual parameters of the layers:

$$\begin{aligned} J_1 &= \frac{j}{j+1} J; & J_2 &= \frac{1}{j+1} J \\ \Delta P_1 &= \frac{p(j+1)}{q+1} \Delta P; & \Delta P_2 &= \frac{j+1}{q+1} \Delta P \end{aligned} \tag{24}$$

## 3 Results

**Examples of real fields.**
Let us consider the application of Bayesian regularization for the factorization problem (3) on the example of three fields. At each of them, productive formation consists of two conditionally distinguished layers, designated as C2vr (upper layer) and C2b (lower layer). For some wells of the fields, a production logging tests were carried out to determine the flow profile. If there were several studies for the same well, then only the latest flow profile results were used.

In total, 15, 87 and 65 flow profile results were used for fields No. 1, 2, and 3. Figure 2 shows the distributions for all three fields, obtained using kernel density estimation (KDE) method. The bandwidth was 3 for $q$ and 0.2 for $\log q$. All distributions also have 0.25, 0.5 and 0.75-quantiles.

As it can be seen, among all fields, the distribution shape $\log q$ for field No. 3 is closest to the normal curve, and its center (median) is located above zero. At the same time, field No. 2 demonstrates a two-modal distribution type $\log q$, i.e. perhaps there are two groups of wells, for one of which the most probable values of $\log q$ will be less than zero, and for the second, on the contrary, they will be above zero. Regarding field no. 1, the only thing that can be said is that for about 75% of the samples there is $\log q > 0.5$. In addition, for all field the percentage of observations, for which $\log q > 0$, is higher than 60%.

The results of the core study to determine the open porosity and absolute permeability of core samples for various formations of the fields are shown in figure 3. Dots denote pairs of values $(\log k,\ \varphi)$, which are assumed to have a joint normal distribution. Thus, according to (16) we use:

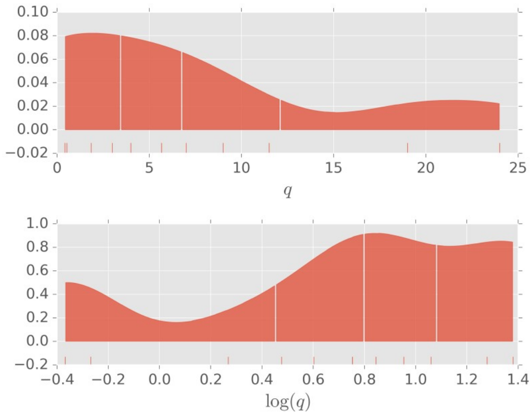
$$\begin{bmatrix} \log \mathrm{K} \\ \Phi \end{bmatrix} \sim \mathrm{N}(\boldsymbol{\mu},\ \boldsymbol{\Sigma}) \tag{25}$$

Where $\boldsymbol{\mu}$ is the vector of mean values for $\log k$ and $\varphi$; $\boldsymbol{\Sigma}$ is covariance matrix 2×2; $\mathrm{N}(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$ is the probability density function of the bivariate normal distribution.

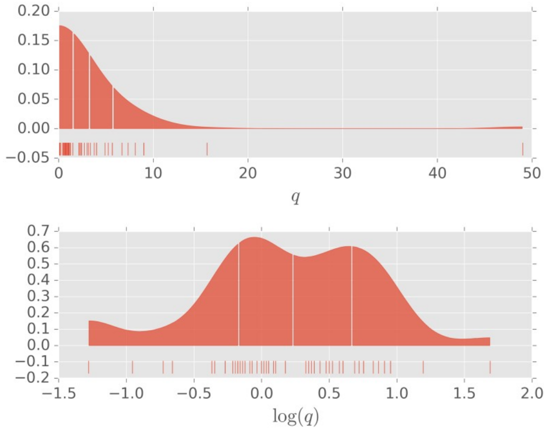
To determine the parameters of the normal distribution (25), a Bayesian inference procedure similar to (18) was used. The following prior distributions of parameters were used:

$$\begin{gathered} \boldsymbol{\mu} \sim \begin{bmatrix} \mathrm{N}(\hat{\mu}_{\log k},\ \sigma_\mu) \\ \mathrm{N}(\hat{\mu}_{\varphi},\ \sigma_\mu) \end{bmatrix} \\ \boldsymbol{\sigma} \sim \begin{bmatrix} \mathrm{Exp}(\lambda) \\ \mathrm{Exp}(\lambda) \end{bmatrix} \\ \mathbf{C} \sim \mathrm{LJK}(\eta) \end{gathered} \tag{26}$$

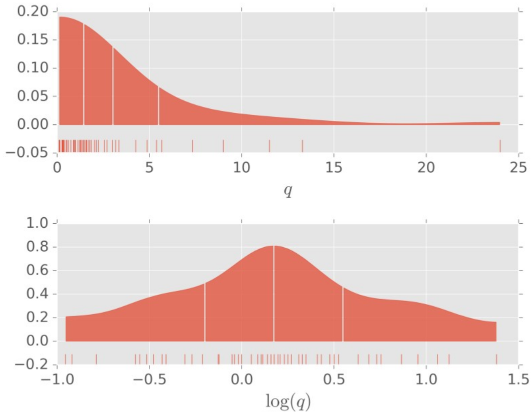
Where $\hat{\mu}_{\log k}$, $\hat{\mu}_{\varphi}$ are sample means for $\log k$ and $\varphi$, respectively; $\mathrm{Exp}(\lambda)$ is the probability density function of the exponential distribution with a parameter $\lambda > 0$; $\mathrm{LJK}(\eta)$ is the probability density function of the LKJ-distribution with a parameter $\eta \geq 1$; $\mathbf{C}$ is the correlation matrix 2×2 for which $\boldsymbol{\Sigma} = \mathbf{D}_\sigma \mathbf{C} \mathbf{D}_\sigma$; $\boldsymbol{\sigma}$ is the vector of standard deviations for $\log k$ and $\varphi$; $\mathbf{D}_\sigma$ is the diagonal matrix with values $\boldsymbol{\sigma}$ on the main diagonal. A reasonable choice is $\sigma_\mu = 1$, $\lambda = 1$ and $\eta = 2$ [18].




a) field number 1

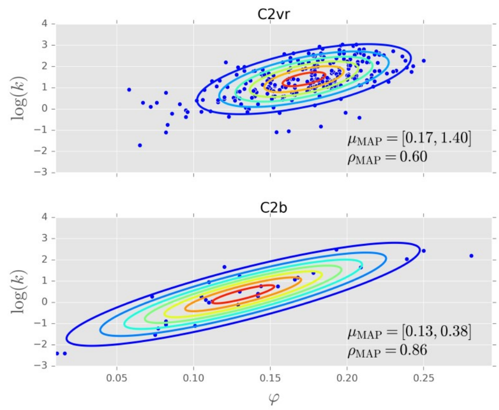



b) field number 2

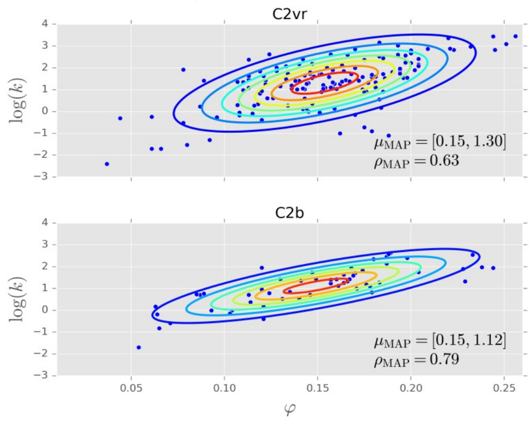



c) field number 3

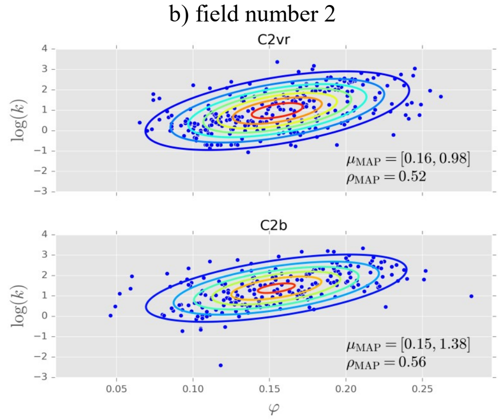
**Fig. 2.** Distributions of $q$ for fields.




a) field number 1



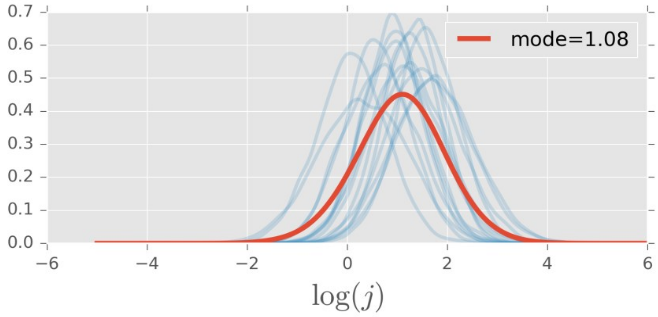

b) field number 2



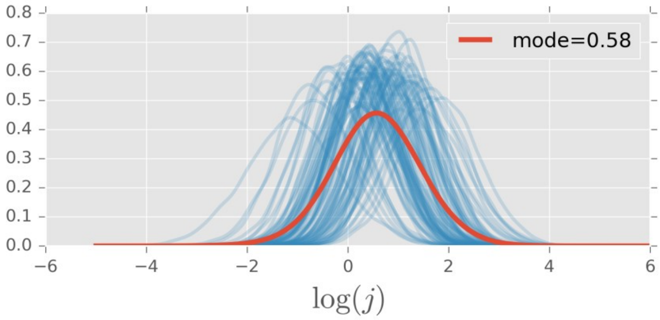

c) field number 3

**Fig. 3.** Distributions of $(\log k,\ \varphi)$ of core samples for fields.

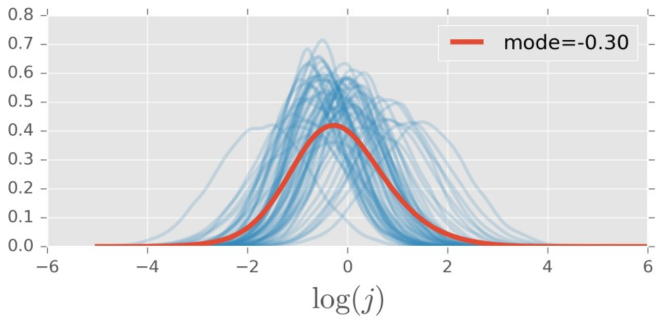
Figure 3 also shows the values of the coordinates of the distribution center $\boldsymbol{\mu}_{\mathrm{MAP}}$ and the correlation coefficient $\rho_{\mathrm{MAP}}$ corresponding to the maximum of the posterior distributions $\boldsymbol{\mu}$ and $\mathbf{C}$. For fields No. 1 and 2, core samples from the C2vr formation have, on average, higher porosity and permeability than for the C2b formation. In turn, for field No. 3, the core samples of the C2vr formation, although on average have a higher porosity, their average permeability is lower than for the C2b formation. It is interesting to note that for all fields, the correlation coefficient between porosity and permeability is highest for the C2b reservoir.




a) field number 1




b) field number 2




c) field number 3

**Fig. 4.** Distributions of $\log j$ for fields.

Then, for each well, the following sequence of actions was implemented:

- According to the data on perforations for each layer, a group of interlayers completed on the date of the flow profile was formed;
- According to well log interpretation data, for each interlayer, the value of porosity was taken and the distribution $\log k$ was calculated based on (25) and (17);
- According to well log interpretation data, for each interlayer, distribution *kh* was calculated based on the obtained distribution $\log k$ and interlayer's thickness;
- For each layer, the distribution of the sum of all interlayer's *kh* obtained at the previous step was calculated;
- The distribution of $\log j$ was calculated based on (15).

Figure 4 shows the results of calculating the distribution of $\log j$ separately for each well, as well as the averaged distribution curve for the entire field. The plots also show the value of the mode of the averaged distribution of $\log j$. Thus, for fields No. 1 and No. 2, the productivity of the C2vr formation is on average about 10 and 4 times higher than the productivity of the C2b formation, respectively. At the same time, for field No. 3, the productivity of C2vr is on average approximately 2 times less than the productivity of the C2b formation. For all fields, the maximum value of the distribution mode of $\log j$ does not exceed 2, while the minimum value of the distribution mode of $\log j$ is higher than -2.

## 4 Discussion

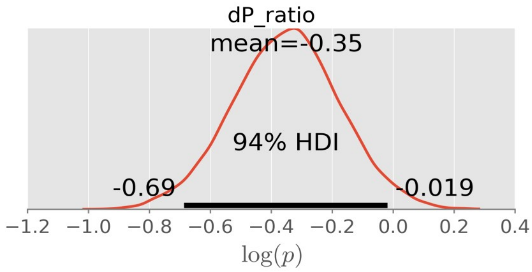
As a result of Bayesian inference base on (18), the posterior distributions of $\log p$ for all the fields are obtained and are shown in figure 5. The mean value and highest density interval (HDI) of 94% are also indicated there. Thus, for fields No. 1 and No. 2 $p_{\text{MAP}} < 1$, i.e. the drawdown for the C2vr formation is on average almost 2.5 times lower than the drawdown for the C2b formation. At the same time, it is the other way around for deposit No. 3, for which $p_{\text{MAP}} > 1$, i.e. it is possible that even at a lower productivity of the C2vr formation, the reservoir pressure in it is higher than in the C2b formation. As a result of which, a greater inflow from / injection into C2vr is observed. It is also interesting to note that the range of the posterior distribution of $\log p$ is the highest for field No. 1, which is possibly due to the small volume of observations of $\log q$.

Figure 6 shows the results of testing the convergence of the procedure for Bayesian inference of the posterior distribution. Trace diagrams on the right in Fig. 6 looks like white noise, which indicates a good mixing of the statistical inference engine. In this case, the NUTS sampler was used, which generated four parallel chains of independent samples from the posterior distribution. The size of each sample was 5000 elements. In addition to them, 2000 elements in the chain were intended for automatic tuning of the sampling algorithm.

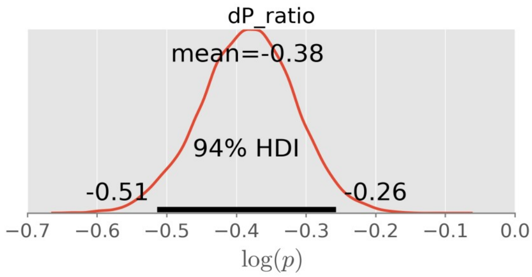
# 5 Conclusion

The problem of factorization of the results of flow profile of a two-layered reservoir using well test data is formulated. It is shown that this problem is ill-posed, and additional restrictions on its conditions (regularization) are required.

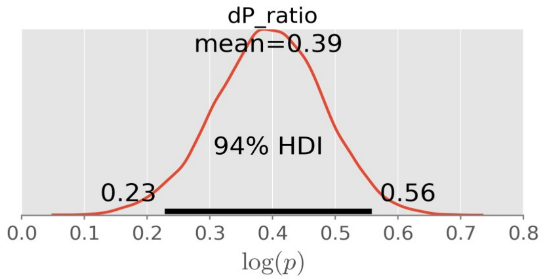
Several methods of deterministic regularization based on the minimization of the $\ell_2$-norm of the vector of unknowns are described. It is shown that such approaches do not take into account the petrophysical properties of formations and have limitations on the coverage of possible factorization solutions.




a) field number 1

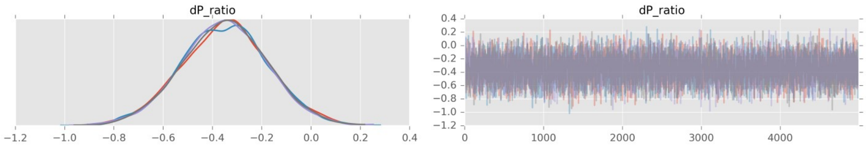



b) field number 2

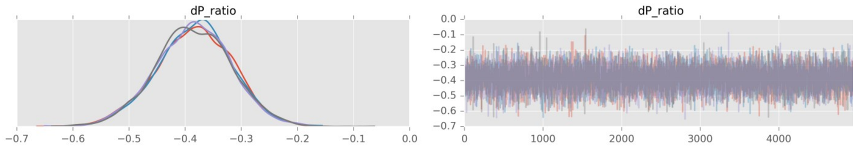



c) field number 3

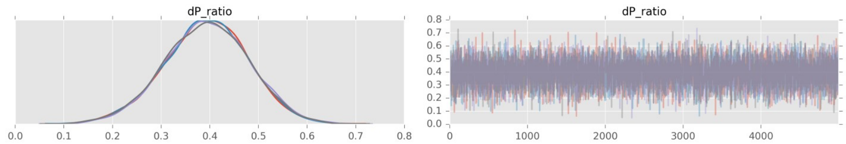
**Fig. 5.** Posterior distributions of $\log p$ for fields.




a) field number 1




b) field number 2




c) field number 3

**Fig. 6.** Convergence results of the Bayesian inference procedure for fields.

A Bayesian regularization is proposed for factorizing the results of the flow profile. According to it, for all relative factors the corresponding probability distribution functions are formed. For this, data from core studies are used to determine the permeability and porosity of the samples. Calculation of the distributions of the ratio of reservoir productivity is carried out separately for each well, taking into account the well log interpretation data. Bayesian statistical inference is used to obtain the general distribution of the drawdown ratio for different reservoirs.

This approach was tested on the example of three real fields. All calculations were carried out in Python using the PyMC3 probabilistic programming library. Visualization and exploratory analysis of the results were carried out using the Matplotlib and ArviZ libraries.

## References


1. Martin O.: Bayesian Analysis with Python: Introduction to Statistical Modeling and Probabilistic Programming Using PyMC3 and ArviZ 2nd Edition. Birmingham-Mumbai: Packt (2018).
2. Downey A.B.: Think Bayes. Sebastopol: O'Reilly Media (2013).
3. Korde A.A.: Probabilistic Decline Curve Analysis in Unconventional Reservoirs Using Bayesian and Approximate Bayesian Inference Thesis. Fairbanks: University of Alaska (2019).
4. Li L.: Bayesian Statistical Inference Applied to Reservoir Modelling and Earthquake Scaling Thesis (Ph.D.). Edinburgh: University of Edinburgh (2006).

5. Booth R., Morton K., Onur M., Kuchuk F.: Grid-based Inversion of Pressure Transient Test Data with Stochastic Gradient Techniques Int. J. for Uncertainty Quantification, 2, 4 323–339 (2012).
6. Christen A., Sanso B., Santana Cibrian M., Velasco-Hernandez J.: 2015 Bayesian Deconvolution of Oil Well Test Data Using Gaussian Processes J. Appl. Statistics, 43, 422 (2016).
7. Eltahan E., Ganjdanesh R., Yu W., Sepehrnoori K.: Assisted History Matching Using Bayesian Inference: Application to Multiwell Simulation of a Huff 'n' Puff Pilot Test in the Permian Basin Proc. of Unconventional Resources Technology Conf. (Austin) (SPE) (2020).
8. Darwis S., Fitriyati N., Gunawan A.Y., Marwati R.: 2012 Bayesian Reservoir Simulation Int. Conf. on Statistics in Science, Business and Engineering. Langkawi, IEEE, 1–5. (2012).
9. Masoudi P., Asgarinezhad Y., Tokhmechi B.: Feature Selection for Reservoir Characterisation by Bayesian Network Arabian J. of Geosciences, 8(5), 3031–43 (2015).
10. Dou Z.: Bayesian Global Optimization Approach to the Oil Well Placement Problem with Quantified Uncertainties Thesis (M.S.). West Lafayette: Purdue University (2015).
11. Anraku T.: Discrimination Between Reservoir Models in Well Test Analysis Thesis (Ph.D.). Stanford: Stanford University (1993).
12. D'Windt A.: Reservoir Zonation and Permeability Estimation: A Bayesian Approach Proc. of 48th Annual Logging Symp. Society of Petrophysicists and Well-Log Analysts (2007).
13. Loures L.L.: Porosity Bayesian Inference from Multiple Well-Log Data. CREWES Research Report, 14 (2002).
14. Lysenko V.D.: Oilfield Development: Theory and Practice. Moscow: Nedra, 367 (1996).
15. Ipatov A.I., Kremenetsky M.I.: Geophysical and hydrodynamic control of the development of hydrocarbon deposits. Moscow: Izhevsk: Institute of Computer Research, 780 (2005).
16. Numaier A.: Solving Ill-conditioned and Singular Linear Systems: A Tutorial on Regularization. SIAM Review, 40(3), 636–66 (1998).
17. Avriel M.: Nonlinear Programming: Analysis and Methods. Mineola: Dover Publishing (2003).
18. LKJ Cholesky Covariance Priors for Multivariate Normal Models https://docs.pymc.io/notebooks/LKJ.html, last accessed 2020/10/29.