**Dynamic Analysis of Factors Affecting Electricity Losses**

Fazliddin Khojayorov1,2,a),Abdusaid Isakov1, Taxir Bayzakov1, Abror Botirov1, Aziz Abduganiev1, Kamoliddin Obidov3

*1“Tashkent Institute of Irrigation and Agricultural Mechanization Engineers” National Research university, Tashkent, Uzbekistan*

*2Tashkent state technical university named after Islam Karimov, Tashkent, Uzbekistan*

*3Bukhara State Technical University, Bukhara, Uzbekistan*

*a) Corresponding author:* [*fazliddinxojayorov09@gmail.com*](mailto:fazliddinxojayorov09@gmail.com)

**Abstract.** This study investigates the factors influencing electricity losses by applying statistical correlation and regression methods to time-series data. The Pearson correlation coefficient was utilized as a primary tool for quantifying linear relationships between electricity losses and multiple influencing variables, including active and reactive electricity consumption, temperature parameters, relative humidity, visibility, and transformer load factor. To increase accuracy, the dataset was divided into 10-day intervals, forming 72 groups, with correlation coefficients computed for each group. The analysis revealed that the strength and direction of correlations vary seasonally, emphasizing the dynamic nature of influencing factors. Regression models up to the eighth degree were developed to capture these variations, with model selection based on the lowest Root Mean Square Error (RMSE). Among the tested models, the third-degree polynomial regression for temperature demonstrated the highest accuracy (RMSE = 2.1%). The findings highlight the limitations of relying solely on static correlation values and confirm the importance of dynamic and regression-based modeling for more reliable forecasting of electricity losses. This approach provides a robust methodological framework for improving energy efficiency analysis and supporting decision-making in power systems management.

**INTRODUCTION**

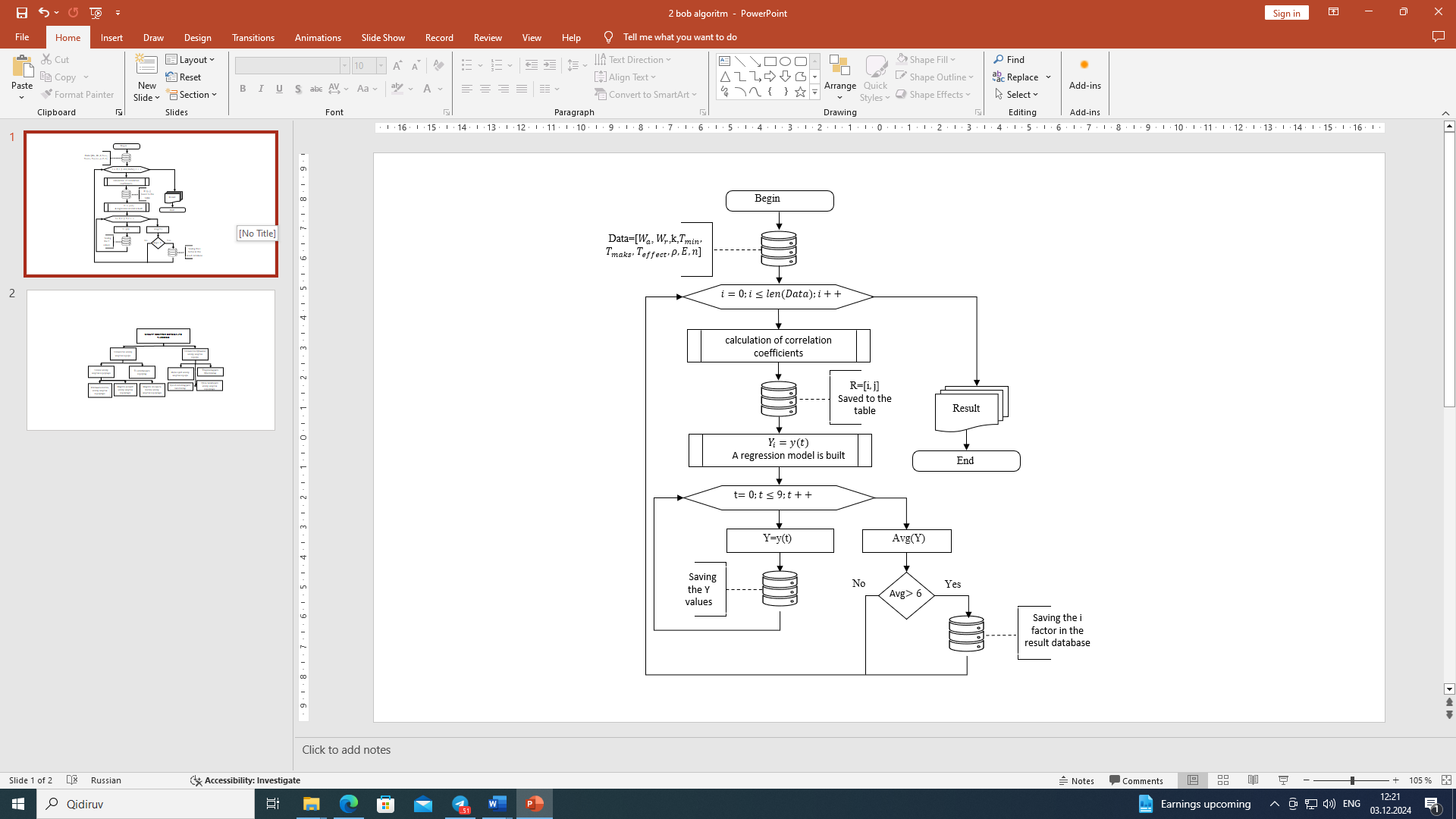
To date, a wide range of statistical methods, such as Pearson, Spearman, Kendall's tau, Phi coefficient, partial correlation, and others, have been extensively used to identify relationships between variables. Among these, the Pearson correlation method stands out as a statistical approach that enables fast and accurate measurement of linear relationships between continuous variables [1,2]. Therefore, within the scope of this scientific study, factors influencing electricity losses were identified using this method. Based on this, an algorithm for determining correlation coefficients was developed during the research to perform a correlation analysis of the influencing factors (Figure 1).

***Development of energy production based on renewable energy sources.*** In 2020-2030, special attention will be paid to the production of electricity from renewable energy sources, especially the development of solar energy. These projects are implemented exclusively at the expense of investors – independent electricity producers.

In order to achieve the indicators of the development of renewable energy sources in 2020-2030, it is planned to build 3 GW of wind and 5 GW of solar power plants. The target parameters of the capacity of renewable energy sources put into operation annually have been established [1-5].

(1)

Here, and represent individual levels, while X and Y are the variables. The correlation coefficient (r) ranges between -1 and 1. If r=1, it indicates a perfect positive linear relationship between the variables, meaning that as one variable increases, the other variable also increases proportionally in a linear manner.



**FIGURE 1.** The graph of the development of installed wind energy (GW) since 2010(Indicators from left to right:Worldwide, EU, Germany, America)

**EXPERIMENTAL RESEARCH**

If the correlation coefficient equals −1, it demonstrates a perfect negative linear relationship, implying that as one variable increases, the other decreases in a linear fashion. Conversely, if r=0, it signifies that no linear relationship exists between the variables.

The strength of the correlation is classified as follows:

* or : Weak positive or negative correlation.
* or : Moderate positive or negative correlation.
* or : Strong positive or negative correlation.

This classification is based on the observed ranges of r [source].

**Table 1.** Correlation Coefficients of Factors Affecting Electricity Losses Based on the Data for the First 10 Days of January

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Electricity Loss (kWh) | Active Electricity Consumption (kWh) | Reactive Electricity Consumption (kVAR·h) | Day of the Week | Minimum Temperature (°C) | Maximum Temperature (°C) | Effective Temperature (°C) | Visibility | Relative Humidity (%) | Load Factor |
| F1 | 1 |  |  |  |  |  |  |  |  |  |
| F2 | 0,92 | 1 |  |  |  |  |  |  |  |  |
| F3 | 0,89 | 0,95 | 1 |  |  |  |  |  |  |  |
| F4 | 0,012 | -0,01 | -0,008 | 1 |  |  |  |  |  |  |
| F5 | 0,65 | 0,75 | 0,72 | 0,02 | 1 |  |  |  |  |  |
| F6 | 0,662 | 0,76 | 0,74 | 0,06 | 0,93 | 1 |  |  |  |  |
| F7 | 0,65 | 0,75 | 0,73 | 0,08 | 0,92 | 0,99 | 1 |  |  |  |
| F8 | -0,81 | -0,71 | -0,74 | 0,03 | 0,42 | 0,15 | 0,14 | 1 |  |  |
| F9 | -0,21 | -0,69 | -0,70 | -0,07 | 0,48 | 0,25 | 0,23 | 0,58 | 1 |  |
| F10 | 0,81 | 0,75 | 0,72 | 0,53 | 0,65 | 0,69 | 0,67 | 0,66 | -0,04 | 1 |

The factors influencing electricity losses were identified based on the data from the first 10 days of January 2022, using the algorithm presented in Figure 1. High-impact factors were selected based on the given criteria. These include active and reactive electricity consumption, temperatures, visibility, and the transformer load factor.

**Table 2.** Correlation Coefficients of Factors Affecting Electricity Losses Based on the Data for the First 10 Days of August.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Electricity Loss (kWh) | Active Electricity Consumption (kWh) | Reactive Electricity Consumption (kVAR·h) | Day of the Week | Minimum Temperature (°C) | Maximum Temperature (°C) | Effective Temperature (°C) | Visibility | Relative Humidity (%) | Load Factor |
| F1 | 1 |  |  |  |  |  |  |  |  |  |
| F2 | 0,9 | 1 |  |  |  |  |  |  |  |  |
| F3 | 0,88 | 0,95 | 1 |  |  |  |  |  |  |  |
| F4 | 0,01 | -0,02 | -0,02 | 1 |  |  |  |  |  |  |
| F5 | 0,68 | 0,73 | 0,67 | 0,02 | 1 |  |  |  |  |  |
| F6 | 0,64 | 0,77 | 0,69 | 0,08 | 0,33 | 1 |  |  |  |  |
| F7 | 0,68 | 0,75 | 0,56 | 0,07 | 0,74 | 0,87 | 1 |  |  |  |
| F8 | -0,42 | -0,28 | -0,25 | 0,07 | 0,32 | 0,45 | 0,45 | 1 |  |  |
| F9 | -0,41 | -0,45 | -0,56 | -0,09 | 0,38 | 0,42 | 0,3 | 0,28 | 1 |  |
| F10 | 0,53 | 0,65 | 0,52 | 0,13 | 0,75 | 0,73 | 0,64 | 0,54 | -0,1 | 1 |

The factors influencing electricity losses were identified based on the data from the first 10 days of August 2022 using the algorithm presented in Figure 1. For August, the most impactful factors were primarily active and reactive electricity consumption, as well as temperatures.

The results indicate that the factors affecting electricity losses during the first 10 days of January and August differ. To ensure reliability, correlation coefficients were calculated for data from 72 groups (Table 3).

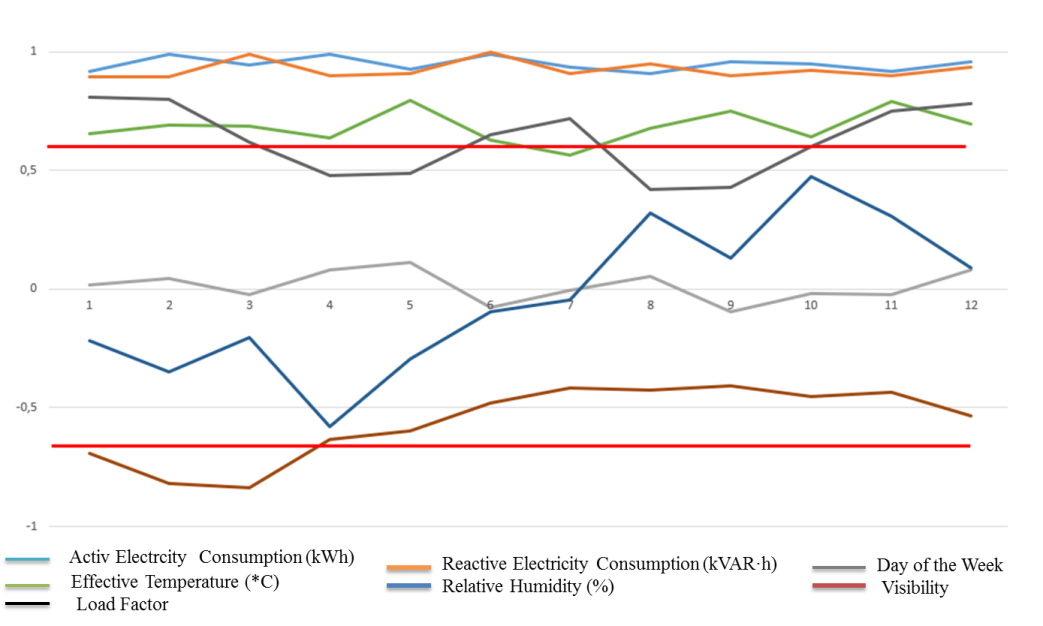
**Table 3.** Correlation Coefficients of Factors Influencing Electricity Losses

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Correlation Coefficient (r) | | | | | | | | |
| № | Electricity Loss (kWh) | Active Electricity Consumption (kWh) | Reactive Electricity Consumption (kVAR·h) | Day of the Week | Minimum Temperature (°C) | Maximum Temperature (°C) | Effective Temperature (°C) | Visibility | Relative Humidity (%) |
| 1 | 0.92 | 0.89 | 0.12 | 0.65 | 0.66 | 0.65 | -0.69 | -0.21 | 0.81 |
| 2 | 0.93 | 0.88 | 0.19 | 0.68 | 0.68 | 0.66 | -0.67 | -0.28 | 0.79 |
| 3 | 0.91 | 0.891 | 0.15 | 0.63 | 0.65 | 0.67 | -0.62 | -0.27 | 0.7 |
| 4 | 0.97 | 0.87 | 0.4 | 0.63 | 0.64 | 0.69 | -0.81 | -0.34 | 0.69 |
| 5 | 0.98 | 0.86 | 0.35 | 0.65 | 0.68 | 0.67 | -0.87 | -0.37 | 0.67 |
| 6 | 0.94 | 0.875 | 0.42 | 0.61 | 0.62 | 0.62 | -0.82 | -0.39 | 0.65 |
| 7 | 0.94 | 0.89 | 0.6 | 0.74 | 0.75 | 0.68 | -0.8 | -0.2 | 0.62 |
| 8 | 0.97 | 0.91 | 0.47 | 0.70 | 0.72 | 0.64 | -0.75 | -0.28 | 0.69 |
| .. |  |  |  |  |  |  |  |  |  |
| .. |  |  |  |  |  |  |  |  |  |
| 70 | 0.97 | 0.91 | 0.7 | 0.69 | 0.7 | 0.69 | -0.5 | 0.9 | 0.7 |
| 71 | 0.94 | 0.93 | 0.69 | 0.67 | 0.69 | 0.64 | -0.48 | 0.89 | 0.72 |
| 72 | 0.98 | 0.92 | 0.72 | 0.71 | 0.78 | 0.71 | -0.54 | 0.78 | 0.73 |

Based on the data in Table 3, the varying correlation coefficients suggest that the influencing factors have a seasonal nature. This indicates that solely relying on the results of statistical correlation analysis does not allow for accurate conclusions about the influencing factors [5,6].

**RESEARCH RESULTS**

As a result, to improve the accuracy of the correlation analysis, it becomes necessary to analyze the dynamic changes of the calculated coefficients. Dynamic analysis enables the evaluation of the changes in correlation coefficients obtained for each group or season, providing deeper insights into their variability (Figure 2).



**Figure 2.** Dynamic Changes of Factors Influencing Electricity Losses

From the graph in Figure 2, depicting the dynamic changes of factors affecting electricity losses, it can be observed that some factors exhibit strong correlations during specific periods while showing weak or moderate correlations during others, or vice versa.

In general, the high-impact correlation analysis suggests that determining factors influencing electricity losses based on one month's data or even a year's historical data may lead to inaccuracies. Therefore, to identify factors influencing future electricity losses using past data, the correlation coefficients (r) for each factor should be collected in a matrix, and a model of their variations should be constructed [6,7]. Analyzing these dynamic changes allows for more accurate conclusions and predictions.

For this reason, models of correlation coefficients for the next three months will be developed using regression analysis (Table 4) to enhance precision.

It is known that, in general, a regression model takes the following form [source]:

(2)

If the sum of the intercept terms is considered as , then expression (1) takes the following form:

(3)

There are several methods for determining the terms in a regression model, with one of the most widely used being the linear regression method [source].

Linear regression, in turn, consists of:

1. **Simple Linear Regression** – constructing a model of the relationship between a single independent variable and a single dependent variable.
2. **Multiple Linear Regression** – analyzing and modeling the relationship between a single dependent variable and multiple independent variables [source].

Based on the second expression (2), the linear regression model takes the following form:

(4)

Where: Y- dependent variable (electricity losses), independent variables (factors influencing electricity losses), - model coefficients, intercept term, -error term.

The most commonly used method for calculating the coefficients of a linear regression model is the **least squares method**. This method minimizes the error, i.e., the difference between the observed values and the values predicted by the model [source].

The coefficients of a simple linear regression model are calculated as follows:

(5)

(6)

In the case of multiple linear regression, matrices are used to calculate the coefficients. The formula is as follows:

= (7)

In the development of models, regression models were created based on the algorithm for selecting an electricity loss model with the least error, as presented in Figures 1 and 2 of subsection 3.1 of the dissertation. These models included single-factor linear regression and polynomial regression models up to the 8th degree. Among the constructed models, the one with the least error was selected.

For example, regression models and RMSE (Root Mean Square Error) values were determined for the effect of the influencing temperature factor on electricity losses.

**Table 4.** Regression Models for Influencing Temperature and Their RMSE Errors

|  |  |  |
| --- | --- | --- |
| № | Regression Model | RMSE % |
| 1 |  | 2.3 |
| 2 |  | 4.5 |
| 3 |  | 2.1 |
| 4 |  | 3.8 |
| 5 |  | 5.2 |
| 6 |  | 6.1 |
| 7 |  | 5.8 |
| 8 |  | 4.6 |

Table.4 demonstrates that the third-degree regression model for the influencing temperature factor has the smallest RMSE error. Based on this observation, regression models up to the 8th degree were constructed for the remaining influencing factors as well. The model with the smallest RMSE error was selected for each factor. These operations were performed automatically (see Table.5).

Using the regression models for the identified factors influencing electricity losses, their values for the next three months were predicted. These predicted values underwent correlation analysis, and factors with a high correlation (r>0.6) were selected. Conversely, factors with low (0<r<0.3) or moderate (0.3≤r≤0.6) correlation were considered non-influential for the current season.

It is important to note that the algorithm retains all identified factors, as these factors could be highly influential in other seasons. This approach ensures a comprehensive consideration of potential influencing variables across different time periods.

**Table 5.** Regression Models for Factors Influencing Electricity Losses and Their RMSE Errors

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № | Influencing Factor | Regression Model | Model Degree | RMSE % |
| 1 | Active Electricity Consumption (kWh) |  | 6 | 2.4 |
| 2 | Reactive Electricity Consumption (kVAR·h) |  | 4 | 3.2 |
| 3 | Day of the Week | + | 5 | 4.5 |
| 4 | Minimum Temperature (°C) | + | 4 | 2.8 |
| 5 | Maximum Temperature (°C) | + | 4 | 2.6 |
| 6 | Effective Temperature (°C) | + | 4 | 2.1 |
| 7 | Relative Humidity (%) |  | 5 | 5.2 |
| 8 | Visibility |  | 4 | 5.8 |
| 9 | Load Factor |  | 5 | 4.4 |

**CONCLUSIONS**

The research demonstrated that electricity losses are influenced by a complex interplay of technical, climatic, and operational factors, whose significance changes dynamically across seasons. The Pearson correlation method, applied to grouped 10-day datasets, allowed for the identification of both strong and weak correlations between losses and independent variables. However, the results showed that correlation coefficients alone cannot provide stable and reliable insights due to seasonal fluctuations and variability. To address this limitation, regression analysis was employed, with models ranging from linear to eighth-degree polynomial functions. Selection of the optimal model was guided by minimizing RMSE values, with the third-degree regression model for temperature emerging as the most accurate predictor. The combined use of correlation analysis, dynamic monitoring of coefficients, and regression modeling enhances the precision of identifying high-impact factors on electricity losses. This methodological framework not only ensures seasonally adaptive forecasting but also supports the development of effective energy-saving strategies and operational measures. Consequently, the study contributes to advancing analytical tools for electricity loss reduction, strengthening the reliability of energy systems, and improving efficiency in line with modern power sector requirements.

**REFERENCES**

1. Li, X., Paster, M., & Stubbins, J. (2015). *The dynamics of electricity grid operation with increasing renewables and the path toward maximum renewable deployment*. Renewable and Sustainable Energy Reviews, 47, 1007-1015. <https://doi.org/10.1016/j.rser.2015.03.039>
2. Saeed, M. S., Mustafa, M. W., Hamadneh, N. N., & Alshammari, N. (2020). Detection of non-technical losses in power utilities: A comprehensive systematic review. *Energies, 13*(18), 4727. <https://doi.org/10.3390/en13184727>
3. Liu, X., Zhang, L., Zou, L., Wang, J., Li, Y., Kurbonov, N. N., & Rakhmonov, I. U. (2025). A unified wind power prediction framework combined with individual wind turbine operation status and error correction. *Energy Reports, 13,* 6117–6131. <https://doi.org/10.1016/j.egyr.2025.05.065>
4. Correia, E., Calili, R., Pessanha, J. F., & Almeida, M. F. (2023). Definition of regulatory targets for electricity non-technical losses: Proposition of an automatic model-selection technique for panel data regressions. *Energies, 16*(6), 2519. <https://doi.org/10.3390/en16062519>
5. Mbasso, W. F., Molu, R. J. J., Harrison, A., Pushkarna, M., Nguemo Kemdoum, F., Donfack, E. F., Jangir, P., & Tiako, P. (2024). Hybrid modeling approach for precise estimation of energy production and consumption based on temperature variations. *Scientific Reports, 14*, 24422. <https://doi.org/10.1038/s41598-024-75244-0>
6. Rakhmonov, I. U., Niyozov, N. N., Ushakov, V. Ya., Kurbonov, N. N., & Najimova, A. M. (2024). Forecasting electricity consumption using the Principal Component Analysis method. *Bulletin of the Tomsk Polytechnic University: Geo Assets Engineering, 335*(12), 198–209. <https://doi.org/10.18799/24131830/2024/12/4731>
7. Coma-Puig, R., López, X., Frías, P., & Morales-Álvarez, R. (2022). Non-technical loss detection in energy consumption using a predictive regression system. *Machine Learning, 111*(6), 1963–1983. https://doi.org/10.1007/s10994-021-06051-1