



Taqsimlovchi elektr tarmoqlarida elektr energiyasi isroflarini prognozlash modellarini ma'lumotlar murakkabligi mezon asosida tanlash

Xilola Z. Nazirova

PhD, dotsent, Toshkent davlat texnika universiteti, Toshkent, 100095, O'zbekiston; tstu_energy@mail.ru
<https://orcid.org/0009-0002-7735-4862>

Dolzarbligi: 0,4 kV taqsimlovchi tarmoqlarda elektr energiyasi isroflarini prognozlashda bir xil model barcha vaqt qatorlari uchun bir xil samaradorlikni ta'minlamaydi. AMI va smart-meter ma'lumotlari hajmining ortishi ma'lumotlar tuzilishiga mos modelni tanlash masalasini alohida ilmiy vazifaga aylantiradi.

Maqsad: ma'lumotlar hajmi va variativligini birlashtiruvchi integral mezon asosida Regression, GMDH, ARIMA, MLP va DBSCAN-MLP modellarining ratsional qo'llanish diapazonlarini aniqlash.

Usullari: 2021-2023-yillardagi 912 ta real ekspluatatsion kuzatuv, korrelyatsion tahlil, Fisher o'zgartirishi va Z-mezon, polinomial regressiya, ARIMA, GMDH, MLP, DBSCAN-MLP hamda MAPE baholash mezonidan foydalanildi. Model tanlash uchun $\theta = \sqrt{\lg(M)^2 + CV^2}$ mezon qo'llanildi.

Natijalar: 20 ta bir omilli va 20 ta ko'p omilli tanlanma tahlili asosida model tanlashning empirik chegaralari olindi. Umumiy MAPE qiymatlari Regression, ARIMA, GMDH, MLP va DBSCAN-MLP uchun mos ravishda 5,01; 3,15; 5,05; 2,58 va 2,10 % ni tashkil etdi. Ko'p omilli ma'lumotlarda gibrid modelga o'tish chegarasi $\theta \geq 3,9$, bir omilli ma'lumotlarda $\theta \geq 4,2$ deb aniqlandi.

Kalit so'zlar: elektr energiyasi isroflari, 0,4 kV taqsimlovchi tarmoq, prognozlash, model tanlash, ma'lumotlar murakkabligi, MLP, DBSCAN.

Выбор моделей прогнозирования потерь электроэнергии в распределительных сетях на основе критерия сложности данных

Хилола З. Назирова

For citation: Kh.Z. Nazirova. Data-complexity-based selection of electricity-loss forecasting models in 0.4 kV distribution networks. Scientific and technical journal of Problems of Energy and Sources Saving, 2026, no. 3, pp. 194-202.

<https://doi.org/10.5281/zenodo.23241972>

Received: 17.02.2026

Revised: 26.03.2026

Accepted: 18.04.2026

Published: 25.09.2026

Copyright, Khilola Z. Nazirova, 2026. Submitted to Problems of Energy and Sources Saving for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

PhD, доцент, Ташкентский государственный технический университет, Ташкент, 100095, Узбекистан; tstu_energy@mail.ru <https://orcid.org/0009-0002-7735-4862>

Актуальность: рост объемов AMI и smart-meter данных делает научно-обоснованный выбор прогнозной модели самостоятельной задачей: одна и та же модель не обеспечивает одинаковую точность для временных рядов различной сложности.

Цель: определить рациональные диапазоны применения Regression, GMDH, ARIMA, MLP и DBSCAN-MLP по интегральному критерию, объединяющему объем и вариативность данных.

Методы: использованы 912 эксплуатационных наблюдений за 2021-2023 гг., корреляционный анализ, преобразование Фишера и Z-критерий, полиномиальная регрессия, ARIMA, GMDH, MLP, DBSCAN-MLP и показатель MAPE. Для предварительного выбора модели применен критерий $\theta = \sqrt{\lg(M)^2 + CV^2}$.

Результаты: по 20 однофакторным и 20 многофакторным выборкам получены эмпирические границы выбора моделей. Общий MAPE составил 5,01; 3,15; 5,05; 2,58 и 2,10 %, соответственно для Regression, ARIMA, GMDH, MLP и DBSCAN-MLP.

Ключевые слова: потери электроэнергии, распределительная сеть 0,4 кВ, прогнозирование, выбор модели, сложность данных, MLP, DBSCAN.

Data-complexity-based selection of electricity-loss forecasting models in 0.4 kV distribution networks

Khilola Z. Nazirova

PhD, Associate Professor, Tashkent State Technical University, Tashkent, 100095, Uzbekistan; tstu_energy@mail.ru
<https://orcid.org/0009-0002-7735-4862>

Relevance: the growing volume of AMI and smart-meter data makes model selection a separate methodological problem in electricity-loss forecasting. A single forecasting method does not provide uniform accuracy across time series that differ in sample size, variability, seasonality, and multivariate structure.

Aim: to determine rational applicability ranges of Regression, GMDH, ARIMA, MLP and DBSCAN-MLP forecasting models using an integral criterion that combines data volume and relative variability.

Methods: the study used 912 real operational observations collected in 2021-2023, correlation analysis, Fisher transformation and Z-testing, polynomial regression, ARIMA, GMDH, MLP, DBSCAN-MLP, and MAPE-based error assessment. Data complexity was characterized by $\theta = \sqrt{\lg(M)^2 + CV^2}$.



Results: experiments on 20 univariate and 20 multivariate datasets yielded empirical model-selection boundaries. Overall MAPE values were 5.01, 3.15, 5.05, 2.58 and 2.10% for Regression, ARIMA, GMDH, MLP and DBSCAN-MLP, respectively. The hybrid-model threshold was $\theta \geq 4.2$ for univariate and $\theta \geq 3.9$ for multivariate data.

Keywords: electricity losses, 0.4 kV distribution network, forecasting, model selection, data complexity, MLP, DBSCAN.

1. Introduction

Electricity losses in low-voltage distribution networks remain an important technical and economic indicator because they reflect not only the condition of network elements but also the quality of metering, loading regimes, and the efficiency of operational control. The 0.4 kV level is particularly difficult for data-driven analysis owing to the large number of consumers, highly branched feeders, rapidly changing loading conditions, and sensitivity to weather and behavioral factors. Previous studies by the author have addressed loss calculation in 10/0.4 kV networks, statistical characteristics of electrical loads, and the applicability of neural-network and GMDH models to loss forecasting [1–4].

The author's recent work has progressively examined individual forecasting classes. Polynomial regression was used to obtain an interpretable nonlinear approximation of electricity losses [5], ARIMA was investigated for temporal dependence in 0.4 kV loss series [6], and ANN optimization was studied as a way to increase predictive accuracy for nonlinear data [7]. In addition, direct comparison of ANN and GMDH demonstrated that the relative performance of forecasting methods changes with data structure and validation regime [2]. These results motivate a broader question: rather than optimizing a single model in isolation, can the model class itself be selected from measurable properties of the input data?

This question is consistent with the current international trend toward data-dependent forecast-model selection. Smart-meter data have already been used for machine-learning-based loss forecasting in secondary distribution networks [9]. Studies of forecasting-model selection show that alternative selection criteria may disagree even for the same time series [10], while cluster-based approaches can reduce the candidate model space by grouping time series according to their characteristics [11]. Complexity-aided time-series modeling treats series complexity as an explicit source of information for model construction [12], and dynamic selection strategies show that using one fixed model may be suboptimal when time-series characteristics change over time [13].

For power-system applications, these observations are especially relevant because the statistical structure of the data changes with feeder loading, external temperature, humidity, seasonality, data volume, and the number of explanatory variables. A low-variability univariate series may not justify a computationally expensive nonlinear architecture, whereas a large multivariate series with pronounced structural heterogeneity may require neural-network or hybrid preprocessing-and-learning schemes.

Accordingly, the scientific problem addressed in this paper is not the development of another isolated forecasting model. The focus is the development of a preselection criterion that links measurable data characteristics to a rational forecasting-model class. The proposed approach combines the logarithmically scaled data volume and the coefficient of variation into an integral criterion θ and uses extensive comparative experiments to establish empirical selection boundaries for Regression, GMDH, ARIMA, MLP, and DBSCAN-MLP.

The main contributions of the study are: (i) statistically justified selection of influential variables for electricity-loss forecasting; (ii) an integral data-complexity criterion based on sample size and relative variability; (iii) comparative evaluation on 20 univariate and 20 multivariate datasets; (iv) empirical model-selection ranges for five forecasting approaches; and (v) a practical interpretation of θ as a preselection rather than an absolute decision rule.

2. Methods and materials

Study object and data preparation. The study object was the 0.4 kV distribution network of the Yunusabad district electricity network in Tashkent. The experimental database consisted of 912 operational observations covering 2021–2023.

The raw dataset contained electricity-loss values together with operational and meteorological variables. Before modeling, the data were synchronized, missing observations were imputed, and the series were aligned to obtain a comparable analytical dataset.

Because data cleaning can alter distributional properties, the loss series was checked before and after preprocessing. Figure 1 reproduces the real distributions used in the dissertation analysis. After correction, skewness and kurtosis were evaluated to verify that the cleaned sample remained suitable for subsequent statistical modeling and factor screening.

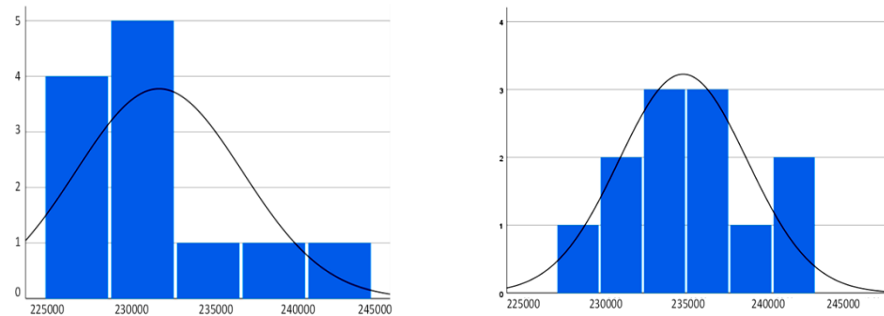


Fig. 1. Real distributions of electricity-loss data: (a) before preprocessing; (b) after preprocessing

Statistical selection of influential factors. Five candidate explanatory variables were initially considered: active power P (F1), reactive power Q (F2), ambient temperature T (F3), number of working days (F4), and relative humidity RH (F5). Pearson correlation was selected after comparing Pearson, Spearman, and Kendall measures because the dataset size was large and the cleaned distributions were suitable for parametric correlation analysis.

Statistical significance of the correlations was additionally verified by applying Fisher's transformation and a Z-test. For a correlation coefficient r , the Fisher-transformed value z and its standard error SE_z are written as:

$$z = \frac{1}{2} \ln \frac{1+r}{1-r}; \quad (1)$$

$$SE_z = \frac{1}{\sqrt{n-3}}; \quad (2)$$

$$Ze = \frac{z - z_0}{SE_z}. \quad (3)$$

At $\alpha=0.05$, a relationship was considered statistically significant when $|Ze| > 1.96$. The strongest significant relationships with electricity losses were obtained for active power, ambient temperature, and relative humidity. Their correlation coefficients and Z-statistics are summarized in Table 1.

Table 1. Statistically significant determinants of electricity losses

Factor	r	z	SE_z	Ze
Active power, P (F1)	0.829	0.73976	0.03317	22.30348
Ambient temperature, T (F3)	0.762	0.79640	0.03317	24.01010
Relative humidity, RH (F5)	-0.703	0.55331	0.03317	16.68222

All three factors substantially exceeded the critical Ze value, while reactive power and the number of working days exhibited weak relationships in the correlation matrix. Therefore, P , T , and RH were retained as the principal external variables in multivariate forecasting.

Forecasting models and validation design. Five model classes were compared: polynomial regression, ARIMA, GMDH, multilayer perceptron (MLP), and the hybrid DBSCAN-MLP model. These methods deliberately represent different levels of structural complexity, from interpretable regression to nonlinear learning with cluster-based preprocessing.

Polynomial regression was used as the low-complexity baseline and provides an explicit functional relationship between losses and the selected variables [5]. ARIMA was used to capture autoregressive dependence and temporal structure; the dissertation experiments identified ARIMA(1,0,1) as the rational configuration for the investigated loss series [6]. GMDH provides self-organizing polynomial structures and is particularly useful for small to medium datasets where automatic argument selection is desirable [1,2]. MLP captures nonlinear multivariate relationships, while the hybrid DBSCAN-MLP procedure first separates dense data structures and anomalous observations before neural-network learning [2,7].

For neural-network experiments, the input array was divided into training, validation, and test subsets in proportions of 67%, 10%, and 23%, respectively. This separation was retained in the comparative experiments so that the selected model was assessed not only by fit to training observations but by generalization to unseen data.

Forecasting accuracy was evaluated by the mean absolute percentage error (MAPE):

$$MAPE = \frac{100}{n} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right|; \quad (4)$$

where y_i and $\hat{y}_i$ denote actual and forecast values, respectively. MAPE was used consistently across all candidate methods and all representative datasets.

Integral data-complexity criterion. A central assumption of the proposed approach is that fore-

casting-model requirements depend jointly on the amount of available data and on the relative variability of the observed process. Data volume alone is insufficient: two datasets containing the same number of observations may have substantially different dispersion and structural complexity.

Relative variability was characterized by the coefficient of variation:

$$CM = \left(\frac{\sigma}{\bar{x}} \right) \times 100\%; \quad (5)$$

where σ is the standard deviation and $\bar{x}$ is the sample mean. Since the number of observations M may reach thousands or millions while CV remains on a much smaller numerical scale, M was logarithmically transformed. The two components were then combined using the Euclidean norm:

$$\theta = \sqrt{\lg(M)^2 + CV^2}. \quad (6)$$

The resulting θ is not interpreted as a universal entropy measure. It is an empirical integral index designed specifically for forecast-model preselection. To calibrate the relationship between θ and model performance, 20 univariate and 20 multivariate experimental datasets were generated from different time intervals, variability levels, and combinations of influential variables.

3. Results and discussion

The first comparison was performed on the complete experimental framework. The minimum MAPE values obtained for the candidate models are given in Table 2 and Fig. 2.

Table 2. Overall forecasting error of the investigated models

Model	MAPE, %
Regression	5.01
ARIMA	3.15
GMDH	5.05
MLP	2.58
DBSCAN-MLP	2.10

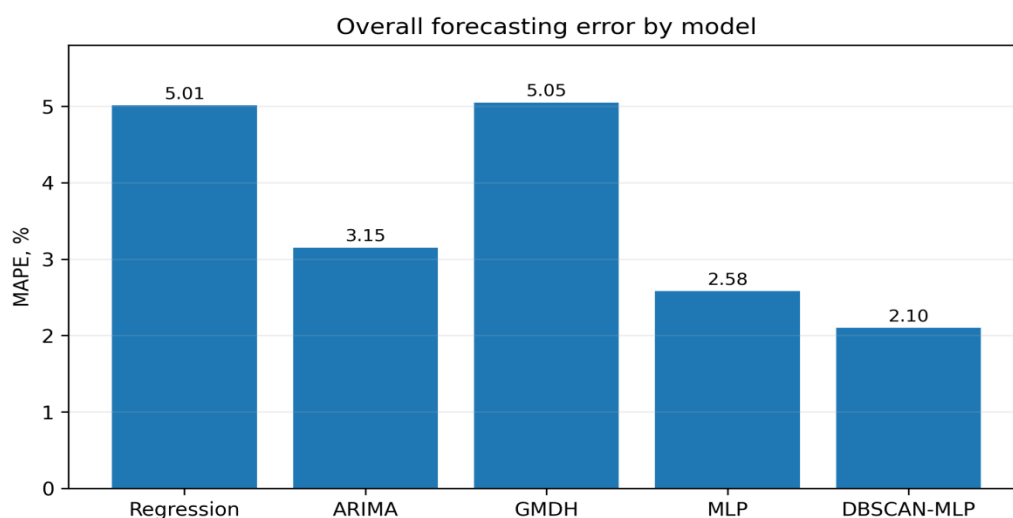


Fig. 2. Overall MAPE obtained by Regression, ARIMA, GMDH, MLP, and DBSCAN-MLP

The hybrid DBSCAN-MLP model achieved the lowest overall error (2.10%), followed by MLP (2.58%) and ARIMA (3.15%). Regression and GMDH yielded 5.01% and 5.05%, respectively. These values confirm the potential of nonlinear and hybrid approaches, but they do not justify applying the most complex architecture to every dataset. The representative experiments below demonstrate why a data-dependent selection rule is required.

Four representative datasets were selected to illustrate the transition from lower to higher data complexity. Their sample sizes, coefficients of variation, and integral θ values are shown in Table 3.

Table 3. Data-complexity characteristics of representative datasets

Dataset	Structure	M	CV, %	θ
DATA1	Univariate	60	2.845	3.355
DATA13	Univariate	350	3.850	4.610
DATA5	Multivariate	83	2.816	3.407
DATA9	Multivariate	385	3.045	3.994

DATA1 and DATA13 represent the univariate case at $\theta=3.355$ and $\theta=4.610$. DATA5 and DATA9 represent multivariate structures at $\theta=3.407$ and $\theta=3.994$. The values show that increasing data volume together with variability moves the problem into a different complexity region even before a forecasting model is trained.

Figure 3 shows the actual and forecast trajectories for DATA1. Although all methods broadly track the time-series pattern, the magnitude of deviations differs by model and by training/validation/test stage. The corresponding MAPE values are given in Table 4.

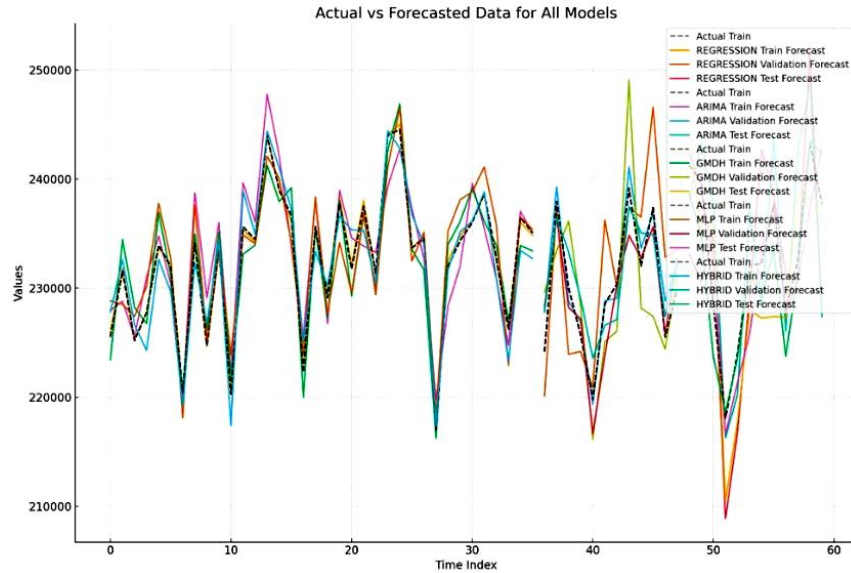


Fig. 3. Real actual-vs-forecast trajectories for univariate DATA1 across the investigated models

Table 4. MAPE (%) for univariate DATA1

Subset	Regression	ARIMA	GMDH	MLP	Hybrid
Train	0.22	1.02	1.14	2.25	2.11
Validation	3.85	1.35	3.21	2.99	2.65
Test	4.27	3.98	4.33	3.87	2.42

At $\theta=3.355$, the simple models remain competitive and the validation differences are moderate, although the hybrid model gives the lowest test error. This is precisely the type of low-complexity region where an automatic system should avoid concluding that a complex model is mandatory solely because it yields the smallest isolated test error.

The picture changes for DATA13, where θ rises to 4.610. Figure 4 shows markedly larger dispersion among the statistical forecasts on the validation and test portions, while the neural-network and hybrid trajectories remain closer to the actual sequence.

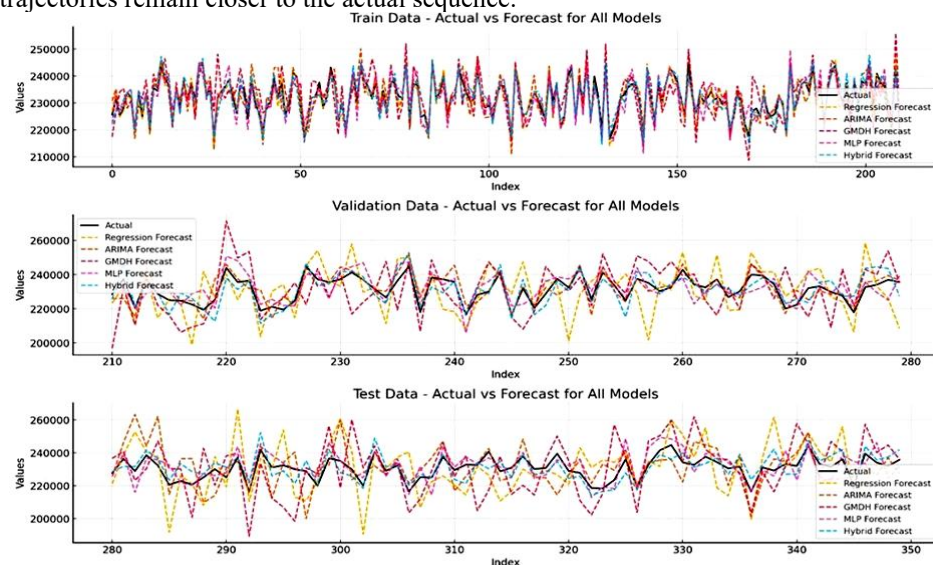


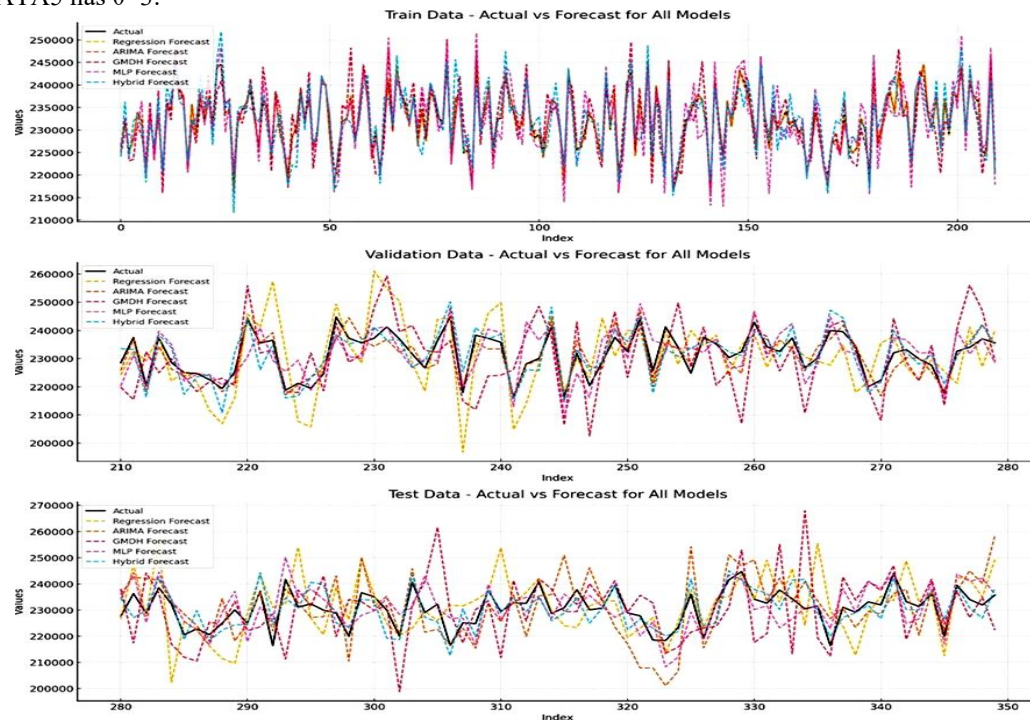
Fig. 4. Real train, validation and test forecasts for the higher-complexity univariate DATA13 dataset

Table 5. MAPE (%) for univariate DATA13

Subset	Regression	ARIMA	GMDH	MLP	Hybrid
Train	0.74	1.51	2.31	1.29	1.06
Validation	4.83	2.13	5.40	2.19	2.15
Test	5.71	5.39	6.10	2.51	2.45

For DATA13, regression and GMDH test MAPE increases to 5.71% and 6.10%, whereas MLP and the hybrid model maintain errors of 2.51% and 2.45%. This provides direct experimental evidence that a model class suitable for a lower- θ univariate series may become inadequate as data volume and variability increase.

The multivariate case incorporates the significant factors P, T, and RH and therefore contains interaction patterns that are absent from a single loss series. Figure 5 presents the real DATA5 forecasts; DATA5 has $\theta=3$.

**Fig. 5.** Real train, validation and test forecasts for multivariate DATA5**Table 6.** MAPE (%) for multivariate DATA5

Subset	Regression	ARIMA	GMDH	MLP	Hybrid
Train	0.027	0.25	1.35	1.55	1.46
Validation	4.11	1.53	4.27	2.27	1.86
Test	4.39	4.25	4.91	2.86	2.19

At $\theta=3.407$, DATA5 lies close to the transition between the regression and GMDH regions. The relatively small training error of regression does not translate into equally strong generalization, demonstrating why the selection procedure must incorporate validation/test behavior rather than training error alone.

DATA9 increases the multivariate complexity index to $\theta=3.994$. The real trajectories in Fig. 6 show that the neural-network and hybrid forecasts follow the actual signal more consistently than the statistical alternatives, especially on the validation and test intervals.

Table 7. MAPE (%) for multivariate DATA9

Subset	Regression	ARIMA	GMDH	MLP	Hybrid
Train	0.41	1.28	1.35	1.79	1.31
Validation	4.74	1.89	4.94	2.67	2.27
Test	5.13	4.67	5.36	2.69	2.33

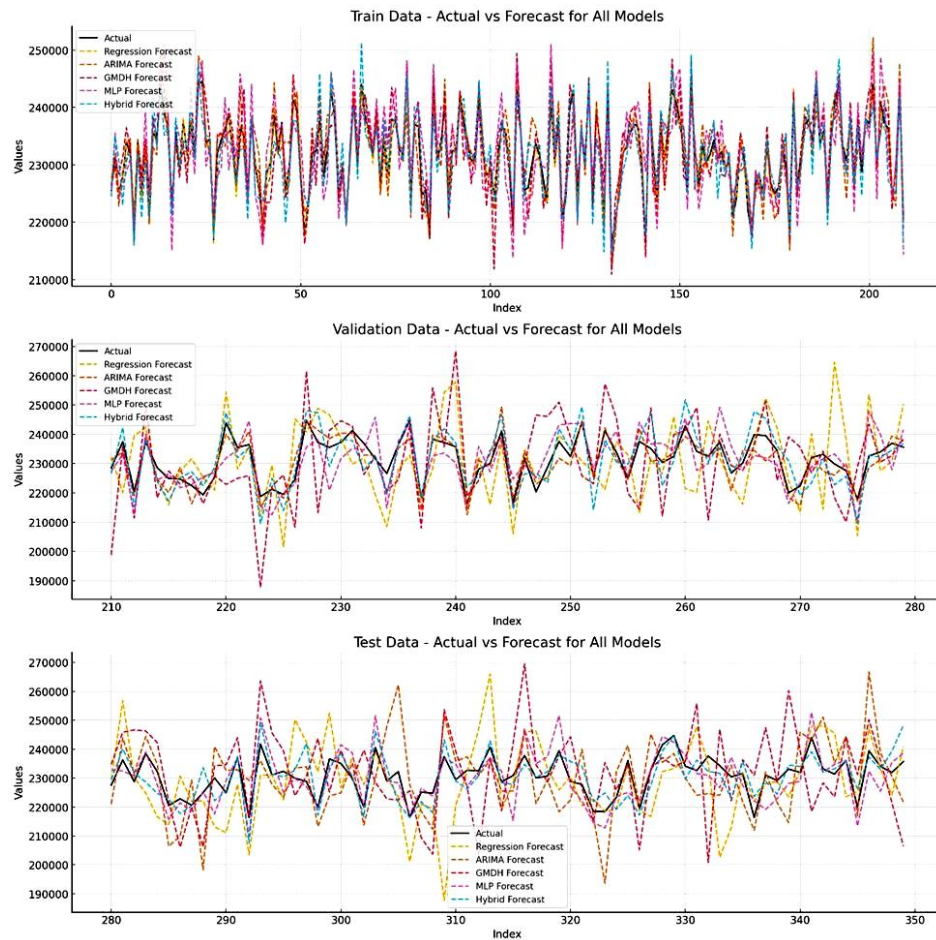


Fig. 6. Real train, validation and test forecasts for multivariate DATA9

In DATA9, the hybrid model reduces test MAPE to 2.33%, compared with 5.13% for regression and 5.36% for GMDH. MLP also remains accurate at 2.69%. The fact that the hybrid transition occurs at a lower θ in the multivariate case indicates that interactions among explanatory variables contribute additional effective complexity that is not fully represented by sample size alone.

Aggregating the 40 experiments produced the model-selection map shown in Table 8 and Fig. 7. The ranges must be interpreted as empirical preselection intervals derived from the investigated 0.4 kV data rather than universal theoretical constants.

Table 8. Empirical model-selection ranges as a function of θ

Model	Univariate data	Multivariate data
Polynomial regression	$0 \leq \theta < 3.6$	$0 \leq \theta < 3.4$
GMDH	$0 \leq \theta < 3.6$	$3.4 \leq \theta < 3.6$
ARIMA	$3.6 \leq \theta < 3.8$	$3.6 \leq \theta < 3.8$
MLP	$3.8 \leq \theta < 4.2$	$3.8 \leq \theta < 3.9$
DBSCAN-MLP	$\theta \geq 4.2$	$\theta \geq 3.9$

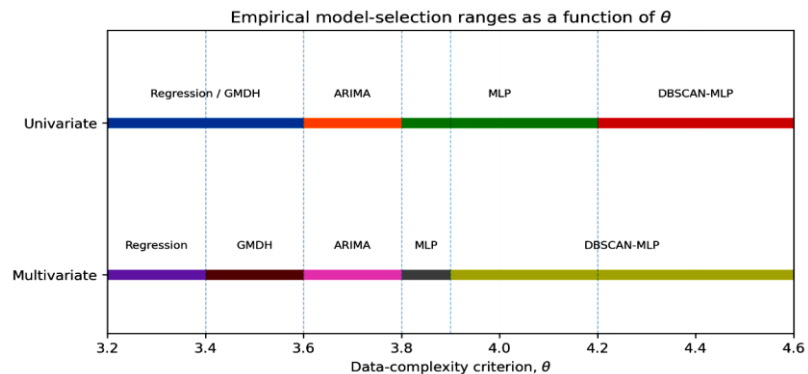


Fig. 7. Empirical transition from statistical to neural-network and hybrid models as data complexity increases

For univariate data, polynomial regression and GMDH overlap in the low-complexity region $0 \leq \theta < 3.6$. Therefore, θ does not force a unique decision there; instead it narrows the candidate set and validation error determines the final choice. ARIMA is favored in the $3.6 \leq \theta < 3.8$ interval, MLP in $3.8 \leq \theta < 4.2$, and DBSCAN-MLP when $\theta \geq 4.2$. For multivariate data, the ranges are more differentiated: regression below 3.4, GMDH from 3.4 to 3.6, ARIMA from 3.6 to 3.8, MLP from 3.8 to 3.9, and DBSCAN-MLP at $\theta \geq 3.9$.

4. Conclusion

The results support the general premise of modern forecast-model selection research: there is no single uniformly optimal forecasting model for all time series. Spiliotis et al. [10] showed that selection criteria themselves may disagree, while Ganzevoort and van Vuuren [11] demonstrated the value of grouping series before selecting a forecasting method. The present study translates this general principle into a compact rule tailored to electricity-loss forecasting in low-voltage distribution networks.

The proposed criterion differs from a conventional model-comparison table in two important ways. First, it moves model selection before full model training by using only M and CV . Second, it expresses the transition from simple to complex models continuously through θ . In this sense, the approach is consistent with complexity-aided forecasting [12] and dynamic model selection [13], but it is operationalized using parameters that can be calculated directly from an incoming AMI dataset.

The representative datasets explain the physical-statistical logic of the ranges. At lower θ , the information structure is sufficiently regular that regression or GMDH can provide an acceptable description with high interpretability and low computational cost. As θ approaches the intermediate region, temporal organization becomes more important and ARIMA is effective. At larger θ values, nonlinear dependence and local heterogeneity become increasingly significant, making MLP and DBSCAN-MLP more appropriate.

The multivariate threshold for the hybrid model ($\theta \geq 3.9$) is lower than the univariate threshold ($\theta \geq 4.2$). This is an important finding. Even when the numerical value of CV is moderate, the simultaneous presence of P , T , and RH introduces nonlinear interactions and local structures. Therefore, effective model complexity increases faster in the multivariate case. The real DATA9 trajectories provide direct evidence: regression and GMDH exhibit substantially larger validation/test deviations than MLP and the hybrid model.

DBSCAN contributes to the hybrid approach by separating dense structures and reducing the influence of atypical observations before MLP training. This preprocessing improves the stability of the learned nonlinear relationship and contributes to the overall MAPE of 2.10%. The result is consistent with the author's previous ANN optimization work [7] and with contemporary data-driven loss-management studies [14].

An important methodological limitation is that θ is empirical. The boundaries were calibrated on operational data from one 0.4 kV distribution area, and their transferability to other regions, climates, customer compositions, sampling frequencies, and voltage levels must be independently tested. Consequently, the criterion should be treated as a rational preselection mechanism, not as a universal law. A practical implementation should calculate θ , restrict the candidate model set, and then make the final decision using validation/test performance.

The approach is well suited to an automated information-and-analytical system in which many feeders are processed repeatedly. Instead of training every candidate model for every incoming dataset, the system can estimate M and CV , calculate θ , identify a small group of plausible algorithms, and then perform detailed validation only for that subset. This reduces unnecessary computational work while retaining a data-driven basis for model choice.

1. A data-complexity-based forecasting-model selection methodology was developed for electricity losses in 0.4 kV distribution networks using 912 real operational observations from 2021-2023.

2. Active power, ambient temperature, and relative humidity were identified as statistically significant determinants of losses, with correlation coefficients 0.829, 0.762, and -0.703 and Z_e values far above the 1.96 critical threshold.

3. The integral criterion $\theta = \sqrt{\lg(M)^2 + CV^2}$ was introduced to combine sample size and relative variability into a single preselection variable.

4. Comparative experiments on 20 univariate and 20 multivariate datasets showed that forecasting requirements shift from statistical to neural-network and hybrid models as θ increases. Overall MAPE values were 5.01% for Regression, 3.15% for ARIMA, 5.05% for GMDH, 2.58% for MLP, and 2.10% for DBSCAN-MLP.

5. For univariate datasets, ARIMA was associated with $3.6 \leq \theta < 3.8$, MLP with $3.8 \leq \theta < 4.2$, and DBSCAN-MLP with $\theta \geq 4.2$. For multivariate datasets, GMDH was associated with $3.4 \leq \theta < 3.6$, ARIMA with $3.6 \leq \theta < 3.8$, MLP with $3.8 \leq \theta < 3.9$, and DBSCAN-MLP with $\theta \geq 3.9$.

6. The lower hybrid-model threshold in the multivariate case indicates that interactions among operational and meteorological variables increase effective data complexity. The proposed criterion can therefore be used as a model-preselection module in automated forecasting and decision-support systems, provided that final selection is confirmed by validation and test results.

REFERENCES

1. Nazirova, H.Z., Badalov, A.A., Jo'rayev, R.R. Comparative analysis of the calculation of electricity losses in 10-6-0.4 kV distribution networks. AIP Conference Proceedings, 2024. DOI: 10.1063/5.0218854.
2. Allaev, K.R., Nazirova, X.Z. Comparative assessment of electricity losses in electrical networks using ANN and GMDH methods. Problems of Energy and Sources Saving, 2024, No. 3, pp. 1-10. DOI: 10.5281/zenodo.14032515.
3. Allaev, K.R., Nazirova, X.Z. Comparison of various methods for predicting electricity losses in networks. Problems of Energy and Sources Saving, 2024, No. 3, pp. 41-50.
4. Nazirova, X.Z. Methods for determining static characteristics in industrial electrical networks. E3S Web of Conferences, 2023, Vol. 384, 01046. DOI: 10.1051/e3sconf/202338401046.
5. Allaev, K.R., Nazirova, H.Z., Badalov, A.A. Forecasting of electricity losses using the polynomial regression method. AIP Conference Proceedings, 2025, Vol. 3331. DOI: 10.1063/5.0305943.
6. Allayev, K.R., Nazirova, H.Z., G'oibova, S. Forecasting electricity losses in 0.4 kV distribution networks using the autoregressive integrated moving average method. AIP Conference Proceedings, 2025, Vol. 3331. DOI: 10.1063/5.0305940.
7. Nazirova, H., Nazirova, O., Toshxo'jayeva, M., et al. Optimization of electricity loss forecasting using ANN. AIP Conference Proceedings, 2025, Vol. 3331. DOI: 10.1063/5.0305944.
8. Rasulov, A., Nazirova, K., Badalov, A., Saidova, M. Study on graphs of electrical load and analysis of power consumption modes. E3S Web of Conferences, 2024, Vol. 497, 01005. DOI: 10.1051/e3sconf/202449701005.
9. Matijasevic, T., Antic, T., Capuder, T. Machine learning-based forecast of secondary distribution network losses calculated from the smart meters data. 2022 7th International Conference on Smart and Sustainable Technologies (SpliTech), 2022. DOI: 10.23919/SpliTech55088.2022.9854276.
10. Spiliotis, E., Petropoulos, F., Assimakopoulos, V. On the disagreement of forecasting model selection criteria. Forecasting, 2023, 5(2), 487-498. DOI: 10.3390/forecast5020027.
11. Ganzevoort, R.C., van Vuuren, J.H. A two-phased cluster-based approach towards ranked forecast-model selection. Machine Learning with Applications, 2023, 13, 100482. DOI: 10.1016/j.mlwa.2023.100482.
12. Song, M., Wang, R. Complexity-aided time series modeling and forecasting under a decomposition-aggregation framework. Information Sciences, 2024, 664, 120352. DOI: 10.1016/j.ins.2024.120352.
13. Hananya, R., Katz, G. Dynamic selection of machine learning models for time-series data. Information Sciences, 2024, 665, 120360. DOI: 10.1016/j.ins.2024.120360.
14. Blinov, I., Radziukynas, V., Shymaniuk, P., et al. Smart management of energy losses in distribution networks using deep neural networks. Energies, 2025, 18(12), 3156. DOI: 10.3390/en18123156.
15. Ding, C., Zhou, Y., Ding, Q., Wang, Z. Loss prediction of ultrahigh voltage transmission lines based on EEMD-LSTM-SVR algorithm. Frontiers in Energy Research, 2022, 10, 811745. DOI: 10.3389/fenrg.2022.811745.
16. Rakhmonov, I.U., Kurbonov, N.N. Forecasting power consumption of industrial enterprises using an autoregressive integrated moving average model. Bulletin of MPEI, 2021, No. 6, pp. 11-19.
17. Benidis, K., Rangapuram, S.S., Flunkert, V., et al. Deep learning for time series forecasting: tutorial and literature survey. ACM Computing Surveys, 2022. DOI: 10.1145/3533382.
18. Makridakis, S., Spiliotis, E., Assimakopoulos, V. The M4 Competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting, 2020, 36(1), 54-74. DOI: 10.1016/j.ijforecast.2019.04.014.