

The object of the study is the power supply system of the mining enterprise Nova-Zinc LLP, located in Central region of Republic of Kazakhstan, which specializes in the extraction and enrichment of lead-zinc ores. Reliable power supply in mining is challenged by heavy mechanical loads, vibration, and severe weather; this case study is used only as an example of conditions common in many mining regions worldwide. This study proposes a diagnostic-and-forecasting method for distribution power systems based on routinely available operational records and climatic indicators. The method was tested using outage data from 2020–2024. Using least squares, a multivariate regression model was obtained for feeder emergency outage duration as a function of cable damage (F5), transformer failure (F6), and the climatic factor (Climate). The model is significant overall (F-test $p < 0.01$) and explains 68.7% of downtime variation ($R^2 = 0.687$); residual diagnostics indicate normality and no autocorrelation. The average marginal effects are 7.561 h for cable failures, 3.314 h for transformer failures, and 2.418 h for climatic impacts, highlighting cables as the dominant driver of prolonged outages. To assess energy performance, a separate model was built for the loss share in the power system as a function of outage duration, phase clashing (F3), and Climate. This loss model has low explanatory power ($R^2 = 0.2013$) and non-significant factor coefficients ($p > 0.05$). Finally, bivariate regressions show that Climate strongly affects phase clashing (F3) ($R^2 = 0.793$) and moderately affects ground faults (F1) and insulator chipping (F2) ($R^2 = 0.533$ each). The proposed method supports maintenance prioritization, climate-mitigation measures, and continuous updating as new outage records are added, strengthening decision-making and system robustness

Keywords: power supply, diagnostics, failure prediction, reliability, mining enterprise, regression analysis

DEVELOPMENT OF A METHOD FOR DIAGNOSING AND FORECASTING POWER SUPPLY SYSTEMS FOR MINING ENTERPRISES

Aizada Kuanyshbaeva

Corresponding author

PhD Student*

E-mail: k_aizada@mail.ru

ORCID: <https://orcid.org/0000-0003-3718-1899>

Yevgeniy Kotov

PhD, Assistant Professor*

ORCID: <https://orcid.org/0000-0003-2660-2010>

Karshiga Smagulova

PhD, Assistant Professor*

ORCID: <https://orcid.org/0000-0001-6834-8490>

Fariza Abilzhanova

Master of Engineering Sciences, PhD Student

Department of Energy Systems**

ORCID: <https://orcid.org/0000-0002-7723-6930>

*Department of Automation of Production Processes**

**Abylka Saginov Karaganda Technical University

N. Nazarbayev ave., 56, Karagandy,

Republic of Kazakhstan, 100000

Received 30.11.2025

Received in revised form 26.01.2026

Accepted 12.02.2026

Published 27.02.2026

How to Cite: Kuanyshbaeva, A., Kotov, Y., Smagulova, K., Abilzhanova, F. (2026). Development of a method for diagnosing and forecasting power supply systems for mining enterprises.

Eastern-European Journal of Enterprise Technologies, 1 (8 (139)), 15–26.

<https://doi.org/10.15587/1729-4061.2026.352513>

1. Introduction

A stable and resilient electrical supply is a global engineering priority for mining and mineral-processing facilities, especially in remote regions with severe climate and intensive operating conditions. The Republic of Kazakhstan is used in this study only as an illustrative example of such environments.

This challenge is intensified by the harsh environments common in remote mining regions. Mines are frequently situated distant from main power grids, resulting in long transmission lines with limited redundancy and prolonged recovery times after outages. Additionally, equipment faces extreme weather-fluctuating temperatures, intense wind, dust accumulation, moisture, icing, vibrations, and impact loads—which accelerates wear on insulating materials and connection points. These factors increase the likelihood of temporary and persistent failures, further complicating diagnostics and repair, particularly when issues manifest intermittently or develop gradually.

Simultaneously, the industry is transitioning from reactive and scheduled maintenance approaches to proactive strategies based on real-time system health assessment and predictive modeling enabled by digital sensors and advanced data analysis. This evolution aligns with national initiatives promoting energy conservation and improved efficiency. Notably, Kazakhstan's government adopted the National Concept for Energy Efficiency and Conservation 2023–2029 (Resolution No. 264, March 28, 2023), setting ambitious goals to curb wasteful energy usage and drive tangible enhancements in energy performance across key economic sectors [1].

Simultaneously, diagnostic and forecasting techniques cannot be directly applied to mining enterprises from other power system environments, as failure mechanisms, operational modes, environmental stresses, and recovery limitations significantly diverge from those found in urban distribution grids or transmission networks. This necessitates ongoing development of frameworks capable of modeling multi-fault failure patterns under extreme conditions, accounting for time-varying

and delayed degradation effects, and remaining feasible with regular monitoring and maintenance routines. Hence, creating specialized diagnostic and reliability prediction models suited to power supply systems in mining operations continues to be an essential area of study.

2. Literature review and problem statement

The literature demonstrates substantial progress in applying data-driven techniques to equipment diagnostics and reliability improvement in power systems; however, important gaps remain for mining-enterprise distribution networks characterized by limited observability, heterogeneous assets, and strong environmental exposure. Paper [2] highlights artificial intelligence for transformer fault detection, classification, and predictive analytics, enabling higher automation. Unexplored part: translation of transformer-level AI diagnostics into feeder-level outage-duration forecasting for industrial distribution systems, including the explicit role of climate and heterogeneous failure causes (cables, switchgear, overhead lines). The reason for this may be the study scope is transformer condition monitoring, typically supported by specialized transformer datasets and measurements, while feeder downtime analysis requires system-level incident logs and multi-cause modeling.

In paper [3], a SCADA-based prediction solution for inverter modules is validated on large photovoltaic plants and multiple technology brands, achieving early warning up to 7 days. Unexplored part: applicability to mining power supply systems (feeders/substations) where failure mechanisms, operating regimes, and data structure differ from PV inverters, and where the target is often downtime duration and restoration behavior rather than incipient device faults. The reason for this may be the method is tailored to PV inverter SCADA data (high-frequency, component-specific), which is not always available or comparable in mining distribution networks.

Reference [4] proposes a real-time digital diagnosis method for high-voltage switchgear using dual coupling (thermal/electric fields) with ANFIS. Unexplored part: scalable use of such diagnostics for fleet-level reliability management and quantitative prediction of outage duration at feeder or substation level. The reason for this may be the approach relies on detailed field modeling and measurement inputs that are difficult to deploy across heterogeneous assets, and it focuses on fault diagnosis rather than outage-time forecasting.

In paper [5], a hybrid system combining DGA and fuzzy logic improves transformer fault diagnosis efficiency and reliability. Unexplored part: integration of transformer diagnostics with system-wide outage analytics, including cables and protection-related events, and quantification of climatic influence on failure causes and recovery time. The reason for this may be DGA-based methods require transformer-specific sampling/monitoring and are primarily designed for transformer health assessment, not for multi-component feeder downtime modelling.

Study [6] discusses real-time AI in industry to accelerate intelligent maintenance in mining operations. Unexplored part: provision of an implementable quantitative method with validated models that link failure causes + climate to feeder outage duration and practical decision metrics. The reason for this may be the paper is positioned as a high-level perspective on adoption and implementation rather than a focused modeling study with validated predictors and effect sizes.

For smart grids, [7] emphasizes AI-based diagnostics and automated substation restoration. Unexplored part: explicit mod-

eling of root-cause categories and climatic sensitivity that drive prolonged outages in industrial networks, and validation under mining-specific operating constraints. The reason for this may be the focus is on restoration automation logic and intelligent substation operation rather than mining distribution reliability with climatic exposure.

Paper [8] targets robust, scalable, and transparent diagnostics for power generation infrastructure. Unexplored part: transfer of these diagnostic principles to distribution feeders in mining enterprises, where outage processes and dominant failure modes (e.g., cable faults) differ from generation assets. The reason for this may be asset classes, measurements, and operational objectives are different across generation and mining distribution systems.

In paper [9], a unified hybrid approach diagnoses sensor and actuator faults in digital twins for remote operations. Unexplored part: modeling of electrical component failures (cables/transformers/switchgear) and their impact on downtime and energy loss, beyond instrumentation anomalies. The reason for this may be the study targets cyber-physical measurement and actuation faults within digital-twin frameworks, not physical power component failures.

Review [10] summarizes machine-learning potential in power system protection and control. Unexplored part: rigorous, interpretable models connecting failure causes and external conditions to outage duration and loss indicators for industrial distribution systems. The reason for this may be the reviewed methods mainly address protection/control functions (classification, settings, detection), whereas downtime and restoration modeling require different data and evaluation criteria.

In paper [11], a fault diagnosis method leveraging high-fidelity digital twins and deep learning is proposed. Unexplored part: feasibility in settings with limited monitoring and sparse labeled data, and direct estimation of outage duration and actionable maintenance prioritization. The reason for this may be high-fidelity digital twins and deep learning typically require substantial modeling effort, sensor coverage, and labeled datasets, which are often constrained in industrial mining networks.

Finally, [12] proposes reliability assessment for distribution networks up to 110 kV considering failure rates, restoration time, and topology. Unexplored part: mining-enterprise-specific modeling using real incident logs, cause categories, and explicit climatic factors to explain downtime variability and to test whether the same drivers also explain loss-share variability. The reason for this may be planning-level reliability assessment often relies on aggregate rates and generic assumptions, while mining networks require data-driven, context-specific cause-effect quantification.

Across the reviewed literature, most diagnostics focus on individual assets or generic smart-grid functions, while mining distribution systems are rarely studied using operational logs, explicit root-cause separation, climatic sensitivity quantification, and validated, interpretable models of feeder outage duration and loss indicators. Therefore, the unresolved problem is the lack of a consistent data-driven method to diagnose and forecast feeder outage duration in mining power supply systems under limited observability, accounting for dominant failure causes and climate impacts to support maintenance and modernization decisions.

3. The aim and objectives of the study

The aim of this study is to develop a method for diagnosing and forecasting the performance of power supply systems in mining enterprises. This will enable a reduction in the duration

of feeder emergency outages caused by failures of power system components. The results of the study will make it possible to forecast economic losses and develop a more effective strategy for the operation and modernization of power grids.

To achieve this aim, the following objectives were accomplished:

- to select significant exogenous factors that may affect feeder outage duration and to examine the interrelationships among them in order to detect multicollinearity;
- to investigate the effect of the selected explanatory variables on feeder outage duration and to quantify their explanatory power;
- to evaluate the impact of the identified factors on system-wide energy losses, to validate the adequacy of the proposed model, and to interpret the results for practical application;
- to determine which failure causes are most sensitive to climatic conditions and to quantify the magnitude of climatic influence.

4. Materials and methods

The object of the study is the power supply system of the mining enterprise Nova-Zinc LLP, located in Central region of Republic of Kazakhstan, which specializes in the extraction and enrichment of lead-zinc ores. The enterprise is supplied with electricity from the 110/35/6 kV Akzhal-Novaya substation, which is fed by a 110 kV power line from the 220/110 kV Akchataw substation.

The study hypothesizes a direct relationship between feeder emergency outage duration and the underlying causes of the outage.

In this study, it was assumed that feeder emergency outage duration depends on the following factors: ground faults, insulator chipping, phase conductor clashing, water ingress, cable failure, transformer failure, as well as the climatic factor.

A simplifying assumption of this study is that the analysis is based only on monthly statistical data for the period 2020–2024, since detailed statistical information on the factors considered is not available for earlier periods.

The study covered all key structural divisions of the enterprise: the underground mine, the enrichment plant, the motor transport and electrical engineering workshops, water supply facilities, and treatment facilities.

Diagnostic analysis of the technical condition of the power system was carried out based on a comprehensive monitoring methodology with elements of intelligent analysis. The methodology was based on correlation analysis and regression analysis.

Under conditions of intensive equipment operation, the impact of drilling and blasting, vibration loads, and frequent movement of mining fronts, the power supply system is prone to failure. According to operational monitoring data, 556 emergency shutdowns of various types were recorded at the enterprise during the period 2020–2024, which necessitated the development of an effective diagnostic system [13, 14].

The following data was used as source data:

- production and technical reports of Nova-Zinc LLP for 2020–2024;
- logs of emergency shutdowns and equipment failures;
- statistics on fuel and energy resource consumption;

- regulatory documentation for distribution and substation equipment.

Table 1 shows the generalized distribution of accidents by type for the period under review.

Table 1

Distribution of accidents by type of cause (2020–2024)

No.	Type of accident	Share, %
1	Ground fault	17.2
2	Insulator chip	26.44
3	Phase overlap	39.08
4	Water ingress	4.60
5	Cable failure	6.90
6	Transformer failure	5.17

The following analysis methods were used:

- expert assessment of failures by type of cause (ground fault, insulator chip, phase overlap, water ingress, cable failure, transformer failure);
- determination of the impact of failures on the overall level of energy reliability;
- development of a model for predicting the number of failures using linear approximation.

The data obtained formed the basis for the development of methodological approaches to diagnosing the technical condition and assessing the remaining service life of equipment.

5. Results of the power supply system diagnostics method

5.1. Selection of significant factors influencing feeder outages

Fig. 1 presents a chart of the number of failures and the total number of hours of feeder emergency outages for 2020–2024.

Analysis of the statistical data presented in Fig. 1 revealed the dynamics of changes in the number of failures in the power supply system for the analyzed period of 2020–2024. Between 2021 and 2022, there was a steady downward trend in the number of emergency shutdowns: from 39 cases recorded in 2020 to 29 cases in 2021 and 30 in 2022, which corresponds to a 26% and 23% reduction in accidents, respectively. This result demonstrates the positive effect of the organizational and technical measures taken to improve reliability and optimize equipment maintenance.

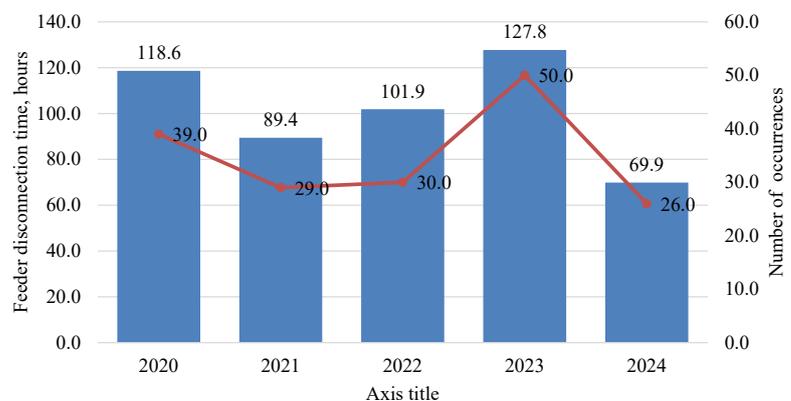


Fig. 1. Number of failures and total emergency hours feeder shutdowns

However, in 2023, there was a sharp increase in the number of emergencies, which amounted to 50 cases, exceeding the figures for previous years by more than 67%. This trend may be due to a number of factors, the key ones being: increased wear and tear of key elements of the power supply system, especially power transformer equipment, as well as the possible insufficient effectiveness of existing preventive and planned preventive measures.

At the same time, in 2024, there was a significant decrease in the total time of emergency downtime, reaching 69.9 hours. This figure is almost half that of the previous period and demonstrates a significant improvement in the technical condition of individual elements of the power supply system. Transformer units remain the most critical elements, as confirmed by accident statistics by type of equipment, according to which transformers are the main cause of prolonged downtime and unplanned outages.

A sharp increase in the duration of emergencies in 2024 may indicate the accumulation of hidden defects, progressive insulation wear, and the insufficient effectiveness of current maintenance and diagnostic schemes. This fact underscores the need to modernize existing diagnostic approaches, introduce more accurate failure prediction methods, and optimize the structure of preventive maintenance, taking into account the actual operating conditions of the equipment. To predict feeder failure times and identify the causes of these failures, I selected monthly data from January 2020 to December 2024.

Table 2 presents the variables for the regression analysis of the causes of line failures.

The Climate variable shows the number of damage incidents related to weather conditions, such as rain, snow, strong winds, and temperature fluctuations.

Using the STATISTICA software package, let's construct a correlation matrix of independent variables.

Table 2

Variables for regression analysis of line failure causes

Y	Feeder outage time, hours
F1	Ground fault
F2	Insulator chip
F3	Phase overlap
F4	Water ingress
F5	Cable failure
F6	Transformer failure
Climate	Climatic conditions
Energy	Total consumption, thousand kWh
Loss	Share of losses, %

Table 3 presents the matrix of pairwise correlation coefficients among the variables under consideration.

According to the results, the explanatory variable Y (feeder downtime) has a confirmed correlation with the factor variables F5 (cable failure), F6 (transformer failure), Climate (climatic factor), and Loss (loss ratio). The relationship between the explanatory variables F5, F6, Climate, and Loss is either insignificant or weaker than the relationship between these variables and the explanatory variable. It is possible to conclude that the effect of multicollinearity is insignificant.

The climate factor has the greatest influence on factor F3 (phase overlap), with a high correlation coefficient between these factors of 0.7495. The correlation coefficients between the climate factor and factors F1 (ground fault), F2 (insulator chip), F5 (cable failure), and loss (loss ratio) are significant.

There is a confirmed correlation between the factor loss (loss ratio), Y (feeder downtime), F3 (phase overlap), and the climatic factor.

Table 3

Correlation matrix of factor variables

Marked correlations are significant at $p < 0.05000$ $N = 55$ (Casewise deletion of missing data)										
Variables	Y	F1	F2	F3	F4	F5	F6	Climate	energy	loss
Y	1.0000	0.1411	0.1336	0.0447	-0.0564	0.5605	0.4782	0.3106	0.0386	0.3023
p-value	$p = ---$	$p = 0.304$	$p = 0.331$	$p = 0.746$	$p = 0.682$	$p = 0.000$	$p = 0.000$	$p = 0.021$	$p = 0.780$	$p = 0.025$
F1	0.1411	1.0000	0.4137	0.1194	-0.1026	-0.4075	-0.2645	0.4349	0.0139	0.1504
p-value	$p = 0.304$	$p = ---$	$p = 0.002$	$p = 0.385$	$p = 0.456$	$p = 0.002$	$p = 0.051$	$p = 0.001$	$p = 0.920$	$p = 0.273$
F2	0.1336	0.4137	1.0000	0.0622	-0.0454	-0.3772	-0.0148	0.4328	-0.0076	0.2591
p-value	$p = 0.331$	$p = 0.002$	$p = ---$	$p = 0.652$	$p = 0.742$	$p = 0.005$	$p = 0.915$	$p = 0.001$	$p = 0.956$	$p = 0.056$
F3	0.0447	0.1194	0.0622	1.0000	-0.1669	-0.3716	-0.1660	0.7495	-0.1764	0.2903
p-value	$p = 0.746$	$p = 0.385$	$p = 0.652$	$p = ---$	$p = 0.223$	$p = 0.005$	$p = 0.226$	$p = 0.000$	$p = 0.198$	$p = 0.032$
F4	-0.0564	-0.1026	-0.0454	-0.1669	1.0000	-0.0396	0.0211	0.0214	0.1106	-0.0024
p-value	$p = 0.682$	$p = 0.456$	$p = 0.742$	$p = 0.223$	$p = ---$	$p = 0.774$	$p = 0.879$	$p = 0.877$	$p = 0.422$	$p = 0.986$
F5	0.5605	-0.4075	-0.3772	-0.3716	-0.0396	1.0000	0.3690	-0.3742	-0.0017	0.0290
p-value	$p = 0.000$	$p = 0.002$	$p = 0.005$	$p = 0.005$	$p = 0.774$	$p = ---$	$p = 0.006$	$p = 0.005$	$p = 0.990$	$p = 0.834$
F6	0.4782	-0.2645	-0.0148	-0.1660	0.0211	0.3690	1.0000	-0.0181	0.2153	0.1417
p-value	$p = 0.000$	$p = 0.051$	$p = 0.915$	$p = 0.226$	$p = 0.879$	$p = 0.006$	$p = ---$	$p = 0.895$	$p = 0.114$	$p = 0.302$
Climate	0.3106	0.4349	0.4328	0.7495	0.0214	-0.3742	-0.0181	1.0000	-0.0479	0.4029
p-value	$p = 0.021$	$p = 0.001$	$p = 0.001$	$p = 0.000$	$p = 0.877$	$p = 0.005$	$p = 0.895$	$p = ---$	$p = 0.728$	$p = 0.002$
energy	0.0386	0.0139	-0.0076	-0.1764	0.1106	-0.0017	0.2153	-0.0479	1.0000	-0.1165
p-value	$p = 0.780$	$p = 0.920$	$p = 0.956$	$p = 0.198$	$p = 0.422$	$p = 0.990$	$p = 0.114$	$p = 0.728$	$p = ---$	$p = 0.397$
loss	0.3023	0.1504	0.2591	0.2903	-0.0024	0.0290	0.1417	0.4029	-0.1165	1.0000
p-value	$p = 0.025$	$p = 0.273$	$p = 0.056$	$p = 0.032$	$p = 0.986$	$p = 0.834$	$p = 0.302$	$p = 0.002$	$p = 0.397$	$p = ---$

Using the results of the correlation analysis, it is possible to move on to regression analysis.

5. 2. Influence of factor variables on the duration of feeder shutdown

Using the results of the correlation analysis, it is possible to obtain a model of the dependence of feeder downtime (Y) on F5 (cable failure), F6 (transformer failure), and Climate (climatic factor) using the least squares method. The simulation results are presented in Table 4.

In the initial model obtained, the variable loss (loss coefficient) turned out to be statistically insignificant, since its p-value is significantly higher than 0.05. Let's exclude this variable from the regression equation.

Table 5 presents the final model describing the dependence of feeder downtime on the considered factors.

The model is written as

$$Y = 1.907 + 7.561 * F5 + 3.314 * F6 + 2.418 * Climate. \quad (1)$$

Let's check the fulfillment of the Gauss-Markov assumptions for the least squares method.

Let's check the model residuals for normality of distribution. Let's construct a histogram of the residual distribution and a normal probability plot of the residuals.

Fig. 2, 3 show the residual distribution histogram and the normal probability plot of the residuals.

As can be seen from Fig. 2, the frequency histogram is relatively symmetrical, and the actual values of the residuals do not systematically deviate from the theoretical normal line (Fig. 3), so the hypothesis of the normality of the distribution of residuals is not rejected.

To check for autocorrelation, let's use the Durbin-Watson statistic.

In Table 6, the Durbin-Watson statistic for Model (1) is presented.

According to the table of critical points for $n = 60$ and $m = 1$ at a significance level of 0.05: $dl = 1.48$, $du = 1.689$. $DW > du$, therefore there is no first-order autocorrelation in model (1).

Table 4

Initial model of the dependence of feeder downtime on factors obtained using the least squares method

N = 57	Regression Summary for Dependent Variable: Y R = 0.82887148 R ² = 0.68702794 Adjusted R ² = 0.66295316 F(4.52) = 28.537 p < 0.00000 standard error of estimate: 2.9102					
	b*	standard error Of b*	b	standard error Of b	t(52)	p-value
Intercept	-	-	1.53094	2.01879	0.758344	0.451668
F5	0.679925	0.091468	7.52979	1.01295	7.433496	0.000000
F6	0.230178	0.084832	3.29254	1.21346	2.713343	0.009012
Climate	0.566166	0.092507	2.38393	0.38952	6.120237	0.000000
loss	0.017709	0.086410	18.87961	92.12248	0.204940	0.838419

Table 5

Final model of the dependence of feeder downtime on factors, obtained using the least squares method

N = 57	Regression Summary for Dependent Variable: Y R = 0.82871898 R ² = 0.68677515 Adjusted R ² = 0.66904544 F(3.53) = 38.736 p < 0.00000 standard error of estimate: 2.8838					
	b*	standard error Of b*	b	standard error Of b	t(53)	p-value
Intercept	-	-	1.906573	0.838482	2.273840	0.027052
F5	0.682729	0.089618	7.560844	0.992465	7.618245	0.000000
F6	0.231660	0.083755	3.313746	1.198064	2.765918	0.007796
Climate	0.574150	0.083142	2.417549	0.350082	6.905662	0.000000

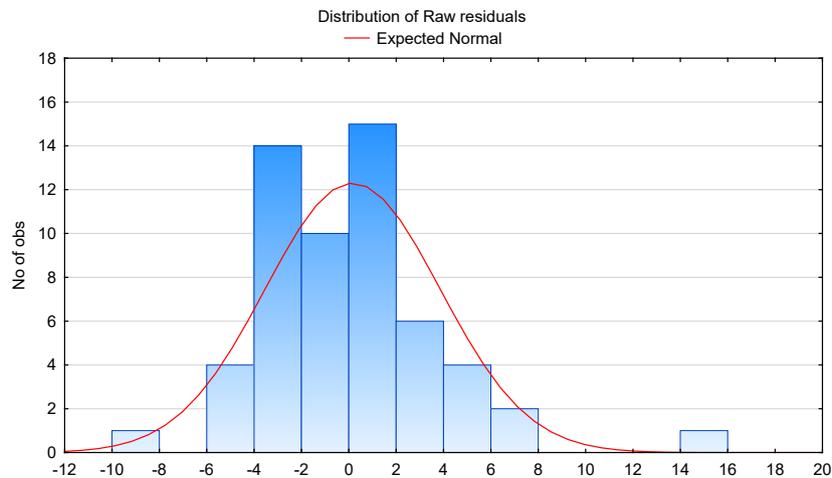


Fig. 2. Histogram of the distribution of model residuals (1)

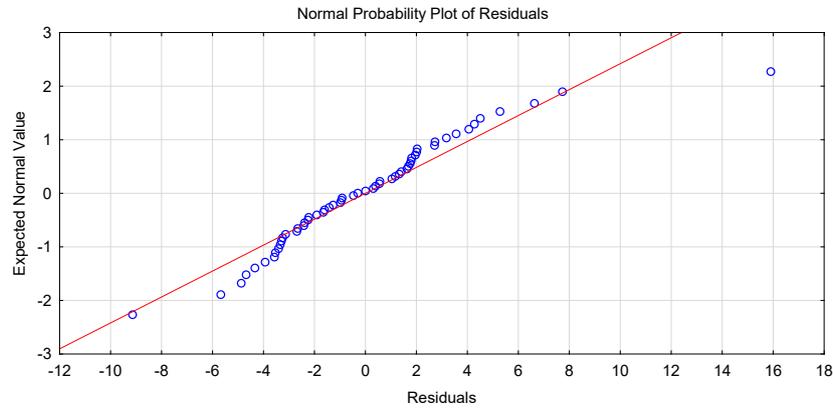


Fig. 3. Normal probability plot of the model residuals (1)

Durbin-Watson statistic for Model (1)

N = 60	Durbin-Watson <i>d</i> and serial correlation of residuals	
	Durbin-Watson <i>d</i>	Serial correlation
Estimate	1.987999	-0.018767

The resulting model (1) can be considered adequate for the initial data. All variables in it are statistically significant, since the p-value for all of them is greater than 0.05. The p-value for Fisher’s criterion is much less than 0.01, i.e., the model is significant as a whole. The coefficient of determination is high at 0.687, meaning that 68.7% of the change in feeder downtime depends on the factor variables included in the model. The model residuals are normally distributed and do not correlate with each other.

Based on the model obtained, cable failure leads to an average feeder shutdown of 7.561 hours, transformer damage leads to a feeder shutdown of 3.314 hours, and climatic factors increase the feeder shutdown time by 2.418 hours.

Thus, to ensure the stable operation of the system, it is necessary to pay special attention to the condition of power cables and, if necessary, replace worn sections, as this type of damage has a significant impact on the downtime of the feeder. Attention should be paid to the condition of transformers. It is also necessary to use elements that reduce the impact of climatic factors.

5. 3. Assessment of the impact of factors on energy losses in the entire system

Above, a confirmed correlation was established between the share of losses in the power system loss, the downtime of feeders *Y* and *F3* (phase overlap), and climatic factors.

Table 6

Let’s construct a multiple regression model in which the dependent variable is loss.

Table 7 presents a model of energy system losses as a function of feeder outage duration, the phase overlap coefficient, and a climatic factor.

In the resulting model, only the constant is significant; the coefficients for the factor variables are not significant, since their p-values are greater than 0.05. The coefficient of determination has a low value of 0.2013. That is, the obtained model explains only 20.13% of the change in the power system loss indicator. This determination value is low.

Therefore, it is possible to conclude that the duration of emergency feeder shutdowns, phase overlap, and climatic factors do not have a significant impact on energy system losses.

5. 4. Influence of climatic factors on the causes of failures

Above, a confirmed pairwise correlation was established between the climate factor and the following factors: *F3* (phase overlap), *F1* (ground fault), *F2* (insulator chip), *F5* (cable failure), and loss (share of losses). Let’s examine the influence of climatic factors on factors confirmed by paired correlation.

Let’s start with factor *F3* (phase overlap), as this factor has the highest correlation coefficient with the climate factor.

Table 8 presents a model of the dependence of the phase overlap coefficient on the climatic factor.

Since the constant was not significant in the initial regression model, let’s exclude it.

The model will be written as

$$F3 = 0.639 * Climate. \tag{2}$$

Let’s check the fulfillment of the Gauss-Markov assumptions for the least squares method.

Table 7

Model of the dependence of losses in the power system on the feeder shutdown time, phase overlap factor, and climatic factor

N = 58	Regression Summary for Dependent Variable: loss <i>R</i> = 0.44865392 <i>R</i> ² = 0.20129034 Adjusted <i>R</i> ² = 0.15691758 <i>F</i> (3.54) = 4.5363 <i>p</i> < 0.00659 standard error of estimate: 0.00442					
	<i>b</i> *	standard error Of <i>b</i> *	<i>b</i>	standard error Of <i>b</i>	<i>t</i> (54)	p-value
Intercept	–	–	0.019243	0.001361	14.13635	0.000000
<i>Y</i>	0.219762	0.133429	0.000215	0.000131	1.64704	0.105359
<i>F3</i>	0.089348	0.192715	0.000387	0.000835	0.46363	0.644779
Climate	0.267300	0.202119	0.001096	0.000829	1.32249	0.191580

Table 8

Model of the dependence of the phase overlap factor on the climatic factor

N = 58	Regression Summary for Dependent Variable: F3 R = 0.89026812 R ² = 0.79257732 Adjusted R ² = 0.78893833 F(1.57) = 217.80 p < 0.0000 standard error of estimate: 0.73882					
	b*	standard error Of b*	b	standard error Of b	t(57)	p-value
Climate	0.890268	0.060324	0.639175	0.043310	14.75809	0.000000

To check for autocorrelation, let's use the Durbin-Watson statistic:

In Table 9, the Durbin-Watson statistic for Model (2) is presented.

Table 9

Durbin-Watson statistic for Model (2)

N = 60	Durbin-Watson d and serial correlation of residuals	
	Durbin-Watson d	Serial correlation
Estimate	1.311079	0.294124

According to the table of critical points for $n = 60$ and $m = 1$ at a significance level of 0.05: $dl = 1.549$, $du = 1.616$. $DW < dl$, therefore, first-order autocorrelation is present in model (2). For a more accurate regression estimate, it is necessary to correct the autocorrelation in model (2). However, the aim of this study is to establish the influence of the climatic factor. A coefficient of determination of 0.79 indicates a significant influence.

Let's check the model residuals for normality of distribution. Let's construct a histogram of the residual distribution and a normal probability plot of the residuals.

As can be seen from Fig. 4, the frequency histogram is relatively symmetrical, and the actual values of the residuals do not systematically deviate from the theoretical normal line (Fig. 5), so the hypothesis of normality of the residual distribution is not rejected.

The coefficient of determination is high, at 0.793, which means that the climatic factor influences the phase overlap factor F3 by 79.3%. The model residuals have a normal distribution but correlate with each other.

Let's consider the influence of the climatic factor on factor F1 (ground fault).

Table 10 presents a model of the dependence of the ground-fault coefficient on the climatic factor.

Since the constant was not significant in the initial regression model, let's exclude it.

The model will be written as

$$F1 = 0.256 * Climate. \tag{3}$$

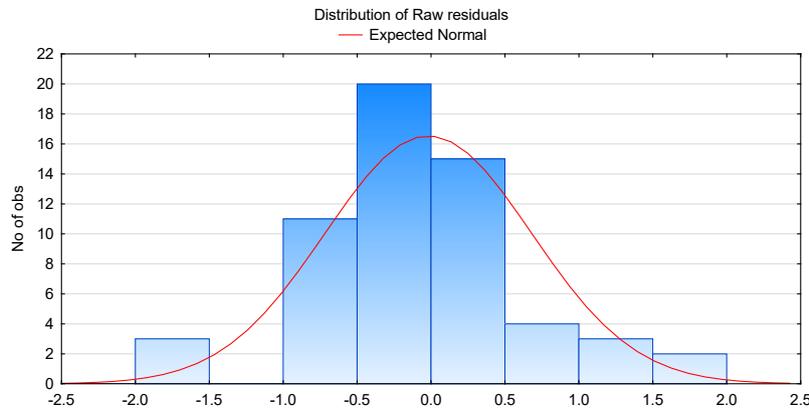


Fig. 4. Histogram of the distribution of model residuals (2)

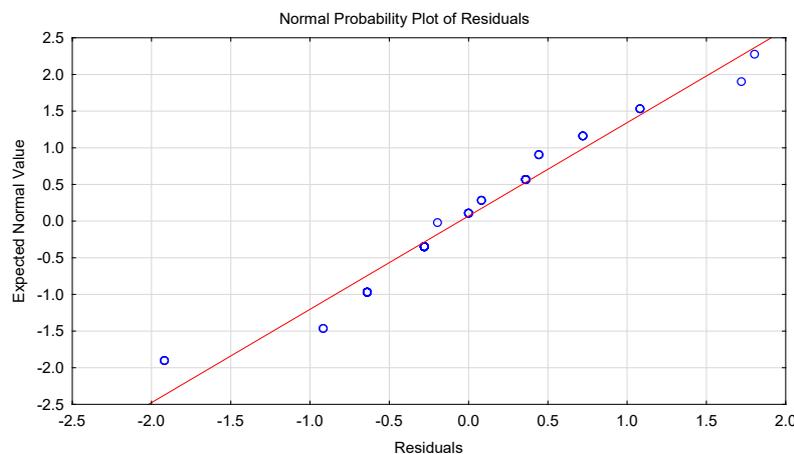


Fig. 5. Normal probability graph of model residuals (2)

Table 10

Model of the dependence of the ground fault factor on the climatic factor

N = 60	Regression Summary for Dependent Variable: F1 R = 0.73025780 R ² = 0.53327645 Adjusted R ² = 0.52536588 F(1.59) = 67.413 p < 0.00000 standard error of estimate: 0.53365					
	b*	standard error Of b*	b	standard error Of b	t(59)	p-value
Climate	0.730258	0.088941	0.255973	0.031176	8.210552	0.000000

Let's check the fulfillment of the Gauss-Markov assumptions for the least squares method.

To check for autocorrelation, let's use the Durbin-Watson statistic.

In Table 11, the Durbin-Watson statistic for Model (3) is presented.

Table 11

Durbin-Watson statistic for Model (3)

N = 60	Durbin-Watson d and serial correlation of residuals	
	Durbin-Watson d	Serial correlation
Estimate	1.960380	0.010100

According to the table of critical points for n = 60 and m = 1 at a significance level of 0.05: dl = 1.549, du = 1.616.

DW > du, therefore there is no first-order autocorrelation in model (3). Let's check the model residuals for normality of distribution. Let's construct a histogram of the residual distribution and a normal probability plot of the residuals.

Fig. 6, 7 show the residual distribution histogram and the normal probability plot of the residuals.

The actual values of the residuals do not systematically deviate from the theoretical normal line (Fig. 7), so the hypothesis of the normality of the distribution of residuals is not rejected. The coefficient of determination is not high and is 0.533, therefore the climatic factor affects factor F1 (ground fault) by 53.3%.

Let's consider the influence of the climatic factor on factor F2 (insulator chipping).

Table 12 presents a model of the dependence of the insulator chipping coefficient on the climatic factor.

Since the constant was not significant in the initial regression model, it is possible to exclude it.

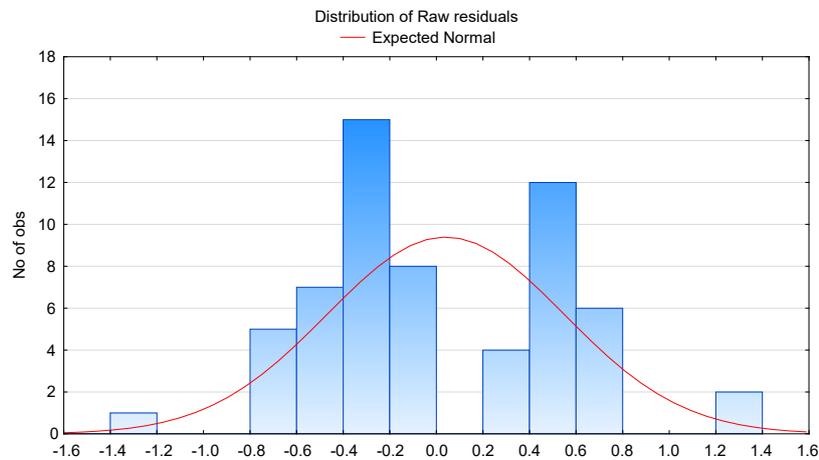


Fig. 6. Histogram of the distribution of model (3) residuals

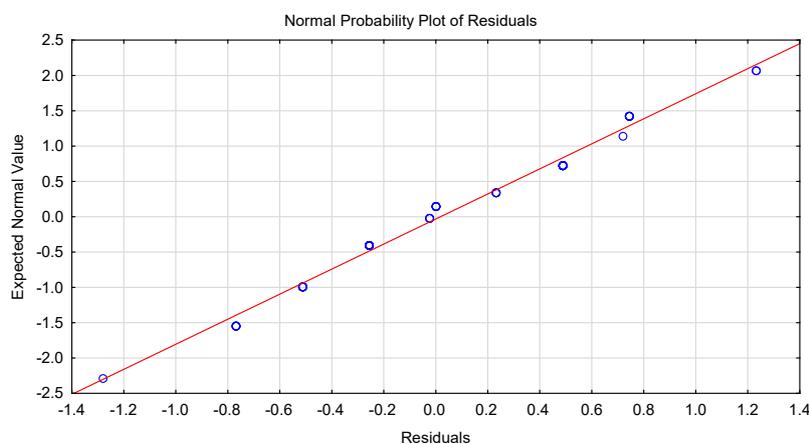


Fig. 7. Normal probability plot of the residuals of model (3)

Table 12

Model of the dependence of the insulator chipping factor on the climatic factor

N = 60	Regression Summary for Dependent Variable: F2 R = 0.73154192 R ² = 0.53515358 Adjusted R ² = 0.52727483 F(1.59) = 67.924 p < 0.00000 standard error of estimate: 0.79391					
	b*	standard error Of b*	b	standard error Of b	t(59)	p-value
Climate	0.731542	0.088762	0.382253	0.046381	8.241580	0.000000

The model will be written as

$$F2 = 0.382 * Climate. \tag{4}$$

Let's check the fulfillment of the Gauss-Markov assumptions for the least squares method.

To check for autocorrelation, let's use the Durbin-Watson statistic.

In Table 13, the Durbin-Watson statistic for Model (4) is presented.

Table 13

Durbin-Watson statistic for Model (4)

N = 60	Durbin-Watson d and serial correlation of residuals	
	Durbin-Watson d	Serial correlation
Estimate	1.707809	0.136810

According to the table of critical points for n = 60 and m = 1 at a significance level of 0.05: dl = 1.549, du = 1.616. DW > du, therefore there is no first-order autocorrelation in model (4).

Let's check the model residuals for normality of distribution. Let's construct a histogram of the residual distribution and a normal probability plot of the residuals.

Fig. 8, 9 show the residual distribution histogram and the normal probability plot of the residuals.

The actual values of the residuals do not systematically deviate from the theoretical normal line (Fig. 9), so the hypothesis of the normality of the distribution of residuals is not rejected.

The coefficient of determination is not high and is 0.535, therefore the climatic factor affects factor F2 (insulator chipping) by 53.5%.

It has been established that the climatic factor affects ground faults (F1), insulator chipping (F2), and phase overlap (F3).

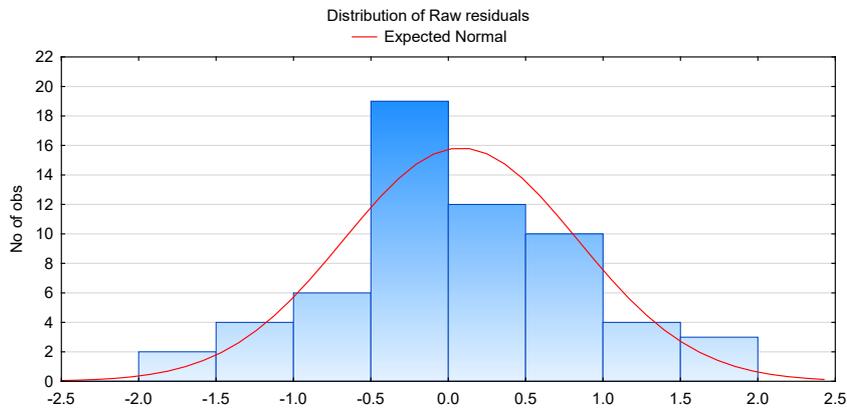


Fig. 8. Histogram of the distribution of model residuals (4)

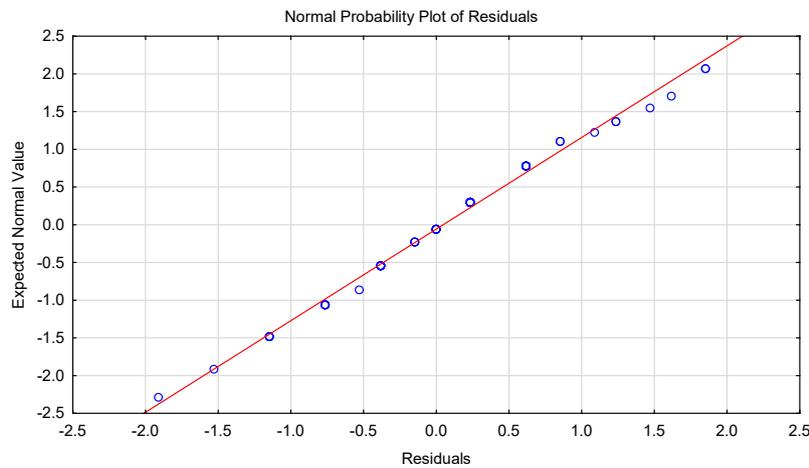


Fig. 9. Normal probability graph of the model residuals (4)

6. Discussion of the development of a method for diagnosing and forecasting power supply systems of mining enterprises

The study showed that the main factors affecting the duration of feeder outages are cable line damage and transformer failure. According to the developed model, which describes the dependence of feeder outage time on cable line damage and transformer failure, these factors significantly affect the feeder downtime. Based on the model obtained, cable damage leads to an average feeder outage of 7.561 hours, while transformer failure results in a feeder outage of 3.314 hours.

A distinctive feature of the proposed method is the analysis of statistics on the root causes of feeder failures and their contribution to the duration of emergency downtime. Having a quantitative assessment of how specific failure causes affect outage duration makes it possible to estimate the economic benefits of modernizing the power supply system [15, 16].

To improve the reliability and stability of the power supply system, a comprehensive diagnostic method has been developed based on a combination of econometric failure analysis, correlation analysis of factor characteristics, and the application of regression analysis, which makes it possible to identify the key determinants of equipment failures, assess their temporal dependence, and justify the need to expand the set of diagnostic factors to improve the accuracy of predicting failures of power supply system components.

The method includes the formation of a system of exogenous indicators of equipment failures, the construction of a correlation matrix to assess the relationships between factor characteristics, the development of regression models, and the analysis of their adequacy based on statistical criteria and the coefficient of determination, which provides comprehensive diagnostics and forecasting of failures of power supply system components.

The study showed that the key factors influencing feeder emergency outage duration are cable line damage and transformer failure [11, 15]. According to the derived model (1), factor F5 (cable line damage) and factor F6 (transformer failure) have a statistically significant effect on feeder emergency downtime. Based on model (1), cable damage results in an average feeder outage of 7.561 hours, while transformer failure results in an average outage of 3.314 hours. This can be explained by the substantial labor requirements associated with cable replacement and transformer repair [5,6].

A distinctive feature of the proposed approach is the analysis of statistics on feeder failure causes and their impact on the duration of emergency downtime [6,9]. Given a quantitative assessment of how failure causes affect emergency outage duration, the economic effect of modernizing the power supply system can be estimated [3, 6, 11].

A limitation of this study is the climatic factor. When constructing models (2), (3), and (4), it was established that the climatic factor has a significant influence on the occurrence of faults in the power supply system. Therefore, in regions with different climatic characteristics, the magnitude and direction of this effect may differ.

The main drawback of the present study is the lack of more complete and detailed statistical information on damage causes prior to 2020. Before this period, the data records contained gaps and not all events were captured. Improvements introduced since 2020 in the statistical accounting of feeder failure causes enabled the present analysis to be conducted.

Future work may involve identifying additional explanatory variables that were not included in the current analysis.

Furthermore, as time progresses, the time series of the explanatory variables will become longer, which will allow a more robust assessment of their effects. In the case of large-scale modernization of the system, structural (dummy) variables should be introduced to quantify the impact of the implemented measures.

A limitation of this study is the climatic factor. It is found that it significantly affects the occurrence of faults in the power supply system. Therefore, for regions with different climatic conditions, this factor may have a different impact.

The main drawback of the study is the lack of comprehensive and detailed statistical information on damage causes prior to 2020. Before this period, data recording contained gaps and not all events were captured. The improvement of statistical accounting for feeder failure causes since 2020 made the present analysis feasible.

Future work may include identifying additional explanatory variables that were not considered in the present analysis. Over time, longer time series of the explanatory variables will become available, enabling a more robust assessment of their effects. Moreover, large-scale system modernization will require the introduction of structural (dummy) variables to quantify the impact of the implemented measures. The obtained results can be used for the further development of intelligent diagnostic and forecasting systems, as well as for expanding the factor space by incorporating operational, climatic, and operational parameters, which represents a promising direction for future research.

7. Conclusion

1. The analysis of the correlation matrix showed that the variable Y (feeder outage time) has a confirmed correlation relationship with the factor variables F5 (cable failure), F6 (transformer failure), Climate (climatic factor), and loss (loss share). The relationship among the explanatory variables F5, F6, Climate, and loss is either insignificant or weaker than the relationship of these variables with the dependent variable. Therefore, multicollinearity has a negligible effect.

The climatic factor Climate has the strongest influence on factor F3 (phase conductor clashing), with a high correlation coefficient of 0.7495. Significant correlation coefficients are also observed between the climatic factor and factors F1 (ground fault), F2 (insulator chipping), F5 (cable failure), and loss (loss share).

A confirmed correlation relationship exists among loss (loss share), Y (feeder outage time), F3 (phase conductor clashing), and the climatic factor.

2. Using the least squares method, a model was obtained describing the dependence of feeder emergency outage time on cable damage (F5), transformer failure (F6), and the climatic factor (Climate).

Model can be considered adequate for the original data. All variables in the model are statistically significant, since their p-values are greater than 0.05. The p-value for the Fisher test is much less than 0.01, i.e., the model is significant overall. The coefficient of determination is high and equals 0.687: 68.7% of the variation in feeder downtime is explained by the factor variables included in the model. The model residuals are normally distributed and do not correlate with each other.

Based on the obtained model, cable failure increases the feeder outage time by an average of 7.561 hours, transformer damage increases the feeder outage time by 3.314 hours, and climatic factors increase the feeder outage time by 2.418 hours.

Thus, to ensure stable operation of the system, special attention should be paid to the condition of power cables; if necessary, worn sections should be replaced, since this type of fault has a significant impact on feeder downtime. Attention should also be paid to the condition of transformers. In addition, elements that reduce the influence of climatic factors should be used.

3. In the resulting model, only the intercept is significant; the coefficients of the factor variables are not significant because their p-values are greater than 0.05. The coefficient of determination is low and equals 0.2013, meaning that the model explains only 20.13% of the variation in power system losses. This R^2 value is small.

Therefore, it can be concluded that feeder emergency outage time, the phase clashing factor, and the climatic factor do not have a significant effect on the loss share in the power system.

4. The dependence of F3 on the climatic factor was assessed; Climate explains 79.3% of the variation in F3 ($R^2 = 0.793$).

The assessment of the influence of the climatic factor on F1 (ground fault) showed a not very high coefficient of determination ($R^2 = 0.533$), meaning that the climatic factor explains 53.3% of the variation in F1.

The assessment of the influence of the climatic factor on F2 (insulator chipping) also showed a not very high coefficient of determination ($R^2 = 0.533$), meaning that the climatic factor explains 53.3% of the variation in F2.

Based on the obtained coefficients of determination and statistically significant regressions, it can be concluded that the climatic factor has a significant influence on ground faults (F1), insulator chipping (F2), and phase conductor clashing (F3). The climatic factor has the greatest influence on phase conductor clashing (F3).

Based on the correlation analysis, it is found that the significant factors affecting feeder failure duration are factor F5 (cable failure), factor F6 (transformer failure), and Climate (climatic factor). This is evidenced by the highest values of the corresponding pairwise correlation coefficients and the statistical significance of these coefficients assessed using Student's t-test. Factor inflation is negligible, as indicated by weak correlation relationships among the selected explanatory factors.

The effect of cable-line and transformer failures, as well as the climatic factor, on the feeder downtime duration was evaluated. The resulting model was considered adequate with respect to the initial data. All variables in the model are statistically significant because their p-values are greater than 0.05. The p-value for the Fisher test is much less than 0.01; therefore, the model is significant overall. The coefficient of determination is high and equals 0.687. Changes in feeder downtime are 68.7% explained by the factor variables included in the model. The residuals are normally distributed and are not correlated with each other.

It was found that damage-related and climatic factors do not affect energy losses in the system. The influence of total energy consumption in the system is not associated with the level of emergency events. For this purpose, a model was built to represent the dependence of power-system losses on feeder outage duration, the phase overlap coefficient, and climatic factors. In the obtained model, only the intercept was significant; the coefficients for the factor variables were not significant because their p-values were greater than 0.05. The coefficient of determination is low (0.2013), meaning that the model explains only 20.13% of the variation in the power-system loss indicator.

The influence of the climatic factor on ground faults (F1), insulator chipping (F2), and phase overlap (F3) was assessed as significant. The climatic factor has the greatest effect on the phase overlap coefficient F3. This conclusion was drawn based on the development and evaluation of simple (bivariate) regression models. The dependence of the phase overlap coefficient F3 on the climatic factor was assessed; the coefficient of determination is high (0.793), indicating that the climatic factor explains 79.3% of the variation in F3. The assessment of climatic effects on factor F1 (ground fault) yielded a low coefficient of determination (0.533), meaning that climatic factors explain 53.3% of the variation in F1. The assessment of climatic effects on factor F2 (insulator chipping) also yielded a low coefficient of determination (0.533), meaning that the climatic factor explains 53.3% of the variation in F2.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this study, whether financial, personal, authorship or otherwise, that could affect the study and its results presented in this paper.

Financing

The study was performed without financial support.

Data availability

Data will be made available on reasonable request.

Use of artificial intelligence

AI tool: ChatGPT (model: GPT-5.2 Thinking, OpenAI).

The tool was used in the Literature Review / Background section to: help structure the literature review, summarize selected publications, and improve academic wording of the review text.

Verification: the authors independently checked all AI-assisted statements against the original sources (articles, standards, reports), verified the correctness of terminology and interpretation, and ensured that all references and factual claims are supported by the cited literature. AI outputs were treated as drafts and were edited by the authors.

Impact on conclusions: AI assistance did not affect the study results or conclusions; conclusions were based solely on the authors' analysis of operational data and calculations. No AI tools were used for data processing, modeling, calculations, or generation of figures/tables.

Authors' contributions

Aizada Kuanyshbaeva: Methodology; Formal analysis; Investigation; Visualization; Writing – original draft; Writing – review & editing; **Yevgeniy Kotov:** Conceptualization; Formal analysis; Investigation; Writing – original draft; Writing – review & editing; **Karshiga Smagulova:** Resources; Data Curation; Writing – original draft; Writing – review & editing; **Fariza Abilzhanova:** Conceptualization; Data Curation; Writing – original draft; Writing – review & editing.

References

1. Government of the Republic of Kazakhstan. Resolution No. 264 of March 28, 2023: On the approval of the Concept for the development of energy saving and energy efficiency in the Republic of Kazakhstan for 2023–2029. Available at: https://adilet.zan.kz/rus/docs/P2300000264?utm_source=chatgpt.com
2. Khan, M. A. M. (2025). AI and machine learning in transformer fault diagnosis: a systematic review. *American Journal of Advanced Technology and Engineering Solutions*, 1 (01), 290–318. <https://doi.org/10.63125/sxb17553>
3. Betti, A., Lo Trovato, M. L., Leonardi, F. S., Leotta, G., Ruffini, F., Lanzetta, C. (2019). Predictive Maintenance in Photovoltaic Plants with a Big Data Approach. *arXiv*. <https://doi.org/10.48550/arXiv.1901.10855>
4. Ding, F., Ma, Y., Li, X., Huang, J. (2025). Research on digital twin diagnosis model for the thermal-electric field of high-voltage switch-gears. <https://doi.org/10.21203/rs.3.rs-6651594/v1>
5. Baker, E., Nese, S. V., Dursun, E. (2023). Hybrid Condition Monitoring System for Power Transformer Fault Diagnosis. *Energies*, 16 (3), 1151. <https://doi.org/10.3390/en16031151>
6. Rojas, L., Peña, Á., Garcia, J. (2025). AI-Driven Predictive Maintenance in Mining: A Systematic Literature Review on Fault Detection, Digital Twins, and Intelligent Asset Management. *Applied Sciences*, 15 (6), 3337. <https://doi.org/10.3390/app15063337>
7. Kim, H., Kim, J.-K., Song, J., Lee, J., Han, K., Shin, J. et al. (2019). Smart and Green Substation: Shaping the Electric Power Grid of Korea. *IEEE Power and Energy Magazine*, 17 (4), 24–34. <https://doi.org/10.1109/mpe.2019.2909792>
8. Maican, C. A., Pană, C. F., Pătrașcu-Pană, D. M., Rădulescu, V. M. (2025). Review of Fault Detection and Diagnosis Methods in Power Plants: Algorithms, Architectures, and Trends. *Applied Sciences*, 15 (11), 6334. <https://doi.org/10.3390/app15116334>
9. Hasan, A., Salvo Rossi, P. (2025). A unified sensor and actuator fault diagnosis in digital twins for remote operations. *Mechanical Systems and Signal Processing*, 222, 111778. <https://doi.org/10.1016/j.ymsp.2024.111778>
10. Porawagamage, G., Dharmapala, K., Chaves, J. S., Villegas, D., Rajapakse, A. (2024). A review of machine learning applications in power system protection and emergency control: opportunities, challenges, and future directions. *Frontiers in Smart Grids*, 3. <https://doi.org/10.3389/frsgr.2024.1371153>
11. Liu, W., Han, B., Zheng, A., Zheng, Z., Chen, S., Jia, S. (2024). Fault diagnosis of reducers based on digital twins and deep learning. *Scientific Reports*, 14 (1). <https://doi.org/10.1038/s41598-024-75112-x>
12. Abilzhanova, F., Bulatbaev, F., Kuanyshbaeva, A. (2025). Development of a method for reliability assessment of distribution power networks up to 110 kV. *Eastern-European Journal of Enterprise Technologies*, 1 (8 (133)), 15–23. <https://doi.org/10.15587/1729-4061.2025.322920>
13. Nova Zinc LLP. Available at: <http://novazinc.kz/>
14. Nova Zinc LLP. Kazakhstan Mining Group. Available at: <https://kazmininggroup.kz/company/partners/zn-pb/>
15. Telbayeva, S., Nurmaganbetova, G., Avdeyev, L., Kaverin, V., Issenov, S., Janiszewski, D. et al. (2024). Development of mathematical models of power consumption at coal plants. *Eastern-European Journal of Enterprise Technologies*, 5 (8 (131)), 22–32. <https://doi.org/10.15587/1729-4061.2024.313932>
16. Breido, I. V., Sichkarenko, A. V., Kotov, E. S. (2013). Emergency control of technological environment and electric machinery activity in coal mines. *Journal of Mining Science*, 49 (2), 338–342. <https://doi.org/10.1134/s1062739149020172>