Machine learning in reservoir permeability prediction and modelling of fluid flow in porous media

To cite this article: A B Zolotukhin and A T Gayubov 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **700** 012023

IOP Conf. Series: Materials Science and Engineering **700** (2019) 012023 doi:10.1088/1757-899X/700/1/012023

Machine learning in reservoir permeability prediction and modelling of fluid flow in porous media

A B Zolotukhin[1,2,3*] **and A T Gayubov**[1]

[1] Gubkin Russian State University of Oil and Gas (National Research University), Leninsky prospect 65, Moscow, 119991, Russia
[2] Northern Arctic Federal University, Severnaya Dvina Emb. 17, Arkhangelsk, 163002, Russia
[3] The University of Stavanger, P.O. Box 8600 Forus N-4036 Stavanger, Norway

* Corresponding author: anatoly.zolotukhin@gmail.com

Abstract. Reliable data on the properties of the porous medium are necessary for the correct description of the process of displacing hydrocarbons from the reservoirs and forecasting reservoir performance. The true permeability of the reservoir is one of the most important parameters which determination is time-consuming, costly and require skilled labor. The paper describes the methodology for determining the permeability of a porous medium, based on machine learning. The results of laboratory experiments, available in the database (terrigenous reservoirs with permeability in the range from 12 to 1132 md), are used to train the neural network, and then to predict the reservoir permeability. Comparison of the predicted and calculated permeability values showed a fairly good match between them with the determination coefficient of 0.92. The last task considered in this paper is to obtain an analytical expression describing a fluid flow in a porous medium using machine learning. This procedure enabled to obtain a resultant equation of fluid flow in a wide range of reservoir parameters and pressure gradients, which can be used in reservoir simulators.

1. Introduction

The determination of the absolute permeability of a rock sample based on standard laboratory studies requires special attention at the stage of processing experimental data. The traditional technique is based on the use of the Forchheimer equation [1], which takes into account the deviation from the Darcy law at high flow velocities caused by inertial forces:

$$|grad\ p| = \frac{\mu}{k}u + \beta\rho u^2 \quad (1)$$

where $|grad\ p|$ – pressure difference, μ – dynamic viscosity coefficient, k – permeability coefficient, u – velocity, β – non-Darcy coefficient, ρ – fluid density. Although equation (1) seems to be a simple relation for determining permeability, estimating the inertial (high velocity) coefficient β is not a simple task [2,3]. To evaluate it, data from experiments conducted at high pressure gradients, i.e. under conditions when inertial forces become comparable with viscous forces and when the deviation from the (linear) Darcy law becomes obvious and can be fixed.

It has been established by numerous experiments (for example, [2,3]) that the coefficient β is not constant and can change its value thousands of times depending on the type of porous medium. It has

also been noted in numerous literature [4-8] that β depends on porosity and permeability and that its determination is not an easy task.

At low pressure gradients, the gas exhibits a slip property (superconductivity of the porous medium), and Darcy's law also ceases to correctly describe the fluid flow. This phenomenon was first described and investigated by Klinkenberg [9]. Thus, the flow of fluids in a porous medium is quite complex and exhibits nonlinearity properties over the entire range of pressure gradients.

In [10,11], an attempt was made to describe the law of flow of fluids in porous media using dimension theory, which enabled to write it in the following compact form:

$$\Pi = 1 + C \cdot Re - exp(-b \cdot Re^{\gamma}), \tag{2}$$

$$\Pi = \frac{k\Delta p}{\mu u l}, \quad Re = \frac{u\rho\sqrt{k}}{\mu\phi^{1.5}}. \tag{3}$$

where l – characteristic length, ϕ – porosity. Let us note that parameter Re is a macroscopic Reynolds number in porous media. Although equation (2) captures the main characteristics of flow in the entire range of pressure gradients. Coefficients C, b and γ are to be determined. However, this task is not a simple one.

In this paper, an approach based on Machine Learning (ML) is used in order to identify these coefficients, compose a methodology of determining the permeability of porous media and obtain a universal equation of fluid flow through porous media.

2. Methodology

The solution of the above task based on the use of ML will be carried out in stages, namely:

1. Determination of the permeability of samples of a porous medium based on available experimental data using artificial neural networks.
2. Obtaining the universal law of fluid flow in a porous medium by normalizing dimensionless parameters obtained using dimensional analysis.
3. Using the machine learning method - a multiple regression model to test and refine the universal law of fluid flow in a porous medium.

2.1. Artificial Neural Network Model

The significant benefit of ML predictive models is that one may not have to perform core analysis (CA) measurements for all the core samples, but can conduct prediction of the results [12]. Once a predictive model of any CA results is trained it could be effectively used for future forecasting. There are quite a number of ML algorithms to build a predictive model [13]. In our work, we used the neural network to predict the permeability of porous media.

The choice of algorithm is highly dependent on the problem at hand, data quality and dataset size. For example, it would be unnecessary to build recursive or convolutional neural networks in our problem due to the small dataset size and its structure.

Artificial neural network (ANN) is a mathematical representation of the biological neural network (Figure 1) [14]. It consists of several layers with units that are connected by links. Each connection has its own weight and level of activation. Each node has an input value, an activation function and an output. In ANN information propagates (forward pass) from first (not hidden) layer with inputs to the next hidden layer and then to further hidden layers until the output layer would be achieved. The value in each node of the first hidden layer obtained after calculation of activation function for the dot product of inputs and weights. Next hidden layer receives the output of the previous one and puts its dot product with weights to the activation, and so on. Initially, all weights for all nodes are assigned randomly. ANN calculates first output with random weights. Then compare it to real values, calculate the error and adopts weights to obtain smaller error on the next iteration via backpropagation (ANN training algorithm). After training all weights are tuned, and one may make a prediction for a new data by passing them into an ANN, which will calculate output via forward propagation throw all activations [15]. Scikit-learn Python library provides ANN representation in MLPRegressor method [16]. This implementation

allows to indicate the number of hidden layers, number of nodes in each layer, activation function, learning rate and some other parameters.

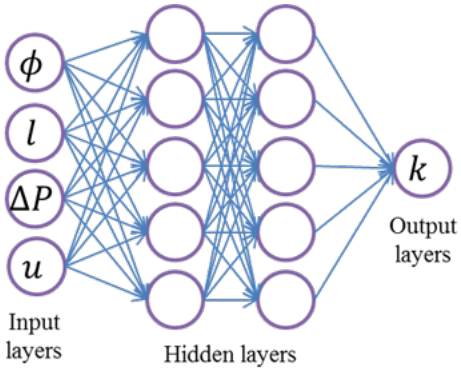


Figure 1. Artificial neural network with two hidden layers.

The objective of this research is to use ANN techniques to build a model to predict reservoir permeability. In the proposed model, the physical characteristics of the rock and fluid sample are used as input parameters, which fully describe the gas flows in the porous medium and are determined prior to the simulation with a high degree of reliability from laboratory studies. As the output (predicted) parameter of the model, the permeability characteristic of the rock sample was used. The choice of this predicted characteristic was based on the fact that when determining the coefficient C in the Forchheimer equation (2), it is necessary to preset the approximate permeability from the obtained physical parameters of the sample and fluid. Sometimes databases contain low-quality data on laboratory measured permeability values. This may be due to the lack of competence of the stuff, inaccuracy of instruments, and other reasons. In this case, the use of a neural network can serve as an additional tool that evaluates the quality of the obtained permeability values. In some cases, inaccurate data maybe replaced with the values obtained by ANN. For permeability predictive models we have used 4 features. All 4 features accounting in machine learning algorithm are presented in Table 1. In our model, we did not consider such input parameters as sample diameter, gas viscosity and density, and ambient temperature, because in all rock samples they had the same value.

Table 1. All features used in predictive model.

No.	Feature	Unit
1	Sample porosity, ϕ	–
2	Length of the sample, l	m
3	Pressure difference, ΔP	Pa
4	Flow rate, u	m/s

The creation of an ANN model was implemented in the programming language Python. Ninety percent of the total data were used for the training of ANN models while the remaining data points were used for testing. Data were randomly mixed for training and testing. The training sample consisted of 203 experimental measurements on four physical parameters of the rock sample (terrigenous reservoirs with permeability in the range from 12 to 1132 md). The network architecture was chosen on the basis of the dependence of the number of hidden layers and the number of neurons in each hidden layer shown in Figure 2. The figure shows various models from 1 to 4 hidden layers, and in each layer consisting from 5 to 20 neurons. For training ANN, we chose a model consisting of 4 hidden layers. The neurons in the hidden layer were 11. Rectified Linear Unit activation function was used as a transfer function between input and hidden layer and linear type activation functions were used between hidden and output layer.

As follows from Figure 2, the best (most reliable) results are coming from the ANN with a limited number of hidden layers (up to four) a number of neurons in each not exceeding high numbers. This is a well-known thesis (not yet formally proven): The number of hidden layers and neurons in each layer should be somehow proportional to the available dataset.

Performance of the ANN model could be demonstrated by plotting predicted values (via cross-validation) of permeability after ablation versus the actual values (Figure 3). The permeability values are set in a normalized form since the values of low-permeability rock samples are not visually visible against the background of other high-permeability samples. The determination coefficient (R^2) between the estimated permeability using the developed ANN model and the actual core permeability was 0.92. Based on these results, we decided to select the ANN model for permeability prediction.

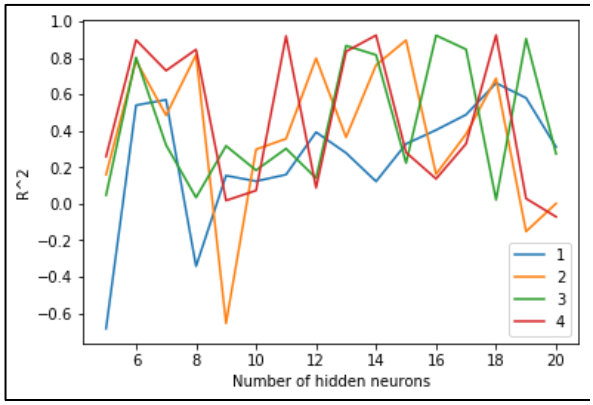


Figure 2. Neural network architecture graph.

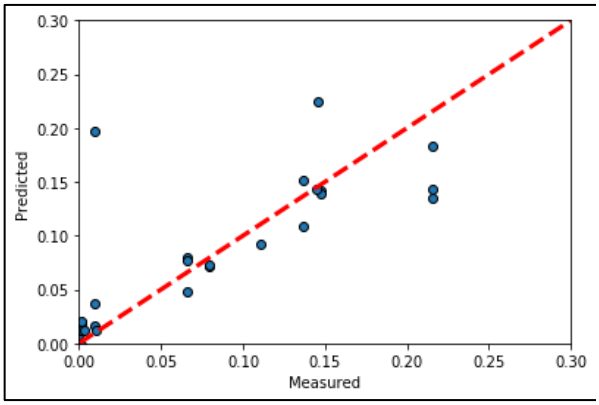


Figure 3. Comparison of real and predicted core permeability by ANN model.

2.2. Method of data normalization

By data normalization in this work, we mean the translation (mapping) of a series of curves into one universal straight line. In accordance with the results of laboratory experiments and previous studies, for further analysis, we select dimensionless criteria described in (2). Using the research results [10,11,17], it is possible to determine two areas in the $\boldsymbol{\Pi - Re}$ diagram (Figure 4), where the dependence has non-linear (interval 1) and linear (interval 2) parts.

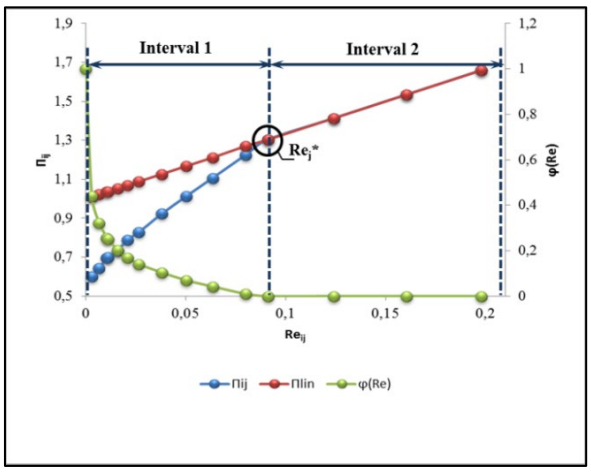


Figure 4. Typical characteristics of the dependency $\Pi = f(Re)$.

2.2.1. Linear area. Let us start with the interval 2 as the initial study area. As follows from earlier works, the fluid flow in the linear interval 2 of the $\Pi - Re$ diagram can be described using the following equation [11,17]:

$$\Pi = 1 + C \cdot Re \tag{4}$$

In accordance with the accepted research algorithm the following dependency graph based on laboratory experiments can be built:

$$\Pi_{ij} = 1 + C_j \cdot Re_{ij} \tag{5}$$

for the linear interval 2 of all laboratory experiments in expiremental database. Here index j indicates experiment number, and index i – measurement number in the j-th experiment. On Figure 5 experimental points of the same color belong to the same experiment.

Analysis of the data obtained from processing each of the experimental straight lines (Figure 6) showed that the value of the coefficient C_j depends on the permeability, porosity and length of a rock sample measured in the j-th experiment and can be approximated by the following dependency:

$$ln\left(\frac{k_j}{\phi_j \cdot {l_j}^2}\right) = 1.9679\, ln\left(\frac{C_n}{C_j}\right) - 20.109 \tag{6}$$

where $C_n = 1.2085$, coefficient obtained for a rock sample with a permeability of 1132 mD.

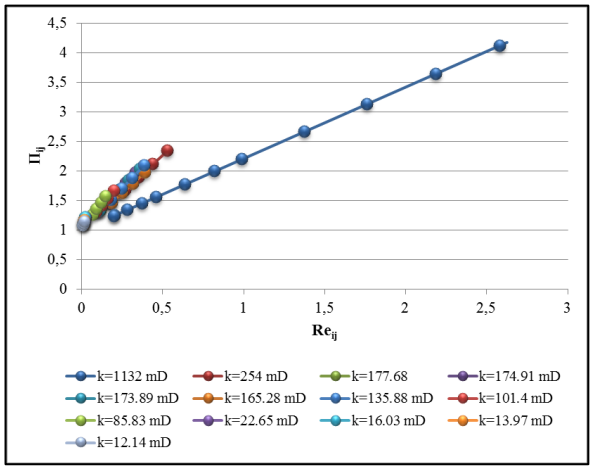


Figure 5. Building the dependency $\Pi = f(Re)$ on experimental data.

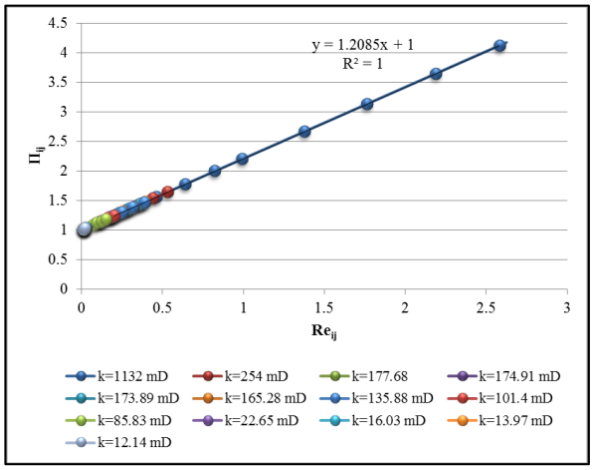


Figure 6. Normalization of experimental data and their translation into a universal dependence.

A dimensionless group $\frac{\phi \cdot l^2}{k}$ can be interpreted as tortuosity τ of porous channels, i.e.

$$\tau_j = \frac{\phi_j \cdot {l_j}^2}{k_j} \tag{7}$$

This definition of tortuosity is referred to [18], where it is stated that "Tortuosity of the streamline S is given by $\tau(S) = {}^{l}/_{l_S}$ for the length of the porous medium l divided by the length of the streamline l_S.

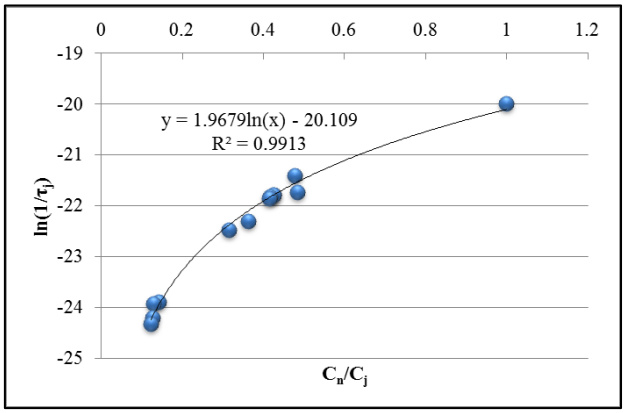


Figure 7. Dependence of $\frac{C_n}{C_j}$ on τ_j.

We believe that more natural and similar to the definition of the term tortuosity will be the ratio of the inverse of the above, i.e. $= {}^{l_s}/_{l}$, which is further converted to (7). With the last result dependency of $\frac{C_n}{C_j}$ on τ_j can be illustrated as shown on Figure 7. Taking into account the last relation, dependence (6) can be rewritten as follows:

$$C_j = B_1 \cdot (\tau_j)^{\alpha} \tag{8}$$

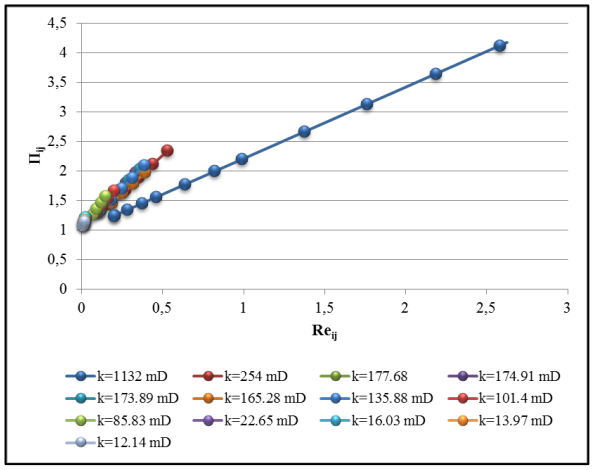
where $B_1 = 4.41 \cdot 10^{-5}$; $\alpha = 0.5$. In view of the last result, the universal dependence takes the following form:

$$\Pi_{ij} = 1 + B_1 \cdot (\tau_j)^{\alpha} \cdot Re_{ij} \tag{9}$$

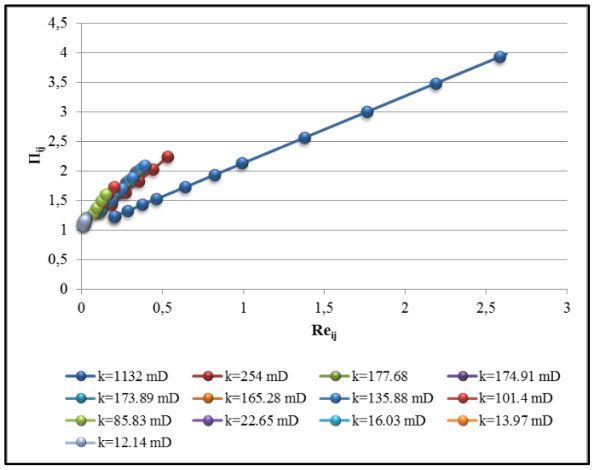
Using the new equation (9) the new values of the normalized coefficient $C_{j,\mathrm{n}}$ were obtained (Table 2) and the dependence $\Pi_{ij}(Re_{ij})$ was built (Figure 8).

Table 2. The results of calculating the coefficients $C_{j,\mathrm{n}}$ and $b_{j,\mathrm{n}}$ based on the normalization method

No.	k_j, mD	ϕ_j	l_j, m^2	$1/\tau$	$ln(1/\tau)$	Re_j^*	C_j	$C_{j,\mathrm{n}}$	b_j	$b_{j,\mathrm{n}}$
1	1132	0.204	0.051	2.10E-09	-19.983	0.20	1.209	1.133	5.000	3.043
2	254	0.188	0.051	5.11E-10	-21.395	0.18	2.531	2.323	7.177	5.657
3	165.28	0.173	0.051	3.65E-10	-21.730	0.19	2.494	2.754	6.326	6.553
4	177.68	0.192	0.051	3.50E-10	-21.774	0.17	2.832	2.816	7.076	6.680
5	135.88	0.147	0.051	3.49E-10	-21.775	0.18	2.832	2.818	4.970	6.683
6	173.89	0.192	0.052	3.37E-10	-21.811	0.11	2.900	2.870	7.617	6.790
7	174.91	0.202	0.051	3.27E-10	-21.840	0.11	2.914	2.913	6.137	6.878
8	101.4	0.184	0.051	2.09E-10	-22.288	0.09	3.331	3.657	6.620	8.372
9	85.83	0.201	0.049	1.73E-10	-22.475	0.07	3.830	4.023	7.309	9.090
10	22.65	0.202	0.051	4.25E-11	-23.880	0.02	8.444	8.214	14.714	16.843
11	16.03	0.149	0.051	4.07E-11	-23.926	0.01	9.300	8.405	17.956	17.181
12	13.97	0.169	0.051	3.12E-11	-24.189	0.01	9.450	9.610	20.439	19.288
13	12.14	0.162	0.052	2.76E-11	-24.315	0.01	9.854	10.242	20.819	20.379



a) Dependence $\Pi_{ij}(Re_{ij})$ on the experimental data



b) Dependence $\Pi_{ij}(Re_{ij})$, calculated according to equation (9)

Figure 8. Graphs $\Pi_{ij}(Re_{ij})$, obtained by using the data normalization method.

2.2.2. Nonlinear area. Now let us consider interval 1 (see Figure 4), in which the effect of nonlinearity of the gas flow is observed. This effect is due to a so-called gas slippage effect, first described by Klinkenberg [9]. To take into account the slippage effect, we can use the following approximation, in which the function $\varphi(Re)$ describes interval 1:

$$\Pi_{ij} = 1 + C_j \cdot Re_{ij} - \varphi(Re_{ij}) \tag{10}$$

where $\varphi(Re)$ can be approximated by the following function used in equation (2) (Figure 9) [11]:

$$\varphi(Re_{ij}) = e^{-b_j \cdot Re_{ij}{}^{\gamma}} \tag{11}$$

In accordance with the above algorithm of the data normalization, we build dependency graphs. The generalized straight line takes the following form (Figure 10):

$$\ln(\Pi_{ij} - 1 - C_j Re) = b_j \cdot Re_{ij}{}^{\gamma} \tag{12}$$

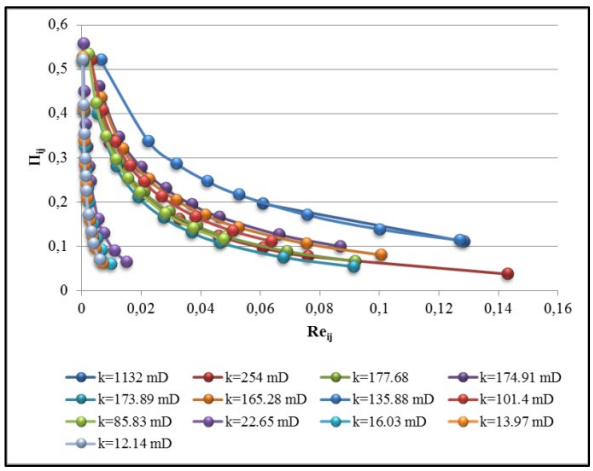


Figure 9. Dependence $\Pi_{ij}(Re_{ij})$, calculated according to equation (11).

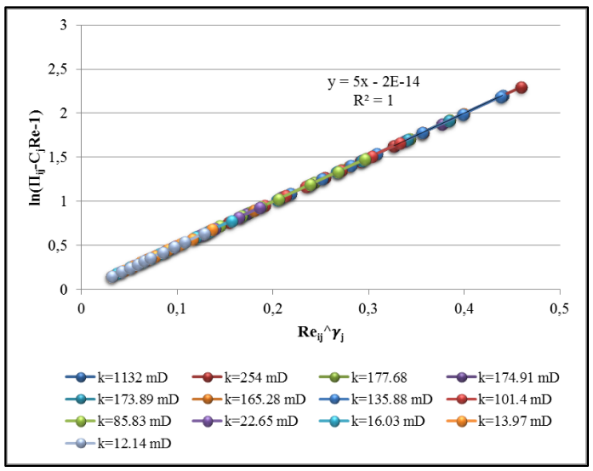


Figure 10. Normalization of experimental data and their translation into a universal dependence $\Pi_{ij}(Re_{ij}{}^{\gamma})$.

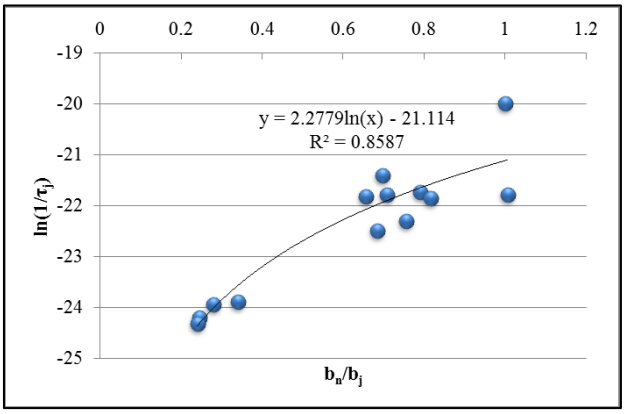


Figure 11. Dependence of coefficient $\frac{b_n}{b_{ij}}$ on τ_j.

Let us now determine the value of the universal constant (or functional) b_j in equation (11). Analysis of the available experimental data showed that the coefficient b_j depends on permeability and porosity and the length of the rock sample, as well as the coefficient C_{ij}, and can be approximated by the following dependency (Figure 11):

$$\ln\left(\frac{1}{\tau}\right) = 2.2779 \ln\left(\frac{b_n}{b_j}\right) - 21.114 \tag{13}$$

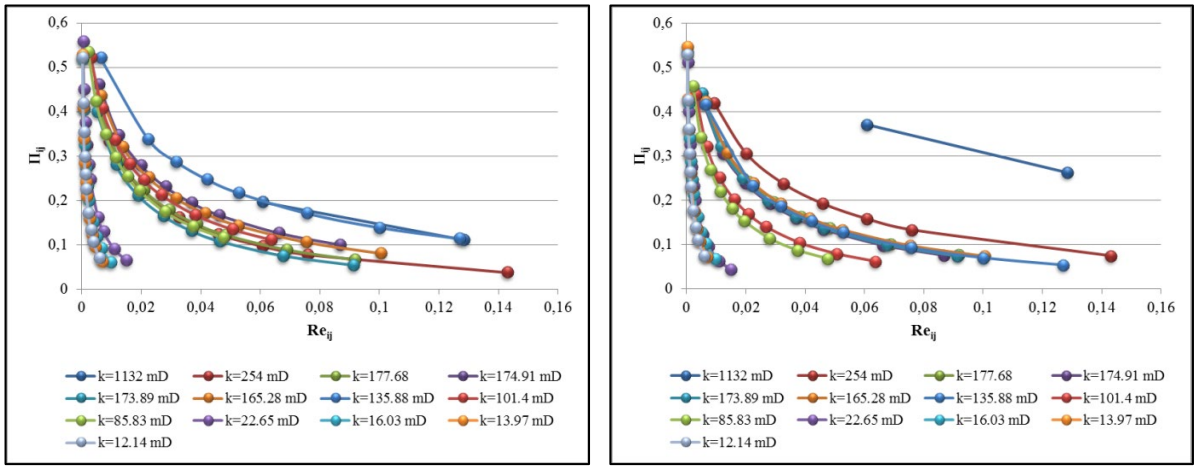
where $b_n = 5$, coefficient obtained for a rock sample with a permeability of 1132 mD.
Equation 13 can be rewritten as:

$$b_j = 4.72 \cdot 10^{-4} (\tau_j)^{0.44} = B_2 \cdot \tau_j^{\gamma_1} \tag{14}$$

where $B_2 = 4.72 \cdot 10^{-4}$, $\gamma_1 = 0.44$. With the last result (11) can be written in the following form:

$$\varphi(Re_{ij}) = e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot Re_{ij}^{\gamma_2}} \tag{15}$$

Using the new equation (15) we obtained the values of the normalized coefficient $b_{j,n}$ (Table 2) and constructed the dependence $\varphi(Re_{ij})$ (Figure 12).



a) Dependence $\varphi_{ij}(Re_{ij})$ based on the experimental method

b) Dependence $\varphi_{ij}(Re_{ij})$, calculated according to the equation (15)

Figure 12. Graphs $\varphi_{ij}(Re_{ij})$, obtained using the data normalization method.

2.2.3. Combining linear and nonlinear areas. Based on the newly obtained universal coefficients C, b and γ, we can rewrite equation (10) as follows:

$$\Pi = 1 + B_1\tau^{\alpha} \cdot Re_{ij} - e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot Re^{\gamma_2}} \tag{16}$$

where $B_1 = 4.41 \cdot 10^{-5}$, $B_2 = 4.72 \cdot 10^{-4}$, $\alpha = 0.5$, $\gamma_1 = 0.44$, $\gamma_2 = 0.4$

2.2.4 The border Reynolds number.
Let us note that the border Reynolds number Re_j^* defined as:

$$\varphi\left(Re_j^*\right) \le \varepsilon \,, \tag{17}$$

where ε is a small number, is the number dividing intervals 1 and 2 on the dependency $\Pi = f(Re)$ (Figure 1). Dependency of Re_j^* on tortuosity τ is shown on Figure 13.

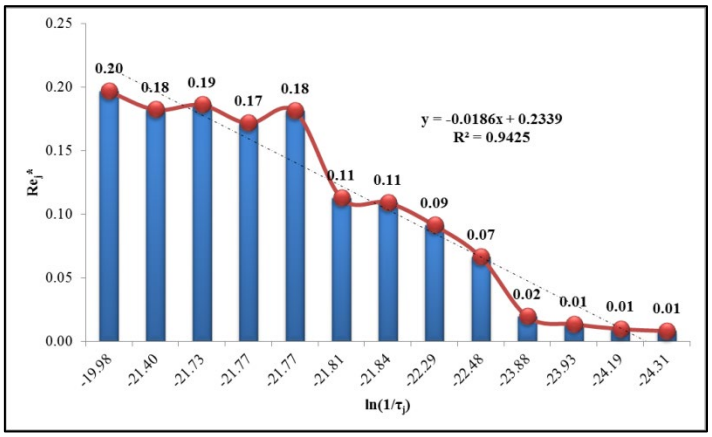


Figure 13. Dependency of the border Reynolds number Re_j^* on tortuosity τ.

2.3. Linear regression model
As was shown above, the dependence of Π_{ij} values on the parameters Re and τ (tortuosity) is generally nonlinear, approaching the linear one only asymptotically (for large Re values). To simplify the subsequent analysis, we divide the entire area of variation of the parameter Re into two, namely, the interval 1 of small Re values (non-linear part) and the interval 2 of large Re values (linear part). We first consider the linear part of the equation, leaving the border number Re_j^* separating both zones for further analysis in more detail.

2.3.1. Linear part of the equation. For the linear part, as was shown above, the following equation holds:

$$\Pi_{ij} = 1 + C_j \cdot Re_{ij}$$

or, in terms of the tortuosity parameter τ introduced above:

$$\Pi_{ij} = 1 + B_1 \cdot \tau_j^{\alpha} \cdot Re_{ij} \tag{18}$$

where $\tau_j = \frac{\phi_j {l_j}^2}{k_j}$. We rewrite equation (18) in a more convenient form for subsequent analysis:

$$\ln\left(\Pi_{ij} - 1\right) = \alpha \cdot \ln\left(\tau_j\right) + \ln\left(Re_{ij}\right) + \ln(B_1) \tag{19}$$

or in terms commonly used in regression analysis:

$$y = w_0 \cdot x_0 + w_1 \cdot x_1 + b_0 \tag{20}$$

where the following notation is used: $w_0 = \alpha$, $w_1 = \lambda$, $b_0 = \ln(B_1)$. The input data for the regression model: $x_0 = \ln(\tau_j)$, $x_1 = \ln(Re_{ij})$, the output is: $y = \ln(\Pi_{ij} - 1)$. The regression model produces the following parameters:

$$\ln(\Pi_{ij} - 1) = 0.471 \cdot \ln(\tau_j) + 0.977 \cdot \ln(Re_{ij}) - 9.244,$$

or

$$\Pi_{ij} = 1 + B_1(\tau_j)^{\alpha} Re_{ij}{}^{\lambda}, \tag{21}$$

where $B_1 = 9.67 \cdot 10^{-5}$, $\alpha = 0.47$, $\lambda = 0.98$.

2.3.2. Linear and nonlinear parts of the equation. The generalized equation for a linear and non-linear zone has the following form:

$$\Pi_{ij} = 1 + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda} - e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot Re_{ij}{}^{\gamma_2}}, \tag{22}$$

which can be rewritten as:

$$e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot Re_{ij}{}^{\gamma_2}} = 1 + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda} - \Pi_{ij}$$

$$B_2 \cdot \tau_j^{\gamma_1} \cdot Re_{ij}{}^{\gamma_2} = -ln(1 + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda} - \Pi_{ij}).$$

In the final form, convenient for using the regression model, equation (22) will take the form:

$$\ln(B_2) + \gamma_1 ln(\tau_j) + \gamma_2 ln(Re_{ij}) = ln[-ln(1 + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda} - \Pi_{ij})], \tag{23}$$

or in a form convenient for regression analysis:

$$y = w_0 \cdot x_0 + w_1 \cdot x_1 + b_0$$

The input data for the regression model: $x_0 = ln(\tau_j)$, $x_1 = \ln(Re_{ij})$, the output data is: $y = ln[-ln(1 - \Pi_{ij} + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda})]$. The regression model produces the following parameters: $w_0 = \gamma_1$, $w_1 = \gamma_2$, $b_0 = \ln(B_2)$.

$$ln[-ln(1 - \Pi_{ij} + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda})] = \gamma_1 \cdot \ln(\tau_j) + \gamma_2 \cdot \ln(Re_{ij}) + \ln(B_2)$$

where $\gamma_1 = 0.34$, $\gamma_2 = 0.36$, $B_2 = 3.82 \cdot 10^{-3}$. Finally, the general formula generated by the method of the regression model is:

$$\Pi_{ij} = 1 + B_1 \cdot \tau_j{}^{\alpha} \cdot Re_{ij}{}^{\lambda} - e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot Re_{ij}{}^{\gamma_2}}$$

where $B_1 = 9.67 \cdot 10^{-5}$, $B_2 = 3.82 \cdot 10^{-3}$, $\alpha = 0.47$, $\lambda = 0.98$, $\gamma_1 = 0.34$, $\gamma_2 = 0.36$

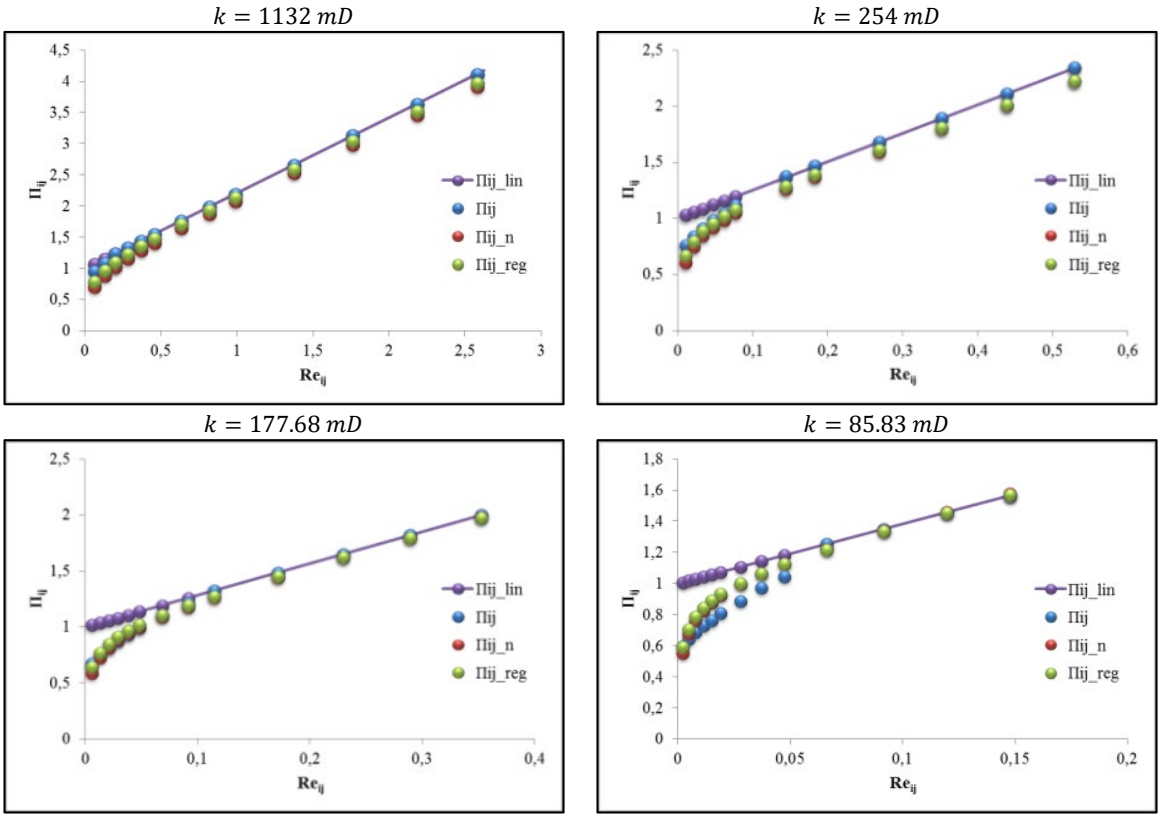
3. Analysis of the results

This section provides a brief analysis of the results of two methods employed for obtaining the universal equation of fluid flow in a porous medium, namely: 1) Normalization method and 2) Regression model. Table 3 lists the formulas obtained using these approaches. As the determination coefficients show, the normalization method showed a slightly better result than the regression model of machine learning. The obtained high values of the determination coefficient confirm the correct choice of the initial parameters and their presentation in a dimensionless form. Note that during the data processing using the normalization method it was found that some data fall far beyond the general trend. The response to the request to the laboratory where the tests were carried out confirmed that the initial porosity of the two samples instead of the correct 0.149 and 0.127 was recorded as 0.008 and 0.010, the correction of which put them back into the general trend.

When obtaining the universal gas flow equation using the two above methods, the corrected values were used, which ensured high determination coefficients between the experimental and calculated data. This result is illustrated by Figure 14, where experimental (blue points) and forecasted data (red points for normalization method and green points for regression method, respectively) are perfectly matched it a wide range of reservoir permeability and Reynolds number values. Due to the lack of space, we managed to place in Figure 14 only a part of the results.

Table 3. Results of using the normalization method and regression model

Method for determining the coefficients.	Formula for calculating	R^2
Normalization method	$\Pi_{ij} = 1 + B_1 \cdot {\tau_j}^{\alpha} \cdot Re_{ij} - e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot Re^{\gamma_2}}$ $B_1 = 4.41 \cdot 10^{-5}$, $B_2 = 4.72 \cdot 10^{-4}$, $\alpha = 0.5$, $\gamma_1 = 0.44$, $\gamma_2 = 0.4$	0.98
Regression model	$\Pi_{ij} = 1 + B_1 \cdot {\tau_j}^{\alpha} \cdot {Re_{ij}}^{\lambda} - e^{-B_2 \cdot \tau_j^{\gamma_1} \cdot {Re_{ij}}^{\gamma_2}}$ $B_1 = 9.67 \cdot 10^{-5}, B_2 = 3.82 \cdot 10^{-3},$ $\alpha = 0.47, \lambda = 0.98, \gamma_1 = 0.34, \gamma_2 = 0.36$	0.97



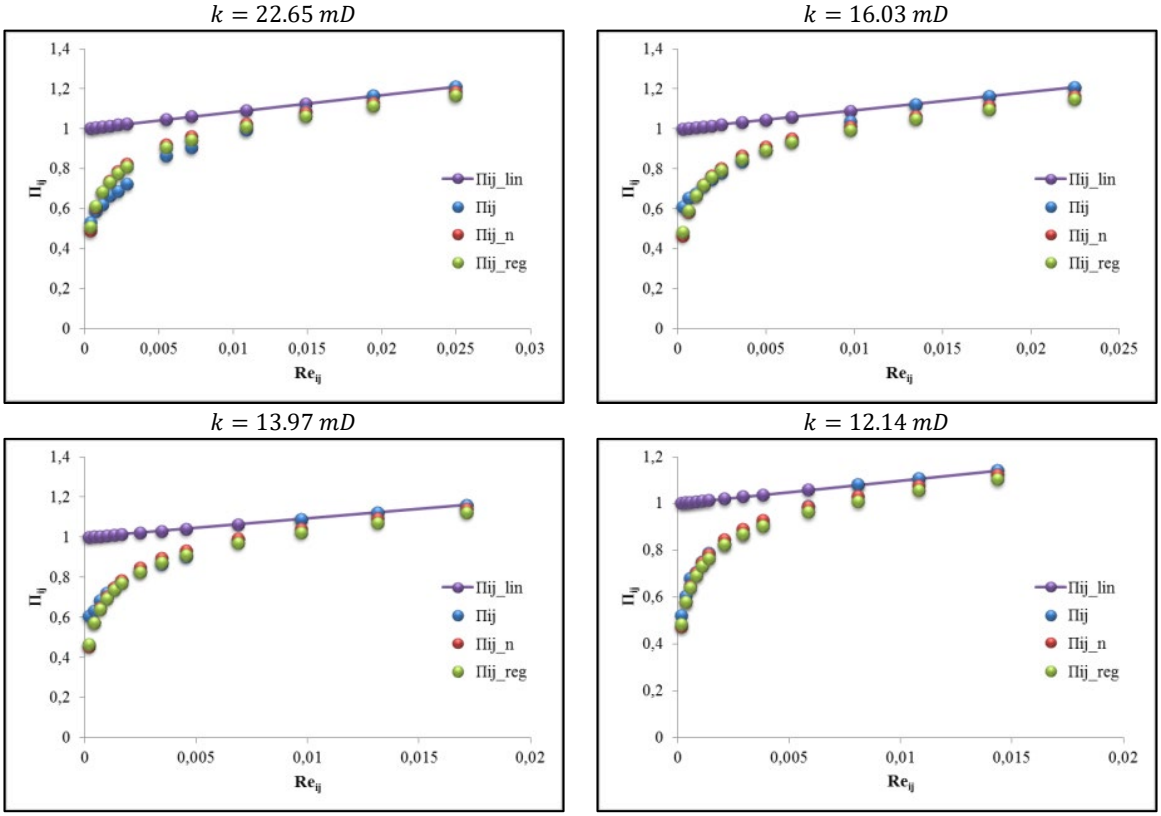


Figure 14. Dependence plots $\Pi - Re$ obtained using: 1. Π_{ij_lin} – the linear part of the equation; 2. Π_{ij} – experimental data; 3. Π_{ij_n} – normalization method; 4. Π_{ij_reg} – regression model.

4. Conclusions

The following general conclusions can be derived from the previous analysis:

- The use of the dimensional analysis method can significantly reduce the number of variables involved in the analysis, improve the quality of the resulting solutions and at the same time simplify the obtained equations and increase the reliability of the analysis.
- In particular, this approach made it possible to evaluate the permeability of porous medium samples using artificial neural networks with a very high correlation of predicted values with the available data in the sample, even with a very limited number of available experimental data.
- The use of dimensional theory methods in combination with regression analysis made it possible to define more precisely the constants included in the universal semi-empirical law of fluid flow. In its turn, it enables high accuracy and reliability of predictive flow regimes over the entire range of pressure gradients at reservoirs and experimental conditions for core samples.
- The use of the universal equation of fluid flow in porous media makes it possible to significantly simplify and refine computer modeling of the oil and gas field development forecast.

Acknowledgments

We are grateful to Dr. Abbakumov V. for valuable advice given in the preparation of this work.

References

[1] Forchheimer P 1901 *Wasserbewegung durch Boden* Zeit ver Deutsch Ing.
[2] Tessem R 1980 *High Velocity Coefficient's Dependence on Rock Properties: A Laboratory Study*, Norway: Thesis Pet. Inst. NTH. pp 26-30.
[3] Torsæter O, Tessem R. and Berge B 1981 *High Velocity Coefficient's Dependence on Rock Properties*" SINTEF Report (Norway: NTH Trondheim).
[4] Cornell D and Katz D L 1953 Flow of gases through consolidated porous media, *Ind. Eng. Chem.* **45**(10), 2145–52.
[5] Geertsma J 1974 Estimating the coefficient of inertial resistance in fluid flow through porous media *SPE J.* **14**(05), 445–50.
[6] Janicek J D and Katz D L 1955 *Applications of unsteady state gas flow calculations,* paper presented at Research Conference on Flow of Natural Gas Reservoirs, Proc. U. of Mich. Res. Conf. (June 30, 1955).
[7] Jones S C 1987 *Using the inertial coefficient, β, to characterize heterogeneity in reservoir rock*. Paper presented at 62nd SPE Annual Tech. Conf. and Exhibit. (Texas: Dallas).
[8] Liu C Y, Civan F and Evans R D 1995 Correlation of the non-Darcy flow coefficient, *J. Can. Petrol. Technol.*, **34**(10), pp 50–53.
[9] Klinkenberg L J 1941 *Permeability of Porous Media to Liquids and Gases*, API Drill. and Prod. Practice.
[10] Zolotukhin A B 2002 *A new gas flow equation and permeability determination technique* Reg. SPE offshore Bergen Conf. on Drilling, Completion and Reservoir Management (Norway: Bergen).
[11] Zolotukhin A B 2006 *Equation of gas flow in porous media and ghe technoque of determining the permeability of a porous medium* Report presented at an Int. Conf. dedicated to the 50th anniversary of TATNIPIneft (Russia: Bugulma).
[12] Unsal E, Dane J and Dozier G V 2005 A genetic algorithm for predicting pore geometry based on air permeability measurements, *Vadose Zone J.* **4**(2), 389–97.
[13] Hastie T, Tibshirani R, Friedman J 2001 *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (New York: Springer–Verlag).
[14] Haykin S 1994 *Neural networks: a comprehensive foundation* (New York: Prentice Hall).
[15] Rumelhart D E, Hinton G E and Williams R J 1986 Learning representations by back-propagating errors*, Nature* **323**(6088), 533–36.
[16] Pedregosa F et al. 2012 Scikit-learn: Machine learning in python, *J. Mach. Learn. Res.* **12,** 2825–30.
[17] Zolotukhin A B, Shulev V E 2018 *Fluid Flow in Porous Media, Dimensional Analysis and Permeability Determination*. Proc. of Int. scientific-technical conf. "Development of exploration technique and exploitation of onshore and offshore hydrocarbon fields" ("Geopetrol-2018") (Poland: Zakopane-Kostelisco) pp 573–80.
[18] Berg C F 2014 Permeability Description by Characteristic Length, Tortuosity, Constriction and Porosity, Transp. Porous Media **103** pp 381–400.