

Forecasting electricity demand in Ukraine using machine learning methods

Irina Kalinina[†] and Aleksandr Gozhyj^{*,†}

Petro Mohyla Black Sea National University, 68 Desantnykiv 10, 54000, Mykolaiv, Ukraine

Abstract

The article studies the solution of the problem of forecasting electricity demand in Ukraine. The sequence of data processing stages in solving the forecasting problem using machine learning methods is presented. It consists of the following stages: data collection, data research and preparation, construction and training of forecasting models, selection of the best model and calculation of forecasts, evaluation and verification of quality indicators of forecasts. A general methodology for solving forecasting problems is proposed. The methodology for solving the forecasting problem on time series is considered. The forecasting process consists of five stages. The first stage includes the collection, analysis and interpretation of data. The next stage includes the procedures of data research and preparation. The third stage - the modeling stage consists of three parts: preparation of a data set for modeling, selection and training of models and evaluation of their quality. The fourth stage is the forecasting stage and calculation of quality indicators of forecasts. At the fifth stage, procedures for improving the efficiency of the selected forecasting model are performed. The following models were used at the modeling stage: ARIMA, GAM, ANN and BSTS. The analysis of the models was carried out and forecasts were built based on each model. For the constructed models with the best quality indicators, the predictive values were calculated. The forecasts were compared with the data of the validation sample. The following indicators were used to select the optimal model: MAPE, MAE, MSE, RMSE. The BSTS model showed the best results.

Keywords

Forecasting, Electricity demand in Ukraine, Machine learning methods, Data processing, Methodology of forecasting, ARIMA, GAM, ANN, BSTS.

1. Introduction

Today, energy consumption around the world is growing rapidly due to the increasing demand for electricity due to the ever-growing global population, large-scale development of industry and technology, rising living standards, large-scale industrialization in developing countries, and the need to maintain high rates of sustainable development. The electric power sector is the basis of the economic potential of any country. This industry belongs to the critical infrastructure industries, which must respond very quickly to external changes, such as natural disasters, military actions, as well as to changing conditions in the electricity market.

Forecasting the demand for electricity in Ukraine today is a strategically important issue, since in the conditions of war and constant attacks on energy infrastructure, it is necessary to promptly distribute energy resources to meet the needs of various types of consumers.

Considering these factors, the construction of adequate and accurate electricity demand forecasting models is necessary and important for accurate planning of investments in electricity generation and distribution. A common difficulty in developing quality forecasts is determining a sufficient amount of information to build forecast models. If there is not enough data, then the forecast will be inaccurate. Similarly, if information is imprecise or redundant, pre-processing the

CIAW-24: Computational Intelligence Application Workshop, October 10-12, 2024, Lviv, Ukraine

* Corresponding author.

[†] These authors contributed equally.

✉ irina.kalinina1612@gmail.com (I. Kalinina); alex.gozhyj@gmail.com (A. Gozhyj)

ORCID 0000-0001-8359-2045 (I. Kalinina); 0000-0002-3517-580X (A. Gozhyj)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

data and building models for forecasting will be difficult. Therefore, there is a need to increase the accuracy of predictive models due to the use of modern effective methods and approaches.

Machine learning methods significantly increase the efficiency of solving machine learning problems such as classification, regression and prediction. But the application of each method requires taking into account the features of the data set, the method of data presentation, and the features of the problem being solved. Therefore, one approach to solving forecasting problems is to systematically use several forecasting models and then select the model that gives the best results.

Let's consider some of them. The paper [1] presents an algorithm for forecasting electricity demand, which is associated with the maximum load in the power grid. The authors substantiated and applied the SVR method. The data set and parameters of the method were configured and optimized by a hybrid method. This approach allows us to reduce the overall forecast error. The hybrid method is based on a combination of a neural network model, the ARIMA method, and a modified SVR method.

In [2], an artificial bee colony algorithm was used to process the initial data. To forecast the demand for electricity, the algorithm was used in combination with ensemble models. A number of independent input variables were used to create homogeneous ensemble models. The ensemble model proposed by the authors provides more accurate predictions.

The authors of [3] propose to combine the feedforward neural network model with the convolutional neural network architecture to forecast the demand for electricity. This approach turned out to be the most effective. Deep learning methods were also applied in [4]. The paper forecasted the energy demand based on statistical data for previous years. For a more in-depth study of the data, the cluster analysis method was used. The load was classified by certain periods and presented as clusters. The forecast of the demand for electricity was assessed using neural network models and SVM.

The article [5] presents an approach to forecasting electricity demand based on a hybrid model. The model is built using a combination of ARIMA and LSSVM. The forecasting results show that this approach to building a model allows for abnormal values in the data to be taken into account. In the work [6], the authors used a regression analysis model to forecast electricity demand in various industries.

To estimate the peak monthly demand for electricity, the following methods were used in [7]: ANFIS method, a special data processing method and various neural network models. In combination with the proposed models, these methods were found to be better suited for determining the peak demand for electricity. In [8], he published long-term forecasts of electricity demand in Greece, and also used the relationship between time series and effective several criteria. The cost estimation model is investigated using data collected between 1999 and 2013. The impact of electricity production in European countries during the quarantine is studied in the works [9,10].

In works [11,12] the effectiveness of various approaches and strategies for predicting daily energy consumption was investigated. The authors of [13] investigated the approach of forecasting the load in the electrical network using artificial intelligence methods. The authors of [14] developed a model for forecasting electricity demand for residential and commercial buildings based on ensemble methods. Short-term forecasting was considered.

In [15], the authors use SVR models in combination with WOA, which includes learning based on elite and chaotic opposition (ECWOA) to improve forecasting results. The results of experiments show that taking into account information about electricity prices leads to higher forecasting accuracy. In works [16,17] approaches based on machine learning algorithms are considered to increase the accuracy of short-term forecasts. The following methods were used: SVM, LSTM, SVR and ensemble structures.

In [18], seasonally adjusted regression was used to obtain forecast values for electricity. In [19], the authors demonstrated the advantage of LSTM over SVM in the task of forecasting electricity demand using specific examples. In [20–22], various neural network architectures were investigated in combination with heuristic algorithms for forecasting electricity demand in different countries.

The authors of [23] used regularized Lasso Lars and RF models to forecast electricity consumption in Brazil. An energy forecasting model based on a deep learning approach [24] combines CNN, LSTM, and an autoencoder (AE) for time series with different lengths. In the work [25], modeling based on neural network models was used to estimate and forecast the demand and consumption of electricity in transport. Thus, it was proven that a combination of different methods and algorithms of machine learning allows to effectively solve the problem of forecasting the demand for electricity.

Problem statement: To study different approaches to forecasting the demand for electricity in Ukraine. To develop a methodology for solving forecasting problems based on machine learning methods. To compare the effectiveness of different machine learning methods in solving the problem of forecasting the demand for electricity.

2. Methodology for solving forecasting problems

The sequence of data processing stages when solving a forecasting problem is shown in Fig. 1. It consists of the following stages: data collection, research and preparation of data, construction and training of forecast models, selection of the best models and calculation of forecasts, evaluation and verification of the quality of forecasts [26-28].

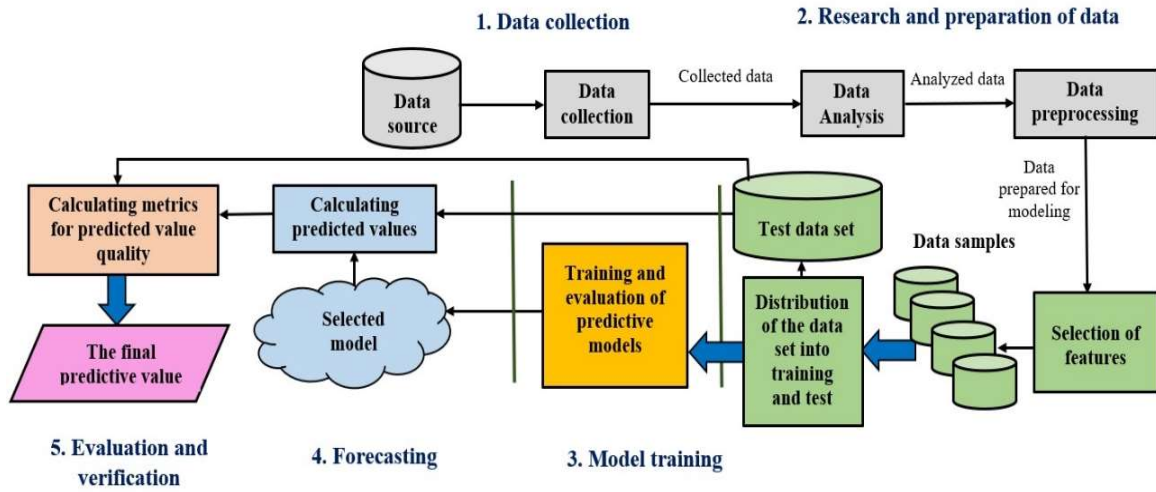


Figure 1: Sequence of stages for solving machine learning problems.

An information model for solving the problem of forecasting using machine learning methods is described using the following set of elements:

$$ML_T = \{NData_T, NL_{Data}, NS_{Data}, MA_T, Mod_T, MP_T, Us_T\}$$

where $NData_T$ is a set of data sets that are processed when solving a machine learning problem; NL_{Data} – a set of data nonlinearities that are taken into account when solving a machine learning problem; NS_{Data} – a set of non-stationary processes that are taken into account when solving a machine learning problem; MA_T – set of methods of analysis and preliminary processing of data sets; Mod_T – set of model building methods and modeling methods for solving the problem of machine learning; MP_T – a set of forecasting methods based on probabilistic statistical analysis, taking into account nonlinearities and non-stationarity of data; Us_T is a set of uncertainties when solving a machine learning problem.

The dataset set $NData_T = Data_{TR} \cup Data_{TST}$ combines two subsets. The first subset $Data_{TR}$ combines training data, the second subset $Data_{TST}$ is a set of test data. The set of uncertainties Us_T includes uncertainties of statistical type, uncertainty due to lack of observations, uncertainty of model parameters, uncertainty of model structure, uncertainty of amplitude and probability type.

Based on the developed information model, a description of the stages of solving forecasting problems was created, and a methodology for solving the problem of predictive modeling on time series was developed [29,30], which is presented in Figure 2.

The methodology is presented as a sequence of the following stages. The *first stage* is necessary for collecting, analyzing and interpreting the initial data. When loading, the data set is analyzed, its structure and features of individual attributes are determined. As a result of preliminary data processing, the set is prepared for subsequent analysis - intelligence analysis of data.

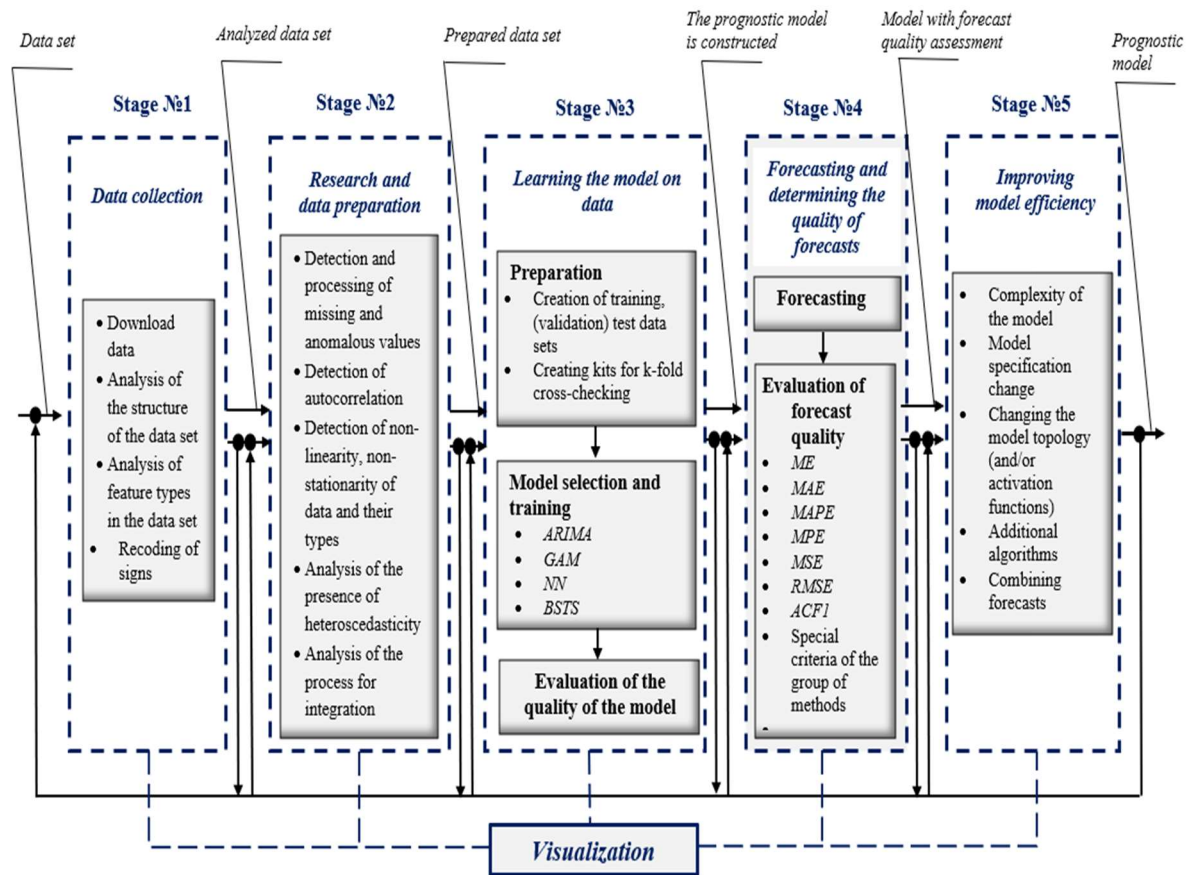


Figure 2: Methodology for solving the problem of predictive modeling on time series

The *second stage* involves research procedures and data preparation. Descriptive statistics for each variable are analyzed, missing and abnormal values are identified, the level of autocorrelation is determined, nonlinearity and nonstationarity of data and their types are identified, heteroscedasticity is analyzed, and the process is analyzed for integrability.

The *third stage* - the modeling stage consists of three parts: preparation of the data set for modeling, selection and training of models and assessment of their quality. Before starting the modeling, the prepared dataset is split into training and test samples, and cross-validation sets are created. When choosing a model, simulation algorithms are tested on a training sample and the best one is selected according to certain quality criteria.

The *fourth stage* is the stage of constructing forecasts and assessing their quality. At the *fifth stage*, procedures are performed to improve the efficiency of the selected forecasting model. For different modeling methods, the following methods can be used to improve quality: complicating the model structure, changing its specifications, changing the model topology (and/or activation functions), using additional algorithms, combining forecast values.

An important feature of the presented method is visualization. With the help of visualization, at each stage, it is possible to adjust the sequence of actions and return to previous stages. The stages of the method have features that reflect the subject area of the problem solution.

3. Forecasting Electricity Demand

3.1. Dataset description and pre-processing

A data set for forecasting electricity demand in Ukraine is presented on the web resource of the state operator of the electricity market [31]. The observation time interval covers the period from 01.07.2019 to 04.10.2024. The set contains hourly data on the volumes of electricity, sales of electricity, demand for it in MWh in the power grids of Ukraine and prices (Fig. 3).

```
'data.frame': 69385 obs. of 10 variables:
 $ date      : chr  "2019-07-01" "2019-07-01" "2019-07-01" "2019-07-01"
 $ hour      : int   1  2  3  4  5  6  7  8  9 10 ...
 $ energy_system : chr  "Burshtyn peninsula" "Burshtyn peninsula" "Burshtyn peninsula" "Burshtyn peninsula"
 $ price      : num   939 939 939 939 939 939 939 939 939 939 ...
 $ amount_sales : num   313 241 259 153 153 ...
 $ amount_purchase : num   313 241 259 153 153 ...
 $ amount_sales_nominated : num   313 241 259 153 153 ...
 $ amount_purchase_nominated : num   364 334 320 320 317 ...
 $ demand      : num  -50.7 -93.3 -60.8 -167.9 -164.5 ...
 $ price_cap    : num   959 959 959 959 959 ...
```

Figure 3: Structure of the data set "electricity".

The variable with *hour* marks in a day is included in the *datetime* indexing variable, and the *demand* variable is selected from the set as the resulting variable. It is noted that the *energy_system* variable takes only three values: *Burshtyn peninsula*, *IPS of Ukraine*, *IPS of Ukraine (synchronized with ENTSO-E systems)*. The values of this variable up to 24 February 2022 were divided into two separate subsystems: the "United Energy System of Ukraine" and the "Burshtyn Energy Island". Since 24 February 2022, the Ukrainian energy system has been synchronized with the European ENTSO-E energy system. Thus, up to 24 February 2024, the dataset contains demand data separately for the two subsystems, and to calculate the all-Ukrainian indicators, the values for these subsystems are summed up (Fig. 4).

```
> # Aggregation of the demand of individual power grid subsystems into one time series
> energy_demand <- data.frame(datetime = unique(data$datetime), demand=0)
> for(i in 1:length(energy_demand$datetime))
+ {
+   energy_demand$demand[i] = sum(data$demand[which(data$datetime==energy_demand$datetime[i])])
+ }
```

Figure 4: A fragment of the code for aggregating the demand of individual power grid subsystems into one time series.

After checking, 5 missing values were found in the data set. The missing values have a yearly interval and are recorded in late March or early April. The presence of gaps is due to the transition of clocks to daylight saving time. Given that the total volume of observations is 69,385, the LOCF method [32] was used to fill in the missing values.

After the time series analysis, the sampling frequency was reduced. Hourly observations of electricity demand in the set were aggregated into daily averages (Figure 5).

```
> # Aggregation by date
> energy_demand_daily <- energy_demand %>%
+   group_by_key() %>%
+   index_by(date = ~ as_date(.)) %>%
+   summarize(avg_demand = mean(demand))
```

Figure 5: Code snippet to down sample a dataset.

To build forecast models, we used a portion of the data from 01.06.2022 to 04.10.2024, excluding the peak sections of the series around 24.02.2024 (Fig. 6). The figure shows that demand can take negative values - this corresponds to a situation where electricity sales volumes exceed purchase volumes.

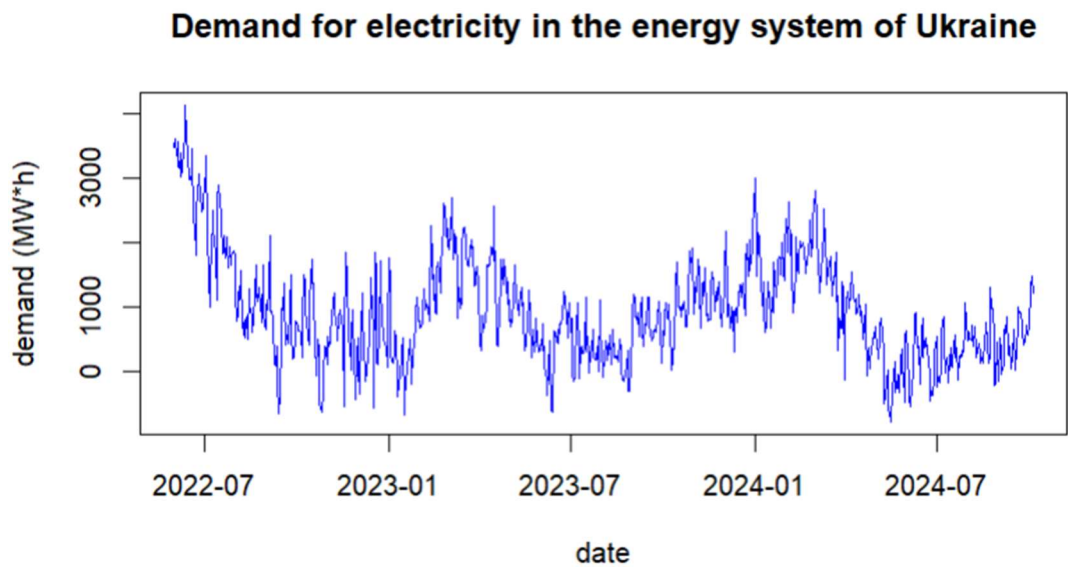


Figure 6: Graph of a fragment of the time series energy_demand.

An important condition for constructing reliable forecast models is understanding the structure of the time series. Decomposition of the series using the STL [33,34] method allowed us to determine the basic principles of modeling (Fig. 7). The figure demonstrates the dominant influence of seasonal components (annual and weekly seasonality), as well as the presence of a nonlinear trend component. The hypothetical non-stationarity of the time series is confirmed by the Ljung-Box tests for independence, the extended Dickey-Fuller test, the KPSS test for the level of stationarity, and the Phillips-Perron test for a unit root.

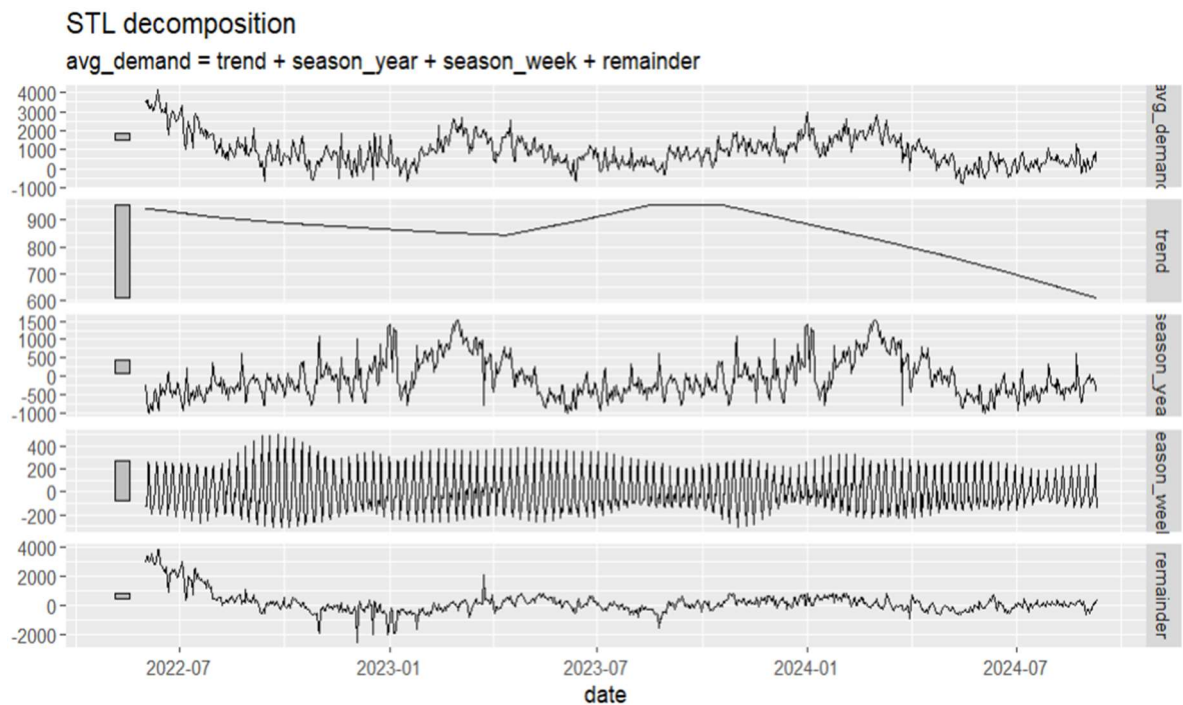


Figure 7: Time series decomposition.

The time series was tested for nonlinearity using a set of statistical tests. The test results confirmed the visual nonlinearity of the series, since the p-value is less than 0.05. The heteroscedasticity of the series was confirmed by the McLeod-Lee test (p-value is less than 0.05). The number of necessary differentiations and seasonal differentiations to obtain a stationary time series is determined as a result of the tests. The first differences are recommended, and the seasonal differentiation is optional.

To test the time series for autocorrelation, the Durbin-Watson and Breusch-Gottfrey tests were used. For both tests, the obtained p-values are much less than 0.05, thus confirming the autocorrelation of the time series. The graphs of the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) in Figure 8 correspond to expectations: the autocorrelation function decreases monotonically with increasing time shift. The figure confirms the presence of significant correlation and weekly seasonality.

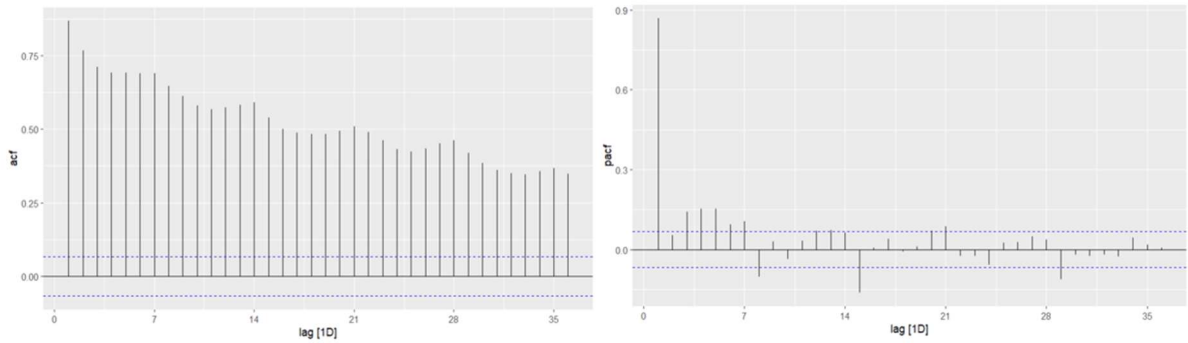


Figure 8: ACF and PACF graphs.

The results of the preliminary analysis confirm that the process under study belongs to the class of nonlinear and nonstationary.

3.2. Construction and evaluation of forecast models

The modeling stage begins with dividing the data set into two parts: training and test samples. The last 14 observations (two-week range) are kept as test observations, which corresponds to a forecast horizon of 14 days for short-term forecasting (Figure 9).

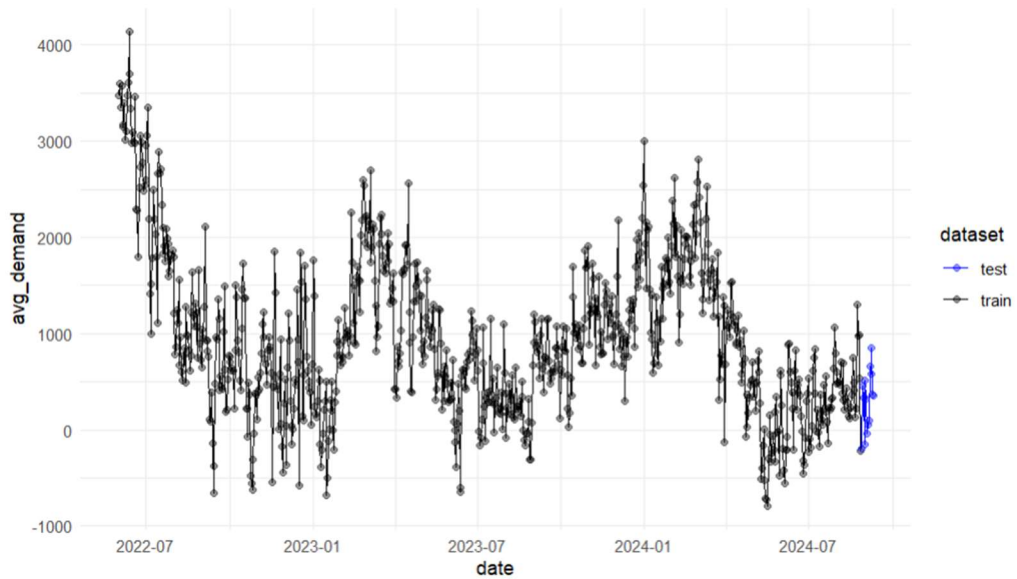


Figure 9: Training (black) and test (blue) data on the dynamics of average daily demand for electricity.

At the modeling stage, the following statistical models were used as basic forecast models: ARIMA, the method of fitting additive regression models (GAM), artificial neural networks of direct propagation (NNAR), and Bayesian structural models of time series (BSTS). The choice of these models is due to their ability to recognize complex patterns in real time series, taking into account the nonlinearity and non-stationarity of the process under study [35].

Modeling based on the ARIMA method. ARIMA models are the result of combining three components: autoregressive (AR), integration (I) and moving average (MA) [33,35,36]. The Box-Jenkins algorithm [37] helps in choosing the best model based on the graphs of the autocorrelation function and partial autocorrelation. However, identifying the best model is a problem with an ambiguous approach to the solution, since one data series can be represented by different parameters of the ARIMA model. However, compared to others, this methodology is easy to use and has good forecast accuracy. Alternative ARIMA models were selected automatically and using manual selection. Automatic selection is based on the methods of complete enumeration, quick enumeration, enumeration with smoothing of the input data set [35]. Table 1 provides a comparison of ARIMA models by quality metrics for the studied time series

Table 1

Comparison of ARIMA models by quality metrics for the *energy_demand* time series

Model type ARIMA (p, d, q) (P, D, Q)	AIC	AICc	BIC	R ²	DW
ARIMA (1, 1, 2)(2, 0, 0)[7]	1254,65	1254,66	1257,50	0,801	2,003
ARIMA (0, 1, 4)(0, 0, 2)[7]	1256,46	1256,47	1259,78	0,797	2,000
ARIMA (0, 1, 2)(0, 0, 2)[7]	1615,17	1615,27	1643,68	0,691	1,783

Based on the results obtained after automatic selection of the ARIMA model parameters, it can be concluded that additional smoothing (the third model) did not give the expected results, and the full and quick enumeration of the model parameters showed different results, but they have a small difference in the information criteria, so both models are suitable for forecasting. The manual method of selecting parameters did not allow obtaining the best model according to the quality criteria.

To select the qualitatively best ARIMA model from alternative options, the assessment was carried out not only according to the Akaike and Bayes information criteria, but also according to the values of the determination coefficient and the Durbin-Watson criterion. For the first two models, the coefficient of determination is 0.8, and the multiple correlation coefficient of these models exceeds 90% in absolute value. The DW = 2 criterion indicates the absence of autocorrelation in the residuals.

Because transformations were used in constructing the models, the residuals are visualized on the transformed scale. These so-called "innovation residuals" are useful for checking whether the model has adequately captured the information in the data (Figure 10). The figure shows the residuals of the best ARIMA model from the alternatives presented in the table. The result of the portmanteau test for this model is 0.066, which means that the residuals for the model are independent.

GAM modeling. GAM modeling. GAM models are created based on the procedure of fitting additive regression models [35,38,39]. To estimate the parameters of GAM models, the Bayesian approach is used when finding the posteriori maximum, or Bayesian inference is used. The Stan library was used for calculations. Based on the preliminary data analysis, the seasonal component of the time series is formed from two parts: weekly and annual. To solve the forecasting problem, the monthly components will also be taken into account in GAM models. Alternative models are presented as follows: an additive model with a monthly seasonal component (GAM 1); a multiplicative model with an annual seasonal component (GAM 2); a multiplicative model with annual and weekly seasonal components (GAM 3). Table 2 shows the values of the quality metrics of the GAM models.

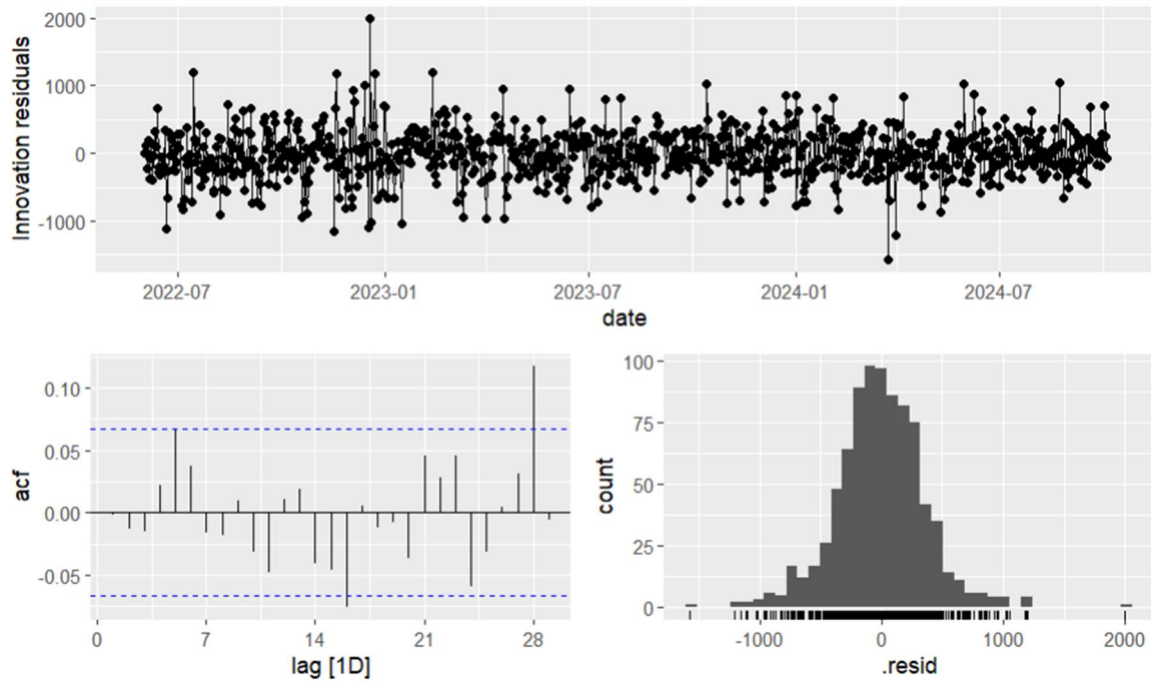


Figure 10: Residuals of the automatically generated model using the exhaustive search method.

Table 2

The values of the quality metrics of GAM models.

Model type GAM	R2	DW
Model with monthly seasonality component	0.992	2.011
Model with annual seasonality component	0.991	1.841
Model with annual and weekly seasonal components	0.994	2.059

Feedforward neural network models. Neural networks can be used to model various time series with complex structures without additional knowledge of the process features reflected in the data. When using neural network models to forecast time series, the following features must be taken into account: 1) the first differences are fed to the model input, not the original series; 2) the number of lags that are significant for describing a specific process is determined; 3) long-term trends are not modeled. The best architecture of the neural network model (3, 15, 1) was selected experimentally. It is shown in Figure 11.

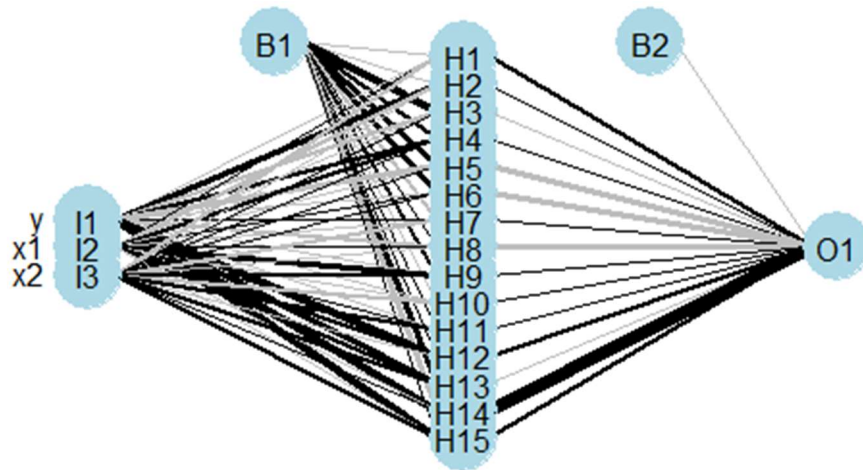


Figure 11: Structure of the neural network model.

The neural network model used was a multilayer direct propagation neural network with one hidden layer, a variable number of neurons and a sigmoid activation function. To prepare time series for modeling and forecasting, the method of constructing neural network models was used. The method consists of the following steps.

Step 1. Feeding the data set to the input of the neural network model based on 1-, 2- and (if necessary) 3-differences and the resulting vector.

Step 2. Splitting the data set into two parts (for training and for testing).

Step 3. Training the neural network.

Step 4. Visualizing the structure of the neural network.

Step 5. Calculating forecasts and evaluating forecast decisions.

Step 6. Determining the type of distribution of the model residuals.

Step 7. Determining the optimal parameters of the forecast model.

Step 8. Returning to the original data (inverting scaling + inverting differentiation).

Step 9. Analysis of the absence of autocorrelation in the residuals (ACF, PACF, Portmanteau test).

Step 10. Determining the final predictive solution.

Figure 12 shows a fragment of the input layer data of the neural network after preprocessing.

	df	y	x1	x2	x3
1	122.35833	0.4169840	NA	NA	NA
2	-255.50417	0.3211952	0.4169840	NA	NA
3	227.02917	0.4435182	0.3211952	0.4169840	NA
4	-403.41250	0.2837002	0.4435182	0.3211952	0.4169840
5	-20.05417	0.3808822	0.2837002	0.4435182	0.3211952
6	249.02500	0.4490942	0.3808822	0.2837002	0.4435182

Figure 12: Fragment of data from the input layer of a neural network.

Bayesian Structural Time Series Model. Structural time series models have three main advantages for modeling and forecasting complex time series [40,41]:

- Ability to account for uncertainty in forecasts, which can then help quantify future risks.
- Open structure of the model.
- Ability to include external information for factors where there is no obvious relationship in the data.

The BSTS model training algorithm consists of the following stages:

1. Defining the model structure and prior probabilities.
2. Using the Kalman filter to calculate state parameters based on current data.
3. Selecting variables in the structural model based on the splash-and-slab method.
4. Combining the results based on averaging over the Bayesian model to calculate the forecast.

The flexibility of the algorithm is based on the selection of components for each alternative BSTS model. This is evident in the first two stages of the algorithm. In the following stages, the model was trained on the available data using a Bayesian method that updates the parameter estimates over time. When solving the forecasting problem for the data set, several alternative BSTS models were compiled based on the results of the preliminary analysis and data processing. Table 3 presents the list of models.

The values of the BSTS model quality characteristics are presented in Table 4. The *residual.sd* characteristic is the mean of the posterior distribution of the standard deviation of the model residuals, and the *predict.sd* characteristic is the standard deviation of the next step errors determined based on the training data. The R^2 characteristic is the determination coefficient. The next characteristic, *relation.gof*, is the Harvey statistic.

Table 3

BSTS models complete set

Name of models	Model components
M1	Local linear trend + weekly seasonality component
M2	Local linear trend + trigonometric seasonality with two Fourier components (2p and sin 2 cos)
M3	Local linear trend + autoregressive component
M4	Local linear trend + monthly seasonality component
M5	Robust local linear trend + autoregressive component
M6	Robust local linear trend + annual seasonality component
M7	Robust local linear trend + autoregressive component + weekly seasonality component
M8	Local level component + autoregressive component

Table 4

Quality assessment of BSTS models

Name of models	residual.sd	prediction.sd	R2	relative.gof
M1	201.292	371.129	0.9397	0.186
M2	221.318	414.288	0.9272	-0.015
M3	87.265	378.960	0.9887	0.151
M4	202.237	421.484	0.9392	-0.050
M5	88.308	377.987	0.9888	0.116
M6	128.550	418.538	0.9754	-0.035
M7	100.053	357.875	0.9851	0.243
M8	107.622	376.483	0.9828	0.162

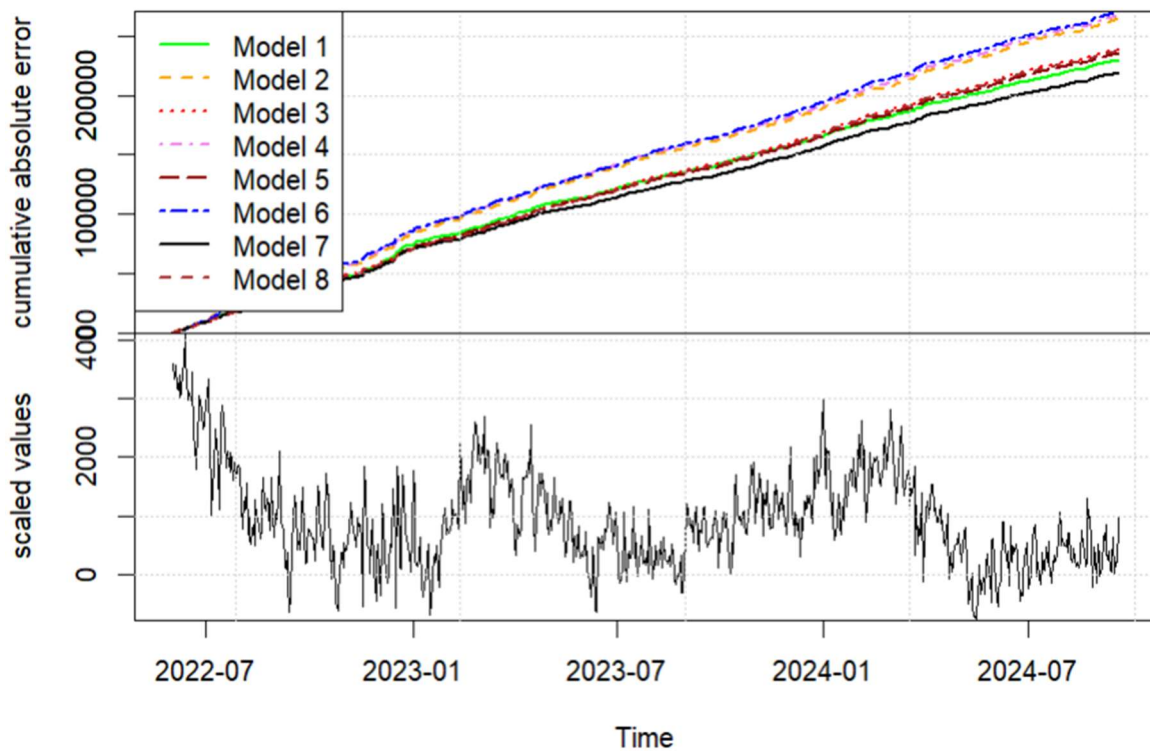
**Figure 13:** Comparisons of the quality of alternative BSTS models.

Figure 13 shows a graph that displays the quality of BSTS models. The training data set is shown below the graph of the accumulated errors curves. This allows us to better understand where exactly the model fails to describe the data. In the figure, the curve (M7), which is located below the other models, confirms the higher quality of this model.

The adequacy of BSTS models was assessed by how well they described the training data. This approach carried the risk of selecting an overfitted model as the optimal one. Evaluating models using the next-step errors partially helps to avoid the error of overfitting models.

3.3. Forecasting and evaluating results

The models with the best quality indicators were the basis for calculating the predicted values (Table 5). The quality indicators of the models were determined on the test sample. The following metrics were used to select the optimal model: MAPE, MAE, MSE and RMSE.

Figure 14 shows a visualization of the forecast values for the time series on electricity demand in Ukraine constructed using the BSTS model (M7). The black line represents 50 training data, the blue line represents the predicted values of the time series. The yellow dots on the graph are the data from the test sample. The green dotted lines limit the 95% confidence interval of the predicted values.

Table 5

Comparative table of forecast quality criteria for different models

Types of models	RMSE	MAE	MAPE	MSE
ARIMA (1, 1, 2)(2, 0, 0)7	360.778	320.009	3.578	130160.77
GAM (annual and weekly seasonal components)	387.664	339.114	3.704	150283.38
BSTS (robust local linear trend + autoregressive component + weekly seasonality component)	289.527	215.634	1.002	83825.84
NNAR (n=15, Sigmoid, maxit=5000)	352.032	314.993	3.289	123926.34

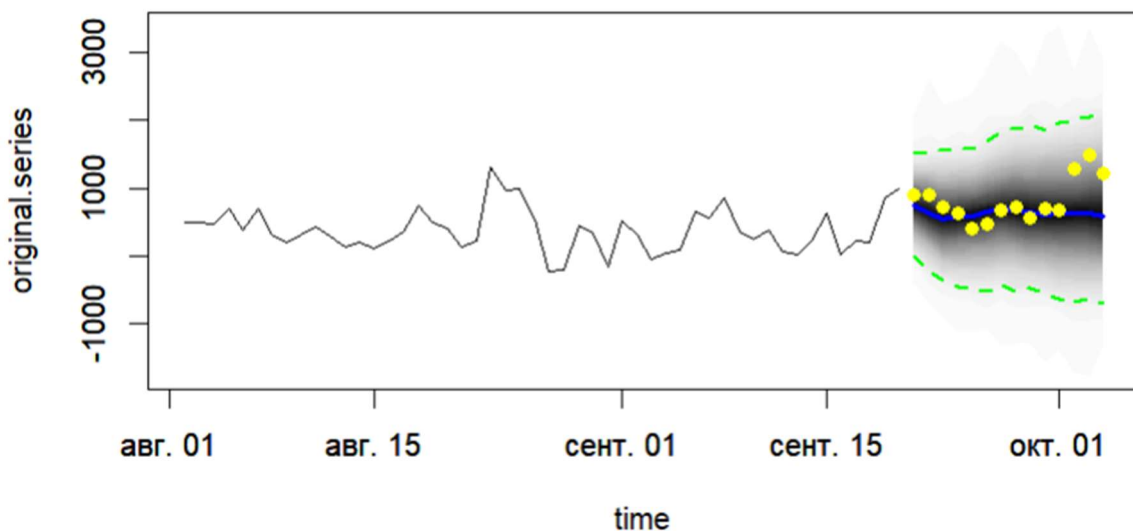


Figure 14: Visualization of forecast values for a fragment of a time series

The presented results demonstrate the effectiveness of using BSTS models to solve forecasting problems. Further improvement of forecasting results is possible through the use of combined forecasts [35].

4. Conclusions

The paper presents a study of forecasting the demand for electricity in Ukraine using machine learning methods. Research was conducted on the basis of data from 2019–2024. The sequence of stages of data processing when solving the problem of forecasting using machine learning methods is developed and presented. The sequence includes the following steps: data collection, data research and preparation, building and training forecast models, selecting the best model and calculating forecasts, evaluating and checking the quality of forecasts. A general methodology for solving forecasting problems is proposed. Methodology for solving the problem of forecasting time series based on machine learning methods. The solution to the forecasting problem consists of five stages. At the first stage, data collection, analysis and interpretation is carried out. At the second stage, research and data preparation procedures are carried out. The third stage - the modeling stage consists of three parts: preparation of the data set for modeling, selection and training of models and assessment of their quality. The fourth stage is the stage of forecasting and determining the quality of forecasts. At the fifth stage, procedures for increasing the effectiveness of the selected forecasting model are performed. At the modeling stage, the following models were used: ARIMA, GAM, ANN and BSTS. A detailed analysis of the models was carried out and predictions were made based on each model. Predictive values were calculated for the built models with the best quality indicators. The forecast was developed for 2 weeks. The forecasts were compared with the data of the validation sample. The following indicators were used to select the optimal model and evaluate it: MAPE, MAE, MSE, RMSE. The BSTS model showed the best results. This confirms the effectiveness of the BSTS model when forecasting on real data.

References

- [1] M. R. Kazemzadeh, A. Amjadian, T. Amraee. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy* 2020, 204, 117948.
- [2] J. Hao, X. Sun, Q. Feng. A Novel Ensemble Approach for the Forecasting of Energy Demand Based on the Artificial Bee Colony Algorithm. *Energies* 2020, 13, 550.
- [3] A. J. del Real, F. Dorado, J. Durán. Energy Demand Forecasting Using Deep Learning: Applications for the French Grid. *Energies* 2020, 13, 2242.
- [4] J. Bedi, D. Toshniwal. Deep learning framework to forecast electricity demand. *Appl. Energy* 2019, 238, pp. 1312–1326.
- [5] F. Kaytez. A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption. *Energy* 2020, 197, 117200.
- [6] S. Di Leo, P. Caramuta, P. Curci, C. Cosmi. Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models. *Energy* 2020, 196, 117058.
- [7] P. Ramsami, R. T. A. King. Neural Network Frameworks for Electricity Forecasting in Mauritius and Rodrigues Islands. In *Proceedings of the 2021 IEEE PES/IAS PowerAfrica*, Nairobi, Kenya, 2021, pp. 1–5.
- [8] D. Angelopoulos, Y. Siskos, J. Psarras. Disaggregating time series on multiple criteria for robust forecasting: The case of long-term electricity demand in Greece. *Eur. J. Oper. Res.* 2019, 275, pp. 252–265.
- [9] U. Şahin, S. Ballı, Y. Chen. Forecasting seasonal electricity generation in European countries under COVID-19-induced lockdown using fractional grey prediction models and machine learning methods. *Appl. Energy* 2021, 302, 117540.
- [10] R. Hou, S. Li, M. Wu, G. Ren, W. Gao, M. Khayatnezhad. Assessing of impact climate parameters on the gap between hydropower supply and electricity demand by RCPs scenarios and optimized ANN by the improved Pathfinder (IPF) algorithm. *Energy* 2021, 237, 121621.
- [11] A. Baba. Advanced AI-based techniques to predict daily energy consumption: A case study. *Expert Syst. Appl.* 2021, 184, 115508.

- [12] M. Pegalajar, L.G. B. Ruíz, M. P. Cuéllar, R. Rueda. Analysis and enhanced prediction of the Spanish Electricity Network through Big Data and Machine Learning techniques. *Int. J. Approx. Reason.* 2021, 133, pp. 48–59.
- [13] N. M. M. Bendaoud, N. Farah, S. Ben Ahmed. Applying load profiles propagation to machine learning based electrical energy forecasting. *Electr. Power Syst. Res.* 2022, 203, 107635.
- [14] R. Porteiro, L. Hernández-Callejo, S. Nesmachnow. Electricity demand forecasting in industrial and residential facilities using ensemble machine learning. *Rev. Fac. De Ing.* 2022, 102, pp. 9–25.
- [15] Y. Lu, G. Wang. A load forecasting model based on support vector regression with whale optimization algorithm. *Multimed. Tools Appl.* 2023, 82, pp. 9939–9959.
- [16] S. Li, X. Kong, L. Yue, C. Liu, M. A. Khan, Z. Yang, H. Zhang. Short-Term Electrical Load Forecasting Using Hybrid Model of Manta Ray Foraging Optimization and Support Vector Regression. *J. Clean. Prod.* 2023, 388, 135856.
- [17] J. Huang, M. Algahtani, S. Kaewunruen. Energy Forecasting in a Public Building: A Benchmarking Analysis on Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and Extreme Gradient Boosting (XGBoost) Networks. *Appl. Sci.* 2022, 12, 9788.
- [18] C. E. Velasquez, M. Zocatelli, F. B. Estanislau, V. F. Castro. Analysis of time series models for Brazilian electricity demand forecasting. *Energy* 2022, 247, 123483.
- [19] F. Pallonetto, C. Jin, E. Mangina. Forecast electricity demand in commercial building with machine learning models to enable demand response programs. *Energy AI* 2022, 7, 100121.
- [20] E. C. May, A. Bassam, L. J. Ricalde, M. E. Soberanis, O. Oubram, O. M. Tzuc, A. Y. Alanis, A. Livas-García. Global sensitivity analysis for a real-time electricity market forecast by a machine learning approach: A case study of Mexico. *Int. J. Electr. Power Energy Syst.* 2022, 135, 107505.
- [21] W. J. Niu, Z. K. Feng, S. S. Li, H. J. Wu, J. Y. Wang. Short-term electricity load time series prediction by machine learning model via feature selection and parameter optimization using hybrid cooperation search algorithm. *Environ. Res. Lett.* 2021, 16, 055032.
- [22] R. Luzia, L. Rubio, C. E. Velasquez. Sensitivity analysis for forecasting Brazilian electricity demand using artificial neural networks and hybrid models based on Autoregressive Integrated Moving Average. *Energy* 2023, 274, 127365.
- [23] P. C. Albuquerque, D. O. Cajueiro, M. D. C. Rossi. Machine learning models for forecasting power electricity consumption using a high dimensional dataset. *Expert. Syst. Appl.* 2022, 187, 115917.
- [24] R. Rick, L. Berton. Energy forecasting model based on CNN-LSTM-AE for many time series with unequal lengths. *Eng. Appl. Artif. Intell.* 2022, 113, 104998.
- [25] M. Maaouane, M. Chennaif, S. Zouggar, G. Krajaci'c, N. Duic, H. Zahboune, A. K. ElMiad. Using neural network modelling for estimation and forecasting of transport sector energy demand in developing countries. *Energy Convers Manag.* 2022, 258, 115556.
- [26] B. Lantz. *Machine Learning with R. Expert techniques for predictive modeling*, 3rd Edition, Packt Publishing, 2019. 458 p.
- [27] J. D. Kelleher, B. Mac Namee, A. D'Arcy. *Fundamentals of Machine Learning for Predictive Data Analytics. Algorithms, Worked Examples, and Case Studies.* Second Edition. The MIT Press, Cambridge, Massachusetts, London, England, 2020. 798 p.
- [28] G. James, D. Witten, T. Hastie, R. Tibshirani. *An Introduction to Statistical Learning with Applications in R.* Springer New York Heidelberg Dordrecht London, 2013. 441 p. doi: 10.1007/978-1-4614-7138-7.
- [29] I. Kalinina, A. Gozhyj. *Methodology for Solving Forecasting Problems Based on Machine Learning Methods. Lecture Notes on Data Engineering and Communications Technologies (Switzerland).* 2023. Vol. 149, pp. 105–125.
- [30] P. Bidyuk, A. Gozhyj, I. Kalinina, V. Vysotska, *Methods for forecasting nonlinear non-stationary processes in machine learning.* In: *Data Stream Mining and Processing. DSMP 2020. Communications in Computer and Information Science.* 2020, Vol. 1158, pp. 470–485. Springer, Cham, (2020). doi: 10.1007/978-3-030-61656-4 32.

- [31] Energy Map. URL: <https://map.ua-energy.org/uk/resources/5a616fba-fbc9-4073-9532-9161592faca8/>
- [32] P. Bidyuk, I. Kalinina, A. Gozhyj. An Approach to Identifying and Filling Data Gaps in Machine Learning Procedures. Lecture Notes on Data Engineering and Communications Technologies (Switzerland). 2022. Vol. 77, pp. 164-176.
- [33] R. J. Hyndman, G. Athanasopoulos. Forecasting: Principles and Practice 3rd ed. Edition. Publisher: OTexts. 2021, 442 p.
- [34] T. Aggarwal. Master the Power of Seasonal Decomposition of Time Series (STL): Unveiling the Essence of Time. 2023. URL: https://medium.com/@tushar_aggarwal/master-the-power-of-seasonal-decomposition-of-time-series-stl-unveiling-the-essence-of-time-26c19a910314
- [35] I. Kalinina., P. Bidyuk., A. Gozhyj, P. Malchenko. Combining Forecasts Based on Time Series Models in Machine Learning Tasks. CEUR-WS. 2023. Vol. 3426. Pp. 25-35. CEUR-WS.org/Vol-3426/paper2.pdf.
- [36] A. Nielsen. Practical Time Series Analysis. Prediction with Statistics and Machine Learning. O'Reilly Media, Inc., 2019. 504 p.
- [37] J. Brownlee. A Gentle Introduction to the Box-Jenkins Method for Time Series Forecasting. 2020. URL: <https://machinelearningmastery.com/gentle-introduction-box-jenkins-method-time-series-forecasting/>
- [38] N. Clark. How to interpret and report nonlinear effects from Generalized Additive Models. 2024. URL: <https://ecogambler.netlify.app/blog/interpreting-gams/>
- [39] J. Lai, J. Tang, T. Li, A. Zhang, L. Mao. Evaluating the relative importance of predictors in Generalized Additive Models using the gam.hp R package. Plant Diversity, vol. 46, issue 4, 2024, pp. 542-546. DOI: 10.1016/j.pld.2024.06.002.
- [40] I. Kalinina, P. Bidyuk, A. Gozhyj. Construction of Forecast Models based on Bayesian Structural Time Series. International Scientific and Technical Conference on Computer Sciences and Information Technologies. CSIT_2022. 2022. Pp. 180-184. doi: 10.1109/CSIT56902.2022.10000484.
- [41] A.M. Almarashi, K. Khan. Bayesian Structural Time Series. 2020. Nanoscience and Nanotechnology Letters, 12(1), pp. 54-61. doi:10.1166/nnl.2020.3083.