Machine learning in environmental sustainability factor analysis in the agricultural sector

Vladislav Kukartsev[1,2], *Anastasia* Kozlova[2,3*], *Svetlana* Kukarceva[1]

[1]Moscow Timiryazev Agricultural Academy, Russian State Agrarian University, 127550 Moscow, Russia
[2]Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia
[3]Bauman Moscow State Technical University, 105005 Moscow, Russia

Abstract. The study employed several key data analysis methods aimed at enhancing the understanding of relationships between variables and improving prediction accuracy. The primary tool used was correlation analysis, which allowed for the identification of the degree of association between two variables by determining how changes in one variable relate to changes in another. This established a foundation for further in-depth data analysis. For a deeper understanding and simplified interpretation of the data, factor analysis was utilized. This method helped to identify latent factors that explain the relationships between observed variables and to reduce the number of variables by grouping them. This made the analysis easier and facilitated the identification of key components affecting the data. Logistic regression was applied to build data models. This method is used to model the probability of a specific event occurring based on independent variables, allowing for the classification and prediction of categorical outcomes. The logistic function was used to estimate probabilities and the relationship between the dependent variable and predictors. To enhance the performance of the logistic regression model, a Weight of Evidence (WoE) analysis was conducted. This method converts categorical and continuous variables into numerical formats, simplifying data interpretation and improving the model's predictive capabilities. WoE analysis helps to identify significant factors, improve the linear relationship between predictors and the dependent variable, and reduce the impact of outliers, which is particularly important in areas such as credit scoring. The results of applying these methods showed that the model based on correlation and factor analysis explained 27.51% of the information on the training set and 76.04% on the test set.

1 Introduction

In recent decades, there has been increasing attention to environmental protection issues and their impact on biological processes. One of the key research areas is the study of the relationship between environmental indicators and plant growth. Plants, being fundamental

* Corresponding author: ankoz9@yandex.ru

components of ecosystems, play a critically important role in maintaining ecological balance, and their development and productivity are closely related to environmental conditions.

Analyzing data on environmental indicators such as temperature, humidity, light intensity, and soil composition provides a unique opportunity to understand how these factors influence plant growth and development. This study examines a dataset containing information about various environmental parameters in which plants grew, as well as their growth and condition indicators.

The aim of this research is to conduct a detailed analysis of the collected data to identify key environmental factors that have the greatest impact on plant development. Such an analysis will not only enhance understanding of ecological relationships but also provide recommendations for optimizing growth conditions for plants in various environmental contexts.

The study focuses on identifying patterns and dependencies that will help predict the impact of environmental changes on plant communities and develop strategies for their sustainable management and conservation.

2 Materials and methods

One of the methods used in this study was correlation analysis. This analysis is applied to examine the degree of association between two variables. It determines how strongly and in what direction changes in one variable are related to changes in another. This method is useful for identifying potential dependencies and patterns in the data, which can serve as a foundation for further in-depth analysis and the development of more complex models. Correlation helps to quickly assess whether a relationship exists between variables and is an important tool in preliminary data processing [1-4].

The second method used to explore relationships between the data was factor analysis, which is employed to identify latent factors that explain the relationships between observed variables. This method allows for the reduction of the number of variables by grouping them into clusters that reflect the underlying structures in the data. Factor analysis helps to uncover the main components affecting the data and simplifies the interpretation of results. This is especially useful in situations where it is necessary to understand complex multidimensional data and identify key variables that should be focused on.

For constructing the data model, linear regression was used. Logistic regression is applied to model the probability of an event occurring based on one or more independent variables [5-7]. This method is useful for classification and predicting categorical outcomes, such as the presence or absence of a specific attribute. Logistic regression evaluates the relationship between the dependent variable and predictors using the logistic function to forecast probabilities, which allows for the interpretation of results in terms of probabilities and makes it suitable for analyzing binary data [8].

To improve the performance of the model built using logistic regression, Weight of Evidence (WoE) analysis was utilized. This analysis is used to convert categorical and continuous variables into a format that simplifies model interpretation and enhances its predictive capabilities. The method involves calculating the ratio between the probability of an event occurring and the probability of it not occurring for each category of the variable [9-14]. This conversion is achieved by computing the logarithm of the odds ratio (WoE) for each category, which transforms categorical data into numerical values that are easier to interpret and use in the model. WoE analysis helps to identify important factors, improve the linear relationship between predictors and the dependent variable, and reduce the impact of outliers. This approach is often used in credit scoring for risk assessment, as it provides a clear relationship between variables and their effect on the target variable.

3 Data structure

The data being studied consists of information about the environmental conditions of the plants under observation. The dataset includes seven fields, the first of which is Soil_Type, which contains information about the type of soil in which the plants are growing. This field has three categories: clay, loam, sandy.

Clay soil has very fine particles of clay, making it dense and heavy. It retains water well, but this can lead to waterlogging issues, sometimes resulting in root rot. However, clay soil contains many nutrients due to its ability to hold them. Nevertheless, its density can impede aeration, making it difficult for roots to get enough oxygen [15-18].

Loam soil is a mixture of clay, sand, and silt, making it one of the most fertile and workable types of soil. It has a good structure: not too dense and not too loose. Loam soil retains water well while also having good drainage capabilities, helping to prevent waterlogging. It also provides good aeration, allowing plant roots to get sufficient oxygen and nutrients.

Sandy soil consists of large sand particles and has a loose structure. This soil drains easily, allowing water to pass through quickly, but it also loses moisture rapidly, which may require frequent watering. Sandy soil does not retain many nutrients, so additional fertilization is often needed to support plant growth. Its loose structure provides good aeration, promoting a healthy root system.

The next field, Sunlight_Hours, represents the amount of time per day that plants receive sunlight: the range starts from 4.033 and ends at 9.914.

The amount of sunlight plants receive significantly affects their health and development. Plants receiving around 4 hours of sunlight per day generally suit shade-loving species that can adapt to lower light conditions. Such conditions are often found in the shadow of large trees or in northern window openings.

With 6-7 hours of sunlight per day, most plants thrive and grow healthily. This duration is ideal for most garden and indoor plants, as it provides sufficient photosynthesis for active growth and productivity.

At 9 hours of sunlight per day, plants receive maximum light, which is beneficial for sun-loving crops. Such conditions can be favorable for many vegetables, fruits, and flowering plants but require careful control of temperature and humidity to prevent overheating and soil drying.

The information on watering includes the field Water_Frequency, which has three data options: daily, weekly, bi-weekly. The frequency of watering plants can vary depending on plant species, soil type, climate conditions, and the time of year.

Daily watering is often required for plants due to high temperatures and dry conditions, where the soil dries out quickly. Daily watering is particularly relevant for plants in greenhouses, on balconies, or in containers, where the soil may dry out faster. It may also be necessary for some heat-loving crops and in regions with hot climates.

Weekly watering suits most garden plants and trees, especially in temperate climates. This is a frequent enough schedule to provide plants with the necessary moisture without causing waterlogging. It is important to consider weather conditions and occasional rains, which can reduce the need for watering.

Bi-weekly watering may be appropriate for more drought-tolerant plants or when the soil has good water-holding capacity. This schedule is often suitable for the fall-winter period, when plants may have lower water needs. It can also be an option for some indoor plants or in moderately humid climates.

Fertilizer_Type refers to the type of fertilizers used to enhance plant growth: organic, chemical, none.

Organic fertilizers include natural substances such as compost, manure, and humus. They improve soil structure, promote the accumulation of organic matter, and support microorganism activity, which ultimately helps plants better absorb nutrients. This type of fertilizer is more sustainable and environmentally friendly.

Chemical fertilizers, on the other hand, contain synthetic components like nitrogen, phosphorus, and potassium in concentrated forms. They provide quick and efficient delivery of essential nutrients to plants but can cause harmful substances to accumulate in the soil and water if used improperly. Chemical fertilizers can also negatively impact soil health and ecosystems overall [19, 20].

In the absence of additives, plants rely solely on natural soil conditions, which may be suitable for hardy plant species or fertile soils. However, without fertilizers, the soil may become depleted over time, leading to reduced productivity and overall plant health.

Temperature refers to the air temperature in which the plants are growing: the range starts from 15.2 and ends at 34.8. At 15 degrees Celsius, plants generally feel quite comfortable, especially those that prefer cooler conditions, such as lettuce, spinach, and some leafy greens. This range is suitable for most spring and fall plants, as well as those grown in cooler climates.

In the range of 20 to 25 degrees Celsius, plants often grow optimally. This temperature range is standard for most garden and indoor plants, including vegetables, flowers, and ornamental plants. In these conditions, plants develop actively, flower, and produce well.

At temperatures between 30 and 35 degrees Celsius, plants can experience stress. This often occurs on hot summer days or in tropical climates. Although some plants, such as cacti and other succulents, can tolerate such temperatures, many others may suffer from overheating, wilting, or reduced productivity. In such conditions, adequate watering and possibly additional shading are essential to protect plants from overheating and drying out.

Humidity refers to the level of air moisture in which the plants grew: the range starts from ~30.6 and ends at ~79.6. At around 30% humidity, plants may experience stress, particularly those that prefer more humid conditions. This often occurs in dry climates or in spaces with low humidity. Such conditions can lead to rapid moisture loss from the soil and plants, requiring frequent watering and additional measures to maintain humidity.

A humidity level between 40-60% is considered comfortable for most plants. In this range, plants generally grow well, as it provides adequate moisture and maintains optimal evaporation rates. This is particularly important for indoor plants and many garden crops, helping to prevent issues related to excessively dry air [21-23].

At a humidity level around 80% and higher, plants, especially tropical and subtropical species, are in comfortable conditions. High humidity supports healthy growth, enhances photosynthesis, and reduces the likelihood of drying out. However, very high humidity conditions can also pose risks for fungal diseases and mold, requiring regular monitoring and good ventilation.

Growth_Milestone is a field indicating key development stages of the plant: it has values 0 and 1. It represents the target field.

Figure 1 shows the distribution of the data fields Sunlight_Hours, Temperature, and Humidity.

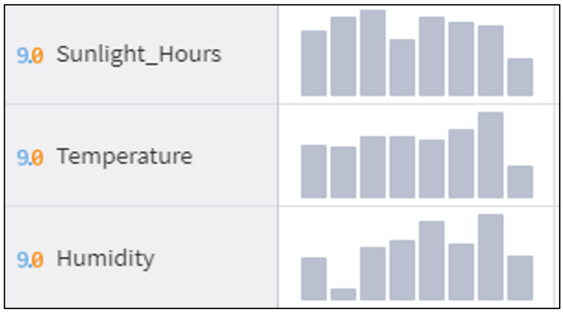

Fig.1. Data distribution

4 Results

Before conducting the data analysis, the information contained in the dataset was subjected to normalization and categorization. Discrete values were replaced with numerical values – classes, and Min-Max Scaling was applied to continuous values.

Correlation analysis revealed a negative moderate relationship between the fields Fertilizer_Type and Growth_Milestone, a negative weak relationship between the fields Humidity and Growth_Milestone, Temperature and Sunlight_Hours, and Sunlight_Hours and Growth_Milestone. A positive weak relationship was identified only between the fields Humidity and Temperature [24, 25].

FertilizerNorm	Growth_Milestone	-0,3328924598
Growth_Milestone	FertilizerNorm	-0,3328924598
HumidNorm	Growth_Milestone	-0,1410764137
Growth_Milestone	HumidNorm	-0,1410764137
TemperNorm	SunHoursNorm	-0,1164709636
SunHoursNorm	TemperNorm	-0,1164709636
SunHoursNorm	Growth_Milestone	-0,1103952838
Growth_Milestone	SunHoursNorm	-0,1103952838
HumidNorm	TemperNorm	0,1115946223
TemperNorm	HumidNorm	0,1115946223

Fig.2. Correlation coefficients

The significant fields for the first factor were air temperature and humidity, as well as watering and sunlight exposure, but with an inverse relationship. For the second factor, the significance of the field related to the type of fertilizers used, as well as the duration of sunlight exposure, can be highlighted [26].

The results of the factor analysis are presented in Figure 3.

 https://doi.org/10.1051/bioconf/202414104050

#	ab Name	9.0 Factor 1	9.0 Factor 2
1	SunHoursNorm	-0,4512486907	0,3577163634
2	WaterFreqNorm	-0,4379676715	0,07643851134
3	FertilizerNorm	0,04071484965	0,9296668002
4	HumidNorm	0,5294656706	-0,04307415108
5	TemperNorm	0,7324123759	0,1772012631

Fig.3. Factor analysis results

After identifying significant features, a data model was built using linear regression. One of the tools used to evaluate the model is the ROC curve (Receiver Operating Characteristic curve).

The ROC curve is a graphical tool for assessing the performance of binary classifiers, including linear regression models when their output is used for classification. It shows the trade-off between true positive and false positive results at various classification thresholds. The X-axis represents the false positive rate (FPR), while the Y-axis represents the true positive rate (TPR), also known as sensitivity [27]. The ROC curve allows for a visual assessment of how well the model distinguishes between positive and negative classes: the higher and more to the left the curve, the better the model. The area under the ROC curve (AUC) is a numerical measure of model quality: an AUC value close to 1 indicates excellent model ability to distinguish between classes, while a value close to 0.5 suggests that the model performs no better than random guessing [28].

The ROC curve is presented in Figure 4.

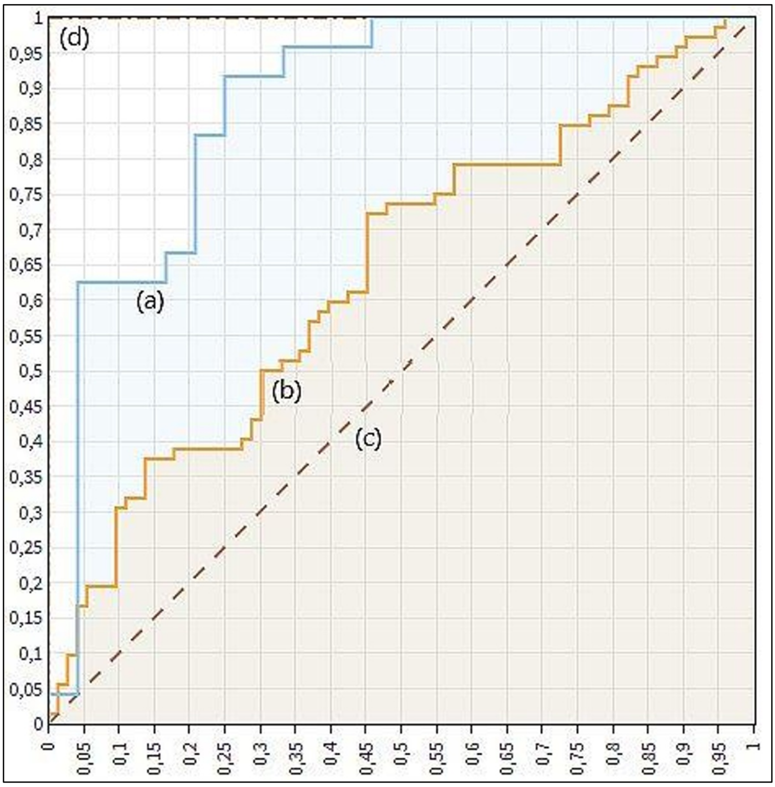

Fig. 4. ROC Curve for the first model: a – Test Set, b – Training Set, c – Baseline, d – Ideal Line

The constructed model, as shown in Figure 5, correctly identified only 86 out of 145 records in the training set and 36 out of 48 records in the test set [29, 30].

As seen in Figure 6, the constructed model explains 27.51% of the data in the training set and 76.04% in the test set. These figures indicate the low quality of the constructed model and the need for further refinement.

Classified	Actual		Total
	Event	Non-Event	
Training	72	73	
Event	40	27	67
Non-Event	32	46	78
Testing	24	24	
Event	13	1	14
Non-Event	11	23	34
Recognized			
Training	86/145		
Testing	36/48		

Fig. 5. Confusion matrices for the first model

Indicator	Sets	
	Training	Testing
Classifier ratings		
AUC ROC	0,6376	0,8802
AUC PR	0,6376	0,8249
Gini coefficient	0,2751	0,7604
KS	27,0167	66,6667
Cutoff threshold		
Value	0,5000	0,5000
TPR (Sensitivity)	0,5556	0,5417
TNR (Specificity)	0,6301	0,9583
FPR (1-Specificity)	0,3699	0,0417
PPV	0,6003	0,9286
F1 Score	0,5771	0,6842
MCC	0,1862	0,5500

Fig. 6. Classification metrics for the first model

To build the second model, the original data were categorized using WoE analysis. Figures 7 and 8 display examples of the output from this analysis [31, 32]. Figure 7 shows that the model has fewer highly significant values compared to less significant ones.

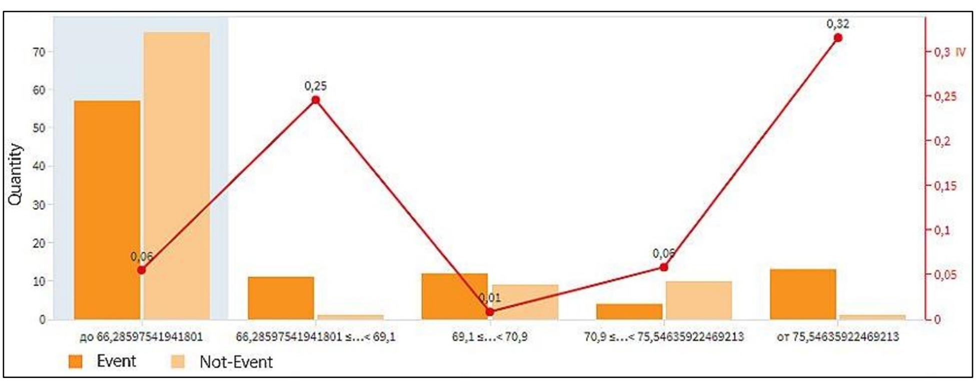

Fig. 7. Importance of data elements in the humidity field for the model

In Figure 8, it can be seen that the contribution assessment of variations in the selected field, which are significant for the model, is lower compared to less significant ones. As a result, the fields with high significance are Humidity, Temperature, and Fertilizer_Type, while Sunlight_Hours has medium significance, and Soil_Type and Water_Frequency have low or no significance, respectively. The significance of attributes when divided into final classes is presented in Figure 9.

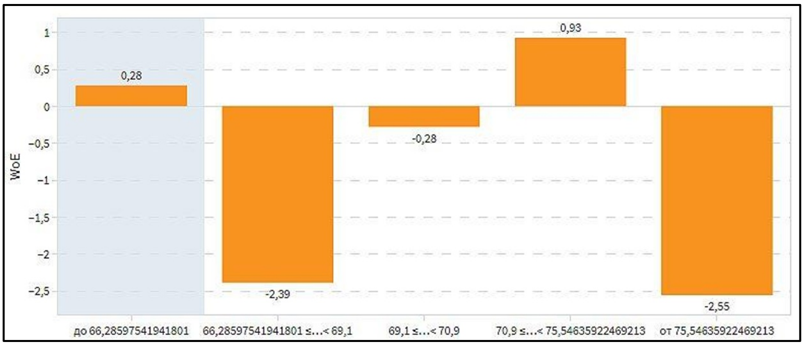

Fig. 8. WoE contribution assessment for variations in the Humidity field

#	Name	Event share	Not-Event share	Information index	Significance
1	Humidity	0,5025906736	0,4974093264	0,6834039494	High
2	Temperature	0,5025906736	0,4974093264	0,4885236901	High
3	Fertilizer_Type	0,5025906736	0,4974093264	0,4801638389	High
4	Sunlight_Hours	0,5025906736	0,4974093264	0,2273767445	Average
5	Soil_Type	0,5025906736	0,4974093264	0,03791815382	Low
6	Water_Frequency	0,5025906736	0,4974093264	0,002746730981	None

Fig. 9. Attribute significance when divided into final classes

Based on the conducted analysis and categorization, a new model was built with better performance compared to the previous model. Figure 10 presents the ROC curve for the second model, which shows an improvement in model quality.

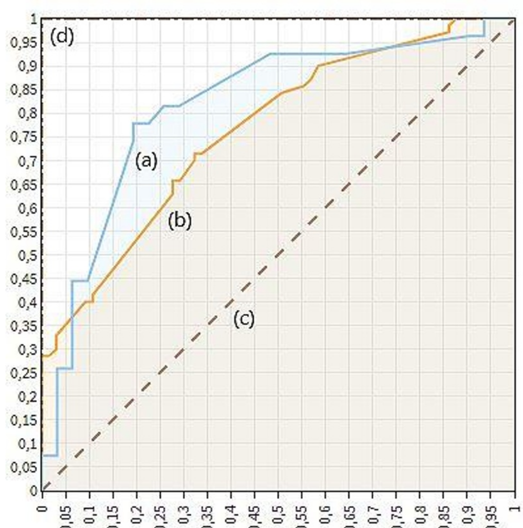

Fig. 10. ROC curve for the second model: a – Test Set, b – Training Set, c – Baseline, d – Ideal Line

As seen in Figure 11, 102 out of 145 values were correctly identified in the training set and 39 out of 48 in the test set. By changing the strategy, it was possible to improve the performance on the training set. As shown in Figure 12, the model explains 52.97% of the data in the training set and 77.74% in the test set.

Classified	Actual		Total
	Event	Non-Event	
Training	74	71	
Event	38	7	45
Non-Event	36	64	100
Testing	23	25	
Event	15	1	16
Non-Event	8	24	32
Recognized			
Training	102/145		
Testing	39/48		

Fig. 11. Confusion matrices for the second model

Indicator	Sets	
	Training	Testing
Classifier ratings		
AUC ROC	0,7648	0,8887
AUC PR	0,7696	0,8356
Gini coefficient	0,5297	0,7774
KS	50,9517	69,9130
Cutoff threshold	эк узла	
Value	0,5000	0,5000
TPR (Sensitivity)	0,5135	0,6522
TNR (Specificity)	0,9014	0,9600
FPR (1-Specificity)	0,0986	0,0400
PPV	0,8273	0,9375
F1 Score	0,6337	0,7692
MCC	0,4535	0,6487

Fig. 12. Classification metrics for the second model

Conclusion

Several key data analysis methods were employed in the study. First, correlation analysis was used to assess the degree of association between two variables. This method allowed for determining how changes in one variable relate to changes in another, which helped to identify potential dependencies and patterns, creating a basis for further in-depth analysis.

Second, factor analysis was applied, focusing on identifying latent factors that explain the relationships between observed variables. This method reduced the number of variables by grouping them, simplifying data interpretation and identifying key components that influence the data.

To build the data models, logistic regression was used, which models the probability of an event occurring based on independent variables. Logistic regression helps classify and predict categorical outcomes by assessing the relationship between the dependent variable and predictors and using the logistic function to forecast probabilities.

To improve the performance of the logistic regression model, a WoE analysis was conducted. This method transforms categorical and continuous variables into a format that simplifies interpretation and enhances the model's predictive capabilities. The transformation involves calculating the logarithm of the odds ratio for each category, converting the data into numerical values. WoE analysis helps identify important factors, improve the linear relationship between predictors and the dependent variable, and reduce the impact of outliers. This method is frequently used in credit scoring to assess risk and establish a clear connection between variables and their impact on the target variable.

As a result of applying these methods, the first model, based on correlation and factor analysis, explained 27.51% of the information in the training set and 76.04% in the test set. The second model, using the WoE method without data normalization, showed higher results: 52.97% in the training set and 77.74% in the test set. It is recommended to expand the dataset to enhance the accuracy and reliability of the models, which will help to better identify patterns and improve generalizability.

Expanding the dataset is recommended to increase the accuracy and reliability of the models. Increasing the volume of data can significantly improve model quality by allowing them to more accurately identify patterns and dependencies. Specifically, expanding the dataset may lead to a more comprehensive coverage of all possible classes and scenarios, ensuring better generalizability and resilience of the models to new, previously unseen situations.

Furthermore, increasing the volume of data may facilitate a more detailed examination of the impact of various factors on the results and provide an opportunity to apply more complex analytical methods, which could further enhance model performance. For the first model, this might involve re-conducting correlation and factor analyses with a broader set of variables, leading to a better understanding of their interrelationships. For the second model, it could mean revisiting and refining the class partitions and reviewing WoE analyses with a more extensive dataset.

Additionally, expanding the dataset may help improve class balance and reduce the risk of overfitting the models. This will allow for more stable and predictable results under different conditions and on new data. As a result, a more extensive and diverse dataset will contribute to creating more accurate and reliable models, which can provide higher-quality predictions and a deeper understanding of the processes being studied.

References

1. I. Bosikov, N. V. Martyushev, R. V. Klyuev, I. A. Savchenko, V. V. Kukartsev, V. A. Kukartsev, Y. A. Tynchenko, Modeling and complex analysis of the topology parameters of ventilation networks when ensuring fire safety while developing coal and gas deposits, Fire, 1-15 (2023), DOI:10.3390/fire6030095.
2. B. V. Malozyomov, N. V. Martyushev, V. A. Kukartsev, V. V. Kukartsev, S. V. Tynchenko, R. V. Klyuev, N. A. Zagorodnii, Y. A. Tynchenko, Study of supercapacitors built in the start-up system of the main diesel locomotive, Energies, 1-15 (2023), DOI:10.3390/en16093909.

3. D. Strateichuk, N. Martyushev, R. Klyuev, V. Gladkikh, V. Kukartsev, Y. Tynchenko, A. Karlina, Morphological features of polycrystalline CdS1−xSex films obtained by screen-printing method, Crystals, 1-15 (2023), DOI:10.3390/cryst13050825.
4. N. V. Martyushev, B. V. Malozyomov, I. H. Khalikov, V. A. Kukartsev, V. V. Kukartsev, V. S. Tynchenko, Y. A. Tynchenko, M. Qi, Review of methods for improving the energy efficiency of electrified ground transport by optimizing battery consumption, Energies, 1-15 (2023), DOI:10.3390/en16020729.
5. Shutaleva, N. Martyushev, Z. Nikonova, I. Savchenko, V. Kukartsev, V. Tynchenko, Y. Tynchenko, Sustainability of inclusive education in schools and higher education: Teachers and students with special educational needs, Sustainability, 1-15 (2023), DOI:10.3390/su15043011.
6. V. A. Kukartsev, V. V. Kukartsev, V. S. Tynchenko, S. O. Kurashkin, R. B. Sergienko, S. V. Tynchenko, I. A. Panfilov, S. V. Eremeeva, T. A. Panfilova, Study of the influence of the thermal capacity of the lining of acid melting furnaces on their efficiency, Metals, 1-15 (2023), DOI:10.3390/met13020337.
7. N. V. Martyushev, D. A. Bublik, V. V. Kukartsev, V. S. Tynchenko, R. V. Klyuev, Y. A. Tynchenko, Y. I. Karlina, Provision of rational parameters for the turning mode of small-sized parts made of the 29 NK alloy and beryllium bronze for subsequent thermal pulse deburring, Materials, 1-15 (2023), DOI:10.3390/ma16093490.
8. T. Panfilova, V. Kukartsev, V. Tynchenko, Y. Tynchenko, O. Kukartseva, I. Kleshko, X. Wu, I. Malashin, Flood susceptibility assessment in urban areas via deep neural network approach, Sustainability, 1-15 (2024), DOI:10.3390/su16177489.
9. Y. Tynchenko, V. Kukartsev, V. Tynchenko, O. Kukartseva, T. Panfilova, A. Gladkov, V. Nguyen, I. Malashin, Landslide assessment classification using deep neural networks based on climate and geospatial data, Sustainability, 1-15 (2024), DOI:10.3390/su16167063.
10. Y. Tynchenko, V. Kukartsev, A. Gladkov, T. Panfilova, Assessment of technical water quality in mining based on machine learning methods, Sustainable Development of Mountain Territories, 1-15 (2024), DOI:10.21177/1998-4502-2024-16-1-56-69.
11. V. Kukartsev, K. Kravtsov, Y. Tynchenko, T. Panfilova, Influence of mountain factors on salt excess and soil toxicity in mountain conditions, Sustainable Development of Mountain Territories, 30 September 2023, DOI: 10.21177/1998-4502-2023-15-3-784-797.
12. K. Yelemessov, M. Isametova, N. Saydinbayeva, V. Kukartsev, Mathematical and computer modeling of gantry crane load-beam system oscillation, Sustainable Development of Mountain Territories, 30 June 2023, DOI: 10.21177/1998-4502-2023-15-2-450-461.
13. V. Brigida, V. I. Golik, E. V. Voitovich, V. V. Kukartsev, V. E. Gozbenko, V. Y. Konyukhov, T. A. Oparina, Technogenic reservoirs resources of mine methane when implementing the circular waste management concept, Resources, 17 February 2024, DOI: 10.3390/resources13020033.
14. T. Panfilova, A. Gladkov, V. Tynchenko, O. Kukartseva, M. Kvesko, Development and optimization of simulation model of decision aid system implementation, 2023 4th International Conference on Intelligent Technologies (CONIT), IEEE, 21 June 2024, DOI:10.1109/conit61985.2024.10626315.
15. N. Fedorova, I. Panfilov, E. Moiseeva, K. Degtyareva, E. Markovskaya, Analytical methods and tools for business process optimization, BIO Web of Conferences, 2024, DOI:10.1051/bioconf/202411305009.

16. A. Boyko, A. Kukartsev, T. Dolgova, K. Lobkov, A. Bagaeva, Simulation model of the urban population taking into account the influence of the manufacturing sector based on big data, 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), IEEE, 15 March 2023, DOI:10.1109/infoteh57020.2023.10094081.
17. N. V. Martyushev, B. V. Malozyomov, I. H. Khalikov, V. A. Kukartsev, V. V. Kukartsev, V. S. Tynchenko, Y. A. Tynchenko, M. Qi, Review of methods for improving the energy efficiency of electrified ground transport by optimizing battery consumption, Energies, 8 January 2023, DOI: 10.3390/en16020729.
18. V. A. Kukartsev, V. V. Kukartsev, V. S. Tynchenko, S. O. Kurashkin, R. B. Sergienko, S. V. Tynchenko, I. A. Panfilov, S. V. Eremeeva, T. A. Panfilova, Study of the influence of the thermal capacity of the lining of acid melting furnaces on their efficiency, Metals, 8 February 2023, DOI: 10.3390/met13020337.
19. B. V. Malozyomov, N. V. Martyushev, V. V. Kukartsev, V. Y. Konyukhov, T. A. Oparina, N. S. Sevryugina, V. E. Gozbenko, V. V. Kondratiev, Determination of the performance characteristics of a traction battery in an electric vehicle, World Electric Vehicle Journal, 12 February 2024, DOI: 10.3390/wevj15020064.
20. R. Alena, V. Kukartsev, A. Glinscaya, I. Pinchuk, A. Gladkov, Designing a system of step-by-step quality control, 2024 12th International Conference on Smart Grid (icSmartGrid), IEEE, 1-6 (2024), DOI:10.1109/icsmartgrid61824.2024.10578172.
21. K. Degtyareva, D. A. Ageev, V. V. Kukartsev, Finding patterns in employee attrition rates using self-organizing Kohonen maps and decision trees, 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), IEEE, 1-6 (2023), DOI:10.1109/icses60034.2023.10465548.
22. V. Tynchenko, K. Kravtsov, N. Podanyov, A. Fedorova, A. Bezvorotnykh, Automation for the sustainable development of agriculture, BIO Web of Conferences, 2024, DOI:10.1051/bioconf/202411305010.
23. K. Degtyareva, V. Tynchenko, V. Kukartsev, V. Khramkov, Use of computer simulation tools to simulate processes at the foundry, 2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH), IEEE, 20 March 2024, DOI:10.1109/infoteh60418.2024.10495977.
24. V. V. Gerasidi, R. G. Dubrovin, V. V. Kukartsev, T. A. Panfilova, G. V. Stas, Boost system diagnostic parameters of coherent gas piston installations of mining enterprises, Series of Geology and Technical Sciences, 15 August 2024, DOI: 10.32014/2024.2518-170x.423.
25. V. O. Gutarevich, N. V. Martyushev, R. V. Klyuev, V. A. Kukartsev, V. V. Kukartsev, L. V. Iushkova, L. N. Korpacheva, Reducing oscillations in suspension of mine monorail track, Applied Sciences, 7 April 2023, DOI: 10.3390/app13084671.
26. V. Kukartsev, K. Kravtsov, O. Stefanenko, N. Podanyov, A. Bezvorotnykh, Using machine learning techniques to simulate network intrusion detection, 2024 International Conference on Intelligent Systems for Cybersecurity (ISCS), 3 May 2024, DOI: 10.1109/iscs61804.2024.10581097.
27. K. Degtyareva, D. A. Ageev, V. V. Kukartsev, Finding patterns in employee attrition rates using self-organizing Kohonen maps and decision trees, 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 14 December 2023, DOI: 10.1109/icses60034.2023.10465548.
28. V. Orlov, V. Tynchenko, A. Nizameeva, D. Shalaeva, D. Ageev, Development of a multifunctional cross-platform system for automation of energy data and resource management, E3S Web of Conferences, 2023, DOI:10.1051/e3sconf/202346007002.

29. K. Kravtsov, V. Tynchenko, E. Semenova, D. Shalaeva, I. Pinchuk, Workflow automation and performance improvement based on PostgreSQL, E3S Web of Conferences, 2023, DOI:10.1051/e3sconf/202345809022.
30. E. Volneikina, O. Kukartseva, A. Menshenin, V. Tynchenko, K. Degtyareva, Simulation-dynamic modeling of supply chains based on big data, 2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH), IEEE, 15 March 2023, DOI:10.1109/infoteh57020.2023.10094168.
31. I.P. Boychuk, A.V. Grinek, N.V. Martyushev, R.V. Klyuev, B.V. Malozyomov, V.S. Tynchenko, V.A. Kukartsev, Y.A. Tynchenko, S.I. Kondratiev, A methodological approach to the simulation of a ship's electric power system, Energies, 16 December 2023, DOI:10.3390/en16248101.
32. T. Panfilova, V. Tynchenko, O. Kukartseva, A. Kozlova, A. Glinscaya, Modernization of electronic document management and systems analysis processes using an automated platform, E3S Web of Conferences, 2024, DOI:10.1051/e3sconf/202454909018.