UDC 004.8, 004.9
IRSTI 28.23.29, 28.23.37

CALCULATION OF FILTRATION CHARACTERISTICS OF WELLS URANIUM DEPOSITS

Yunicheva N.R., Kuchin Y., Mukhamedieva E.L.
Institute of Information and Computational Technologies KN MES RK, Almaty
e-mails: naduni@mail.ru , ykuchin@mail.ru, muhamedijeva@gmail.com
ORCID ID: https://orcid.org/0000-0001-6351-3450,
ORCID ID: https://orcid.org/0000-0002-5271-9071

Abstract. Possession of information about the filtration characteristics of the host rocks allows you to plan the volume of ore production. The existing methodology for calculating the filtration properties of wells for uranium extraction by the method of underground borehole leaching is based on a system of rules that take into account only one geophysical parameter (apparent resistance-CS). The method gives a relatively low accuracy of calculating the filtration coefficient. In addition, it is not applicable in the case of defects in the recording of CS or distortion of values under the action of acid, which is widely used in uranium mining. At the same time, other geophysical parameters can be used to calculate the filtration coefficient, the use of which can increase the accuracy of the calculation by 10-20% when using machine learning methods.

The paper assumes the study of the applicability of machine learning methods for predicting the filtration properties of the host rocks of wells for uranium extraction by the PSV (underground well leaching) method and a numerical assessment of the advantages of such an approach. A machine learning method will be developed to predict the filtration properties of rocks based on inaccurately marked electrical logging data.

Keywords: Machine learning, geophysical studies of wells, filtration coefficient, inaccurate data.

1. Introduction

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 8 countries. We will tell you a little more about the countries with the largest uranium reserves.

Figure 1 shows the countries with the largest uranium reserves according to [1].

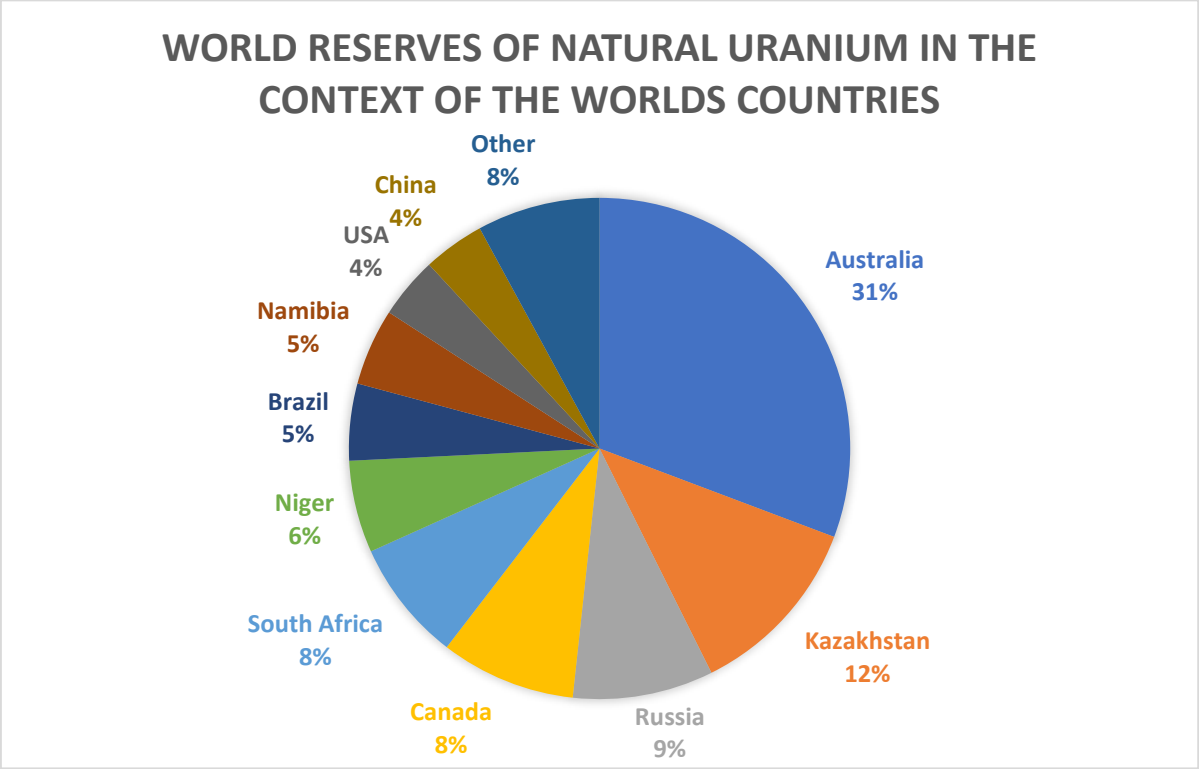

Figure 1 - World reserves of natural uranium in the context of the worlds countries

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.

Kazakhstan accounts for 39% of global uranium production. Production is carried out using the underground well leaching method (UWL). In this method, uranium is extracted through a network of pumping and injection wells, through which the leaching solution is circulated. For uranium production planning, the filtration properties of the host rocks are an important characteristic. Knowledge of the filtration properties in the form of a filtration coefficient allows planning the volume of ore extraction. The currently used method for calculating the filtration properties of wells for uranium production by the in-situ leaching method is based on a system of rules that takes into account only one geophysical parameter (apparent resistance). However, this method is not applicable in the case of a defect in the recording of apparent resistance or distortion of values under the influence of acid, which is widely used in uranium mining. However, other geophysical parameters may also be used to calculate the filtration coefficient, the use of which may improve the accuracy of the calculation. Multiple parameters can be taken into account with actual measurements using machine learning models (MLM).

MLM are divided into two groups [2, 3, 4]:

– *Unsupervised Learning (UL)* [5].

– *Supervised Learning (SL)* [6].

UL solve the clustering problem when a multitude of previously unmarked objects are grouped by an automatic procedure based on the properties of these objects [7, 8].

SL is used when a potentially infinite number of objects distinguish final groups (classes) of some marked objects. The algorithm uses this initial classification as a model and marks out the unmarked objects and assigns them to a suitable group (class) based on the properties of these objects. Although such algorithms are known, it is not trivial to select or develop a classifier for a particular practical task. What is required is a set of computational experiments, significant

preliminary work on data preparation, etc.

For many years, the focus of machine learning researchers has been on petroleum geophysics, where the main task is to isolate reservoirs. In uranium geophysics, the standard set of studies contains far fewer methods, while at the same time it is necessary to determine the properties of many fairly small alternations. It makes the application of MLMs in uranium deposits more difficult than in oil geophysics.

The task of determining filtration properties from log data refers to supervised learning or learning by example. The mathematical problem of example-based learning should be regarded as an optimization task which can be solved by searching for the minimum value of the cost function $J(\theta)$ in all available examples, defined as the sum of squares of the difference between the "predicted" value and the real value of y in a numerious examples *m*. Hypothesis $h_\theta(x)$ is selected, which, with a certain set of parameters $\theta_i \in \Theta$ provides a minimum value $J(\theta)$:

$$J(\theta) = min\frac{1}{2m}\sum_{i=1}^{m}(h_\theta\left(x^{(i)}\right) - y^{(i)})^2 \quad (1.1)$$

where *m* – numerious examples, h_θ – function of the hypothesis, which can be linear ($h_\theta = \theta_0 + \theta_1 x$) or non-linear (example, $h_\theta = \theta_0 + \theta_1 x + \theta_2 x^2$) with a different set of parameters $\theta_i \in \Theta$.

In order to find the optimal function $h_\theta(x)$ a gradient descent algorithm is used, the essence of which is the sequential change of θ_0, θ_1 parameters using an expression

$$\theta_j := \theta_j - \alpha\frac{\partial}{\partial\theta_j}J(\theta_0,\theta_1) \quad (1.2)$$

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

In this process, the algorithm steps are performed in such a way that both parameters are first changed simultaneously based on expression 1.2, and only then new values are assigned to them. In other words, the algorithm sequence of one algorithm step for the case of two parameters, expressed in pseudo-code, will be the next:

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

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

$\theta_0 := temp\,\theta_0$;

$\theta_1 := temp\theta_1$.

With regard to differentiation

$$\theta_0 := \theta_0 - \alpha\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})x_0^{(i)}$$

$$\theta_1 := \theta_1 - \alpha\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})x_1^{(i)}$$

Several algorithms are useful for conducting the experiments, for example, ElasticNet (linear regression with L1 and L2 regularisations), gradient boosting (an ensemble of weak solution trees aggregated in a meta-busting model), and a neural network with a hidden layer.

ElasticNet [9] implements a model that combines two types of regulators: linear and square. Expression of the cost function for ElasticNet

$$J(\theta) = \min \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 \quad + \alpha * l1 * \| \theta \| + 0.5 * \alpha * (1 - l1) * \theta^2$$

In this case, if l1=0, he regularisation is only carried out using the square value of the parameters (l2), if l1=1, only the linear regulator (l1) is used, in other cases, when 0<l1<1 both types are combined. In this case

$$\alpha * l1 * \| \theta \| + 0.5 * \alpha * (1 - l1) * \theta^2$$

can be expressed as a sum

a * l1 + b * l2 ,

where alpha = a + b и l1 = a / (a + b).

The essence of gradient boosting [10] s that, after calculating the optimum regression coefficient values and obtaining the hypothesis function $h_\theta(x)$ c using some algorithm (a), the error is calculated and a new function $h_{b\theta}(x)$ is selected, possibly using another algorithm (b), so that it minimizes the error of the previous one.

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

In other words, it is about minimizing the function

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

where L – error function that takes into account the results of algorithms a and b. If $J_b(\theta)$ is still large, the third algorithm (c) is selected, and etc. Often as algorithms (a), (b), (c) and etc., relatively shallow solution trees are used. The value of the error function gradient is used to find the minimum of function $J_b(\theta)$

$$L(y^{(i)}, h_\theta(x^{(i)}))_{i=1}^{m}$$

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 configured so that the target values are not $(y^{(i)})_{i=1}^{m}$, but the antigradient ($-L'(y^{(i)}, h_\theta(x^{(i)}))_{i=1}^{m}$), i.e. the $(x^{(i)}, -L'(y^{(i)}, h_\theta(x^{(i)}))$ pairs are used instead of $(x^{(i)}, y^{(i)})$ pairs in training algorithm (b).

Multilayer artificial neural networks (multilayer perceptron) are one of the most popular methods of supervised learning, especially for multiple classes. The following cost function is used to set the θ neural network weight:

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

where L – number of neural network layers; S_l – number of neurons in layer l; K – number of classes (equal to the number of neurons in the output layer); Θ – weight matrix, and the hypothesis function is often a sigmoidal (logistic) function $h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}}$,

To minimize the error (learning) function of the multilayer ANN, the Back Propagation Error (BPE) algorithm [11] and its modifications are used to speed up the learning process.

In the execution of the project, it is planned to determine the algorithm and its parameters, which will give the best result in predicting the filtration coefficient, by performing computational experiments using interval inaccurate data. The complexity of this process lies in its high variability. For example, if the researcher has large data volumes and significant computing power at his disposal, deep learning methods that solve complex machine learning tasks with high

precision can be used. On the other hand, if the number of examples is less than the number of properties, it is difficult to use support vector machines (SVM) because they are prone to retraining in this case. Thus, by selecting a certain number of methods to solve a task and changing their parameters (e.g. regularisation coefficient, number of layers of neural networks, etc.), it is necessary to assess the results of their work using one or more indicators, e.g. percentage (fraction) of correctly classified examples (accuracy). For the evaluation of 1st and 2nd type errors, the following are used: precision, recall, and the aggregates are F1 Score, Precision-Recall, and ROC. The following are used to assess the quality of regression dependence: r2_score, mean_absolute_error (absolute difference of target & predicted values), mean_squared_error (squared difference of target & predicted values), median_absolute_error (robust to outliers).

There are currently no systems in place to intelligently interpret well logging data from uranium deposits for the purpose of calculating filtration characteristics. The method developed as a result of the project will make it possible to determine the value of the filtration coefficient with greater accuracy. Preliminary experiments suggest that the use of MLM will increase the accuracy of the calculation by 10-20%.

2. Description of computational experiments and results

As the initial data, the results of hydrogeological studies of wells performed at the exploration stage were used as the only source of actual data on the filtration properties of rocks.

Computational experiments include data preprocessing, the formation of training and test data sets, and the calculation of the parameters of a classification or regression model. Data preprocessing has special features in this case, since interval samples obtained during electrical logging of wells serve as data sources. The formation of a training and test data set also has special features, since they must be formed taking into account the downhole data splitting and data splitting into folds.

To find the best regression parameters and regularization coefficients, a grid-search was performed using the parameters of three algorithms with the recording of the results of testing cross-validation methods.

- A set of models has been formed based on the gdir-search principle.
- Testing was carried out, as a result of which model quality indicators were obtained, including the standard error and determination coefficient, after which the models were sorted by these indicators on a test sample and the top of the best models with optimal meta-parameter settings were selected.
- To train the model, a dataset was formed containing data on the Budennovskoye field.
- Splitting into folds for cross-validation was performed in two ways:

The usual stratified splitting into folds, which ensures that the number of objects with different target values (the value of the filtration coefficient) is proportional to the total sample in the test and training sample

Splitting into folds by wells — with this approach, each individual well was separated into the test sample in order to exclude the possibility of retraining the model for specific physical and filtration properties of individual wells, since a simple stratified splitting into folds cannot provide the calculation of indicative metrics of the assessment quality in such cases.

- The search of regressors and their parameters took place according to the grid search principle, in which the significant parameters of the classifier are selected and their search takes place in a given interval and with a given step, after which the model is selected that has optimal quality indicators on test samples.

Table 1 shows the regressors that showed the best results on different sets of input data.

Table 1. Results of regressors on different sets of input data

	Перрессор	Correlation train	Correlation test	MSE train	MSE test
КС	MLP Regressor - one hidden layer with 201 neurons, 600 max iterations	0,494	0,613	32,389	35,872
КС+Код породы	MLP Regressor - one hidden layer with 201 neurons, 850 max iterations	0,502	0,488	30,968	38,384
КС+ПС	MLP Regressor - one hidden layer with 201 neurons, 600 max iterations	0,596	0,535	25,785	30,472
КС+ПС+ Код породы	MLP Regressor - one hidden layer with 251 neurons, 100 max iterations	0,698	0,570	21,244	30,363
КС	Расчет по принятой методике		**0,261**		**48,32**

Conclusion

Thus, the applied approach allowed us to evaluate the filtration properties of rocks better than the existing method. It is revealed that the correlation with the actual data is higher and less than the standard error. The expansion of the dataset due to logging data for a larger number of wells will increase the accuracy of the model.

Acknowledgements

This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP09562260 "Development of a method for calculating the filtration characteristics of wells of uranium deposits based on machine learning algorithms working with inaccurately marked interval data".)

References:

[1] URL: https://eenergy.media/2018/11/06/zapasy-urana-u-kakih-stran-oni-samye-bolshie (13.07.2021).

[2] Ayodele T. O. Types of machine learning algorithms. New advances in machine learning. 2010. 3. 19-48.

[3] Ibrahim H. A. H. et al. Taxonomy of machine learning algorithms to classify real time interactive applications. *International Journal of Computer Networks and Wireless Communications*. 2012. 2(1). 69-73.

[4] Muhamedyev R. Machine learning methods: An overview. Computer Modelling & New Technologies. 2015. 19(6). 14-29.

[5] Hastie T., Tibshirani R., Friedman J. Unsupervised learning. The elements of statistical learning. Springer, New York, NY, 2009. 485-585.

[6] Kotsiantis S. B., Zaharakis I., Pintelas P. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering. 2007. 160(1). 3-24.

[7] Jain A. K., Murty M. N., Flynn P. J. Data clustering: a review. ACM computing surveys (CSUR). 1999. 31(3). 264-323.

[8] Ashour B. W., Ying W., Colin F. Review of Clustering Algorithms. Non-Standard Parameter Adaptation for Exploratory Data Analysis. Studies in Computational Intelligence. 2009. 249. 7-28.

[9] ElasticNet. [Electronic resource]. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ElasticNet.html.

[10] Friedman J. H. Greedy function approximation: a gradient boosting machine. *Annals of statistics*. 2001. 1189-1232.

[11] Werbos P. J. New tools for prediction and analysis in the behavioral sciences. Ph. D. dissertation, Harvard University. 1974. 38.