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

МОДЕЛИРОВАНИЕ ПОТЕРЬ ЭЛЕКТРОЭНЕРГИИ С ИСПОЛЬЗОВАНИЕМ РЕГРЕССИОННОГО АНАЛИЗА
MODELING ELECTRICITY ENERGY LOSSES USING REGRESSION ANALYSIS

Аннотация: В данной статье рассматривается использование регрессионного анализа как основного метода моделирования потерь электроэнергии. Были разработаны и оценены различные полиномиальные регрессионные модели с использованием данных временных рядов. Алгоритм включает предварительную обработку данных, расчет коэффициентов методом наименьших квадратов и валидацию модели на основе средней абсолютной процентной ошибки (MAPE). Исследование подчеркивает эффективность полиномиального регрессионного анализа, особенно при применении к наборам данных различного размера и длительности. Результаты показывают, что модели высокой степени обеспечивают большую точность для больших объемов данных, но требуют больше времени на вычисления. Работа демонстрирует потенциал регрессионных методов для эффективного и точного моделирования потерь электроэнергии.

Abstract: This paper explores the use of regression analysis as a fundamental method for modeling electricity energy losses. Various polynomial regression models are developed and evaluated using data from time series analysis. The algorithm incorporates data preprocessing, coefficient calculation using the least squares method, and model validation based on the Mean Absolute Percentage Error (MAPE). The study highlights the effectiveness of polynomial regression, particularly when applied to datasets of varying sizes and durations. Results demonstrate that higher-degree models provide greater accuracy for larger datasets, but they also require more computation time. This work demonstrates the potential of regression methods for efficient and accurate modeling of electricity energy losses.

Ключевые слова: Регрессионный анализ, Полиномиальная регрессия, Потери электроэнергии, Ошибка MAPE, Моделирование временных рядов, Точность модели.

Keywords: Regression analysis, Polynomial regression, Electricity energy loss, MAPE error, Time series modeling, Model accuracy.

Introduction. Nowadays, there are numerous methods for modeling electricity energy losses or consumption, among which regression analysis is considered one of the fundamental approaches [1]. Regression analysis allows for the creation of specific models by determining the values of coefficients for independent variables, giving precise outcomes. Based on the literature review conducted during the research, it can be concluded that it is possible to form any model using regression analysis. However, performing calculations for each type of regression model is necessary, which may lead to significant time consumption and an increase in errors. Currently, using computational machines, models up to the 6th degree can be calculated. However, modern technologies, such as software or programming language packages based on artificial intelligence, enable the creation of regression models of any degree. It is known that artificial intelligence-based methods [2,3] are also among the effective approaches. However, these methods are based on iterative techniques and assign different values for various neurons. The downside of such methods is that the model produces different values each time it is retrained.

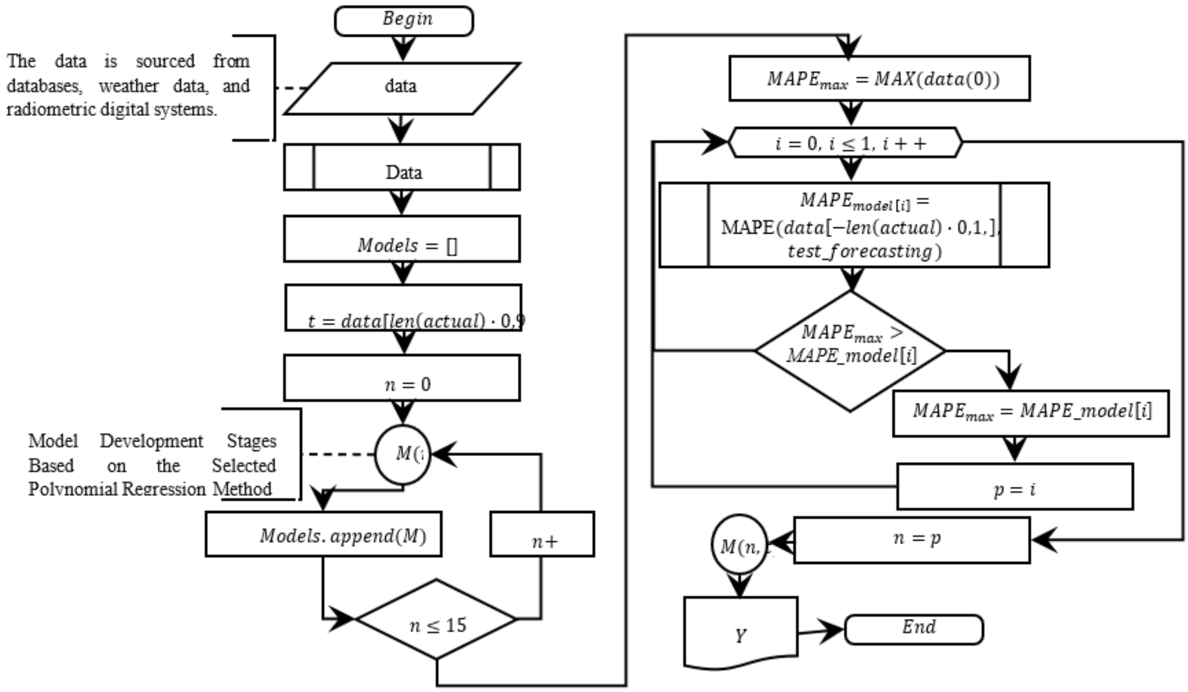

Figure 1. Algorithm for Selecting the Electricity Energy Loss Model Using Multifactor Polynomial Regression Analysis with Minimum Error

This occurs because the model does not accept a single value but rather selects a range corresponding to each value, and during retraining, a specific value is taken from this range. One of the advantages of these methods is the rapid execution of the algorithm, requiring minimal time.

Methods. Using polynomial methods [4] and Python packages, the following algorithm was developed, allowing the selection of the optimal model from regression models of up to the 15th degree (see Figure 1). In the algorithm, data is initially received and processed. Then, an empty cell is reserved for creating models for each input data point, and a model is formed based on 90% of the total data. As is known, in polynomial models, a zero-degree model represents a straight line, a first-degree model is linear, a second-degree model is quadratic, and so on. To construct these models, a zero value is initially declared, and part M of the algorithm is executed (see Figure 2).

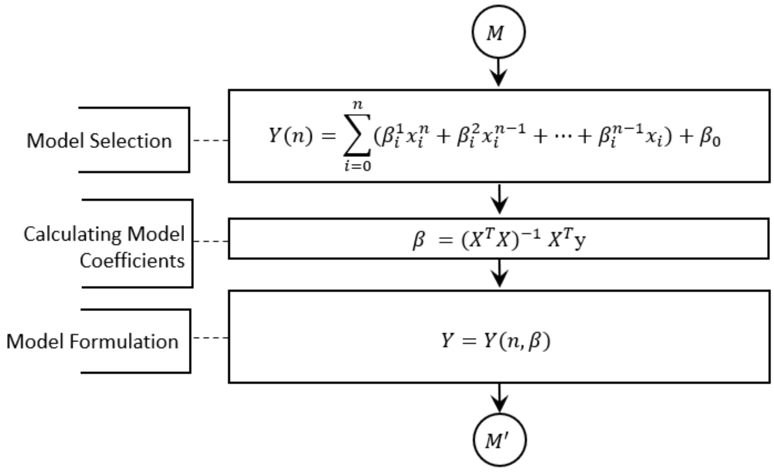

Figure 2. Algorithm for Building an Electricity Energy Loss Model Using Multifactor Polynomial Regression Analysis

According to part M of the algorithm, coefficients are determined using the least squares method, and the model is constructed. The developed model is added to the list of models created earlier. Then, for the next calculation, one higher degree is selected, and a new model is built. This process continues based on the algorithm presented in Figure 2, up to the 15th degree.

Discussion. After constructing the models, the Mean Absolute Percentage Error (MAPE) [5] is calculated, and the model with the smallest MAPE error is printed. Additionally, the test data (10% reserved earlier) is compared with test forecasting results. The degree of the model with the smallest error is determined, and the remaining 100% of the data is fed into the selected model for validation.

Initially, models for one-month, two-month, four-month, one-year, and two-year periods are created for all 15 degrees. Among these, the model with the best training fit (i.e., the lowest error) is identified. The training fit is assessed by comparing the train data's MAPE error with its predictions.

Using data from the research subject for the month of August, the model presented in Figure 3 was obtained with the help of the algorithm in Figure 1.

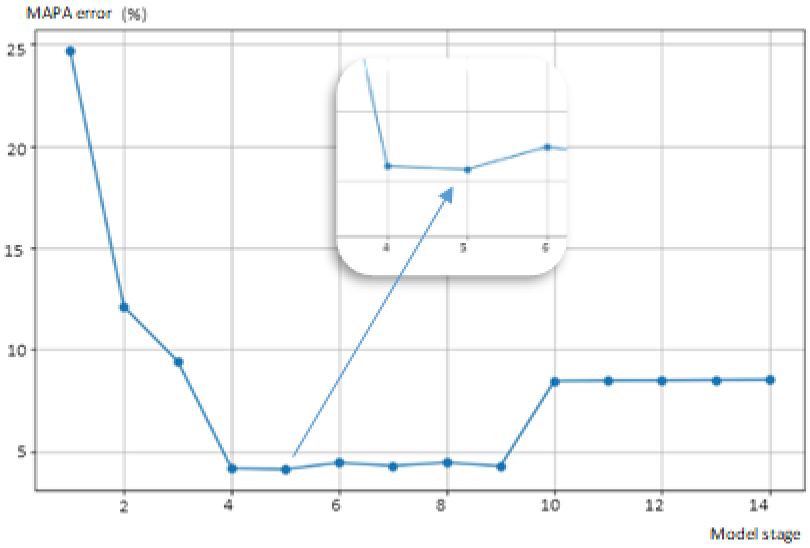

Figure 3. Selecting the Model Degree with the Smallest MAPE Error

From the graph above, it is evident that the most optimal model, with the smallest error, is the 5th-degree model. Based on this model, the adaptation graph was constructed (see Figure 4).

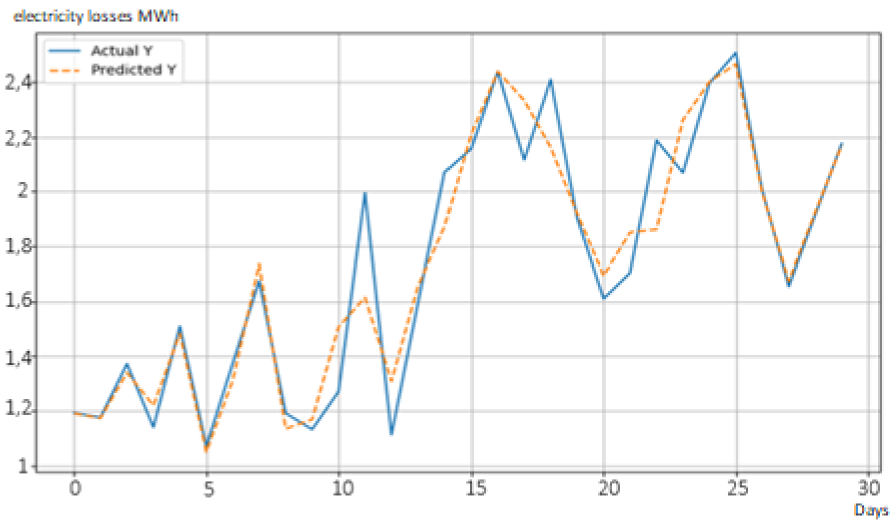

Figure 4. Adaptation Graph of the Polynomial Regression Model

From the adaptation graph, it is clear that the graphs do not overlap perfectly, indicating a 2.04% MAPE error in adaptation.

The laboratory analysis presented in Table 3 demonstrates that different time series data require models of varying degrees. As the quantity of data increases, the model's adaptation performance decreases, and the time required to develop the model proportionally increases.

• For one month of data (30 time series points), a 5th-degree model provides a 4% MAPE error and requires 6 seconds to construct.

• For twelve months of data (366 points), an 11th-degree model gives a higher accuracy, but with an error of 18.2%, and takes 28 seconds to compute.

• For two years of data (732 points), a slightly higher-degree model yields a 23.5% MAPE error, with a significantly longer computation time.

Table 3

Model Degree, MAPE Error, and Computation Time Based on Extracted Time Series Data

№	Data Duration (Months)	Model Degree	MAPE Error (%)	Time (seconds)
1	1 (January)	5	4	6
2	1 (August)	7	4.2	6
3	2 (February-March)	14	5.5	9
4	2 (September-October)	12	5.3	10
5	4 (March-June)	10	8.8	16
6	4 (October-January)	9	8.3	16
7	12	11	18.2	28
8	24	15	23.5	35

Conclusion. The analyses presented above indicate that polynomial degrees of different levels are suitable for time series data from various intervals. Correspondingly, for a two-month analysis during the summer season, a 14th-degree model required 9 seconds to compute, while for two months of data from the winter season, the error differs significantly. Specifically, a 12th-degree model had a 5.3% MAPE error and required 10 seconds for computation. This demonstrates that polynomial regression modeling enables the creation of highly accurate and high-degree models within a short time frame.

Список литературы:

1. Steposhina, Svetlana. (2024). Improving the accuracy of regression models by methods of polynomial. Transport engineering. 2024. 4-12. 10.30987/2782-5957-2024-4-4-12.

2. He, Zhonghao & Achterberg, Jascha & Collins, Katie & Nejad, Kevin & Akarca, Danyal & Yang, Yinzhu & Gurnee, Wes & Sucholutsky, Ilia & Tang, Yuhan & Ianov, Rebeca & Ogden, George & Li, Chole & Sandbrink, Kai & Casper, Stephen & Ivanova, Anna & Lindsay, Grace. (2024). Multilevel Interpretability Of Artificial Neural Networks: Leveraging Framework And Methods From Neuroscience. 10.48550/arXiv.2408.12664.

3. Hanif, Dr. (2024). Gauss- Newton Method for Feedforward Artificial Neural Networks.

4. Steposhina, Svetlana. (2024). IMPROVING THE ACCURACY OF REGRESSION MODELS BY METHODS OF POLYNOMIAL APPROXIMATION. Transport engineering. 2024. 4-12. 10.30987/2782-5957-2024-4-4-12.

5. Ahmar, Ansari. (2020). Forecast Error Calculation with Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). JINAV: Journal of Information and Visualization. 1. 10.35877/454RI.jinav303.