Таслимов Абдурахим Дехконович
д.т.н., проф., ТГТУ, Ташкент
Taslimov Abdurakhim Dekhonovich, TSTU

Султанов Асрор Нурали угли
PhD студент ТГТУ, Ташкент
Sultanov Asror Nurali ugli

ДИНАМИЧЕСКИЙ АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА ПОТЕРИ ЭЛЕКТРОЭНЕРГИИ
DYNAMIC ANALYSIS OF FACTORS AFFECTING ELECTRICITY LOSSES

Аннотация: В данной статье рассматривается динамический анализ факторов, влияющих на потери электроэнергии, с использованием статистических и регрессионных методов. Коэффициенты корреляции Пирсона рассчитываются для определения взаимосвязей между потерями электроэнергии и различными факторами, такими как температура, влажность и коэффициенты нагрузки. Исследование подчеркивает сезонную изменчивость этих факторов и важность динамических изменений коэффициентов корреляции для точного прогнозирования. Разработаны регрессионные модели, включая линейные и полиномиальные модели до 8-й степени, для прогнозирования потерь электроэнергии. Для каждого фактора выбирается модель с наименьшей ошибкой, что обеспечивает надежный метод анализа и снижения потерь электроэнергии.

Abstract: This article investigates the dynamic analysis of factors influencing electricity losses using statistical and regression methods. Pearson correlation coefficients are calculated to determine relationships between electricity losses and various factors such as temperature, humidity, and load factors. The study highlights the seasonal variability of these factors and emphasizes the importance of dynamic changes in correlation coefficients for accurate predictions. Regression models, including linear and polynomial models up to the 8th degree, are developed to forecast electricity losses. The model with the least error for each factor is selected for prediction, providing a robust methodology for analyzing and mitigating electricity losses.

Ключевые слова: Потери электроэнергии, энергоэффективность, коэффициент нагрузки, корреляция Пирсона, регрессионные модели, сезонные факторы, влияние климата, прогнозирование, полиномиальная регрессия, динамический анализ.

Keywords: Electricity losses, energy efficiency, load factor, Pearson correlation, regression models, seasonal factors, climate impact, forecasting, polynomial regression, dynamic analysis.

Introduction. To date, a wide range of statistical methods, such as Pearson, Spearman, Kendall's tau, Phi coefficient, partial correlation [1-3], 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. 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).

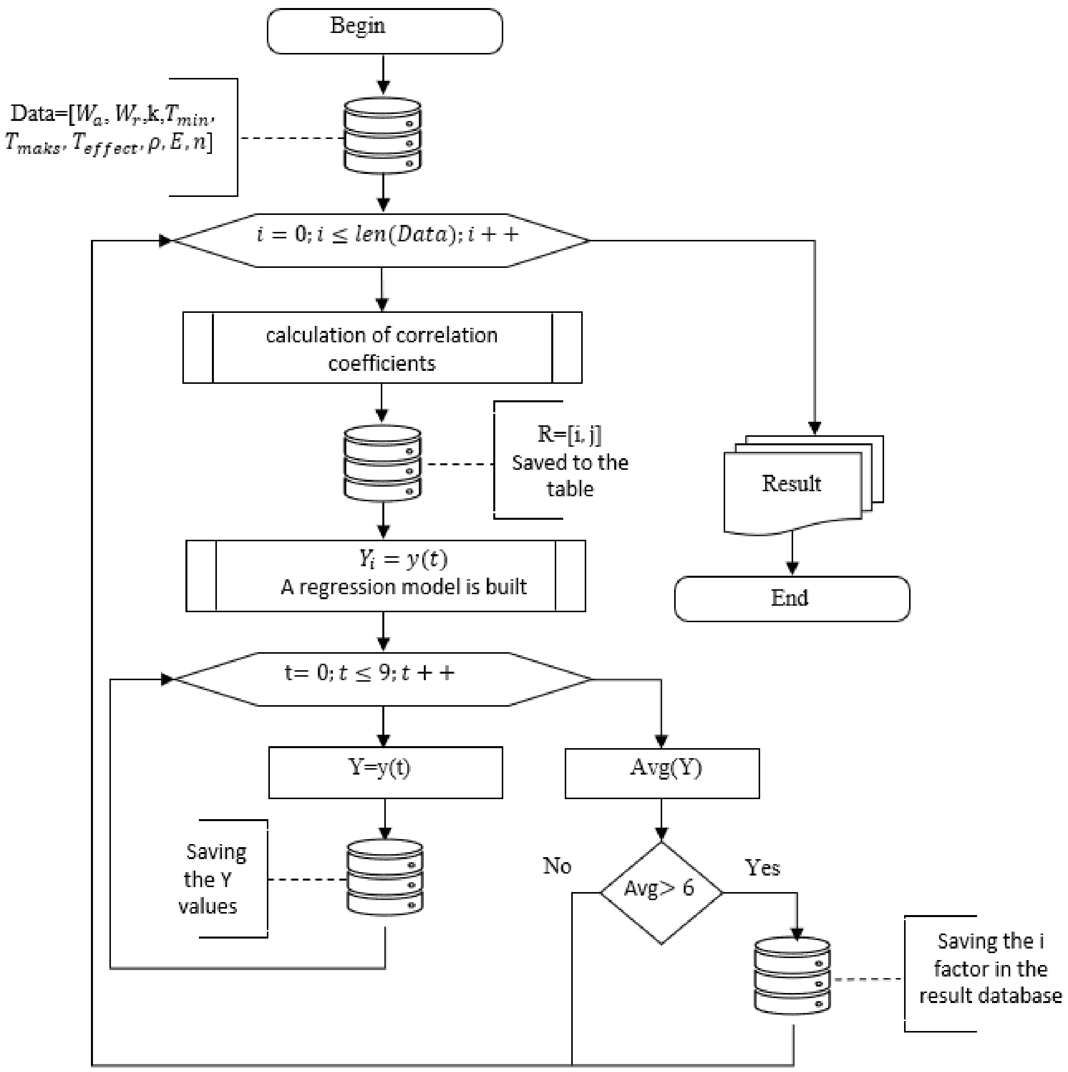

Figure 1. Algorithm for Identifying Factors Affecting Electricity Losses

Methods. Based on the provided algorithm, the daily collected data were analyzed, and the results were grouped into 10-day intervals. This grouping was necessary because the scarcity of time series data increases the accuracy of the obtained results. Based on the 10-day data, 72 groups were formed, and for each group, the Pearson correlation coefficients of the factors affecting electricity losses were calculated using the algorithm shown in Figure 1. The calculations were performed according to the following formula, as detailed in Table 1.

$$r = \frac{\sum(X_i - X)(Y_i - Y)}{\sqrt{\sum(X_i - X)^2 \sum(Y_i - Y)^2}} \qquad (1)$$

Here, X_i and Y_i 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.

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:

- $0 < r < 0.3$ or $0 < r < -0.3$: Weak positive or negative correlation.
- $0.3 \leq r \leq 0.6$ or $-0.3 \leq r \leq -0.6$: Moderate positive or negative correlation.
- $0.6 < r \leq 1$ or $-0.6 < r \leq -1$: Strong positive or negative correlation.

This classification is based on the observed ranges of r [4].

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	Maximum Temperature	Effective Temperature	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	Maximum Temperature	Effective Temperature	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.

Discussion. 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).

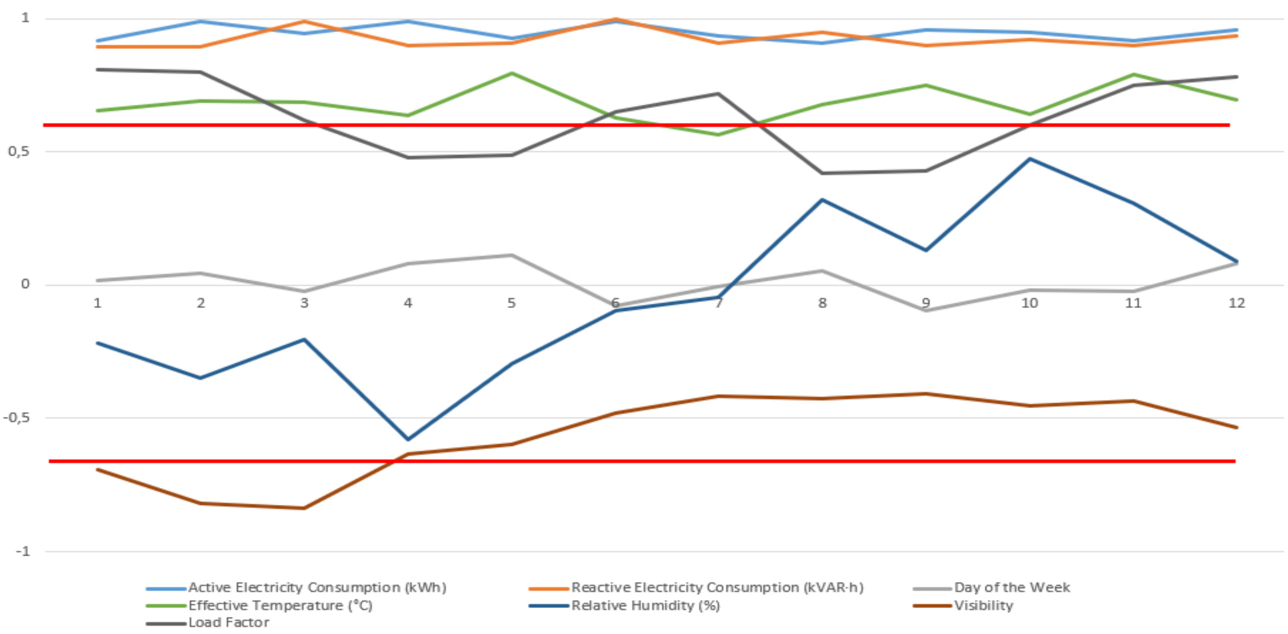

Figure 2. Dynamic Changes of Factors Influencing Electricity Losses

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. 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:

$$Y = \sum_{i=0}^{n} \beta_i^1 x_i^n + \beta_i^2 x_i^{n-1} + \cdots + \beta_i^{n-1} x_i + \beta_i^n \quad \textbf{(1)}$$

If the sum of the intercept terms is considered as $\beta = \sum_{i=0}^{n} \beta_i$, then expression (1) takes the following form:

$$Y = \sum_{i=0}^{n} (\beta_i^1 x_i^n + \beta_i^2 x_i^{n-1} + \cdots + \beta_i^{n-1} x_i) + \beta \quad \textbf{(2)}$$

There are several methods for determining the terms in a regression model, with one of the most widely used being the linear regression method [5].

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 [6].

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

$$Y = \beta_0 + \beta_1 X_1 + \beta_1 X_2 + \cdots + \beta_n X_n + \epsilon \quad \textbf{(3)}$$

Where: Y- dependent variable (electricity losses), $X_1, X_2, X_n -$ independent variables (factors influencing electricity losses), $\beta_1 ... \beta_n$- model coefficients, $\beta_0 -$ intercept term, ϵ-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 [7].

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

$$\beta_1 = \frac{\sum(x_1 - \bar{x})(y_1 - \bar{y})}{\sum(x_1 - \bar{x})^2} \quad \textbf{(4)}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{y} \quad \textbf{(5)}$$

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

$$\begin{bmatrix} n & \Sigma x_{1i} & \Sigma x_{2i} \\ \Sigma x_{1i} & \Sigma x_{1i}^2 & \Sigma x_{1i} x_{2i} \\ -- & -- & -- \\ \Sigma x_{(n-1)i} & \Sigma x_{ni}^2 & \Sigma x_{ni} x_{ni} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ - \\ \beta_{n-1} \end{bmatrix} = \begin{bmatrix} \Sigma y_i \\ \Sigma x_{1i} y_i \\ - \\ \Sigma x_{(n-1)i} y_i \end{bmatrix} \quad \textbf{(6)}$$

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	$y = 0.667 - 0.002x$	2.3
2	$y = 0.641 + 0.014x - 0.002x^2$	4.5
3	$y = 0.611 + 0.046x - 0.010x^2 - 0.001x^3$	2.1
4	$y = 0.6075 + 0.0503x - 0.0120x^2 + 0.0009x^3 - 0.0850x^4$	3.8
5	$y = 0.9375 - 0.5609x + 0.3636x^2 - 0.0996x^3 + 0.0121x^4 - 0.0105x^5$	5.2
6	$y = -0.0375 + 1.5670x - 1.3057x^2 + 0.5226x^3 - 0.1068x^4 + 0.0107x^5 - 0.0104x^6$	6.1
7	$y = 0.900 - 0.747x + 0.867x^2 - 0.509x^3 + 0.165x^4 - 0.029x^5 + 0.013x^6 - 0.058x^7$	5.8
8	$y = 0.611 + 0.039x + 0.021x^2 - 0.026x^3 + 0.014x^4 + 0.023x^5 - 0.041x^6 + 0.012x^7 - 0.078x^8$	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).

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)	$y = -0.5100 + 3.2801x - 2.7710x^2 + \cdots - 0.0041x^6$	6	2.4
2	Reactive Electricity Consumption (kVAR·h)	$y = 0.8627 + 0.0417x - 0.0196x^2 + 0.0029x^3 - 0.0021x^4$	4	3.2
3	Day of the Week	$y = -0.19886 + 0.0131x - 0.0632x^2 + 0.019x^3 - 0.012x^4 + 0.001x^5$	5	4.5
4	Minimum Temperature (°C)	$y = 0.571 + 0.145x - 0.1080x^2 - 0.189x^3 + 0.004x^4$	4	2.8
5	Maximum Temperature (°C)	$y = 0.831 + 0.081x - 0.130x^2 - 0.021x^3 + 0.003x^4$	4	2.6
6	Effective Temperature (°C)	$y = 0.611 + 0.046x - 0.010x^2 - 0.001x^3 + 0.01x^4$	4	2.1
7	Relative Humidity (%)	$y = 0.5625 - 1.4631x + 0.9211x^2 - 0.2610x^3 + 0.0332x^4 - 0.0015x^5$	5	5.2
8	Visibility	$y = -0.343 + 0.2976x - 0.1564x^2 + 0.1274x^3 - 0.045x^4$	4	5.8
9	Load Factor	$y = 0.6475 - 0.517x - 0.3457x^2 + 0.046x^3 + 0.0127 - 0.0035x^5$	5	4.4

Conclusion. 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.

Список литературы:

1. Shaqiri, Mirlinda & Iljazi, Teuta & Kamberi, Lazim & Halili, Rushadije. (2023). DIFFERENCES BETWEEN THE CORRELATION COEFFICIENTS PEARSON, KENDALL AND SPEARMAN.

2. Okoye, Kingsley & Hosseini, Samira. (2024). Correlation Tests in R: Pearson Cor, Kendall's Tau, and Spearman's Rho. 10.1007/978-981-97-3385-9_12.

3. El-Hashash, Essam & Hassan, Raga. (2022). A Comparison of the Pearson, Spearman Rank and Kendall Tau Correlation Coefficients Using Quantitative Variables. Asian Journal of Probability and Statistics. 20. 36-48. 10.9734/AJPAS/2022/v20i3425.

4. T, Saadalov & R, Myrzaibraimov & Abdullaeva, Zhypargul. (2022). МЕТОДИКА РАСЧЕТА КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ ФЕХНЕРА И ПИРСОНА, И ИХ ОБЛАСТИ ПРИМЕНЕНИЯ CALCULATING PROCEDURE FOR THE CORRELATION COEFFICIENT OF FECHNER AND PEARSON AND THEIR APPLICATION AREAS. Bulletin of Science and Practice.

5. De Carvalho, Miguel & Kumukova, Alina & Dos Reis, Gonçalo. (2022). Regression-type analysis for multivariate extreme values. Extremes. 25. 1-28. 10.1007/s10687-022-00446-6.

6. de Carvalho, Miguel & Kumukova, Alina & Dos Reis, Gonçalo. (2022). Regression-type analysis for multivariate extreme values. Extremes. 25. 1-28. 10.1007/s10687-022-00446-6.

7. Wang, Shuai & Wang, Yu & Song, Yiwei. (2024). Least Squares Generalization-Memorization Regression. Advances in Computer and Engineering Technology Research. 1. 41. 10.61935/acetr.2.1.2024.P41.