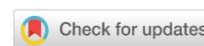


INFORMATION TECHNOLOGY, COMPUTER SCIENCE AND MANAGEMENT ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ



UDC 681.5:004.942

Original Empirical Research

<https://doi.org/10.23947/2687-1653-2026-26-2-2237>

Machine Learning-Based Condition Assessment Method for Shell-and-Tube Heat Exchangers to Improve Energy Efficiency

Larisa G. Tugashova¹ , Andrey V. Zatonskiy² 

¹ State Technological University “Higher School of Petroleum”, Almetyevsk, Russian Federation

² Perm National Research Polytechnic University, Perm, Russian Federation

✉ tugashova@yandex.ru



EDN: JTIHTA

Abstract

Introduction. Shell-and-tube heat exchangers are widely used in oil treatment and refining plants, as well as in heating systems, where their reliable operation is largely determined by the intensity of thermal fouling on the tube surface. Therefore, it is crucial to promptly determine when heat exchange equipment requires maintenance, as deposit accumulation can result in the reduced heat transfer efficiency and increased operating costs. A review of the literature shows that the problem of predicting heat exchanger fouling is being actively researched. However, existing approaches, which focus on process parameters and the physicochemical properties of heat transfer fluids, are not always applicable to the specific conditions of oil refining, where feedstock composition, operating conditions, and deposit nature differ significantly from typical industrial processes. Thus, there remains a gap in scientific knowledge associated with the insufficient development of models that allow for the reliable description of changes in the thermal resistance of contaminants, and the heat transfer coefficient specifically for shell-and-tube heat exchangers of oil refineries. The objective of this study is to determine the dependence of the coefficients of thermal resistance of contaminants and heat transfer on operating parameters, in particular, on the output temperatures of coolants and oil density. For this, it is planned to use machine learning methods to build more adequate predictive models and thereby increase the validity of decisions on the maintenance of heat exchange equipment.

Materials and Methods. A shell-and-tube heat exchanger in a small refinery unit was investigated. Here, circulating diesel reflux (210–235 °C) was the hot medium on the shell side, and oil was the cold medium in the tube bundle. For the calculation, the cross-sectional areas of the tube and annular spaces, linear flow rates, Reynolds, Prandtl, and Nusselt criteria, as well as the heat transfer coefficients of the hot and cold coolants were determined. The heat transfer coefficient was calculated taking into account the thermal resistance of the wall and contaminants, while heat transfer was calculated from heat balance and heat transfer equations. Machine learning and symbolic regression methods, including PySR, SISSO, FROLS, and CatBoostRegressor, were used to estimate the thermal resistance of contaminants and the heat transfer coefficient, based on oil density and coolant temperature data. The verification calculation of the heat exchanger was performed by the method of successive approximations with the solution to a system of nonlinear equations in Matlab.

Results. A verification calculation of a shell-and-tube heat exchanger for oil from three fields showed that with an increase in the thermal resistance of contaminants from 0 to 0.002 (m²·°C)/W, the heat transfer coefficient decreased from approximately 93–95 to 81–83 W/(m²·°C). It was established that changes in the outlet temperatures of the coolants was an information basis for identifying contaminants. Using the SISSO and PySR symbolic regression methods, analytical dependences of thermal resistance on oil density and outlet temperatures were obtained with RMSE 1.25·10^{−8} and 2.34·10^{−8} (m²·°C)/W, respectively. To predict the dynamics of heat transfer, NARX models based on FROLS and CatBoostRegressor were built, validated on industrial data. Ex-post forecast error for the algorithm CatBoostRegressor RMSE = 0.03573 W/(m²·°C), for the algorithm FROLS RMSE = 0.01296 W/(m²·°C).

Discussion. A 13% reduction in heat transfer coefficient with an increase in contaminant thermal resistance to $0.002 \text{ (m}^2 \cdot ^\circ\text{C)/W}$ was consistent with theoretical models and experimental data from other researchers. The use of SISSO and PySR methods provided significantly higher accuracy compared to neural network approaches with small training sets. The FROLS algorithm outperformed CatBoostRegressor in the accuracy of ex-post forecast of heat transfer coefficient dynamics, which is explained by the compactness of the polynomial model and the smooth nature of the process under study. The established threshold for reducing the heat transfer coefficient by 25% complies with industry standards and provides the transition to equipment maintenance based on actual condition.

Conclusions. A method for determining the thermal resistance of heat exchanger fouling based on the outlet temperature of the coolant and the oil density was developed. Analytical relationships with $\text{RMSE } 1.25 \cdot 10^{-8} \text{ (m}^2 \cdot ^\circ\text{C)/W}$ and $2.34 \cdot 10^{-8} \text{ (m}^2 \cdot ^\circ\text{C)/W}$ were obtained using the SISSO and PySR symbolic regression methods. NARX models for predicting heat transfer coefficients ($\text{RMSE } 0.01296$ and $0.03573 \text{ W/(m}^2 \cdot ^\circ\text{C)}$) were built and validated using industrial data. A 25.2% reduction in heat transfer coefficient was adopted as the criterion for the need for cleaning the unit, which provided a transition to condition-based maintenance and increased energy efficiency in oil refining.

Keywords: shell-and-tube heat exchanger, thermal resistance, heat transfer coefficient, symbolic regression

Acknowledgements. The authors would like to thank the editors and reviewers for their time and valuable comments.

For Citation. Tugashova LG, Zatonskiy AV. Machine Learning-Based Condition Assessment Method for Shell-and-Tube Heat Exchangers to Improve Energy Efficiency. *Advanced Engineering Research (Rostov-on-Don)*. 2026;26(2):2237. <https://doi.org/10.23947/2687-1653-2026-26-2-2237>

Оригинальное эмпирическое исследование

Способ оценки состояния кожухотрубчатых теплообменников с применением машинного обучения для повышения энергоэффективности

Л.Г. Тугашова¹ , А.В. Затонский² 

¹ Альметьевский государственный технологический университет «Высшая школа нефти», г. Альметьевск, Российская Федерация

² Пермский национальный исследовательский политехнический университет, г. Пермь, Российская Федерация

✉ tugashova@yandex.ru

Аннотация

Введение. Кожухотрубчатые теплообменники широко применяются на установках подготовки и переработки нефти, а также в системах теплоснабжения, где их надежная работа во многом определяется интенсивностью термического загрязнения поверхности труб. В связи с этим актуальной задачей является своевременное определение момента, когда теплообменному оборудованию требуется техническое обслуживание, поскольку накопление отложений приводит к снижению эффективности теплопередачи и росту эксплуатационных затрат. Анализ литературных источников показывает, что проблема прогнозирования загрязнения теплообменных аппаратов достаточно активно исследуется. Однако существующие подходы, ориентированные на учет режимных параметров и физико-химических свойств теплоносителей, не всегда применимы к специфическим условиям нефтепереработки, где состав сырья, режимы эксплуатации и характер отложений существенно отличаются от типовых промышленных процессов. Таким образом, сохраняется пробел в научных знаниях, связанный с недостаточной разработанностью моделей, позволяющих надежно описывать изменение термического сопротивления загрязнений и коэффициента теплопередачи именно для кожухотрубчатых теплообменников нефтеперерабатывающих установок. Целью настоящего исследования является определение зависимости коэффициентов термического сопротивления загрязнений и теплопередачи от режимных параметров — в частности, от выходных температур теплоносителей и плотности нефти. Для достижения поставленной цели предполагается применение методов машинного обучения, позволяющих построить более адекватные прогностические модели и тем самым повысить обоснованность решений по техническому обслуживанию теплообменного оборудования.

Материалы и методы. В качестве объекта исследования рассмотрен кожухотрубчатый теплообменник установки переработки нефти малой мощности, в котором циркуляционное орошение дизельного топлива ($210\text{--}235 \text{ }^\circ\text{C}$) выступает горячим теплоносителем в межтрубном пространстве, а нефть — холодным теплоносителем в трубном пучке. Для расчета определялись площади сечений трубного и межтрубного пространств, линейные скорости потоков, критерии Рейнольдса, Прандтля и Нуссельта, а также коэффициенты теплоотдачи со

стороны горячего и холодного теплоносителей. Коэффициент теплопередачи рассчитывался с учетом термического сопротивления стенки и загрязнений, теплообмен — по уравнениям теплового баланса и теплопередачи. Для оценки термического сопротивления загрязнений и коэффициента теплопередачи использовались методы машинного обучения и символьной регрессии, в том числе PySR, SISSO, FROLS и CatBoostRegressor, на основе данных о плотности нефти и температуре теплоносителей. Поверочный расчет теплообменника выполнен методом последовательных приближений с решением системы нелинейных уравнений в Matlab.

Результаты исследования. Поверочный расчет кожухотрубчатого теплообменника для нефти трёх месторождений показал, что при росте термического сопротивления загрязнений от 0 до $0,002 \text{ (м}^2 \cdot \text{°C)/Вт}$ коэффициент теплопередачи снижается примерно с 93–95 до 81–83 $\text{Вт/(м}^2 \cdot \text{°C)}$. Установлено, что изменение выходных температур теплоносителей служит информационной основой для идентификации загрязнений. Методами символьной регрессии SISSO и PySR получены аналитические зависимости термического сопротивления от плотности нефти и выходных температур с RMSE $1,25 \cdot 10^{-8}$ и $2,34 \cdot 10^{-8} \text{ (м}^2 \cdot \text{°C)/Вт}$ соответственно. Для прогнозирования динамики теплопередачи построены NARX-модели на базе FROLS и CatBoostRegressor, валидированные на промышленных данных. Ошибка постпрогноза для алгоритма CatBoostRegressor RMSE = $0,03573 \text{ Вт/(м}^2 \cdot \text{°C)}$, для алгоритма FROLS RMSE = $0,01296 \text{ Вт/(м}^2 \cdot \text{°C)}$.

Обсуждение. Снижение коэффициента теплопередачи на 13 % при росте термического сопротивления загрязнений до $0,002 \text{ (м}^2 \cdot \text{°C)/Вт}$ согласуется с теоретическими моделями и экспериментальными данными других исследователей. Применение методов SISSO и PySR обеспечило существенно более высокую точность по сравнению с нейросетевыми подходами при малых обучающих выборках. Алгоритм FROLS превзошёл CatBoostRegressor по точности постпрогноза динамики коэффициента теплопередачи, что объясняется компактностью полиномиальной модели и гладким характером исследуемого процесса. Установленный порог снижения коэффициента теплопередачи на 25 % соответствует отраслевым нормам и обеспечивает переход к обслуживанию оборудования по фактическому состоянию.

Заключение. Разработана методика определения термического сопротивления загрязнений теплообменника по выходной температуре теплоносителей и плотности нефти. Методами символьной регрессии SISSO и PySR получены аналитические зависимости с RMSE $1,25 \cdot 10^{-8} \text{ (м}^2 \cdot \text{°C)/Вт}$ и $2,34 \cdot 10^{-8} \text{ (м}^2 \cdot \text{°C)/Вт}$. Построены NARX-модели прогнозирования коэффициента теплопередачи (RMSE $0,01296$ и $0,03573 \text{ Вт/(м}^2 \cdot \text{°C)}$), валидированные на промышленных данных. Снижение коэффициента теплопередачи на 25,2 % принято критерием необходимости очистки аппарата, что обеспечивает переход к обслуживанию по фактическому состоянию и повышение энергоэффективности нефтепереработки.

Ключевые слова: кожухотрубчатый теплообменник, термическое сопротивление, коэффициент теплопередачи, символьная регрессия

Благодарности. Авторы выражают благодарность редакции и рецензентам за уделенное время и ценные комментарии.

Для цитирования. Тугашова Л.Г., Затонский А.В. Способ оценки состояния кожухотрубчатых теплообменников с применением машинного обучения для повышения энергоэффективности. *Advanced Engineering Research (Rostov-on-Don)*. 2026;26(2):2237. <https://doi.org/10.23947/2687-1653-2026-26-2-2237>

Introduction. Improvement of the energy efficiency of oil processing and pretreatment is a pressing practical challenge, particularly given tighter energy regulations, rising cost pressures, and recent export restrictions on oil and petroleum products. Oil producers have no influence over these restrictions, but they can improve internal processes at their facilities to reduce production costs. Specific opportunities include reducing heat loss, thermal energy consumption, and increasing the efficiency of installations. This can be accomplished through controlled process temperatures and equipment diagnostics aimed at reducing unscheduled downtime. These approaches are closely related to heat exchange processes, which are crucial in the petrochemical industry. A $1\div 3 \text{ mm}$ layer of contaminant deposits on tube surfaces can significantly reduce heat transfer efficiency. Heat losses are estimated at 20–40 %. To compensate for the decrease in heat transfer and maintain the required process parameters, the coolant flow rate must be raised, which in turn increases fuel and electricity consumption. According to experts, a 1 mm fouling of the heat exchanger can result in a 10–15% increase in fuel consumption.

Assessing the condition of shell-and-tube heat exchangers for petroleum products is difficult because of their operating temperatures and the impossibility of directly inspecting internal fouling, corrosion, or deformation. Therefore, the key parameters for monitoring tube wall fouling are the heat transfer coefficient, pressure drop, and outlet coolant temperature.

The cause of deposits on the surfaces of the walls of apparatus is the presence of asphalt-resinous substances, paraffins, water, and salts in the raw materials. To get rid of deposits and scale in the cavities of oil pipelines and heat exchangers, mechanical, thermal and chemical cleaning methods are used. The heat exchanger is cleaned by washing it with a paraffin

solvent. Ultrasonic cleaning of paraffin deposits is used at oil preparation and refining facilities.

Deposits on tube walls impair heat transfer and reduce the heat transfer coefficient. Various methods exist for diagnosing and assessing the fouling layer on the surface of heat exchanger tubes. For example, the characteristics of a clean and fouled heat exchanger can be determined experimentally using monitoring systems. Another approach involves constructing mathematical models. An example is the Ebert-Panchal model and its improved versions.

Article [1] provides an overview of models and examples of quantitative assessment of scale deposit thickness in plate heat exchangers of heating systems. The authors [2] performed an analytical accounting of the impact of fouling on the technical and economic indicators of industrial oil treatment units. In the work, the coefficients of heat transfer, heat emission, and thermal resistance of deposits in heat exchangers were obtained by an experimental method. In [1, 2], the proposed models concern specific types of heat exchangers and deposits; they do not predict the dynamics of fouling.

Software tools, particularly process modeling programs, are widely used to solve heat exchange equipment diagnostic problems. For example, in [3], the authors examine the problem of fouling heat exchangers in diesel fuel hydrotreating units. The problem is solved using the Hysys modeling program. In [4], a numerical solution is obtained by constructing a thermal model based on the finite element method in the Ansys program. In [5], the authors propose simplified formulas for calculating the effective thermal conductivity of porous structures by Mathcad and Ansys. Currently, the use of the above programs is not possible within the framework of import substitution.

The authors [6] use the software package “MiR PiA” to model the quality indicators of oil preparation for transportation. A license is required for use.

A number of studies are aimed at improving the design parameters of equipment used in the industry. Improved design and aerodynamics of tube bundles should provide enhanced heat transfer with high reliability. In [7], this is linked to pressure drops in the heat exchanger, which can cause equipment malfunctions. The implementation of algorithms for automatic adjustment of operating modes is also proposed. The authors [8] propose a neural network model for predicting the thermal-hydraulic efficiency of surface heat exchange intensifiers. Neural network models require large volumes of training data on heat exchanger fouling, which is difficult to collect. Furthermore, such solutions, which require changes to the device design, may not be applicable to the equipment in operation.

The authors [9] use simulation methods to obtain a model of a shell-and-tube heat exchanger and propose techniques for improving the control system of the heat exchanger. The mathematical model of the shell-and-tube evaporator is presented as an object with lumped parameters. A heat balance is compiled in differential form taking into account transport delays.

In [10], a mathematical model of a shell-and-tube heat exchanger is presented in the form of an ARMAX model for participation in a control system. These models do not take into account nonlinearities.

In the operation of heat exchange equipment, the colloidal state of raw materials and products, specifically heavy fractions and residues, plays a major role in increasing the efficiency of heat and mass transfer [11]. This solution is relevant for heavy fractions and also requires the installation of additional equipment.

The literature review shows that predicting thermal fouling and heat transfer coefficients based on feedstock properties, as well as applying small-data methods that handle nonlinearities, remains insufficiently investigated. For approximation and forecasting, a combination of approaches appears to be an interesting approach, such as combining physicochemical (balance) methods with machine learning.

The objective of this research is to create a methodology for calculating the fouling thermal resistance and heat transfer coefficient of a heat exchanger in relation to the outlet temperatures of crude oil and diesel fuel and the density of the oil, as well as to enable prediction of the heat transfer coefficient.

To achieve this goal, the following tasks must be completed:

- perform a verification thermal calculation of a shell-and-tube heat exchanger for a small-capacity oil refining unit, that is, obtain a mathematical model of the heat exchanger in balance form;
- using the shell-and-tube heat exchanger model obtained above, study the effect of the physicochemical properties of the heat transfer fluids and the thermal fouling coefficient on the heat transfer coefficient and output parameters (using oil samples from various fields as an example);
- obtain the dependence of the coefficient of thermal resistance of contaminants and the heat transfer coefficient on the output temperatures of oil and diesel fuel in heat exchangers, as well as on the density of oil, that is, construct approximating dependences through symbolic regression;
- build a nonlinear dynamic model suitable for predicting the heat transfer coefficient using the obtained approximating dependences, experimental data of the operating installation, as well as a description of the object in the form of previously derived ordinary differential equations.

Materials and Methods. To determine the thermal resistance coefficient of fouling and the heat transfer coefficient in shell-and-tube heat exchangers, it is proposed to apply machine learning methods using small sample sizes, such as symbolic regression, since it is difficult to obtain large volumes of information on heat exchanger fouling under real plant operating conditions. A NARX model is proposed for predicting the heat transfer coefficient.

The object of this study is a shell-and-tube heat exchanger for a low-capacity processing unit. The initial design parameters of the heat exchanger were used for the shell-and-tube heat exchanger calculation: heat transfer surface area — 68 m², heat exchanger body length — 7.2 m, and tube length — 6 m.

The total thermal contamination of the tube surface was taken in the range from 0 (clean heat exchanger) to 0.002 (m²·°C)/W (contaminated heat exchanger).

For such units, the circulation reflux (CR) is organized as follows. Diesel fuel from the main column outlet is divided into two streams: product and circulating. The CR, passing through a shell-and-tube heat exchanger, returns to the main column. The CR temperature is 210–235 °C. Diesel fuel (annular space) is the hot coolant, while oil (tube bundle) is the cold coolant.

The first calculation step is to obtain the cross-sectional areas of the tubular and annular space. The second step is to find the linear velocities of the oil and diesel streams.

The flow modes of the coolant and the heat transfer rates are determined by the Reynolds, Prandtl, and Nusselt criteria. The next step is to calculate the heat transfer coefficient from the hot flow (diesel fuel) to the tube wall α_1 and the heat transfer coefficient from the tube surface to the heated feedstock (oil) α_2 . The heat exchanger calculation is performed by standard methods. Therefore, well-known formulas and heat exchanger calculation details are not provided.

The heat transfer coefficient is determined by the formula:

$$K = \frac{1}{\frac{1}{\alpha_1} + \frac{\delta_t}{\lambda_t} + \frac{1}{\alpha_2}}, \quad (1)$$

where δ_t — heat exchange tube wall thickness, m; λ_t — thermal conductivity coefficient, W/(m·°C); α_1, α_2 — heat transfer coefficients, W/(m²·°C).

The heat transfer coefficient depends on the heat transfer coefficients in the heat exchange tubes and the thermal resistance and thermal conductivity of the tube wall.

If there is contamination inside and outside the heat exchange tubes, the heat transfer coefficient is determined by the following formula:

$$K_z = \frac{1}{\frac{1}{\alpha_1} + \frac{\delta_t}{\lambda_t} + \sum r + \frac{1}{\alpha_2}}, \quad (2)$$

where $\sum r$ — total thermal resistance of the contamination layer, (m²·°C)/W.

The heat transfer equation is as follows:

$$Q = K \cdot F \cdot \Delta T, \quad (3)$$

where K — heat transfer coefficient, W/(m²·°C); F — heat exchange surface, m²; ΔT — mean temperature difference, °C.

Heat balance equation for a shell-and-tube heat exchanger:

$$G_{LC} \cdot c_{LC} \cdot (T_{13} - T_{LC}) = G_n \cdot c_n \cdot (T_{n2} - T_{n1}), \quad (4)$$

where c_{LC} — heat capacity of CR diesel fuel, kJ/(kg·°C); c_n — heat capacity of oil, kJ/(kg·°C); G_{LC} — consumption of CR diesel fuel; G_n — oil consumption rate, kg/s; T_{13} — temperature on tray #13 of the diesel fuel outlet, °C; T_{LC} — CR return temperature, °C; T_{n1} — oil temperature at the inlet to the heat exchanger, °C; T_{n2} — oil temperature at the outlet of the heat exchanger, °C.

Using the initial design data, a verification calculation of the shell-and-tube heat exchanger was performed. The heat transfer coefficients and heat transfer coefficient were determined, and the system of equations (3) and (4) was solved.

Figure 1 shows a flowchart of the algorithm for solving equations (3) and (4) for the shell-and-tube heat exchanger.

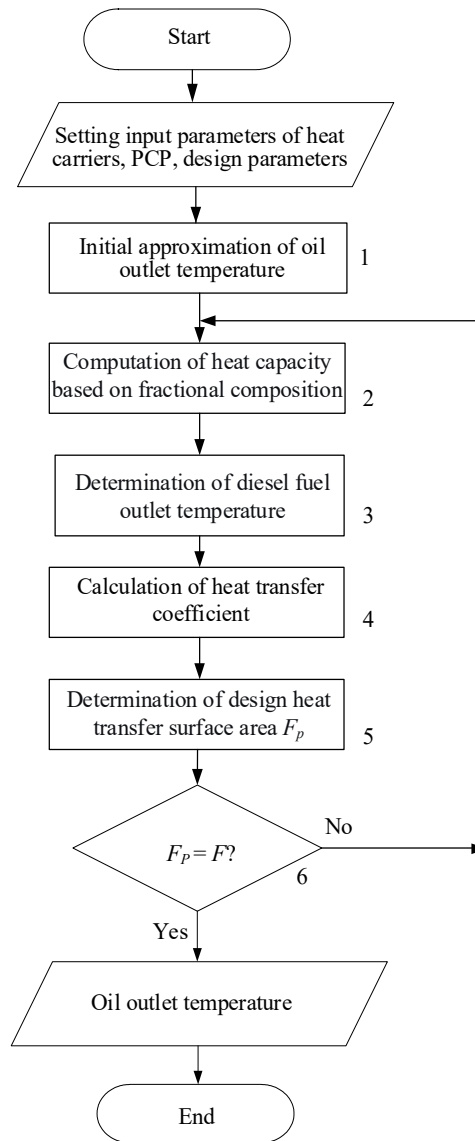


Fig. 1. Flowchart of the heat exchanger equation system solution algorithm

The diesel fuel and oil outlet temperatures are calculated using a method of successive approximations. After entering the initial data in block 1, the first approximation of the hot fluid outlet temperature (diesel fuel) is selected.

Block 2 calculates the heat capacity of oil and diesel fuel, taking into account the fractional composition of the oil. Block 3 calculates the outlet temperature of the diesel fuel through solving the heat balance equation for heat exchanger (4). Block 4 calculates the following parameters: mean temperature difference, heat transfer coefficient of clean heat exchanger (1), and heat transfer coefficient taking into account the thermal resistance coefficient of contaminant layer (2).

In the next block, block 5, the calculated heat transfer area is determined from heat transfer equation (3). In block 6, the obtained calculated area value is verified. A significant discrepancy between the obtained and specified values indicates an incorrect selection of oil and diesel fuel temperature. The iterative calculation continues until the discrepancy between the calculated and specified heat transfer areas is no more than 5%. The solution to the system of nonlinear equations (3) and (4) is implemented by numerical methods in the Matlab environment.

The calculation results are determined for given values of oil and diesel fuel inlet temperatures and oil and diesel fuel flow rates. The oil properties (density, viscosity), its fractional composition at true boiling temperature (TBT), heat exchanger design parameters, thermal load, and thermal resistance of the contaminant layer on the pipe surface are specified.

The efficiency of the heat exchange process is affected by the thermal resistance of the tube surface fouling. Furthermore, the composition and physicochemical properties (PCP) of the feedstock also influence the output parameters, particularly the outlet temperature of the coolant. This study used the PCP of oil from three fields (Table 1).

Table 1

Physicochemical Properties of Oil from Three Fields [12]

Oil field	Density ρ , kg/m ³	Output of fractions up to 350 °C, %	Viscosity η , mPa·s	Asphaltene content, %
Arlanskiy	891	42.8	35.37	5.80
Tuimazinskiy	856	53.4	10.18	3.40
Kuleshovskiy	819	63.6	3.29	0.70

Table 1 lists the physicochemical properties of crude oil from various fields. Density and viscosity are related to asphaltene content. Density and viscosity increase with increasing molecular weight and fraction boiling point. Lower-density oils yield higher light fractions at temperatures up to 350°C. The density and viscosity of oil, based on the PCP listed in Table 1, can be measured in the field and in the laboratory. The data in Table 1 are also used to establish the relationship between the thermal coefficient of contaminant resistance, the heat transfer coefficient, and the density and viscosity of oil.

The following method is proposed for determining the thermal resistance of contaminants. Using the mathematical model of a heat exchanger given in [13] and formulas (1–4), the input and output parameters of the object are determined. These formulas form a training sample. To identify the thermal resistance of contaminants and the heat transfer coefficient in shell-and-tube heat exchangers, it is proposed to use artificial neural networks [14] and machine learning methods associated with symbolic regression and system identification. Examples of symbolic regression methods include SISO [15], TorchSISO [16], PySR [17], SyMANTIC, and RGP. To obtain relationships between heat exchanger parameters, the PySR (High-Performance Symbolic Regression) and SISO (Sure Independence Screening and Sparsifying Operator) methods are selected. They allow finding mathematical expressions that relate the original data. These methods are selected because they are suitable for working with small data sets. Key hyperparameters of the PySR method: `binary_operators` — list of binary operators ['+', '-', '*', '/', '^'], `unary_operators` — list of unary operators ['sin', 'cos', 'exp', 'log', 'sqrt', etc.], `maxsize` — maximum expression size, `iterations` — number of iterations, `loss_function` — loss function, `population_size` — size of population. Key hyperparameters of the SISO method: `n_expansion` — number of feature expansions, `n_term` — number of terms in the final formula, `k` — number of features for L0-regularization, `initial_screening` — initial feature selection, `operators` — a set of mathematical operations for generating features.

For identification and forecasting tasks over a given period, the NARX (Nonlinear Autoregressive with Exogenous Inputs) model is selected. FROLS (Forward Regression Orthogonal Least Squares) and CatBoostRegressor algorithms are used to determine its structure.

In the presented work, the oil density and the outlet temperature of the hot and cold coolants are used as factors. The output parameter of the relationship is the thermal resistance of the contaminants, or the heat transfer coefficient. The target parameter is calculated from the mathematical expression obtained by the PySR or SISO, CatBoostRegressor or FROLS methods and measured process parameter values. It allows for determining the heat transfer coefficient and maintenance intervals for heat exchangers.

Research Results. Table 2 presents the values of heat transfer coefficients for different values of thermal fouling, density and viscosity of oil from different sources (data from Table 1).

Table 2

Heat Transfer Coefficient Values

Thermal resistance, (m ² ·°C)/W	Heat transfer coefficient of clean and contaminated heat exchangers, W/(m ² °C)		
	Arlanskiy oil field	Tuymazinskiy oil field	Kuleshovskiy oil field
0.00000	95.27486	94.24999	93.19691
0.00018	93.97257	92.97539	91.95044
0.00036	92.70541	91.73479	90.73687
0.00055	91.47196	90.52687	89.55492
0.00073	90.27091	89.35035	88.40336
0.00091	89.10098	88.20401	87.28104

0.00109	87.96099	87.08672	86.18686
0.00127	86.84981	85.99737	85.11978
0.00145	85.76635	84.93495	84.07879
0.00164	84.70959	83.89845	83.06296
0.00182	83.67855	82.88695	82.07138
0.00200	82.67231	81.89954	81.1032

In the next step, the output temperatures of the coolants (oil and diesel fuel) are calculated depending on the thermal resistance of the contaminant layer and the physicochemical properties of the oil (density and viscosity). Figures 2 and 3 show the results obtained.

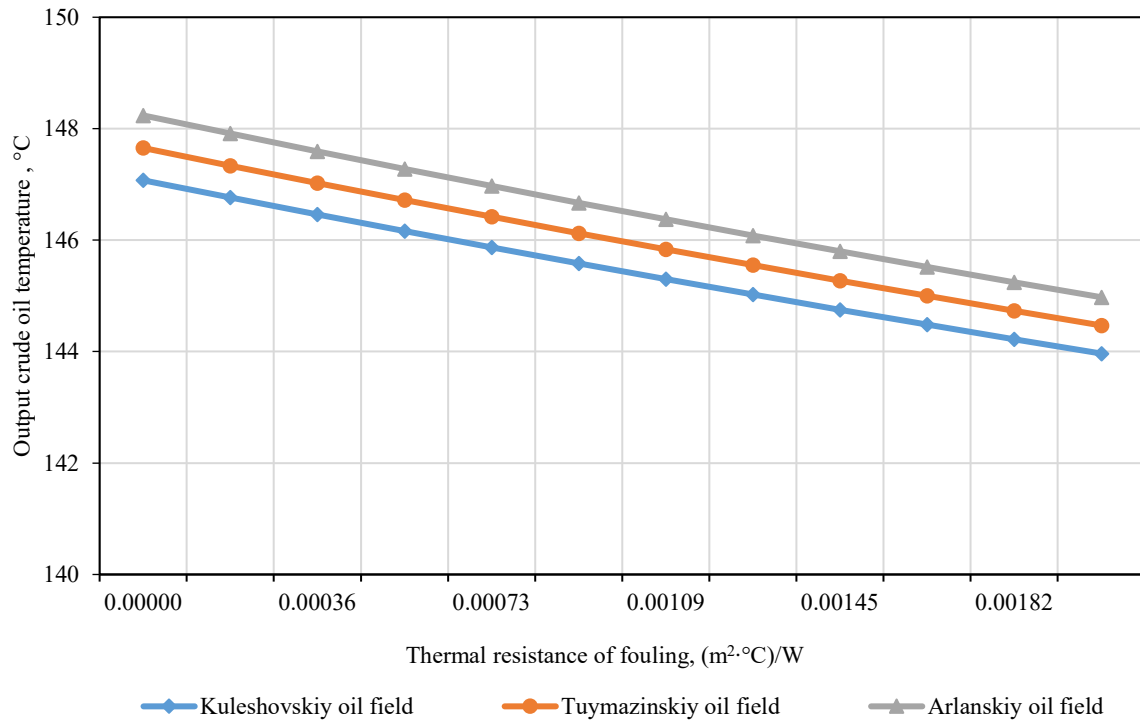


Fig. 2. Oil outlet temperature

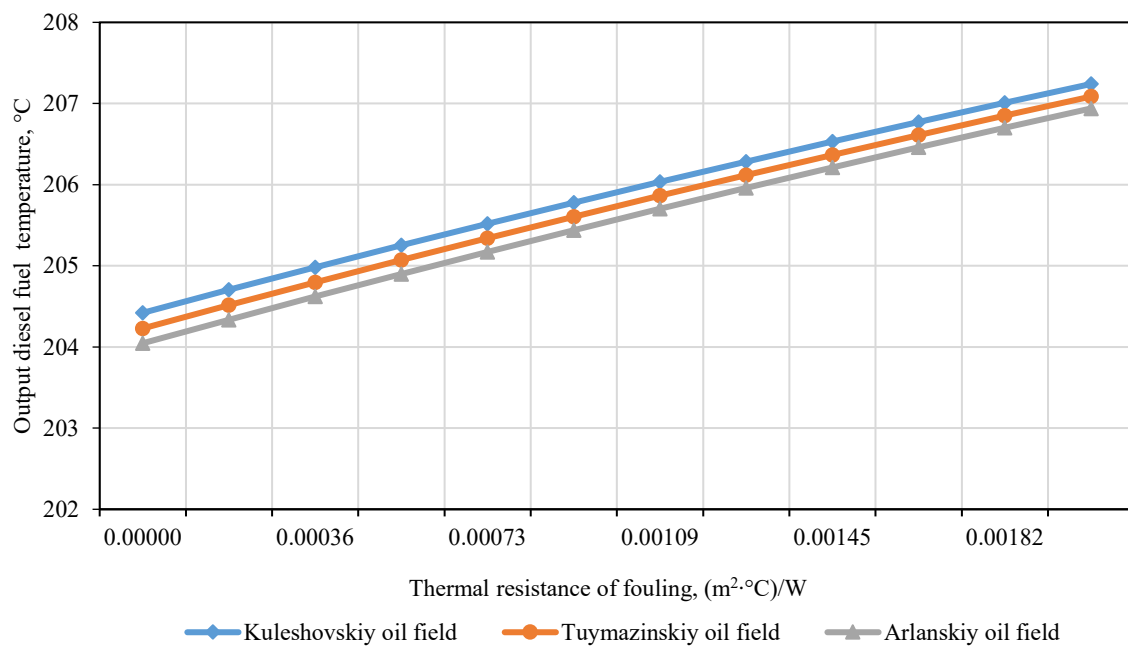


Fig. 3. Diesel fuel outlet temperature

The graphs in Figures 2 and 3 show that thermal fouling of the tube surface reduces heat transfer efficiency. The outlet temperature of the heated oil does not reach the required value, which corresponds to a clean heat exchanger. To maintain the heat flow at the specified level, it is required to increase the flow rate or temperature of the hot coolant.

It can be seen that as the thermal fouling coefficient increases, the difference and ratio between the outlet temperatures of oil and diesel fuel increase. This pattern is taken into account when deriving the thermal fouling coefficient by the SISSO method. The ratio of the outlet temperatures of the fouled heat exchanger to those of the clean heat exchanger also changes. This pattern is taken into account when applying the PySR method.

For training, PySR and SISSO are fed with the oil density and temperature at the outlet, and the diesel fuel temperature at the outlet. The target parameter is the thermal resistance of the contaminants. The Python software package is selected as the implementation tool.

Examples of mathematical relationships for determining the thermal resistance of contaminants obtained using SISSO:

$$y = k_1 \cdot ((x_2 - x_3) - (x_2 / x_1)) - k_2 \cdot \frac{\ln(x_1)}{x_2 - x_3} + k_3 \cdot \frac{x_2 / x_3}{x_2 - x_3} + k_4, \quad (5)$$

where x_1 — oil density, kg/m³; x_2 — diesel fuel outlet temperature, °C; x_3 — oil outlet temperature, °C; y — thermal resistance, (m²·°C)/W; k_1 – k_4 — coefficients.

Using PySR, a mathematical relationship is obtained for determining the thermal resistance of contaminants:

$$y = x_0 \times \log(x_1 + k_1) \cdot k_2 - \frac{k_3}{x_2}, \quad (6)$$

where x_0 — oil density, kg/m³; x_1 — ratio of the outlet temperature of diesel fuel in the contaminated heat exchanger to the outlet temperature of diesel fuel in the clean heat exchanger; x_2 — ratio of the outlet temperature of oil in the contaminated heat exchanger to the outlet temperature of oil in the clean heat exchanger; y — thermal resistance, (m²·°C)/W; k_1 – k_3 — coefficients.

To compare the accuracy of the dependences obtained, the root mean square error (RMSE) is used, which determines the average difference between the predicted and original data. The RMSE for expression (5) is $1.250371 \cdot 10^{-8}$ (m²·°C)/W. For expression (6), RMSE = $2.336612 \cdot 10^{-8}$ (m²·°C)/W.

Mathematical expressions (5, 6) are used to find the thermal resistance values of contaminants based on the measured parameters of the operating unit.

For practical applications, it is also necessary to know the values of the contaminant thermal resistance and heat transfer coefficient, determined not only at a given moment but also over a certain period. To obtain a consistent set of contaminant thermal resistance and heat transfer coefficient values, historical data is required.

The training and test samples are formed using the dynamic model of the heat exchanger given in [18], with dependences (5) or (6) added to the model. After training and testing the models, it is possible to predict the target parameter.

The Python SysIdentPy library is used to model and forecast multiple target parameter values over a given period. The NARX model is suitable for modeling dynamic systems that use both current and previous input and output values.

The FROLS and CatBoostRegressor algorithms are used to select the model structure. Lags of input (xlag) and output (ylag) variables are selected. Lag values are determined by search method.

The parameters selected for the FROLS algorithm are:

method for estimating model parameters — LeastSquares and TotalLeastSquares;

– order_selection=True — automatic selection of the model order (number of significant terms);

– n_terms=12 — maximum number of terms in the final model;

– model_type="NARMAX" — type of model;

– xlag=[1,1,1] — lags for input variables;

– ylag=12 — lags for the target variable;

– Polynomial — base class for generating polynomial functions with a selection of polynomial degree (degree=2);

– information criterion — AIC (Akaike Information Criterion).

For the CatBoostRegressor algorithm (a gradient boosting regressor), the following hyperparameters are defined:

– iterations=500 — number of trees in the ensemble;

– learning_rate=0.1 — learning rate (gradient descent step);

– depth=6 — maximum tree depth;

– xlag=[1,1,1];

– ylag=12.

In this case, a parameter prediction problem is solved, producing results that have not been previously reported. Thus, the ex-post forecast method is the most appropriate. The data are split 80/20 for model building and ex-post forecast.

After training the model, its quality is assessed using ex-post forecast. The post-prediction error for the CatBoostRegressor algorithm is $RMSE = 0.03573 \text{ W}/(\text{m}^2 \cdot ^\circ\text{C})$, for the FROLS algorithm, $RMSE = 0.01296 \text{ W}/(\text{m}^2 \cdot ^\circ\text{C})$. When forming the training and test samples, the range of variation of the thermal resistance of contaminants from 0 to $0.0022 \text{ (m}^2 \cdot ^\circ\text{C})/\text{W}$, and the oil density — from 810 to $910 \text{ kg}/\text{m}^3$ are used. The results obtained for an oil density of $900 \text{ kg}/\text{m}^3$ are shown in Figure 4.

Industrial data for 12 readings is used to validate the models. Time series of measured oil and diesel fuel temperatures are processed by a median filter. Expert assessments by process engineers (regarding the dates of scheduled heat exchanger cleanings) are taken as additional information.

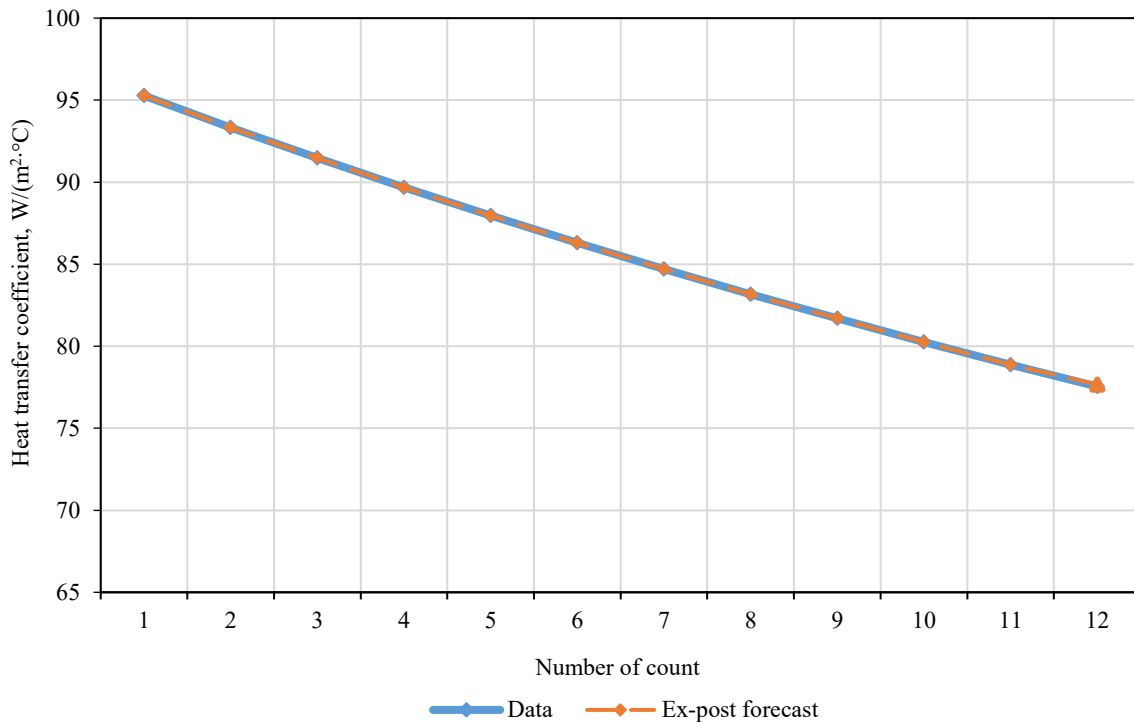


Fig. 4. Actual and predicted values of heat transfer coefficient

For a clean heat exchanger, the heat transfer coefficient is $95 \text{ W}/(\text{m}^2 \cdot ^\circ\text{C})$, and after 12 readings, it is $77 \text{ W}/(\text{m}^2 \cdot ^\circ\text{C})$. During this time, the heat transfer coefficient has decreased by 18.61%. After 18 readings (not shown in Fig. 4), the heat transfer coefficient has decreased by 25.2%, indicating the need for equipment cleaning to maintain operating efficiency. In this way, the target parameter can be predicted in the event of a change in oil density.

A comparative analysis of the heat transfer coefficient for a clean versus a fouled heat exchanger over a specified time interval indicates a loss of thermal efficiency. If the parameter falls by 25–30%, this serves as a signal that cleaning of the heat exchange unit is required.

Discussion. The results of a calibration thermal calculation of a shell-and-tube heat exchanger for a low-power oil refinery confirm the correctness of the constructed mathematical model. When the total thermal resistance of contaminants increases from 0 to $0.002 \text{ (m}^2 \cdot ^\circ\text{C})/\text{W}$, the heat transfer coefficient decreases by approximately 13–14% for all three studies of oil from different sources.

A relationship between heat transfer efficiency and the physicochemical properties of oil is established. As density increases, the heat transfer coefficient increases slightly due to higher heat capacity and thermal conductivity. However, heavier oils contain more asphaltenes, which accelerate the buildup of contaminants. Thus, the effect of density is twofold and requires further study.

The PySR and SISSO symbolic regression methods demonstrated high accuracy in reconstructing analytical relationships: RMSE were $1.25 \cdot 10^{-8}$ and $2.34 \cdot 10^{-8} \text{ (m}^2 \cdot ^\circ\text{C})/\text{W}$, respectively. The expression obtained by the SISSO method appeared to be more practically applicable, since it did not require reference values for the clean heat exchanger temperatures.

In the forecasting task, the FROLS algorithm as part of the NARX model outperformed CatBoostRegressor in accuracy (RMSE = 0.013 versus 0.036 W/(m²·°C), which is explained by the better adaptability of orthogonal regression to small samples with smooth dynamic dependences. The established cleaning criterion (a reduction in the heat transfer coefficient by 25–30%) allows for a transition from a reactive to a proactive maintenance strategy.

Conclusion. This paper presents a method for determining the thermal resistance coefficient of fouling and the heat transfer coefficient of a shell-and-tube heat exchanger for a small-capacity oil refinery by machine learning methods. In the course of the study, a verification thermal calculation of the shell-and-tube heat exchanger was performed using an iterative algorithm for solving a system of heat transfer and heat balance equations. It was shown that as the thermal fouling coefficient of the tube surface increased, the heat transfer coefficient decreased, and the outlet temperatures of the coolants deviated from the values corresponding to a clean unit. The effect of the physical and chemical properties of oil — density and viscosity — on the output parameters of the heat exchanger was also established.

Using the PySR and SISSO symbolic regression methods, analytical dependences of the thermal resistance coefficient of contaminants on the outlet temperatures of oil and diesel fuel, as well as on the oil density, were obtained. The root-mean-square error of the obtained expressions did not exceed $1.25 \cdot 10^{-8}$ and $2.34 \cdot 10^{-8}$ (m²·°C)/W, respectively, demonstrating their high accuracy.

To predict the heat transfer coefficient over time, the FROLS and CatBoostRegressor algorithms were used within the NARX model. According to the ex-post forecast results, the error for the FROLS algorithm was RMSE = 0.013 W/(m²·°C), and for CatBoostRegressor — RMSE = 0.036 W/(m²·°C). Based on the dynamics of the predicted heat transfer coefficient, it was shown that a decrease in this parameter by 25–30% relative to the value of a clean heat exchanger was a justified criterion for scheduling maintenance.

Validation of the developed models was performed using industrial data and expert assessments of technologists, which confirmed the applicability of the proposed methodology under real operating conditions.

The practical significance of this work is that the developed method uses readily available operational data (outlet temperatures and oil density) to evaluate heat exchanger fouling and predict cleaning schedules, thereby improving energy efficiency and reducing unplanned downtime in oil refining.

In the future, it is expected that the application of the method will be expanded to other types of heat exchange equipment, as well as its integration into automated process control systems.

References

1. Elistratova Y, Seminenko A, Minko V. Relevance of Contamination Models for Diagnostics of Plate Heat Exchangers. *Bulletin of Belgorod State Technological University named after V.G. Shukhov*. 2020;5(10):33–40. <https://doi.org/10.34031/2071-7318-2020-5-10-33-40>
2. Antipov AI, Golubev LG, Mukhtarov YaS. The Effect of Deposits of Solids in Heat Exchange Equipment on Thermal Resistances and Power Supply Costs in the Oil Fields Dealing with Oil Preparation. *Power Engineering: Research, Equipment, Technology*. 2003;5–6:28–35.
3. Demin AM, Naumenko AP. Diagnostics of Heat Exchange Equipment on Basis of Operating Parameters of Diesel Fuel Hydrotreating Unit. *Omsk Scientific Bulletin*. 2019;166(4):84–88. <https://doi.org/10.25206/1813-8225-2019-166-84-88>
4. Pozevalkin VV, Polyakov AN. Implementation of a Digital Model of Thermal Characteristics Based on the Temperature Field. *Advanced Engineering Research (Rostov-on-Don)*. 2024;24(2):178–189. <https://doi.org/10.23947/2687-1653-2024-24-2-178-189>
5. Gubareva KV, Eremin AV. Numerical Solution to the Problem of Thermal Conductivity in a Porous Plate with a Topology of Triply Periodic Minimal Surfaces. *Advanced Engineering Research (Rostov-on-Don)*. 2025;25(1):23–31. <https://doi.org/10.23947/2687-1653-2025-25-1-23-31>
6. Konygin SB, Konovalenko DV, Kryuchkov DA. Estimating the Effect of Instrument Accuracy on the Accuracy of Mathematical Modeling of Quality Indicators of Oil Treatment for Transportation. *IOP Conference Series: Materials Science and Engineering*. 2020;862(3):032068. <https://doi.org/10.1088/1757-899X/862/3/032068>
7. Martins T, Spengler AW, Oliveira JLG, de Paiva KV, Seman LO. Active Control System to Prevent Malfunctioning Caused by the Pressure Difference in Gasket Plate Heat Exchangers Applied in the Oil and Gas Industry. *Sensors*. 2022;22(12):4422. <https://doi.org/10.3390/s22124422>

8. Gil'fanov KK, Shakirov RA. Neural Network Modeling of Thermal Hydraulic Efficiency of Promising Surface Heat Transfer Intensifiers. *Russian Aeronautics*. 2021;64(1):61–70. <https://doi.org/10.3103/S1068799821010086>
9. Tugashova LG, Zatonskiy AV. Comparison of Methods for Heat Exchanger Control. *Theoretical Foundations of Chemical Engineering*. 2021;55(1):53–61. <https://doi.org/10.1134/S0040579520060226>
10. Sahoo A, Radhakrishnan TK, Sankar Rao C. Modeling and Control of a Real Time Shell and Tube Heat Exchanger. *Resource-Efficient Technologies*. 2017;3(1):124–132. <https://doi.org/10.1016/j.reffit.2016.12.001>
11. Glagoleva OF, Kapustin VM, Piskunov IV, Usmanov MR. Controlling the Aggregative Stability of Feedstock Blends and Petroleum Products. *Petroleum Chemistry*. 2020;60(5):971–978. <https://doi.org/10.1134/S0965544120090108>
12. Kapustin VM, Rudin MG, Kukes SG. *Handbook of Oil Refiner*. Moscow: Khimiya; 2018. 416 p. (In Russ.)
13. Tugashova LG. *Increasing Energy Efficiency and Productivity at Fuel and Energy Complex Facilities Using Modeling and Optimization Methods*. Monograph. Almet'yevsk: ASTU "HSO"; 2025. 101 p. (In Russ.)
14. Deryabin IA, Korolev VV, Sorokin GS. Using a Recurrent Neural Network for Solving Inverse Heat Conduction Problem with Application to Calculating the Temperature of Equipment of VVER-Based NPPs. *Nuclear Power Engineering*. 2024;(4):144–154. <https://doi.org/10.26583/npe.2024.4.12>
15. Xiaolin Jiang, Guanqi Liu, Jiaying Xie, Zhenpeng Hu. Boosting SISSO Performance on Small Sample Datasets by Using Random Forests Prescreening for Complex Feature Selection. [arXiv:2409.19209](https://arxiv.org/abs/2409.19209)
16. Muthyala M, Sorourifar F, Paulson JA. TorchSISSO: A PyTorch-Based Implementation of the Sure Independence Screening and Sparsifying Operator for Efficient and Interpretable Model Discovery. *Digital Chemical Engineering*. 2024;13:100198. <https://doi.org/10.1016/j.dche.2024.100198>
17. Tonda A. Review of PySR: High-Performance Symbolic Regression in Python and Julia. *Genetic Programming and Evolvable Machines*. 2025;26:7. <https://doi.org/10.1007/s10710-024-09503-4>
18. Zatonskiy AV, Tugashova LG. Identification of Parameters of a Dynamic Model of a Heat Exchange Apparatus. *Perm National Research Polytechnic University Bulletin. Electrotechnics, Information Technologies, Control Systems*. 2020;33:99–114. <https://doi.org/10.15593/2224-9397/2020.1.06>

About the Authors:

Larisa G. Tugashova, Cand.Sci. (Eng.), Associate Professor of the Department of Automation and Information Technologies, Almet'yevsk State Technological University "Higher School of Petroleum" (186 a, Sovetskaya Str., Almet'yevsk, 423462, Russian Federation), [SPIN-код](#), [ORCID](#), [ScopusID](#), [ResearcherID](#), tugashova@yandex.ru

Andrey V. Zatonskiy, Dr.Sci. (Eng.), Professor, Head of the Department of Automation and Information Technologies, Perm National Research Polytechnic University, Berezniki Branch (29, Komsomolsky Pr, Perm, 614990, Russian Federation), [SPIN-код](#), [ORCID](#), [ScopusID](#), [ResearchGate](#), z xenon@narod.ru

Claimed Contributorship:

LG Tugashova: formal analysis, investigation, methodology, writing – original draft preparation, writing – review & editing.

AV Zatonskiy: formal analysis, writing – original draft preparation, writing – review & editing.

Conflict of Interest Statement: the authors declare no conflict of interest.

The authors have read and approved the final version of the manuscript.

Об авторах:

Лариса Геннадьевна Тугашова, кандидат технических наук, доцент кафедры «Автоматизация и информационные технологии» Альметьевского государственного технологического университета «Высшая школа нефти» (423462, Российская Федерация, г. Альметьевск, ул. Советская, 186 а), [SPIN-код](#), [ORCID](#), [ScopusID](#), [ResearcherID](#), tugashova@yandex.ru

Андрей Владимирович Затонский, доктор технических наук, профессор, заведующий кафедрой «Автоматизация и информационные технологии» Березниковского филиала Пермского национального исследовательского политехнического университета (614990, Российская Федерация, г. Пермь, пр. Комсомольский, 29), [SPIN-код](#), [ORCID](#), [ScopusID](#), [ResearchGate](#), z xenon@narod.ru

Заявленный вклад авторов:

Л.Г. Тугашова: формальный анализ, проведение исследования, разработка методологии, написание черновика рукописи, написание рукописи – рецензирование и редактирование.

А.В. Затонский: формальный анализ, написание черновика рукописи, написание рукописи – рецензирование и редактирование.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Авторы прочитали и одобрили окончательный вариант рукописи.

Received / Поступила в редакцию 16.02.2026

Reviewed / Поступила после рецензирования 13.03.2026

Accepted / Принята к публикации 23.03.2026