IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 1, 2022, accepted January 29, 2022, date of publication February 7, 2022, date of current version February 22, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3149625

Estimation of Filtration Properties of Host Rocks in Sandstone-Type Uranium Deposits Using Machine Learning Methods

RAVIL I. MUKHAMEDIEV[1,2], YAN KUCHIN[1,2], YEDILKHAN AMIRGALIYEV[1], NADIYA YUNICHEVA[1], AND ELENA MUHAMEDIJEVA[1]

[1]Institute of Information and Computational Technologies, MES RK, Almaty 050010, Kazakhstan
[2]Institute of Automation and Information Technologies, Satbayev University (KazNRTU), Almaty 050000, Kazakhstan

Corresponding authors: Yan Kuchin (ykuchin@mail.ru), Ravil I. Mukhamediev (ravil.muhamedyev@gmail.com), and Nadiya Yunicheva (naduni@mail.ru)

This work was supported by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan under Grant AP09562260 and Grant BR10965172.

ABSTRACT The nuclear decay of uranium is one of the cleanest ways to meet the growing energy demand. The uranium needed for power plants is mainly extracted by two methods in roughly equal amounts: quarries (underground and open pit) and in-situ leaching (ISL). The effective use of ISL requires, among other things, the correct determination of the filtration characteristics of the host rocks. In Kazakhstan, this calculation is still based on methods that were developed more than 50 years ago, and in some cases, give inaccurate results. At the same time, knowledge of filtration characteristics is necessary for the calculation of recoverable reserves, prediction of production dynamics, calculation of the optimum number of wells, etc. This paper describes a method for calculating the filtration coefficient of ore-bearing rocks using machine learning. The proposed method is based on nonlinear regression models. It also allows the estimation of the filtration properties of rocks within the process acidification zone, where the existing method is not applicable. The proposed method applies to approximately half of the uranium mined in the world and makes it possible to significantly (by 22 %–70%) increase the accuracy of the filtration coefficient determination and, accordingly, improve the accuracy of recoverable reserves calculation and economic indicators of mining processes.

INDEX TERMS Uranium mining, machine learning, regression model, filtration characteristics.

I. INTRODUCTION

Nuclear power, despite the environmental risks involved, remains one of the cleanest ways to meet the growing demand for energy without increasing greenhouse gas emissions. Nuclear power plants require the mining of uranium ore to power them. Uranium is mined in 28 countries, of which ten countries account for more than 90% of the established reserves [1] (Fig. 1).

According to the World Nuclear Association, in 2018, the largest uranium mining companies produced 86% of the world's total uranium production [2], of which NAC Kazatomprom JSC accounted for 21%. Companies use two main mining methods: open pit (underground and open-pit), which accounted for 45.9% of the production, and in-situ leaching (ISL), which accounts for 48.3% of the world's uranium production. Approximately 5.8% of uranium is mined as a byproduct, such as in gold mining [3]. Appendix A provides a detailed overview of uranium mining worldwide (https://www.dropbox.com/s/ijl3my3z9dhfg4z/Appendix_A.pdf?dl=0).

The ISL is considered the most modern and environmentally friendly technology, which is widely used in Kazakhstan, Uzbekistan, the United States, and partly in Canada, Australia, and China. However, despite the merits of ISL, there are several challenges.

First, the application of this method requires a fairly accurate determination of the lithological composition of rocks, because uranium mining in this case is carried out in the ore body, located between the impermeable layers and, as a rule, below the groundwater level.

The associate editor coordinating the review of this manuscript and approving it for publication was Sajid Ali.

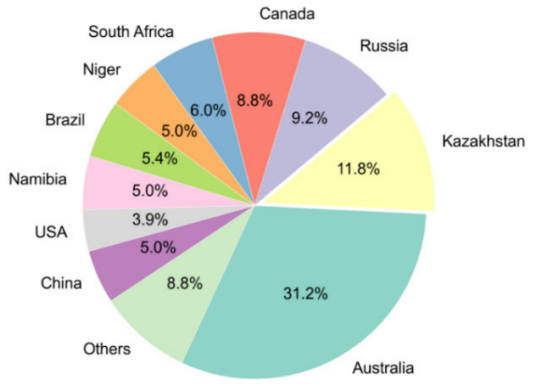

FIGURE 1. Countries with the largest uranium reserves.

Second, the filtration properties of the rocks must be known to estimate the reserves and extract the maximum amount of uranium.

Inaccuracies in determining the lithological composition and filtration characteristics lead to errors in the technological process of filter installation and errors in determining ore reserves. For example, economic losses from incorrect lithological classification in the deposits of Kazakhstan can be estimated to be approximately 1 to 4 million dollars per year [4].

These inaccuracies are caused by both the technological limitations of logging readings and, to a large extent, by the methods used to determine the lithological composition and filtration properties of rocks. When determining the filtration properties of rocks in the field, the key aspect is to determine the filtration coefficient (K_f) at the stage of exploratory drilling, which is further used to calculate the filtration properties of technological wells. However, the accepted methodology, based on analytical methods, has not changed since the end of the last century [5]. Meanwhile, the correct determination of K_f is necessary for the calculation of recoverable reserves, prediction of production dynamics, calculation of the number of wells, and the distance between them (hexagonal cell diameter or distance between well rows), and selection of optimal length and location of filter by depth.

This study considers the application of machine learning methods to estimate the filtration characteristics of ore-bearing rocks. The method is based on the use of nonlinear regression models and has shown results 22 %–70% better than calculations using the existing methodology used in Kazakhstan. The method also allows for the estimation of the filtration properties of rocks within the technological acidification zone. The proposed method concerns approximately half of the mined uranium in the world.

The work consists of the following sections:

- The first section briefly describes the existing techniques for determining the filtration coefficient and its shortcomings.

- In the second section (related works), we provide an overview of the work devoted to the application of machine learning methods to mining problems.

TABLE 1. Rock type depending on prevailing fraction diameter.

clayey-siltstone	less than 0,05
fine grained sand	0,05 - 0,1
close grained sand	0,1 - 0,25
medium sand	0,25 - 0,5
coarse sand	0,5 -1,0
gravel	over 1,0

- In the third section, we present the methodological scheme of the study, describe the machine learning models we applied, and the metrics for evaluating the quality of their performance.

- The fourth section presents and discusses the results obtained.

- The conclusion briefly describes the results obtained, the limitations of the method, and formulates the objectives of future research.

II. THE METHODOLOGY USED IN PRACTICE FOR DETERMINING THE FILTRATION PROPERTIES OF HOST ROCKS AND ITS LIMITATIONS

To determine the relationships between the filtration properties of host rocks and the value of apparent resistivity (AR), hydrogeological studies (pumping) were carried out at the stage of exploration, and sampling for the analysis of rock grain size distribution (GS) was carried out.

Analysis of GS samples of rocks of productive horizons in hydrogenous fields shows that the distribution of particle sizes can be well approximated by the log-normal law, and to characterize the rocks of productive horizons it is advisable to allocate the following lithological types, each of which is characterized by a certain range of particle fractions (in millimeters) (Table 1) [6]:

If the mass fraction of clay-silty particles exceeds 50%, the rock is identified as clay, and sand is subdivided into fine, close, medium, and coarse grained depending on which fraction exceeds 50% by mass fraction; if the mass fraction of gravel particles exceeds 50%, the rock is identified as gravel. If the mass fraction of none of the fractions exceeded 50%, the rock was identified by the fractions whose sum of mass fractions of particles exceeded 50%. For example, if the sum of the mass fractions of the fine- and medium-grained fractions exceeds 50%, the rock is identified as fine-to medium-grained sand. If the sum of the mass fractions of three or more sand fractions exceeds 50%, then the rock is identified as multigrained sand.

As a result of the joint processing of electric logging data, the results of the analysis of particle size distribution and data on filtration properties of rocks of productive horizons in fields of infiltration type, the following regularities are established:

- The most stable parameters that characterize individual lithologic rock types are particle diameters $d_e = d_{0,1}$ (the so-called effective diameter) and $d_{0,6}$ - the average

particle diameters with relative mass fraction of 0,1 and 0,6, respectively;

- sands, as a rule, are characterized by a coefficient of heterogeneity $K_H = d_{0,6}/d_{0,1} \leq 5$, which allows us to classify them as homogeneous rocks.

- effective diameter $d_{0,1}$ (or $d_{0,6}$ - mainly for sands) carries the main information about the belonging of a rock to a particular type.

- parameters $d_{0,1}$ and $d_{0,6}$ are connected by statistical dependences with electric parameters ρ_κ and Spontaneous Polarization (SP)

$$\alpha_{\rho_\kappa} = \frac{\rho_\kappa - \rho_\kappa^{(\min)}}{\rho_\kappa^{(\max)} - \rho_\kappa^{(\min)}} \quad \text{or } \alpha_{sp} = \frac{U_{\pi c} - U_{\pi c}^{(\min)}}{U_{\pi c}^{(\max)} - U_{\pi c}^{(\min)}},$$

where *max* and *min* denote the maximum and minimum values of the corresponding parameters (apparent resistivity and potential) within the productive horizon, respectively. As a rule, at small values of electric parameters, their connection with values $d_{0,1}$, is more stable, and at large, with values $d_{0,6}$;

- filtration coefficient K_f and parameters $d_{0,1}$ ($d_{0,6}$) are statistically connected by the dependence of the form $K_f = Ad_{0,1}^2$ or $K_f = A_1 d_{0,6}^2$, where A and A_1 are constant multipliers for a given productive horizon, which are determined by the results of pilot pumping in hydrogeological wells with known values of $d_{0,1}$ and $d_{0,6}$ of rocks comprising the productive horizon. Appendix B provides a detailed explanation of physical aspects of logging data acquisition (https://www.dropbox.com/s/8d26z0umi8s5gow/Appendix_B.pdf?dl=0).

The statistical relationship between the parameters $d_{0,1}$ and/or $d_{0,6}$ and the electrical parameters $\rho_\kappa, \alpha_{\rho_\kappa}$ or α_{sp} allows to restore the values of $d_{0,1}$ and/or $d_{0,6}$, and therefore, to identify the rock type and evaluate the filtration coefficient K_f.

The obtained dependencies are given in the standards and are used to calculate K_f based on the average apparent resistivity within the allocated lithological interval.

To obtain data on the lithological structure of a sandstone type uranium deposit, the following electrical logging methods were used: induction log (IL), apparent resistance logging (AR), and spontaneous polarization potential (SP). During the logging process, a probe is lowered into the drilled borehole, which, when lifted, provides measurement data in 10-cm increments. Appendix C explains the process of assessment of filtration properties at the exploration stage (https://www.dropbox.com/s/wfwjnk8ufv675oo/Appendix_C.pdf?dl=0).

By interpreting the logging data, the expert, by characteristic points of the AR curve, identifies the boundaries of lithological intervals, within which the average value of the apparent resistivity is determined. Using the dependencies the filtration coefficient is calculated [7]. Appendix D (https://www.dropbox.com/s/vnlg2lfiw40bxb6/Appendix_D.pdf?dl=0) provides an example.

By obtaining information about the rock distribution, filtration properties, and depth of the ore body occurrence, we can proceed to determine the optimal location of filters for acid injection and pumping of the productive solution (the filter position depends on whether the well is injection or pumping) [8].

A significant drawback of K_f determination method used in practice is that it uses data from only one logging method. Consequently, it becomes inapplicable when the record is distorted. This occurs most often in acidified blocks, that is, where rocks have been exposed to acid and have changed their physical properties. To control the process of acidification with an interval of 1-2 years in the wells of the geological site, IL logging was carried out, during which the conductivity of rocks was measured and the degree of acidification was determined by the increment of conductivity, that is, the drop in resistance of rocks. An example of such a sequential conductivity measurement is presented in Fig. 2.

Induction log curves were recorded for 2018, 2020, and 2021. We can see their consistent increment at the acidification interval (205-235m), highlighted in yellow. The increment is maximum (up to five times) for the reservoirs with the highest filtration coefficients. Fig. 3 shows a passport of the well with highlighted acidification intervals, for which the AR curve values were underestimated, IL values were overestimated, and SP curve values were not distorted.

Because the conductivity increment indicates a resistance drop, it is practically impossible to apply the existing method to acidification intervals. Therefore, for these intervals, the filtration coefficient values of the corresponding tracks were deleted. At the same time, such intervals are present in 30%–40% of wells and, as a rule, in the ore-bearing horizon, which is of the greatest interest for the interpreter. It is difficult to use the IL curve to determine filtration properties because such measurements are not taken at the exploration stage; hence, they cannot be compared with the pumping results. In addition, in Kazakhstan, IL is not part of the standard set of measurements, that is, it is not performed in all wells. Thus, the standard method for determining the filtration coefficient has clear drawbacks that prevent accurate determination of reserves and planning of production processes.

III. RELATED WORKS. MACHINE LEARNING METHODS IN MINING TASKS

Machine learning (ML) is a subset of artificial intelligence techniques that allows computer systems to learn from previous experiences (i.e., from data observations) and improve their behavior to perform a particular task [9]. ML solves the problems of regression, classification, clustering, and data dimensionality reduction. ML models are divided into five classes [10], [11]: unsupervised learning (UL) or cluster analysis [12], supervised learning (SL) [13], semi-supervised learning (including self-learning) (SSL), reinforcement learning (RL), and deep learning (DL).

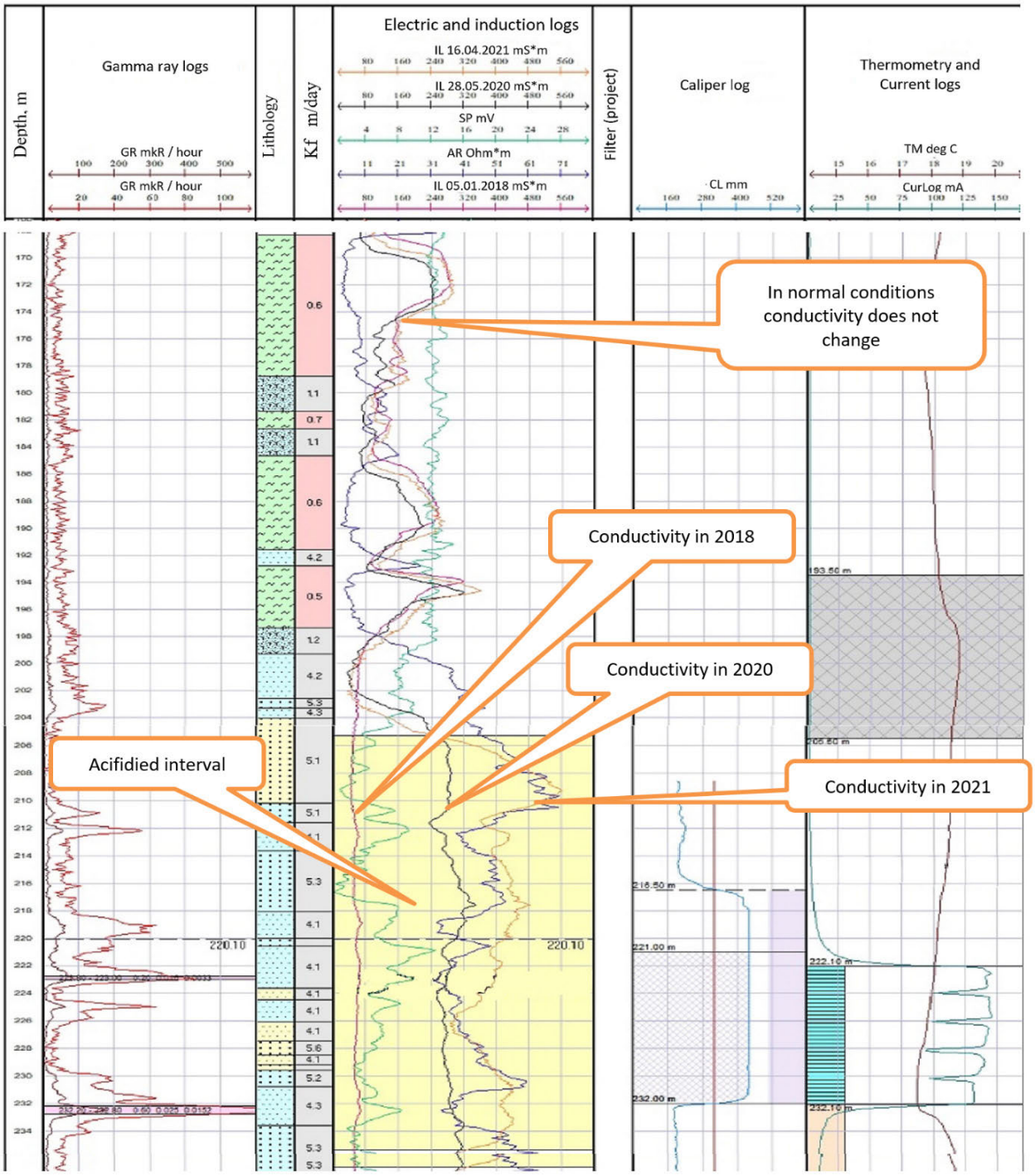

FIGURE 2. Conductivity measurements at different time periods.

UL models solve the problems of clustering and data dimensionality reduction when a set of unlabeled objects is partitioned into groups by an automatic procedure based on the properties of these objects [14], [15].

SL models solve classification or regression problems. A classification problem arises when finite groups of objects in a potentially infinite set of objects are distinguished by labeling [16]. Labeling is often performed by experts. The classification algorithm, using this initial classification as a pattern, must assign the unlabeled objects to this or that group based on the properties of these objects. Regression is the task of predicting a continuous quantity.

The SSL, RL, and DL models are often used for classification and regression tasks. The peculiarity of DL is the possibility of applying end-to-end learning (end-to-end), which, in turn, requires large volumes of marked-up data.

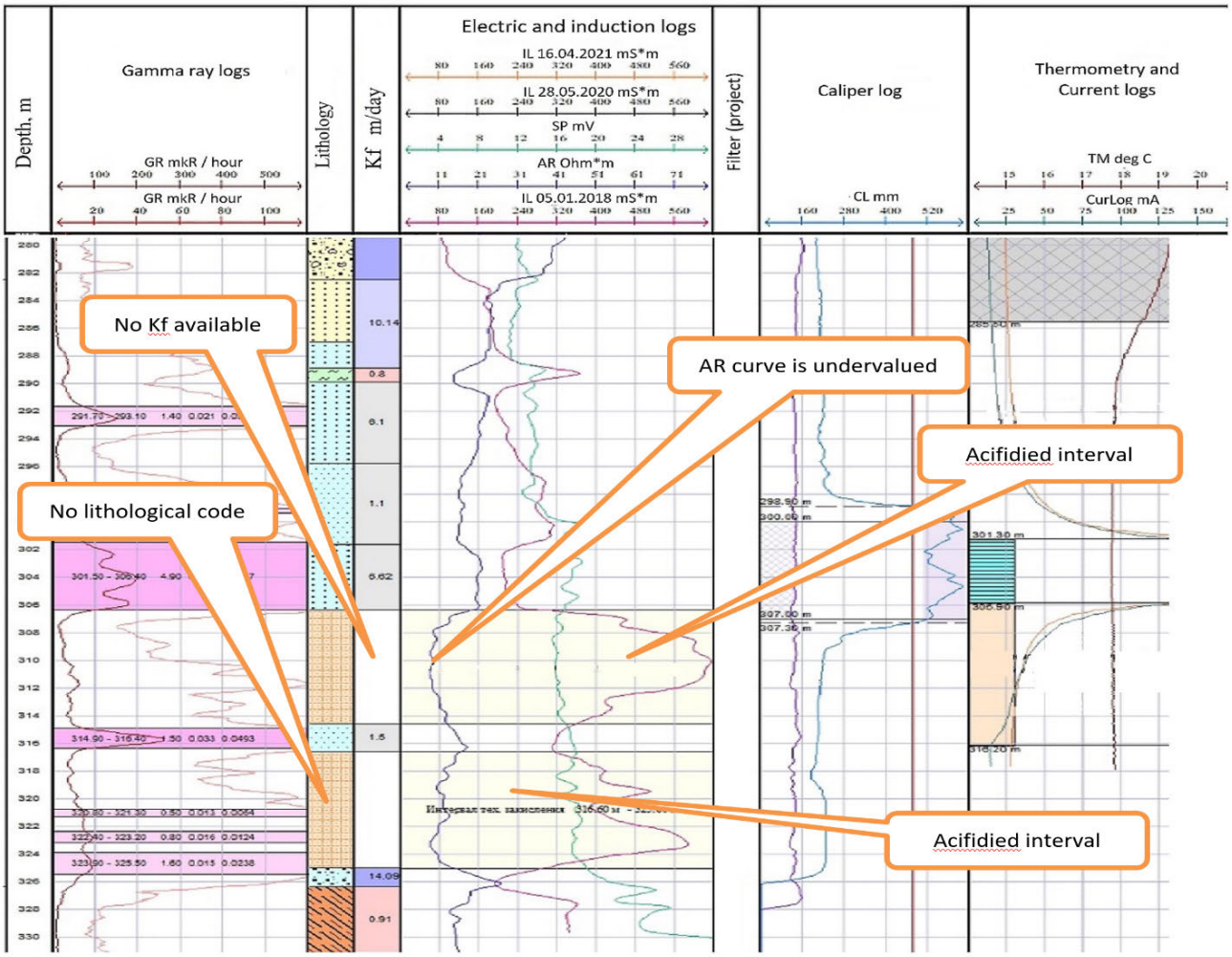

FIGURE 3. Well passport with highlighted acidification intervals.

ML models can be roughly divided into ''classic'' and ''modern'' (Table 2).

ML has been successfully used to solve problems in medicine [32], [33], biology [34], robotics, urban [35] and industry [36], [37], agriculture [38], modeling of environmental [39] and geoecological processes [40], to create a new type of communication system [41], astronomy [42], geological exploration [43], natural language processing [44], [45], and problems in search, evaluation, and mining.

ML methods are used in geological mapping to search for ore deposits, risk assessment, and hydrological and environmental modeling [46]. One of the most popular tasks is lithological classification. For example, the work in [47] deals with the lithological mapping of Hutti province in India using AVIRIS-NG multispectral data. As a result of the comparison, the Support Vector Classifier (SVC) algorithm was chosen to solve this problem. In [48], it was shown that SVC and Ensemble Methods (EM) showed the results of classification of rock types better or equal to those obtained using standard classification methods in 2D or 3D modeling of geological objects. This article emphasizes the high dependence of the results on the quality of expert labeling. The lithological mapping task using SVÑ and remote sensing data for the southern provinces of Morocco was also considered in [49]. The lithological classification of crystalline rocks based on logging data was also considered in [50]. Again, the SVC algorithm was chosen as the best algorithm.

Another popular direction is the analysis of remote sensing data of the earth's surface and the evaluation of mining prospects.

For example, [51] discussed the applications of machine learning to analyze remotely sensed data in mineral prospecting. Classification of the covering surface near uranium ore processing zones for nuclear nonproliferation treaty compliance assessment using machine learning techniques and remotely sensed earth surface data was considered in [52]. In [53], a method for evaluating the prospectivity of tungsten deposits using machine learning and deep learning techniques. RF, SVC, ANN, and CNN were used to solve the classification problem in this study. Based on the results of

TABLE 2. ML models for data analysis.

	ML class methods	ML model
Classic	SL	k-Nearest Neighbor (k-NN) [21,22,23], Logistic Regression, Decision Tree (DT), Support Vector Classifier (SVC) [24], Feedforward Artificial Neural Networks (ANN) [25]
	UL	k-means [32] Principal Component Analysis (PCA)
Modern	UL	Isometric Mapping (ISOMAP) [17], Locally Linear Embedding (LLE) [18], t-distributed stochastic neighbor embedding (t-SNE) [19], kernel principal component analysis (KPCA) [18], multidimensional scaling (MDS) [20]
	SL, RL, Assembling methods (composition of algorithms)	Ensemble methods (EM), boosting [26], random forest (RF) [27], etc.
	SL, SSL, RL, DL	Deep Feedforward Neural Networks [28], Convolutional Neural Networks (CNN) [29], Recurrent Neural Networks (RNN) [30], Long short-term memory (LSTM) [31]

the quality assessment of the methods, RF was chosen as the main prediction method.

The application of machine learning to solve the problems of stratigraphy at uranium deposits in Kazakhstan was considered in [54]. Some papers by the authors of this article are devoted to the problems of classifying lithological types in uranium deposits using ''classical'' algorithms [55], combining the results of several classification models [56], comparative analysis of ''classical'' models and some deep learning methods [8], [57], evaluation of the quality of expert labelling of logging data in solving the problem of lithological classification [58], [59]. In the papers mentioned above, a quantitative assessment of the influence of experts on the solution of the lithological classification problem was carried out. Regression methods were used to estimate the amount of silica and iron in the ore [60]. The authors have shown that boosting methods and EM demonstrate good prediction result (96-98%).

The analysis of the works shows a great interest of researchers in solving the problem of classification and regression in the mining industry using ''classical'' machine learning algorithms. However, we have not been able to identify works devoted to the calculation of the filtration properties of rocks in uranium deposits of formation-infiltration type using machine learning methods, despite the fact that the analytical method used in practice yields inaccurate results. At the same time, the correct estimation of K_f is critical for calculating the amount of recoverable reserves, predicting the dynamics of production, the number of necessary wells, the choice of parameters, and the location of the filter to be installed within the productive horizon. Therefore, improving the accuracy of K_f estimation based on the application of machine learning methods is the goal of this research. This task can be solved using regression models because of the problem of predicting continuous values.

IV. METHOD

Given the shortcomings of the existing K_f estimation methodology, we propose a machine learning model that receives basic logging data as input and generates filtration coefficients as output. Such a model can be trained on data from exploratory wells that have actual (pumped out) K_{fpo}. values. The trained model can then be used to calculate K_f of the production wells.

The problem of estimating filtration properties from logging data belongs to the class of example-based or supervised learning problems. Mathematically, it is reasonable to consider the problem of learning by examples as an optimization problem, which can be solved by searching for the minimum value of the cost function $J(\theta)$ on all available examples, defined as the sum of squares of the difference between the ''predicted'' value and the real value on the set of examples m. In this case, a hypothesis $h_\theta(x)$ is selected that provides the minimum value of $J(\theta)$ on a certain set of parameters $\theta_i \in \Theta$:

$$J(\theta) = min\frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2,$$

where m is the set of training examples, h_θ is the hypothesis function, which can be linear ($h_\theta = \theta_0 + \theta_1 x$) or nonlinear (e.g., $h_\theta = \theta_0 + \theta_1 x + \theta_2 x^2$) with a different set of parameters $\theta_i \in \Theta$.

To find the optimal function $h_\theta(x)$, the gradient descent algorithm is used, the essence of which is to change the parameters θ_0, θ_1 sequentially using the expression:

$$\theta_j := \theta_j - \alpha\frac{\partial}{\partial\theta_j}J(\theta_0, \theta_1),$$

where α is the learning parameter, and $\frac{\partial}{\partial\theta_j}J(\theta_0, \theta_1)$ is the derivative of the cost function by θ_j. The sign $:=$ means assignment as opposed to the equality sign ($=$) in algebraic expressions.

Two algorithms were chosen for the experiments: gradient boosting (an ensemble of weak decision trees aggregated into a meta-model by the boosting method) and a neural network with a hidden layer.

The essence of the gradient boosting [26] is that after the optimal values of the regression coefficients are calculated and the hypothesis function $h_\theta(x)$ is obtained using algorithm (a), the error is calculated and a new function $h_{b\theta}(x)$ is selected, possibly using another algorithm (b) to minimize the error of the previous one.

$$h_\theta(x^{(i)}) + h_{b\theta}(x^{(i)}) - y^{(i)} \to min$$

In other words, we are talking about minimizing the function:

$$J_b = \sum_{i=1}^{m} L(y^{(i)}, h_\theta(x^{(i)}) + h_{b\theta}(x^{(i)}))$$

where L is an error function that considers the results of algorithms a and b. If $J_b(\theta)$ is still large, the third algorithm (c) is chosen. Often, decision trees of relatively small depth are used as algorithms (a), (b), (c), etc. The value of the gradient $L(y^{(i)}, h_\theta(x^{(i)}))_{i=1}^m$ is used to determine the minimum of the function $J_b(\theta)$. Given that the minimization of the function $J_b(h_{b\theta}(x^{(i)}))_{i=1}^m$ is achieved in the direction of the antigradient of the error function, algorithm (b) is adjusted such that the target values are not $(y^{(i)})_{i=1}^m$, but the antigradient $(-L'(y^{(i)}, h_\theta(x^{(i)}))_{i=1}^m$), that is, when training the algorithm (b) pairs $(x^{(i)}, -L'(y^{(i)}, h_\theta(x^{(i)}))$ are used instead of $(x^{(i)}, y^{(i)})$.

Multilayer artificial neural networks (multilayer perseptrons) are one of the most popular methods of supervised learning, especially in the case of multiple classes. To adjust the weights θ of the neural network (network training), a cost function of the following form is used:

$$J(\Theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y_k^{(i)} log(h_\Theta\left(x^{(i)}\right))_k + \left(1 - y_k^{(i)}\right) log(1 - h_\Theta\left(x^{(i)}\right))_k\right] + \frac{\lambda}{2m} \times \sum_{l=1}^{L-1}\sum_{i=1}^{S_l}\sum_{j=1}^{S_{l+1}} (\Theta_{ji}^l)^2,$$

where L is the number of layers of the neural network, s_l- number of neurons in layer l, K is the number of classes (equal to the number of neurons in the output layer), Θ is the weight matrix, and the hypothesis function is often a sigmoid (logistic) function.

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}},$$

To minimize the loss function (learning) of a multilayer ANN, the backpropagation error (BPE) algorithm [61] and its modifications are used to speed up the learning process.

The quality of the constructed regression dependence was assessed using a list of indices.

Coefficient of determination

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}},$$

$$SS_{res} = \sum_{i=1}^{n} (y^{(i)} - h^{(i)})^2$$

$$SS_{tot} = \sum_{i=1}^{n} (y^{(i)} - \bar{y})^2, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y^{(i)},$$

where $y^{(i)}$- actual value for the i-th sample;

$h^{(i)}$– calculated (predicted) value (hypothesis function value) for the *ith* sample of total n samples.

In existing libraries R^2 is denoted by r2_score. The best value of r2_score=1.

Root mean square error

$$RMSE = \sqrt[2]{\frac{1}{n}\sum\nolimits_{i=1}^{n} (y^{(i)} - h^{(i)})^2}.$$

Linear correlation coefficient (or Pearson correlation coefficient)

$$R(y, h) = \frac{\sum_{i=1}^{n} (h_i - \bar{h})(y_i - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (h_i - \bar{y})^2},$$

where

$$\bar{h} = \frac{1}{n}\sum\nolimits_{i=1}^{n} h_i.$$

The methodological scheme of the study consists of the following steps:

- Data collection and preprocessing. This step is necessary to form a set of input variables and to select the target variable.
- Application of machine learning methods in two experiments.

a) Experiment 1: ANN-based regression model based on data from exploratory wells of the Budennovskoye field.

b) Experiment 2. Regression models based on ANN and Extreme Gradient Boosting (XGBoost) use data from the Inkai field.

- Verification of results using RMSE, R^2, R.

V. DATA AND RESULTS

To increase the reliability of K_f estimation, it is desirable to consider as much data as possible. However, during the exploration phase, a limited set of geophysical surveys are usually performed that do not include IL or neutron methods. Therefore, the models had to be limited to rock code and AR and SP log data (as a set of values in 0.1 m increments).

Experiment 1 (Calculation of Filtration Coefficients of ≪Budennovskoye≫ Field):

To train the model, a dataset was generated containing data for ≪Budennovskoye≫ field, part of which is shown in Fig. 4. (AR and SP are given for 90 centimeter intervals, for which, in turn, the actual values K_{fpo}. obtained by pumping out (pump out) was determined. As a result, the input variable set consisted of 19 values, including the rock code (AR, SP). The target column is K_{fpo}. The full dataset with the obtained calculation results is available in the following link [62].

The regression model was based on an ANN with one hidden layer consisting of 31 neurons. K_f values were also

Well number	AR1	AR2	AR3	AR4	AR5	AR6	AR7	AR8	AR9	SP1	SP2	SP3	SP4	SP5	SP6	SP7	SP8	SP9	Lithological code	K_{fpo}	K_{fc}	K_{fr}
1001	14	12,4	11,6	11,2	10,7	10,3	10,1	10,1	10,1	62,9	60,6	59,9	60,2	59,9	58,4	57,7	57,5	57,2	3	1,766	8,3	9,38
1001	10,8	10,8	11	11,1	11,5	11,3	10,8	10,8	10,6	52,3	51,3	51	51,5	52	52,9	52	48,7	45,6	123	15,705	8,09	10,96
1001	9,6	10,1	11,6	12,7	13,3	13,5	13,7	13,4	13,1	58	58,7	56,8	55,4	55,1	55,4	56,2	57,1	57,4	123	17,488	9,29	13,86
1001	12,8	13	12,6	11,9	12,8	14	13,6	12,7	12,3	57,7	57,9	58,8	59,3	60,5	64,1	60,6	58,2	57,8	123	17,488	9,6	12,66
1001	12,1	12,1	11,5	12	12,8	12,8	12,5	12,2	13	57,8	56,8	55,2	55,3	55,1	54,4	54,1	54,8	56,4	123	17,488	9,29	12,53
1001	10,4	10,7	10,9	11,1	10,7	10,4	10,1	10,3	10,6	60,1	58,4	58	58,1	58,4	57,8	57,4	58,6	60,5	3	15,769	7,68	10,57
1001	10,9	10,8	11	10,5	10,3	11,3	12,3	12,5	12,5	61,3	60,5	59,1	57,9	57	53,6	51	49,9	47,7	3	15,769	8,48	11,95
1002	13,9	14,5	14	13,7	14,7	15,9	16,4	18,6	21,4	48,1	48,4	47,9	46,7	46,5	47,8	49,6	53,9	63,7	12	30,151	11,51	18,45
1002	10,2	9,8	10,1	10	10,3	10,4	10,6	10,4	10,3	45	44,9	44,5	44,2	43,5	42,7	42,6	44,7	46,1	3	15	7,32	10,29
1002	7,9	8,9	9,9	9,8	10	10,3	10,3	9,7	9,2	63,1	57,4	49,8	45,1	43,7	43,9	43,3	41,7	39,1	123	15,386	6,62	10,33
1002	12,8	11,3	10,1	9,6	9,8	9,7	10,1	10,7	10,7	48,2	48,3	49,5	52,3	56,9	57,9	57,6	56	53,7	4	15,919	7,63	9,47
1002	9,7	9,5	9,5	9,2	9,9	9,8	10,2	9,9	9,8	45,9	46,5	46,2	46	46,3	46,1	45,4	45,1	45,6	3	14,728	6,79	9,82
1002	13,7	14,3	14,6	14,2	14,1	14,1	14,1	14,5	14,7	39,3	41,4	41,8	40,4	39,4	40,7	42,4	44	46,8	12	20,317	10,43	14,45
10003	12	12,1	14,3	15	16,2	16,2	17,4	18,2	18,9	59,8	61,6	61,1	59,8	58,6	58,6	59,4	60,5	61,8	122	1,886	11,28	18,26
10003	18,1	15,4	13,2	12,2	9,3	8,9	9,1	8,9	8	60,2	64,1	67	69,7	74,5	73,9	68,2	63,4	63,2	3	2,815	8,6	7,05

FIGURE 4. Dataset for ≪Budennovskoye≫ field.

TABLE 3. Evaluation of the quality of K_f calculation by the existing method (calculation) and by regression model (regression).

	RMSE	R
Calculation	6,95	0,261
Regression	5,69	0,46

calculated for all intervals of the specified dataset using the currently used procedure K_{fc},.

The results showed that neither method fully agrees with the actual data (K_{fpo}). However, the correlation of the results of the regression model - K_{fr} with the actual data - K_{fpo} is significantly higher than the correlation between K_{fpo} and K_{fc}. Accordingly, the RMSE value of the regression model was lower (Table 3).

Experiment 2 Calculation of Filtration Coefficients of ≪Inkai≫ Field:

The experiment for the ≪Inkai≫ field was designed so that the models were trained and verified on the data of exploration wells as in Experiment 1. Then, the best model according to R^2 and RMSE estimates was used to calculate K_f of technological wells. The calculation results for the technological wells were compared with the debits of the wells because there was no actual data (K_{fpo}) for the technological wells.

A much larger dataset was used for training, containing approximately 600 intervals for more than 30 exploration wells. Table 4 shows the results of XGBOOST, ANN (hidden_layer_sizes = 91), Support Vector Regressor (SVR) and random forest regressor (RFR) for different input datasets. The datasets differed in terms of input parameters. Below, in Table 4 and Fig.5:

Input - set of input variables,

AR - set of input variables consisting of AR values for the interval in consideration,

SP - set of input variables consisting of SP values,

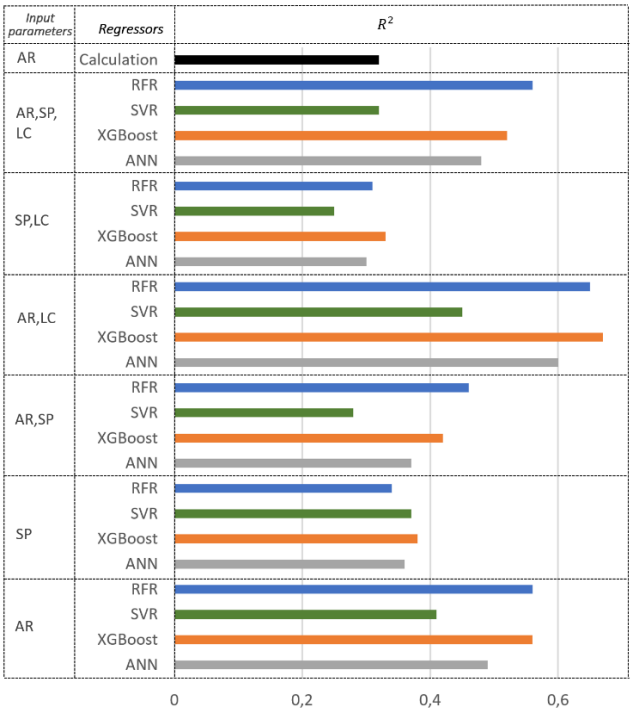

FIGURE 5. Comparative assessments of regressors.

LC - lithological code set by the expert.

It can be seen that the best results were obtained using AR and LC as the input parameters. The worst RMSE results were obtained using the calculations. The model based on the SP showed a weak correlation. It can be assumed that this is due to the poor quality of SP curve recording, as high-quality recording requires strict requirements for drilling fluid preparation, which are often neglected in practice. Because the size of the dataset was small, the training time for

TABLE 4. Assessments of the performance of the models trained on the exploration wells of the ≪Inkai≫ field.

Input	Regressor	R	RSME	R^2	Training time, sec
AR	ANN	0,7464	5,09	0,49	0,46
	XGBoost	0,7810	4,74	0,56	0,06
	SVR	0,7047	5,46	0,41	0,40
	RFR	0,7809	4,74	0,56	0,03
SP	ANN	-0,2454	8,48	0,36	0,56
	XGBoost	-0,2892	8,06	0,38	0,07
	SVR	-0,2769	9,59	0,37	0,33
	RFR	-0,2566	8,16	0,34	0,03
AR, SP	ANN	0,6725	5,67	0,37	0,59
	XGBoost	0,7046	5,42	0,42	0,08
	SVR	0,4599	6,59	0,28	0,37
	RFR	0,7201	5,23	0,46	0,05
AR, LC	ANN	0,7994	4,49	0,60	0,63
	XGBoost	**0,8401**	**4,09**	**0,67**	**0,06**
	SVR	0,7208	5,25	0,45	0,51
	RFR	0,8298	4,21	0,65	0,04
SP,LC	ANN	-0,3120	7,37	0,30	0,58
	XGBoost	0,6173	5,82	0,33	0,08
	SVR	0,4353	7,18	0,25	0,49
	RFR	0,6116	5,90	0,31	0,03
AR,SP,LC	ANN	0,7128	5,13	0,48	0,64
	XGBoost	0,7582	4,90	0,52	0,10
	SVR	0,4927	6,42	0,32	0,33
	RFR	0,7774	4,71	0,56	0,04
AR	Calculation	0,5839	13,89	0,32	

all algorithms was less than 1 s. At the same time, methods based on the use of decision trees learn faster.

Because the XGBoost regressor showed the best results when using (AR, LC) as input parameters (Fig. 5), but for acidified intervals, the values of both AR and LC are incorrect, we propose a hybrid model. In this case, XGBoost (AR, LC) is used for all non-acidified intervals, and XGBoost (SP) for acidified intervals, as only SP preserves true values on acidified intervals. Nevertheless, it is worth noting that because of the poor quality of the SP record, the error in estimating K_f of acidified intervals must be quite high.

The XGBoost-based models were tested on eight technological wells that contained technological acidification intervals. The calculation results for one of the wells are presented in Table 5. The last column shows the results of the hybrid model, in which multiple input variables (AR, SP) were used at all intervals except the acidified intervals, and the SP-based model on the acidified intervals, because the AR and rock code on the acidified intervals are distorted.

The technological acidification interval is highlighted in yellow (rock code 28). It can be seen that for this interval the K_f, values calculated by the existing methodology are significantly underestimated (1.1 m/day).

It is worth noting that this value (1.1 m/day) is actually not the result of calculation, but was forcibly set for these intervals as the minimum possible value for permeable intervals.

Because it is difficult to obtain actual values of filtration coefficients at technological wells, it is possible to estimate the correctness of the calculated K_f only indirectly by comparing it with the well flow rates (maximum volume of injected/outflow fluid). It is logical to assume that the higher the well flow rate, the better the filtration properties of rocks in the near-filter zone and a longer filter length. However, for calculations as a zone of solution movement in Kazakhstan, it is accepted to use not the filter itself, but the so-called zone of active movement of solutions (ZAMS). It extends 2 m upward and 6 m downward relative to the actual filter location. Therefore, the product of the average K_f value within the filter ($\bar{K}_f$) by the filter length (length of filter - LF) and the product of the average K_f value within ZAMS by the length of ZAMS (length of ZAMS -LZAMS) was also used for comparison with well flow rates. Part of the calculation results for the process wells is listed in Table 6. The calculation results for all 46 wells are given in Appendix E (https://www.dropbox.com/s/a8rtrgzohykcrob/Appendix_E.pdf?dl=0).

Column $\bar{K}_{fr}$ shows the results of the XGBoost-based regression model using the set (AR, LC) as input parameters. Column $\bar{K}_{frh}$ shows the results of a hybrid model, which also uses an XGBoost-based regression model, but uses only SP data for areas containing acidic rocks.

Wells 42, 43, 45, and 46 contain zones of technological acidification. It can be seen that the assessment of filtration properties, when calculated according to the current instructions for them, is significantly underestimated, which is clearly visible in the example of well number 45. Overall, the assessment of filtration properties, obtained by means of a hybrid model, correlates significantly better with actual values of well flow rate after development (R = 0.550) in comparison with the calculation based on the existing method (R = 0.164), even if the acidulated wells were considered (the last row of the table).

VI. DISCUSSION

According to the current methodology, the parameters of the rock filtration properties and the actual value of the filtration coefficient (K_f) were identified at the exploration drilling stage. Subsequently, the obtained parameters were used to calculate the filtration properties of the technological wells. Correct calculation of K_f affects the estimation of recoverable reserves and parameters of the production process. However,

TABLE 5. Results of K_f calculating using the existing method and using regression models.

Interval (from) м	Interval (to), м	AR, Ом*м	SP, mV	LC	K_{fc}	K_{fr} , AR	K_{fr} , SP	K_{fr}, AR, SP	K_{fr} AR, LC, SP
337,10	337,60	12,78	16,00	12	25,00	7,84	16,14	8,88	7,62
337,80	338,30	12,95	20,20	123	17,00	5,80	16,14	9,46	5,03
338,30	338,80	10,35	17,50	123	17,00	5,66	16,14	8,48	3,78
338,80	339,30	10,38	18,40	123	17,00	6,00	16,14	8,69	5,04
339,30	339,80	11,31	20,50	123	17,00	6,93	16,14	8,64	6,29
339,80	340,30	11,19	21,30	123	17,00	5,53	16,14	8,48	3,29
340,30	340,80	9,38	22,20	123	17,00	5,53	15,39	7,74	3,59
340,80	341,30	10,02	26,40	7	0,92	5,53	16,37	5,68	3,59
341,30	341,80	9,35	41,10	7	0,92	6,06	17,71	5,84	2,42
342,20	342,60	6,61	54,40	28	1,10	8,07	13,82	5,80	13,82
342,60	343,10	6,68	34,60	28	1,10	2,95	15,12	6,29	15,12
343,10	343,60	5,17	31,90	28	1,10	2,12	15,12	6,07	15,12
343,60	344,10	3,50	32,30	28	1,10	2,38	15,61	5,89	15,61
344,10	344,60	3,03	35,80	28	1,10	2,38	14,93	6,07	14,93
344,60	345,10	2,96	29,20	28	1,10	2,38	12,34	5,21	12,34
345,10	345,60	3,76	30,50	28	1,10	2,38	14,91	5,67	14,91
345,60	346,10	3,43	41,60	28	1,10	2,38	14,04	2,86	14,04
346,10	346,60	4,16	38,90	28	1,10	2,24	17,78	5,32	17,78
346,60	347,10	5,33	38,80	28	1,10	2,12	17,90	2,93	17,90
347,10	347,60	4,36	41,00	28	1,10	2,38	13,77	5,19	13,77
348,20	348,50	3,60	39,30	3	10,24	5,48	13,72	7,41	4,68
348,50	349,00	8,49	39,70	3	10,24	6,93	12,26	4,16	6,17

TABLE 6. Results of applying machine learning models to calculate K_f for technological wells.

Well number	Length of filter	ZAMS	Well flow rate m^3/h	$\bar{K}_{fc}$	$\bar{K}_{fc}$ * LF	$\bar{K}_{fc}$ * LZAMS	$\bar{K}_{frh}$	$\bar{K}_{frh}$ * LF	$\bar{K}_{frh}$ * LZAMS	$\bar{K}_{fr}$	$\bar{K}_{fr}$ * LF	$\bar{K}_{fr}$ * LZAMS
1	8,00	16,00	21,1	20,0	159,9	319,8	15,8	126,2	252,3	15,8	126,2	252,3
2	7,00	15,00	23,0	17,6	123,1	263,9	15,4	107,5	230,3	15,4	107,5	230,3
3	5,00	13,00	19,4	17,8	89,1	231,7	15,8	79,0	205,3	15,8	79,0	205,3
4	7,00	15,00	34,6	18,7	130,7	280,1	17,7	123,8	265,2	17,7	123,8	265,2
…												
22	6,00	14,00	24,0	20,8	124,8	291,2	19,1	114,6	267,4	19,1	114,6	267,4
23	5,00	13,00	20,0	18,9	94,7	246,1	20,2	101,1	262,9	20,2	101,1	262,9
24	6,00	14,00	24,4	18,4	110,1	256,9	18,3	110,0	256,8	18,3	110,0	256,8
25	6,00	14,00	20,5	20,1	120,7	281,5	19,5	116,7	272,3	19,5	116,7	272,3
…												
40	6,00	14,00	21,8	16,1	96,4	224,8	20,2	121,0	282,4	20,2	121,0	282,4
41	6,00	14,00	22,0	28,7	172,2	401,8	16,9	101,1	235,9	16,9	101,1	235,9
42	7,00	15,00	21,0	18,9	132,0	282,8	16,2	113,2	242,6	9,6	67,1	143,7
43	7,00	15,00	24,0	11,4	80,0	171,5	18,3	128,4	275,1	16,9	118,0	252,8
44	8,00	16,00	24,0	29,6	237,0	473,9	17,7	141,4	282,9	17,7	141,6	283,2
45	6,00	14,00	44,0	7,3	43,6	101,8	23,4	140,2	327,0	8,8	52,8	123,2
46	8,00	16,00	31,0	19,6	156,4	312,8	24,1	192,7	385,4	11,7	93,6	187,2
Correlation with well flow rates				**-0,196**	**0,164**	**0,004**	**0,418**	**0,550**	**0,553**	**-0,092**	**0,301**	**0,145**
Correlation with well flow rate (excluding acidified wells)				**0,053**	**0,436**	**0,310**	**0,247**	**0,556**	**0,500**	**0,247**	**0,556**	**0,500**

K_f estimation is inaccurate. A comparison of the calculated data with actual data shows that the RMSE is 13.89 and the linear correlation value is 0.584 (see Table 7). In addition, the calculated values correlated poorly with the debits of the wells (R = 0.164).

Based on the results of the experiments, a two-stage scheme for determining filtration coefficients in the fields of Kazakhstan using machine-learning models was proposed (Fig. 6).

In the first stage, machine learning models were tuned using data from exploratory wells. A hybrid model is formed from the tuned models, which use the ML model for acidified well sections, where the input data are SPs (XGBoost(SP)). For non-acidified sections, the AR and LC data were used (XGBoost(AR,LC)).

Because technogenic acidification intervals are found only in production wells, they are not present in the dataset generated from exploration wells. Therefore, it is impossible to directly teach the correct predictions K_f for acidified intervals. Because of the poor quality of SP curve recording, adding it as an input regression parameter usually slightly worsens the accuracy.

However, in acidified well sections, the AR curve is too distorted, and the lithological code only indicates the acidification interval (not the actual rock type). Moreover, this distortion is dependent on the lithological composition of

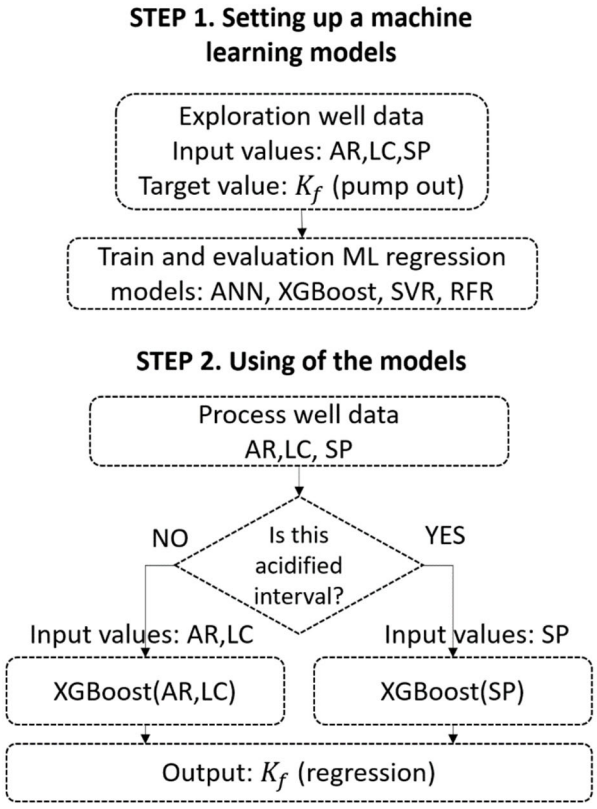

FIGURE 6. Scheme of K_f estimating Kazakhstan ISL fields using machine-learning models.

the rocks and the amount of acid. Therefore, the only option in this case is to use the SP, even though it has low accuracy.

The proposed model for determining the filtration coefficient is not only much more accurate (RMSE = 4.89, R^2 = 0.59), but also correlates much better with the actual well flow rates (R = 0.550) (Table 7).

It can be noted that when using the approved methodology on data from exploration wells (Table 4), very poor results (RSME = 13.89) were obtained, which were significantly inferior to the regression models. At the same time, when using it on technological wells (Table 6), the correlation with well flow rates, if not considering acidified wells, is not much inferior to the results of the regression models. This is due to the fact that the existing methodology is designed to use the average value of resistivity within the allocated lithological interval, and it can be correctly determined only for intervals of at least 1.5 - 2m, as in the fields of Kazakhstan for recording AR used downhole device with a distance between the electrodes in 1m. Because the borehole device was used to record AR in the fields of Kazakhstan, the distance between the electrodes was 1.1 m.

During the training on data from exploration wells, only intervals with a thickness of 0.5 m were used for comparison with data from hydrogeological studies, which led to a high RMSE value. In the technological well data, lithologic intervals were mainly more than 2 m thick, so the result of the approved methodology was relatively good (in

TABLE 7. Comparison of machine learning models and the existing methodology for determining K_f in the ≪Inkai≫ field.

	RMSE	R	R^2	Correlation with well flow rates
Calculation	13,89	0,584	0,32	0,164
Hybrid model	4,89	0,81	0,59	0,550

wells without acidified intervals, almost comparable with regression models). Based on this, we can draw the following conclusions.

1. Regression models work well for all intervals, while the current methodology is only suitable for intervals greater than 1.5-2m.

2. The current methodology is not applicable to wells containing acidified intervals.

3. Hybrid can be applied to wells containing acidified intervals.

VII. CONCLUSION

Uranium mining by in situ leaching requires a fairly accurate assessment of the lithological composition and filtration properties of ore-bearing rocks. The methodology used in Kazakhstan to estimate filtration properties is based on the fact that at the stage of exploratory drilling, the parameters are determined, which are then used to calculate the filtration properties of technological wells. However, the existing methodology yields inaccurate results and cannot be used in technological acidification zones, which account for up to 40% of all considered data. Inaccuracies in determining the filtration coefficient lead to errors in the technological production process and inaccurate calculation of recoverable reserves.

To overcome the shortcomings of the existing approach, we propose a method for calculating the filtration coefficient based on the use of regression models. The proposed model receives electric logging data as an input and the calculated filtration coefficient as an output. To improve the quality of the model, it is made hybrid; that is, it is formed from two models. For non-acidic areas, a model with AR and LC as the input variables was used. For acidified sites, a model with input variables consisting of SP data was used.

The analysis shows that the proposed method yields a much smaller mean square error of filtration coefficient determination, correlates better (by 70%) with well debits, with actual filtration coefficient values (by 27%) applicable for small intervals, and can also be used for calculation of the filtration coefficient in acidified zones.

A. LIMITATION OF THE METHOD

Application of the method may be limited when analyzing data from fields explored 20-30 years ago. In such cases, it is not possible to obtain hydrogeological data, and the logging data are often fragmentary and unsuitable for the formation of a training dataset.

B. FUTURE RESEARCH

In future studies, we plan to analyze the possibility of applying pre-trained models for such cases. In other words, it is planned to investigate the possibility of using transfer learning methods to calculate the filtration coefficient of technological wells in fields where data from exploration wells are incomplete.

The second direction of research is to develop methods for interpolation of K_f in the interwell space, which can improve the accuracy of estimating K_f using data from nearby wells.

APPENDIX A. A BRIEF OVERVIEW OF THE RESERVES AND PRODUCTION OF NATURAL URANIUM IN THE WORLD

All over the world, uranium is the main resource for the operation of nuclear power plants. Deposits of uranium ores are not evenly distributed around the globe. Today, only 28 countries of the world extract valuable raw materials in their bowels. The main world reserves of uranium in the world are located in 10 countries. We will tell you a little more about the countries with the largest uranium reserves.

Table 8 below contains a summary of the uranium reserves, the number and names of deposits in the leading countries in terms of uranium reserves according to [1], [2].

Analysis of information related to the leadership of countries in uranium reserves does not allow us to conclude that these same countries are leaders in uranium mining.

In 2018, the world's largest uranium miners produced 86% of the world's uranium mined, according to the World Nuclear Association. The main uranium mining companies are mining corporations from Kazakhstan, Canada, Australia: they account for two-thirds of the world's production [2] (Fig. 7)

Table 9 shows the top ten uranium mines based on 2018 production results. At least three of these top ten mines (Rössing, Arlit (SOMAÏR) and Ranger), representing 10% of 2018 production, are scheduled/expected to close before the end of the 2020s and will need to be replaced by new mine capacity by then, in order not to cause further reduction of primary uranium production [3].

During the development of the uranium market, the technology of mining processing itself has changed more than once. Basically, uranium ore is mined in two ways - mine or open pit, depending on the depth of the layers with uranium ore. The career path means less radiation and higher safety. Underground (mine) allows you to extract higher quality uranium ore, but it is also more dangerous because of radon - a radioactive gas that accumulates in mines.

Underground leaching of uranium ores is the most advanced uranium mining technology, first used since 1957. The method is the injection of a special chemical solvent underground into the layer of uranium ores, which reacts to uranium compounds. Then this solution is brought to the surface and processed already. The disadvantage of this method is the ability to use it only in sandstone and below the groundwater level.

The method has gained particular popularity in Kazakhstan, Uzbekistan and the United States, although this method of production is used in Canada, Australia and China. The geography of the in-situ leaching method is steadily increasing its share of the total volume, mainly due to Kazakhstan (this method covers more than half of the production). Since 2015, global production has proceeded as follows:

Conventional mines had a mill where the ore was crushed, crushed, and then leached with sulfuric acid to dissolve uranium oxides. In a conventional mine mill or sewage treatment plant with ISL operation, the uranium is then separated by ion exchange before drying and packaging, usually as uranium oxide (U3O8). Some mills and ISL operations (especially in the US) use carbonate leaching instead of sulfuric acid, depending on the orebody.

Today, the approximate distribution of uranium mining methods is as follows [3] (Table 10). According to the table, the method of underground leaching of uranium ores is used approximately equally along with two methods – mine or quarry. The extraction of uranium using underground leaching causes significantly less damage to the environment than the methods described above.

Over time, reclamation processes occur on the developed land plot. The use of this method can reduce economic costs. But it has its limitations. It is not used only in sandstone and below the water table.

Almost all the uranium mining companies listed above have all the existing mining technologies, depending on the characteristics of specific deposits and mines.

APPENDIX B. PHYSICAL ASPECTS OF LOGGING DATA ACQUISITION

Measurements of apparent resistivity (AR) method are performed using a four-electrode AMNB unit. Two electrodes A and B (supply electrodes are connected to the current source). M and N (measuring electrodes) are connected to the meter. The electric probe consists of three electrodes set at a strictly defined distance from each other. The fourth electrode is mounted on the surface and is called a ''fish''. Electrodes AB and MN are called paired electrodes, and electrodes AM, AN, BM and BN are unpaired.

Gradient probes and potential probes are mainly used to measure ρk of rocks in resistivity logging in the sedimentary section. Gradient probes are probes in which the distance between paired electrodes M and N (A and B) is at least 7 times less than the distance between unpaired ones. The distance from the middle of the paired electrodes close together which is called the recording point, to the unpaired electrode is called the length of gradient probe. Potential probes are probes in which the distance between unpaired electrodes is small compared to the distance between paired electrodes. The length of the potential probe is the distance between the unpaired electrodes. The essence of measurements is briefly that a current I is passed through supply electrodes of the probe located in the borehole, which

TABLE 8. Leading countries in terms of natural uranium reserves in the world.

№	Country	Information about uranium reserves, the number of deposits, etc.
1.	Australia	Australia is the undisputed leader in the world's uranium reserves. According to the World Nuclear Association, about 31% of all world uranium reserves are located in this country, which in numerical terms means 661 thousand tons of uranium. There are 19 uranium deposits in Australia. The largest and most famous are Olympic Dam, where about 3 thousand tons of uranium is mined per year, Beaverley (production of 1 thousand tons) and Honemun (900 tons per year). The cost of uranium mining in the country is $ 40 per 1 kg.
2.	Kazakhstan	The second place in uranium reserves belongs to Kazakhstan. The Asian country contains 11.81% of the world's fuel reserves, which equals 629 thousand tons of uranium. There are 16 developed deposits in Kazakhstan, where a valuable resource is extracted. The Chusaray and Syrdarya uranium provinces are home to the largest Korsan, South Inkai, Irkol, Kharasan, West Mynkuduk and Budenovskoe deposits.
3.	Russia	Russia ranks third in uranium reserves. According to experts, there are 487,200 tons of uranium in its depths, which is 9.15% of the world's uranium resources. Despite the size of the country and large reserves of uranium, there are only 7 deposits in Russia, and almost all of them are located in Transbaikalia. More than 90% of the uranium mined in the country comes from the Chita region. This is the Streltsovskoe ore field, which includes more than ten uranium ore deposits. The largest center is the city of Krasnokamensk. The remaining 5-8% of uranium in the country is located in Buryatia and the Kurgan region.
4.	Canada	he leading place in the reserves of uranium ore in North America, and the fourth in the world belongs to Canada. The total uranium reserves in the country are 468,700 tons of uranium, which is 8.80% of the world's reserves. Canada owns unique deposits of the "unconformity" type, the ores of which are rich and compact, the largest of which are MacArthur River and Cigar Lake. The country is developing the Waterbury Project uranium deposit, which consists of several deposits, the area of which is 12,417 hectares. Canada has enjoyed tremendous advantages throughout its history due to its proximity to the United States. The main uranium mining company in Canada is Cameco.
5.	South Africa	South Africa, uranium is mined along the way from gold deposits. The Dominion deposit is the largest in the country with open-pit and underground mining. Large mines include Western Ariez, Palabora, Randfontein and Vaal River, where gold tailings are mainly mined. The average cost of uranium mining in an African country is $ 40 per kg. In uranium production, South Africa lags far behind the leading countries in this industry, producing 540 tons of uranium a year, which is the twelfth figure in the world. According to some estimates, South Africa has 6% of the world's total uranium reserves. However, other sources claim that South Africa has less reserves than Niger and Namibia. The main problems in the country's economy are unemployment, high levels of poverty and inequality. The country is better known for mining gold, platinum and chromium, rather than uranium. There are I nuclear power plants in South Africa, but there are plans to build several more nuclear power plants. Thus, South Africa could become a potentially large market for the use of uranium.
6.	Niger	Uranium reserves account for 5% of the world's total. The largest deposits in the country are Imuraren, Madauela, Arlit and Azelit, there are 12 of them in total in the country. The prime cost of uranium mined in Niger is $ 34-50 per 1 kg. The main player in the uranium market of the country is the French company Areva SA, which is mining at the Arlit deposit, one of the 10 largest uranium deposits in the world. In addition, uranium is Niger's largest export. According to Areva, uranium accounts for about 5% of the country's GDP. At the same time, Niger is a rather poor country and depends on foreign investment in the extraction of natural resources.
7.	Brazil	The country's indicator is equal to 276,700 tons of uranium ore. Production of U per year is 198 tons per year.
8.	Namibia	Uranium reserves in the country are 261 thousand tons. There are four major uranium deposits in Namibia. Namibia's reserves are believed to account for 5% of the world's total. The economy of Namibia, as in the case of Niger, is rather poor, but it is more diversified than in Niger. The country exports diamonds, copper, gold, zinc as well as uranium. In general, mining accounts for 11.5% of GDP.
9.	USA	The total reserves of U in the USA are 207,000 tons.
10.	China	China possesses approximately 5% of the world's uranium reserves. At the same time, various sources estimate the reserves in different I: some indicate that China's reserves are slightly higher than in Namibia and Niger, some put China only in eighth place in terms of reserves. The country has about 20 nuclear power plants under construction.

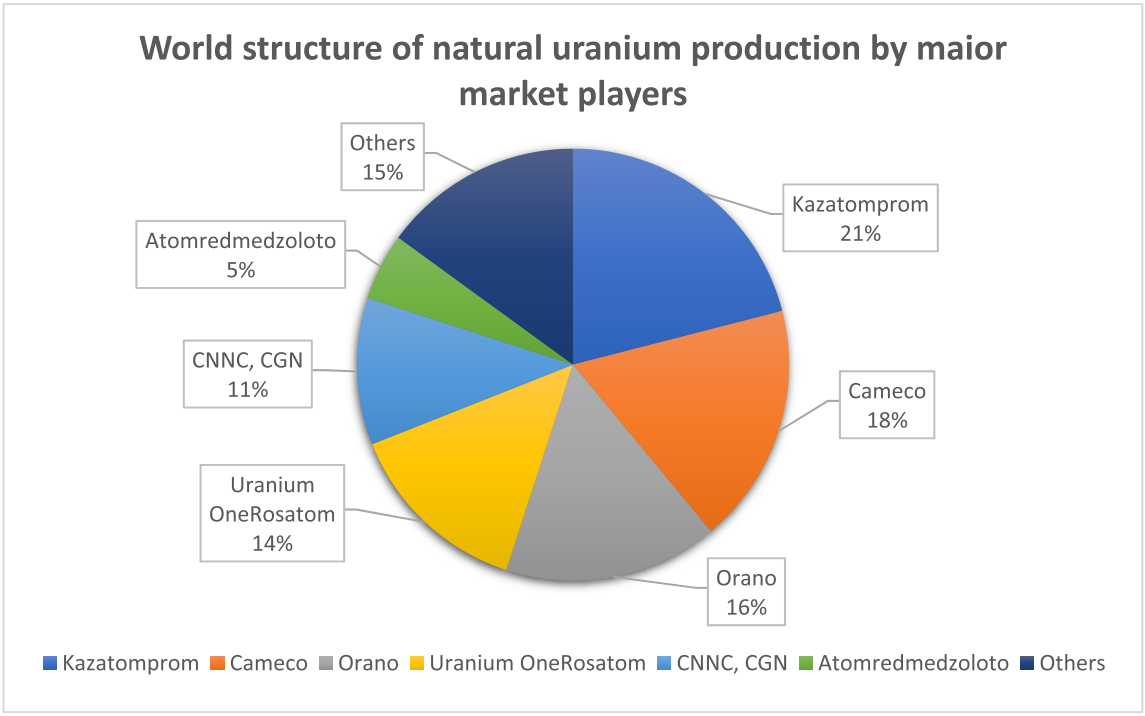

FIGURE 7. World natural uranium production structure by major market players.

TABLE 9. Ten largest world uranium mines, Ranked by 2018 production, TU.

Main	Country	Main owner	Type	tU	% of world
Cigar Lake	Canada	Cameco/Orano	Underground	6,924	12.9
Olympic Dam	Australia	BHP Billiton	By-product	3,159	5.9
Husab	Namibia	Swakop Uranium (CGN)	Open-pit	3,028	5.7
Inkai 1-3	Kazakhstan	Kazatomprom/ Cameco	ISL	2,643	4.9
Rössing	Namibia	Rio Tinto	Open-pit	2,102	3.9
Budenovskoye 2	Kazakhstan	Uranium One/ Kazatomprom	ISL	2,081	3.9
Tortkuduk	Kazakhstan	Orano/Kazatomprom	ISL	1,900	3.6
Arlit (SOMAÏR)	Niger	Orano	Open-pit	1,783	3.3
Ranger	Australia	Rio Tinto/ERA	Open-pit	1,695	3.2
Kharasan 2	Kazakhstan	Kazatomprom	ISL	1,631	3,0
Others				26,553	49.6
Total				**53,498**	**100**

TABLE 10. Evaluation of uranium mining methods.

№	Uranium mining methods	t, U	Specific weight, %
1	In-situ leaching (ISL)	29,197	48,3%
2	Underground and quarrying (except Olympic Dam) *	27,791	45,9%
3	By-product	3,509	5,8%
	Total:	**60,497**	**100%**

creates an electric field in the medium under study. Using measuring electrodes M and N, the potential difference ΔU between two points of this electric field is measured.

Consequently, the resistivity, which is called apparent resistivity, is equal to:

$$\rho_K = K\frac{\Delta U}{I} \tag{1}$$

where ΔU is the potential difference between the measuring electrodes, I is the current supplying the probe, K is the probe coefficient (a constant value depending on the distance between the electrodes). The probe coefficient K is calculated by the formula:

$$K = \frac{4\pi * MA * MB}{AB} \tag{2}$$

where MA, MB and AB are distances between electrodes.

The spontaneous polarization potential (SP) measurement is reduced to the measurement of the natural potential difference between the M electrode moving along the

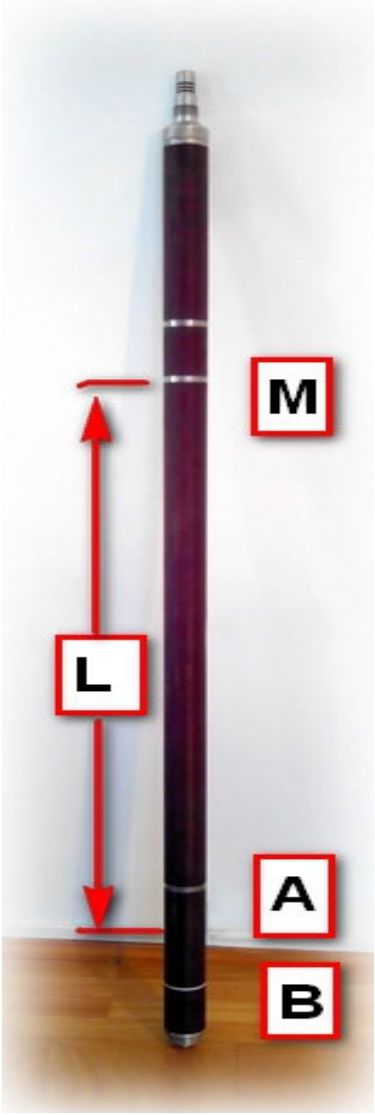

FIGURE 8. A plantar gradient probe.

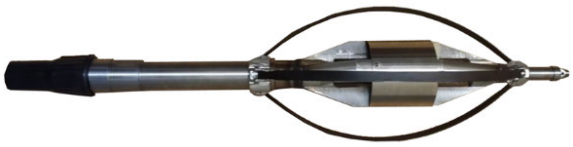

FIGURE 9. Flow meter.

borehole and the N electrode located on the surface near the borehole head.

We used a model of a plantar gradient probe, with a probe length of 1m, which is mainly used for logging in the uranium deposits of Kazakhstan. Its scheme is shown in (Fig. 8). A, B, M- electrodes, electrode N is on surface, L-length of probe (in our case 1m), the point of recording of AR is between electrodes A and B, the point of recording of SP is on electrode M.

Distance between AR and SP recording points is 1m (ten 10-cm interlayers).

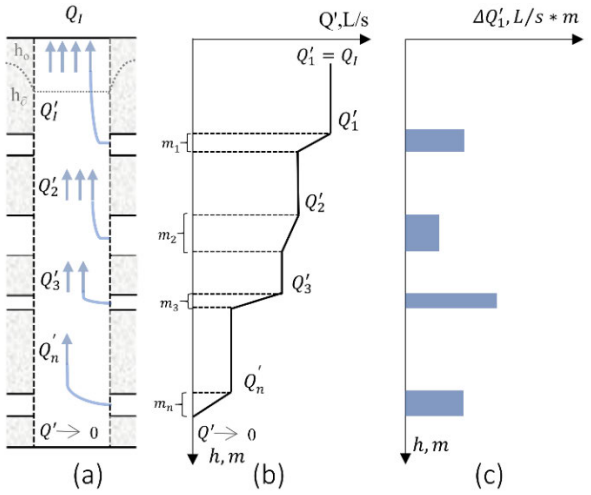

FIGURE 10. The flow-metric graph.

TABLE 11. Dependence of filtration properties on the apparent resistivity of host rocks.

Horizon	Deposit (zone)	Electrical resistance value reduced to a single (standard level)*ρκ,·м Filtration coefficient K_f, m/day	Filtration coefficient K_f, m/day
Mynkuduk	3	0,00	0,00
		4,78	0,60
		5,20	1,00
		7,40	3,50
		7,71	4,00
		9,44	6,50
		11,84	9,00
		15,20	11,00
		22,00	16,00
		22,01	0,00
		100,00	0,00

The registered value of SP depends only on the value at point M, the registered value of AR depends on the values of all the interlayers located between electrode M and the middle between electrodes A and B.

APPENDIX C. ASSESSMENT OF FILTRATION PROPERTIES OF ROCKS AT THE EXPLORATION STAGE

At the stage of exploration in uranium deposits of Kazakhstan, a set of hydrogeological studies is carried out to assess the filtration properties of the host rocks. The average filtration coefficient of the water-bearing horizon is determined by the rate of water level recovery in the well after its pumping. Then, to determine layer-by-layer filtration coefficients, flow velocity is measured with a flow meter (Fig. 9).

The method of determining layer-by-layer filtration properties of rocks in the uranium deposits of Kazakhstan is described in detail in [1]. The essence of flowmetry is that the flow rate of axial water flow measured in the wellbore

[1]Grinbaum I.I., Flow measurement of hydrogeological and engineering-geological wells, Nedra, Moscow, 1975, 271 pp.

TABLE 12. Results of applying machine learning models to calculate K_f for technological wells (full).

Well number	Length of filter	ZAMS	Well flow rate m^3/h	$\bar{K}_{fc}$	$\bar{K}_{fc}$ * LF	$\bar{K}_{fc}$ * LZAMS	$\bar{K}_{frh}$	$\bar{K}_{frh}$ * LF	$\bar{K}_{frh}$ * LZAMS	$\bar{K}_{fr}$	$\bar{K}_{fr}$ * LF	$\bar{K}_{fr}$ * LZAMS
1	8,00	16,00	**21,1**	20,0	159,9	319,8	15,8	126,2	252,3	15,8	126,2	252,3
2	7,00	15,00	**23,0**	17,6	123,1	263,9	15,4	107,5	230,3	15,4	107,5	230,3
3	5,00	13,00	**19,4**	17,8	89,1	231,7	15,8	79,0	205,3	15,8	79,0	205,3
4	7,00	15,00	**34,6**	18,7	130,7	280,1	17,7	123,8	265,2	17,7	123,8	265,2
5	6,00	14,00	**29,0**	20,5	122,8	286,4	18,2	109,4	255,2	18,2	109,4	255,2
6	8,00	16,00	**28,6**	20,8	166,3	332,6	19,1	153,0	306,1	19,1	153,0	306,1
7	8,00	16,00	**45,0**	19,1	152,4	304,8	21,1	168,4	336,8	21,1	168,4	336,8
8	8,00	16,00	**29,0**	17,7	141,8	283,7	16,9	135,4	270,7	16,9	135,4	270,7
9	8,00	16,00	**21,8**	19,0	151,7	303,4	19,4	155,4	310,7	19,4	155,4	310,7
10	6,00	14,00	**23,2**	18,7	112,4	262,4	16,8	101,0	235,6	16,8	101,0	235,6
11	7,00	15,00	**21,3**	18,9	132,2	283,4	15,5	108,6	232,7	15,5	108,6	232,7
12	6,00	14,00	**20,0**	11,4	68,6	160,2	17,1	102,3	238,7	17,1	102,3	238,7
13	6,00	14,00	**21,4**	14,7	88,4	206,2	12,1	72,5	169,3	12,1	72,5	169,3
14	4,00	12,00	**20,6**	18,2	72,9	218,8	16,8	67,1	201,4	16,8	67,1	201,4
15	5,00	13,00	**20,5**	17,2	85,9	223,3	14,3	71,4	185,6	14,3	71,4	185,6
16	6,00	14,00	**22,3**	19,9	119,3	278,5	21,4	128,5	299,7	21,4	128,5	299,7
17	6,00	14,00	**26,0**	19,0	113,9	265,9	18,4	110,6	258,2	18,4	110,6	258,2
18	8,00	16,00	**31,8**	20,8	166,2	332,5	17,4	139,5	279,0	17,4	139,5	279,0
19	6,00	14,00	**28,0**	18,2	109,2	254,8	16,8	100,5	234,5	16,8	100,5	234,5
20	8,00	16,00	**27,0**	22,2	177,2	354,4	20,5	163,7	327,4	20,5	163,7	327,4
21	6,00	14,00	**28,0**	19,0	114,0	266,0	19,1	114,6	267,4	19,1	114,6	267,4
22	6,00	14,00	**24,0**	20,8	124,8	291,2	19,1	114,6	267,4	19,1	114,6	267,4
23	5,00	13,00	**20,0**	18,9	94,7	246,1	20,2	101,1	262,9	20,2	101,1	262,9
24	6,00	14,00	**24,4**	18,4	110,1	256,9	18,3	110,0	256,8	18,3	110,0	256,8
25	6,00	14,00	**20,5**	20,1	120,7	281,5	19,5	116,7	272,3	19,5	116,7	272,3
26	7,00	15,00	**20,5**	19,2	134,1	287,4	19,7	137,9	295,5	19,7	137,9	295,5
27	5,00	13,00	**21,8**	18,1	90,5	235,3	17,1	85,7	222,8	17,1	85,7	222,8
28	5,00	13,00	**19,5**	18,4	91,9	238,9	15,5	77,5	201,5	15,5	77,5	201,5
29	6,00	14,00	**21,1**	20,0	119,7	279,3	17,2	103,4	241,2	17,2	103,4	241,2
30	5,00	13,00	**19,0**	20,9	104,5	271,6	21,1	105,4	274,0	21,1	105,4	274,0
31	4,00	12,00	**20,5**	19,9	79,4	238,2	17,8	71,1	213,2	17,8	71,1	213,2
32	6,00	14,00	**30,0**	21,8	130,9	305,5	19,6	117,4	274,0	19,6	117,4	274,0
33	8,00	16,00	**34,6**	16,5	131,9	263,8	20,0	159,8	319,5	20,0	159,8	319,5
34	6,00	14,00	**29,5**	20,8	125,0	291,8	17,3	103,7	241,9	17,3	103,7	241,9
35	4,00	12,00	**28,8**	17,5	70,1	210,2	17,9	71,8	215,3	17,9	71,8	215,3
36	3,00	11,00	**18,9**	22,2	66,5	243,7	20,4	61,2	224,5	20,4	61,2	224,5
37	6,00	14,00	**23,0**	23,6	141,3	329,7	19,0	114,1	266,3	19,0	114,1	266,3
38	4,00	12,00	**22,9**	13,9	55,6	166,8	15,8	63,0	189,1	15,8	63,0	189,1
39	6,00	14,00	**20,8**	18,2	108,9	254,1	19,4	116,4	271,6	19,4	116,4	271,6
40	6,00	14,00	**21,8**	16,1	96,4	224,8	20,2	121,0	282,4	20,2	121,0	282,4
41	6,00	14,00	**22,0**	28,7	172,2	401,8	16,9	101,1	235,9	16,9	101,1	235,9
42	7,00	15,00	**21,0**	18,9	132,0	282,8	16,2	113,2	242,6	9,6	67,1	143,7
43	7,00	15,00	**24,0**	11,4	80,0	171,5	18,3	128,4	275,1	16,9	118,0	252,8
44	8,00	16,00	**24,0**	29,6	237,0	473,9	17,7	141,4	282,9	17,7	141,6	283,2
45	6,00	14,00	**44,0**	7,3	43,6	101,8	23,4	140,2	327,0	8,8	52,8	123,2
46	8,00	16,00	**31,0**	19,6	156,4	312,8	24,1	192,7	385,4	11,7	93,6	187,2
Correlation with well flow rates				**-0,196**	**0,164**	**0,004**	**0,418**	**0,550**	**0,553**	**-0,092**	**0,301**	**0,145**
Correlation with well flow rate (excluding acidified wells)				**0,053**	**0,436**	**0,310**	**0,247**	**0,556**	**0,500**	**0,247**	**0,556**	**0,500**

in spouting, pumping, filling or injecting mode changes only in intervals of permeable (water-bearing) rocks, and within water-bearing rocks remains constant or equal to zero. As a consequence, the flow-metric graph $Q' = f(h)$, constructed based on the results of a set of water flow measurements in the experimental well, allows to determine the depth, thickness and hydrodynamic characteristics of permeable (water-bearing) formations. The boundaries of formations, which differ in their filtration properties, are fixed by the breakpoints of the flow-metric graph (Fig. 10)

Fig. 10 Scheme of the flow-metric study of the well during pumping: (a) - flow diagram of fluid flows along the wellbore; (b) - flow chart $Q' = f(h)$; (c) - differential flow chart $\Delta Q' = f(h)$; h_o - steady-state combined water level in the well; h_ϑ - dynamic water level in the well during pumping.

The water inflow rate (water absorption) of any permeable layer is determined by the difference between the flow rate of water circulating in the borehole in its top and bottom.

APPENDIX D. THE DEPENDENCE OF K_f ON ρ_K

See Table 11.

APPENDIX E. RESULTS OF APPLYING MACHINE LEARNING MODELS TO CALCULATE K_f FOR TECHNOLOGICAL WELLS

See Table 12.

REFERENCES

[1] (2018). *Uranium Reserves, Which Countries Have the Largest Reserves?* [Online]. Available: https://eenergy.media/2018/11/06/zapasy-urana-u-kakih-stran-oni-samye-bolshie

[2] World Nuclear Association. (Jun. 2020). *Recent Uranium Production, the Nuclear Fuel Report: Expanded Summary—Global Scenarios for Demand and Supply Availability 2019–2040*. Accessed: Jul. 15, 2021. [Online]. Available: https://world-nuclear.org/getmedia/b488c502-baf9-4142-8d12-42bab97593c3/nuclear-fuel-report-2019-expanded-summary-final.pdf.aspx

[3] U. K. Amirova and N. A. Uruzbaeva, "Overview of the development of the world market of Uranium," *Universum, Econ. Law, Electron. Sci. J.*, vol. 6, no. 39, pp. 1–8, Jun. 2017. Accessed: Aug. 23, 2021. [Online]. Available: https://7universum.com/ru/economy/archive/item/4802

[4] R. I. Mukhamediev, Y. I. Kuchin, K. O. Yakunin, E. L. Mukhamedieva, and S. V. Kostarev, "Preliminary results of the assessment of lithological classifiers for uranium deposits of the infiltration type," *Cloud Sci.*, vol. 7, no. 2, pp. 258–272, Jun. 2020.

[5] Energoizdat, Moscow, Russia. (1981). *Guidelines for Determining the Coefficient of Filtration of Water-Bearing Rocks by Experimental Pumping*. Accessed: Aug. 15, 2021. [Online]. Available: https://www.geokniga.org/books/17383

[6] (2018). *Physical Basis for Determining the Lithological and Filtration Properties of Rocks of the Productive Horizon*. [Online]. Available: https://ozlib.com/832945/tehnika/fizicheskie_osnovy_metoda

[7] *Correction Factors and Dependencies in the Interpretation of Geophysical Well Survey Data on Hydrogenous Uranium Deposits*, Standard Kazatomprom ST NAC 15.4-2017, 2017.

[8] Y. I. Kuchin, R. I. Mukhamediev, and K. O. Yakunin, "One method of generating synthetic data to assess the upper limit of machine learning algorithms performance," *Cogent Eng.*, vol. 7, no. 1, Feb. 2020, Art. no. 1718821, doi: 10.1080/23311916.2020.1718821.

[9] G. Nguyen, S. Dlugolinsky, M. Bobák, V. Tran, Á. L. García, I. Heredia, P. Malík, and L. Hluchý, "Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77–124, Jan. 2019.

[10] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019.

[11] R. I. Mukhamediev, A. Symagulov, Y. Kuchin, K. Yakunin, and M. Yelis, "From classical machine learning to deep neural networks: A simplified scientometric review," *Appl. Sci.*, vol. 11, no. 12, p. 5541, Jun. 2021.

[12] T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," in *The Elements of Statistical Learning* (Springer Series in Statistics). New York, NY, USA: Springer, 2009, pp. 485–585.

[13] S. B. Kotsiantis, B. Sotiris, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," in *Emerging Artificial Intelligence Applications in Computer Engineering*. Amsterdam, The Netherlands: IOS Press, 2007, pp. 3–24.

[14] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.

[15] W. A. Barbakh, Y. Wu, and C. Fyfe, "Review of clustering algorithms," in *Non-Standard Parameter Adaptation for Exploratory Data Analysis* (Studies in Computational Intelligence), vol. 249. Berlin, Germany: Springer-Verlag, 2009, pp. 7–28.

[16] R. I. Mukhamediev, E. L. Mukhamedieva, and Y. I. Kuchin, "Taxonomy of machine learning methods and assessment of the quality of classification and learning," *Cloud Sci.*, vol. 2, no. 3, pp. 356–375, 2015.

[17] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000.

[18] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[19] L. van der Maaten and G. Hinton, "Visualizing high-dimensional data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

[20] I. Borg and P. Groenen, "Modern multidimensional scaling: Theory and applications," *J. Educ. Meas.*, vol. 40, no. 3, pp. 277–280, Sep. 2003.

[21] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *Amer. Statist.*, vol. 46, no. 3, pp. 175–185, Feb. 1992.

[22] S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 4, pp. 325–327, Apr. 1976, doi: 10.1109/TSMC.1976.5408784.

[23] (2012). *K-Nearest Neighbor Algorithm*. [Online]. Available: http://en.wikipedia.org/wiki/K-nearest_neighbor_algorithm

[24] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Mar. 1995.

[25] G. P. Zhang, "Neural networks for classification: A survey," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 30, no. 4, pp. 451–462, Nov. 2000, doi: 10.1109/5326.897072.

[26] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.

[27] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[28] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[29] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2010, pp. 253–256, doi: 10.1109/ISCAS.2010.5537907.

[30] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017.

[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[32] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.* Berkeley, CA, USA: Univ. of California Press, 1967, vol. 1, no. 14, pp. 281–297.

[33] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities and challenges," *Briefings Bioinf.*, vol. 19, no. 6, pp. 1236–1246, Nov. 2018.

[34] P. J. Ballester and J. B. O. Mitchell, "A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking," *Bioinformatics*, vol. 26, no. 9, pp. 1169–1175, May 2010.

[35] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. Sheth, "Machine learning for Internet of Things data analysis: A survey," *Digit. Commun. Netw.*, vol. 4, no. 3, pp. 161–175, Aug. 2018, doi: 10.1016/j.dcan.2017.10.002.

[36] C. Farrar and K. Worden, *Structural Health Monitoring: A Machine Learning Perspective*. New York, NY, USA: Wiley, 2012, pp. 295–320, doi: 10.1002/9781118443118.

[37] J. Lai, J. Qiu, Z. Feng, J. Chen, and H. Fan, "Prediction of soil deformation in tunnelling using artificial neural networks," *Comput. Intell. Neurosci.*, vol. 2016, p. 16, Jan. 2016, doi: 10.1155/2016/6708183.

[38] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, Aug. 2018, doi: 10.3390/s18082674.

[39] F. Recknagel, "Applications of machine learning to ecological modelling," *Ecol. Model.*, vol. 146, pp. 303–310, 2001.

[40] V. Tatarinov, A. Manevich, and I. Losev, "A system approach to geodynamic zoning based on artificial neural networks," (in Russian), *Mining Sci. Technol.*, vol. 1, no. 3, pp. 14–25, 2018, doi: 10.17073/2500-0632-2018-3-14-25.

[41] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea, "Applications of machine learning to cognitive radio networks," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 47–52, Aug. 2007.

[42] N. Ball and R. J. Brunner, "Data mining and machine learning in astronomy," *Int. J. Mod. Phys. D*, vol. 19, no. 7, pp. 1049–1106, 2010.

[43] Y. Chen and W. Wu, "Application of one-class support vector machine to quickly identify multivariate anomalies from geochemical exploration data," *Geochem., Explor., Environ., Anal.*, vol. 17, no. 3, pp. 231–238, May 2017.
[44] J. Hirschberg and C. D. Manning, "Advances in natural language processing," *Science*, vol. 349, no. 6245, pp. 261–266, Jul. 2015.
[45] Y. Goldberg, "A primer on neural network models for natural language processing," *J. Artif. Intell. Res.*, vol. 57, pp. 345–420, Nov. 2016.
[46] M. Cracknell, "Machine learning for geological mapping: Algorithms and applications," Ph.D. dissertation, School Earth Sci., Univ. Tasmania, Hobart, TAS, Australia, 2014.
[47] C. Kumar, S. Chatterjee, T. Oommen, and A. Guha, "Automated lithological mapping by integrating spectral enhancement techniques and machine learning algorithms using AVIRIS-NG hyperspectral data in gold-bearing granite-greenstone rocks in Hutti, India," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 86, Apr. 2020, Art. no. 102006.
[48] A. Caté, E. Schetselaar, P. Mercier-Langevin, and P.-S. Ross, "Classification of lithostratigraphic and alteration units from drillhole lithogeochemical data using machine learning: A case study from the Lalor volcanogenic massive sulphide deposit, Snow Lake, Manitoba, Canada," *J. Geochem. Explor.*, vol. 188, pp. 216–228, May 2018.
[49] I. Bachri, M. Hakdaoui, M. Raji, A. C. Teodoro, and A. Benbouziane, "Machine learning algorithms for automatic lithological mapping using remote sensing data: A case study from Souk Arbaa Sahel, Sidi Ifni inlier, western anti-Atlas, Morocco," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 6, p. 248, May 2019.
[50] C. Deng, H. Pan, S. Fang, A. A. Konaté, and R. Qin, "Support vector machine as an alternative method for lithology classification of crystalline rocks," *J. Geophys. Eng.*, vol. 14, no. 2, pp. 341–349, Mar. 2017.
[51] H. Shirmard, E. Farahbakhsh, R. D. Müller, and R. Chandra, "A review of machine learning in processing remote sensing data for mineral exploration," 2021, *arXiv:2103.07678*. Accessed: Aug. 14, 2021.
[52] N. Kim, "Object-based land cover classification for Pyongsan uranium mine and concentration plant using machine learning based classifier," presented at the Trans. Korean Nucl. Soc. Autumn Meeting, Daejeon, South Korea, Dec. 2020.
[53] T. Sun, H. Li, K. Wu, F. Chen, Z. Zhu, and Z. Hu, "Data-driven predictive modelling of mineral prospectivity using machine learning and deep learning methods: A case study from southern Jiangxi Province, China," *Minerals*, vol. 10, no. 2, p. 102, Jan. 2020.
[54] T. Merembayev, R. Yunussov, and A. Yedilkhan, "Machine learning algorithms for stratigraphy classification on uranium deposits," *Proc. Comput. Sci.*, vol. 150, pp. 46–52, Jan. 2019.
[55] E. N. Amirgaliev, S. Iskakov, Y. Kuchin, and R. Mukhamediev, "Machine learning methods in the problems of rock recognition at uranium deposits," *News NAS RK, Physics Math. Ser.*, vol. 3, pp. 82–88, Oct. 2013.
[56] R. Muhamediyev, E. Amirgaliev, S. Iskakov, Y. Kuchin, and E. Muhamedyeva, "Integration of results from recognition algorithms applied to the uranium deposits," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 18, no. 3, pp. 347–352, 2014.
[57] Y. Kuchin, K. Yakunin, E. Mukhamedyeva, and R. Mukhamedyev, "Project on creating a classifier of lithological types for uranium deposits in Kazakhstan," *J. Phys., Conf. Ser.*, vol. 1405, no. 1, Nov. 2019, Art. no. 012001.
[58] Y. Kuchin, R. Mukhamediev, and K. Yakunin, "Quality of data classification under inconsistent expert judgment," *Cloud Sci.*, vol. 6, no. 1, pp. 109–126, 2019.
[59] Y. Kuchin, R. Mukhamediev, K. Yakunin, J. Grundspenkis, and A. Symagulov, "Assessing the impact of expert labelling of training data on the quality of automatic classification of lithological groups using artificial neural networks," *Appl. Comput. Syst.*, vol. 25, no. 2, pp. 145–152, Dec. 2020, doi: 10.2478/acss-2020-0016.
[60] A. Dogan, D. Birant, and A. Kut, "Multi-target regression for quality prediction in a mining process," in *Proc. 4th Int. Conf. Comput. Sci. Eng. (UBMK)*, 2019, pp. 639–644, doi: 10.1109/UBMK.2019.8907120.
[61] P. Werbos, "Backpropagation: Past and future," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 1, Mar. 1988, pp. 343–353.
[62] Y. Kuchin, R. Mukhamediev, N. Yunicheva, and E. Muhamedijeva, "Data for filtration properties estimation of host rocks," IEEE Dataport, Sep. 2021, doi: 10.21227/fw57-ka70.

RAVIL I. MUKHAMEDIEV received the Engineering degree in radio-electronics and the Ph.D. degree in information systems from the Riga Civil Aviation Engineers Institute, Riga, Latvia, in 1983 and 1987, respectively. From 1987 to 1995, he worked as an Assistant Professor, a Lecturer, and an Assistant Professor at RCAEI. From 2004 to 2010, he was an Assistant Professor and the Department Head with the Higher School of Information System Management, Riga. He is currently a Professor with Kazakh National Technical University (Satbayev University). He is the author of three books, more than 200 articles, and a leading researcher of six research projects. His research interests include applications of machine learning, data processing, and decision support systems.

YAN KUCHIN was born in Almaty, Kazakhstan, in 1980. He received the bachelor's degree in physics from Kazakh National University, Almaty, in 2002, the bachelor's degree in computer science from IITU, Almaty, in 2011, and the master's degree in computer science, in 2016. He is currently pursuing the Ph.D. degree with Riga Technical University. From 2004 to 2018, he worked at Kazatomprom. Since 2018, he has been a Programmer-Engineer with the Institute of Information and Computing Technologies and a Researcher with Satbayev University. He was the author of more than 30 articles. His research interests include geophysics, machine learning, data processing, and natural language processing.

YEDILKHAN AMIRGALIYEV is currently the Doctor of technical sciences, a Professor, and the Head of the Laboratory of Artificial Intelligence and Robotics, Institute of Information and Computational Technologies, Science Committee, Ministry of Education and Science of the Republic of Kazakhstan. His research interests include theoretical aspects of artificial intelligence and practical applications.

NADIYA YUNICHEVA was born in Frunze, Kyrgyzstan, in 1968. She received the Science degree in specialty 05.13.01—"Management in technical systems," in 1998. In 2002, she was an Associate Professor with the Higher Attestation Commission of the Republic of Kazakhstan, specializing in 05.13.00. Her research interests include the intersection of computer science, control theory, and data analysis of geophysical and climatic research.

ELENA MUHAMEDIJEVA received the master's degree in computer science from the ISMA University of Applied Sciences, Riga, Latvia, in 2008. Since 2010, she has been a Researcher with the Institute of Information and Computational Technologies, Kazakhstan. Her research interests include machine learning, data processing, and decision support systems.