Article

Neuro-Cybernetic System for Forecasting Electricity Consumption in the Bulgarian National Power System

Kostadin Yotov, Emil Hadzhikolev [ID], Stanka Hadzhikoleva *[ID] and Stoyan Cheresharov [ID]

Faculty of Mathematics and Informatics, University of Plovdiv Paisii Hilendarski, 236 Bulgaria Blvd., 4027 Plovdiv, Bulgaria
* Correspondence: stankah@uni-plovdiv.bg

Abstract: Making forecasts for the development of a given process over time, which depends on many factors, is in some cases a difficult task. The choice of appropriate methods—mathematical, statistical, or artificial intelligence methods—is also not obvious, given their great variety. This paper presented a model of a forecasting system by comparing the errors in the use of time series on the one hand, and artificial neural networks on the other. The model aims at multifactor predictions based on forecast data on significant factors, which were obtained by automated testing of different methods and selection of the methods with the highest accuracy. Successful experiments were conducted to forecast energy consumption in Bulgaria, including for household consumption; industry consumption, the public sector and services; and total final energy consumption.

Keywords: electricity consumption; forecast energy consumption; forecasting system

1. Introduction

The forecasting of the future is extremely important for the effective management of a process or system. Forecasting is about predicting the future as accurately as possible, given all of the information available, including historical data and knowledge of any future events that might impact the forecasts [1]. From a scientific point of view, forecasting is a scientifically based assumption about the future state and development of processes, events, indicators, etc. [2]. Considering the possibility of the existence of many different forecasts for the development of a given process in the future, forecasting can be defined as a reasonable assumption of possible options for development in a given area and the probability that they will be realized.

The synergy between mathematics and computer science has led to the development of a wide variety of algorithms, approaches, methods, and tools for forecasting. Widely used, with application in various fields are mathematical and statistical methods including regression and clustering [1,3], time series [4,5], polynomial approximations [6], fuzzy collaborative methods [7], as well as many methods for artificial intelligence predicting, such as machine learning [8,9], etc. On the one hand, this diversity provides an opportunity to choose a specific approach to solving a given task, but, on the other hand, it makes it difficult to find the most effective solution.

In the process of our work on multifactor and multi-step forecasting of energy consumption in the Republic of Bulgaria, we came to the need to forecast many socio-economic factors through which to make the final forecast. The functions, by which the individual factors change, as well as the energy consumption, can have a variety of linear and nonlinear forms, where the appropriate forecasting methods for each of them may be different. Determining the most accurate forecast values for the factors would have a positive effect on the accuracy of forecasting the target value, which in our case is energy consumption.

The automation of the process of choosing the most effective method for any individual factor or target value contributes to the acceleration of the process and the improvement of

the forecast accuracy. Finding the most effective forecasting method requires experimenting with several different approaches that are combined and compared. Making more effort to solve the prognostic task could save future effort, time, and money. Fast, but essentially inaccurate forecasts often lead to unreasonable investments or unrealized opportunities.

2. Predicting Electricity Consumption—State of the Art

Whether the forecast is long-term, medium-term, or short-term, predicting electricity consumption has a key role to play in investment planning, the introduction of new capacities or the decommissioning of unnecessary ones, and the assessment of the behavior of the entire economic system. Effective modeling of electricity consumption is becoming a vital task aimed at avoiding costly mistakes in unreasonable investments, shutting down important facilities, improperly scheduled repairs or short-sightedness in exports or imports. Therefore, we should not be surprised that in the literature on the subject there are many proposed options for dealing with forecasting problems.

Official studies focused on the development of the energy sector in Bulgaria, factors influencing electricity consumption and approaches for forecasting the consumption and factors have been conducted by teams from the Bulgarian Academy of Sciences (BAS) and Risk Management Lab. Traditional methods such as correlation and regression analyses have been used in both studies.

In the research, conducted in BAS, a basic and in-depth analysis has been made of the impact of individual factors influencing electricity consumption. Taken into consideration are the country's gross domestic product, gross value added by economy sectors, population size, number of employees, income, prices, changes in temperature until 2040 according to the Bulgarian National Institute of Meteorology and Hydrology, etc. Forecasts have been made based on three different scenarios according to different expectation for changes in factors [10].

The team of Risk Management Lab creates mathematical and statistical models for forecasting the electricity balance (including as an element of itself and forecasting electricity consumption). The study examines specific factors in terms of electricity consumptions by households and the industry [11].

Due to the great social and economic importance of forecasting electricity consumption, many scientists have proposed different types of forecasting models to solve the problem in the last few decades. The methods for forecasting electricity consumption can be defined in several categories:

- Statistical models for analysis—correlation methods and regression models—BAS strategy [10], Risk Management Lab [11], Mohamed and Bodger [12];
- Time series—Lee, Gaik and Yee [13], Sun, Zhang et al. [14], Chou and Tran [15];
- Granger gray forecasting systems—Ding, Hipel and Dang [16], Lee and Tong [17], Huang, Wang, et al. [18];
- Neural networks—Chung [19], Chernykh, Chechushkov and Panikovskaya [20], Khosravani, Castilla, et al. [21], Yoo and Myriam [22], Hu and Yi-Chung [19], Jahn [23];
- Advanced machine learning methods—Bouktif, Fiaz, Ouni and Serhani [24], Alamaniotis, Bargiotas and Tsoukalas [25], Alamaniotis [26], Chou, Truong [27], Kong, Dong, Jia, Hill, Xu and Zhang [28], Moradzadeh, Moayyed, Zakeri, Mohammadi-Ivatloo and Aguiar [29];
- Bagging and boosting methods—Khwaja, Anpalagan, Naeem and Venkatesh [30], Khwaja, Zhang, Anpalagan and Venkatesh [31], Cao, Wan, Zhang, Li and Song [32];
- Hybrid methods—Karatasou, Santamouris and Geros [33], Cervená and Schneider [34], Farsi, Amayri, Bouguila and Eicker [35], Chung, Gu and Yoo [36], Sajjad, Khan, Ullah, Hussain et al. [37], Alamo, Medina, Ruano et al. [38] etc.

Each of the considered approaches has its advantages and disadvantages in different specific tasks and situations. This shows that the creation of an automated forecasting system in which multiple forecasting models can be implemented and compared could bring many benefits in forecasting not only electricity consumption but also other quantities.

3. Materials and Methods

3.1. Forecasting Process

In the general case, the solution of forecasting tasks is performed in a similar way and the process can be reduced to several stages—data collection and processing, a study of solution methods, solution, and analysis of the results [1]. The algorithm is iterative and individual stages can be performed and overlap repeatedly over time.

Data collection is often associated with preliminary research in the subject area, which provides additional information related to the set task. At this stage, an idea of the possible factors influencing the predicted values is formed. In the best case, the data on the selected factors are provided by the stakeholder or can be collected from one or more sources. In other cases, new technical equipment and software applications must be built in order to collect them.

The merging and synchronization of data by time, location, seasons, or other criteria is an integral activity when using multiple data sources providing data for various factors.

Data analysis and processing include activities to check and clear incorrect input data; conversion of data into structures suitable for modeling; and graphical representation of the data through which trends, periodicity, seasonality, etc., can be detected. In some cases, data behavior may be affected by various methodological and technological differences in data collection, as well as social, societal, climatic, or other changes. This implies the use of various methods for analysis and subsequent forecasting of the formed segments.

In the stage of research of methods for solving the problem of forecasting, various existing methods and algorithms are considered or specific ones are created. Both standard statistical and artificial intelligence methods and models can be used as forecasting methods. The number of methods and their use may depend on the requirements of the specific task, on factors related to the environment, performers, etc. An important part of this stage is the definition of indicators for evaluating the effectiveness of forecasting methods. The evaluation of efficiency may include various parameters depending on the volume of data, technical capabilities of computer systems, cost, speed, etc.

In the last stages, one or several of the perspective models for forecasting are selected and applied for the solution of the problem. After evaluating the results, individual steps can be repeated many times in order to achieve better results.

This whole process requires a lot of time and resources. This raises the question—is it possible to fully or partially automate the forecasting process, including the conducting of experiments with various mathematical and artificial intelligence methods, comparing errors, and choosing the best solution? This is the purpose of our work that is presented in this article.

3.2. Multifactor Forecasting System through Automated Selection of the Best Methods

The successful solution of the set task for forecasting energy consumption in the national power system is directly related to the correct assessment of the factors influencing energy consumption. These are macroeconomic and demographic indicators, social parameters, weather conditions, and others. This work investigated several factors influencing electricity consumption: gross domestic product; energy intensity; population size; population income; price of electricity; expected temperatures for the respective period; energy efficiency; and electricity consumption in a preceding period.

The automation of various stages and activities of the forecasting process requires careful planning. Such a system must support the following basic capabilities:

- mechanisms for integrating different forecasting methods into the system;
- potential to work with a different number of factors;
- automated search and selection of effective forecasting methods when solving a specific task.

In our proposed model for an automated search of effective forecasting methods, two main approaches are used:

- Complex forecasting, which is multivariate time series forecasting. It is applicable to multifactor forecasting, in which preliminary forecasting of individual factors is performed;
- Simple forecasting, which is univariate time series forecasting.

The long-term goal of our research is to develop a neurocybernetic system for energy consumption forecasting that supports various mathematical and artificial intelligence methods. In the first version, we used a neural network in its basic version. Further work on the system involves adding other forecasting methods.

3.3. Mathematical Model of the Forecasting Task

The main approach to work in the forecasting system through artificial neural networks can be formally presented as follows:

Let us assume we have the following event:

$$S = S\left(\vec{X_1}, \vec{X_2}, \dots \vec{X_k}\right),$$

whose outcome is determined by k influencing factors $\vec{X_1}, \vec{X_2}, \dots \vec{X_k}$. Let a sample of values be provided for each of the factors $\vec{X_i}$

$$\vec{X_i} : \left(x_{i,1}, x_{i,2}, \dots x_{i,j_i}\right), i = 1, 2, \dots k, \ i, j_i \in N.$$

The main steps performed by the system when using complex forecasting are the following:

(1) Automatically searches for and constructs a neural network NN with input vectors $\vec{X_i}$, approximating S with minimal error E_{NN};

(2) Based on the samples $\left(x_{i,1}, x_{i,2}, \dots x_{i,j_i}\right)$ forecasts the future behavior of the factors $\vec{X_i}$ using p in a number of different factors $\{m_t\}_{t=1\dots p}$;

(3) For each of the factors $\vec{X_i}$ the efficiency of the used forecasting methods is compared separately;

(4) For the final forecast value of $\vec{X_i}$, the forecast of the respective most effective method is selected;

(5) Through the forecasted values of the factors $\vec{X_i}$ and the method M, the outcome of the event is predicted.

Various criteria can be used to evaluate the effectiveness of forecasting methods, such as decision error, cost, speed, computer resources used, etc. In our study, we chose to call the method M_1 more effective than M_2 if it has a lesser prediction error, i.e.,

$$E_{M_1} \leq E_{M_2}. \tag{1}$$

Condition (1) allows us to introduce a formula for the efficiency of a forecasting method as the inversely proportional values of the error that occurs when forecasting with it:

$$efficiency = \frac{1}{error}. \tag{2}$$

Absolute error (AE) or root mean square error, obtained when forecasting on an n-tuple data array, can be used. For single point forecasts created on $(n-1)$ the sample element, the absolute forecasting error for the last element is:

$$AE = |Real_n - Forecast(n)|, \tag{3}$$

where

- *Real_n* is the actual value of the measured value,
- *Forecast*(n) is the predicted value proposed by the method used.

For multi-step forecasts for the last *k* elements of the data array, one of the most popular error metrics were used—the mean absolute error (MAE), mean square error (MSE), and symmetric mean absolute percentage error (SMAPE) [39]. They are calculated by the formulas:

$$\text{MAE} = \frac{1}{k} \sum_{i=n-k+1}^{n} |real_i - forecast(i)|, \tag{4}$$

$$\text{MSE} = \frac{1}{k} \sum_{i=n-k+1}^{n} [real_i - forecast(i)]^2. \tag{5}$$

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|Forecast - Real|}{|Forecast| + |Real|} \tag{6}$$

3.4. System Architecture

The modules in a forecasting software system automate the main activities (Figure 1). The user of the system accesses the individual modules through a common interface provided by the Manager module. In addition to the connection with the user, it provides management and control over the other modules.
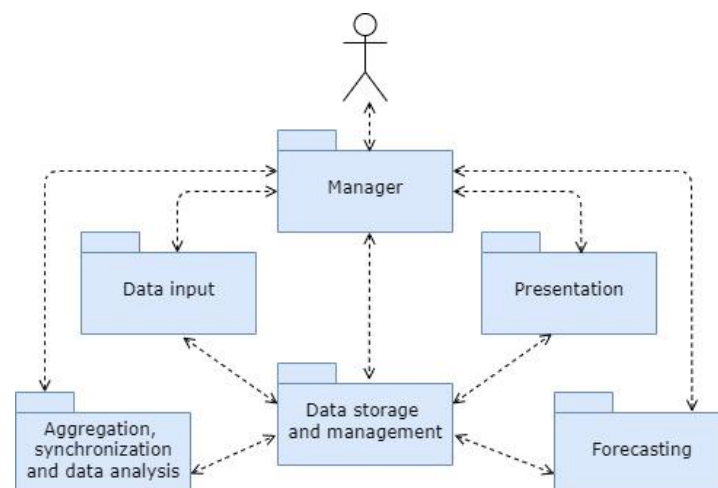

Figure 1. Main modules and the connections between them in an automated forecasting system.

The data entry module assists the user in entering data and their initial classification.

Merging and synchronization tools are useful in cases where data from different sources are used. At the user's choice, through functions or parameters, the data are converted into a format suitable for making forecasts. The data analysis is supported by graphical tools integrated in the module and standard statistical methods, providing the user with opportunities for additional classification and clearing of incorrect data, as well as preparation of various data models.

The forecasting module presents an opportunity to choose one or more forecasting methods, as well as an opportunity to evaluate the most effective method. An important feature that the system must support is easy integration of new forecasting and evaluation methods in the module.

The presentation module contains graphical tools for visualization of the results obtained from the most effective method, as well as from all other methods used.

The data storage and management module provide access to various types of data that can be used in the configuration, training of forecasting methods and their subsequent use:

- primary input–output data and initial parameters;
- data models used by forecasting methods;
- maximum permissible error criteria set by the users;

- number and type of methods used;
- errors obtained in each of the methods;
- identifier of the most effective method;
- results from forecasting tasks, etc.

3.5. Multifactorial Multi-Step Forecasts

The forecasting module (Figure 2) provides us with two different forecasting approaches:

- simple forecasting—direct forecasting with known behavior of the factors, or if we consider the values of the target variable as a time series;
- complex forecasting—the forecasting of the target value is based on additionally created forecast data for the individual factors.
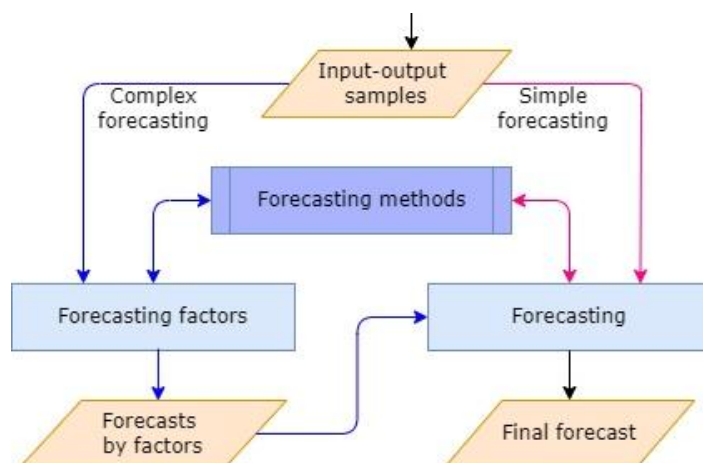

Figure 2. Forecasting module.

During the process of factor forecasting, for each of the factors the most effective method was sought, which could be used repeatedly (Figure 3). It is appropriate to forecast the individual factors in parallel and to use cloud computing to save time and resources.

The factor forecasting module receives a two-dimensional array of input data $\left\{x_{i,\,j}\right\}_{i=1...k}^{j=1...n_i}$, containing the sample for k in number factors, each with n_k values. Other input parameters are the number of desired forecast values $c \in N$ and the lower limit of the desired *efficiency*. For each of the factors, independent forecasts were made using the library of "prediction methods".

The set of forecasting methods $\{m_t\}_{t=1...p}$, $p \in N$ was pre-set in the software system and could be extended. The methods include a variety of time series forecast models, models based on artificial neural networks in which parameters such as the number of neurons, activation functions, learning algorithms and other artificial intelligence algorithms can be changed. Models with time series and artificial neural networks are integrated in the system so far.

Applying the forecasting methods, for each of the factors there are approximating functions with the desired efficiency $\left\{f_t(x_{i,\,j})\right\}_{j=(n_i-c+1)...n_i}^{t=1...p}$, $i = 1, 2, \ldots k$.

In the process of evaluating the effectiveness, using the AE, MAE, and MSE errors, the efficiencies of the tested methods or variants of methods were compared. The most appropriate forecasting model was selected, including the method $m_{t_1} \in \{m_t\}_{t=1...p}$ and its corresponding parameters.
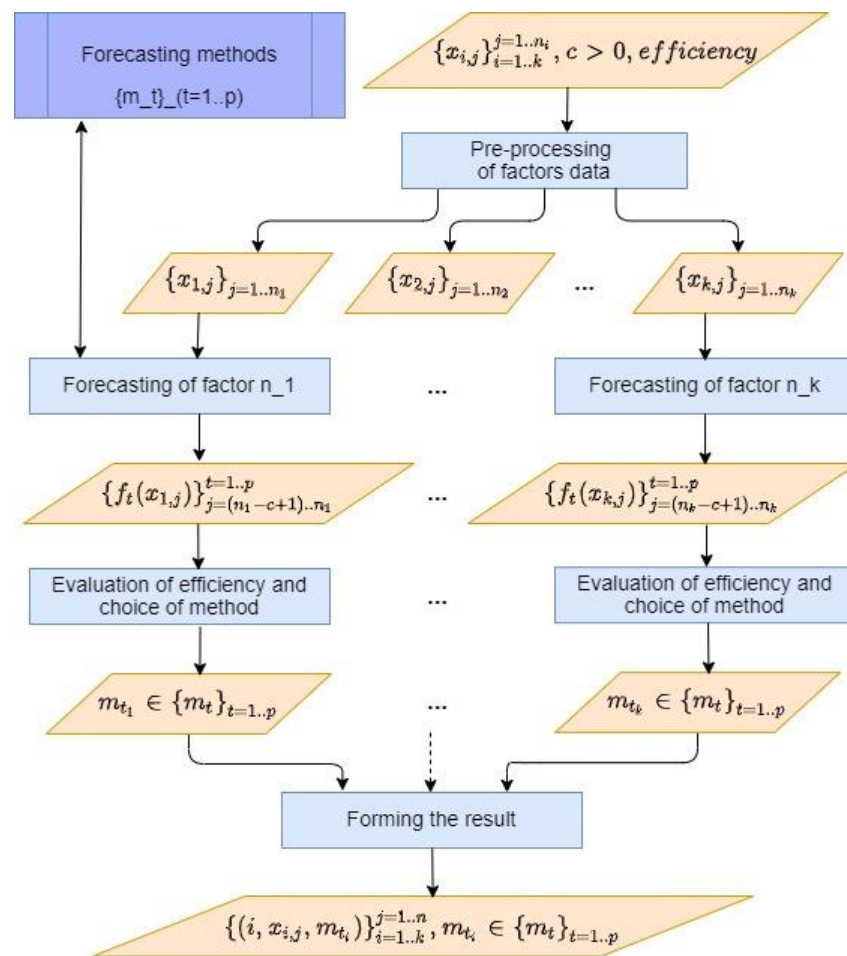
Figure 3. Choice of forecasting method for each factor.

The end result of the process is a set of ordered triples <factor, factor data, selected forecasting method and its corresponding parameters>, i.e., $\left\{\left(i,\ x_{i,\ j},\ m_{t_i}\right)\right\}_{i=1...k}^{j=1...n}$, $m_{t_i} \in \{m_t\}_{t=1...p}$. When the data change, the same selected method can be used, or the process can be restarted to search for a new method.

3.6. Algorithm for Searching for an Optimal Artificial Neural Network

To solve a task, many different neural networks can be built, with different numbers of neurons in the hidden layer, with greater or lesser error and different result functions, which have similar behavior in the input–output samples used. An error is not always an indicator of the complexity of the neural network. The lesser number of neurons implies faster and easier learning of the neural network and, subsequently, faster work. The optimal neural network for solving a certain task has characteristics such as a minimum number of neurons and compliance with the user-specified allowable error in training, testing, and validation with the available input-output samples. Its effectiveness in newcomer input data can be evaluated at a later stage. An algorithm providing capabilities for automated construction of multiple neural networks would help choose the optimal solution.

One of our experimented approaches for creating an optimal neural network is iteration over various parameters (Figure 4) needed to create neural networks: number of neurons in the hidden layer, activation functions (Table 1), training algorithms (Table 2), and number of training epochs. It successively changes the parameters for creating neural networks and examines the efficiency of the current neural network. The first neural network that meets the requirements set by the user is considered optimal. Therefore, an important part of the algorithm is how to change the iterative parameters. Since the creation

and training of each neural network requires a certain amount of computer resources and time, the iterative approach is appropriate to use on single-processor machines only for tasks where finding the appropriate neural network is expected to have a relatively small number of iterations.
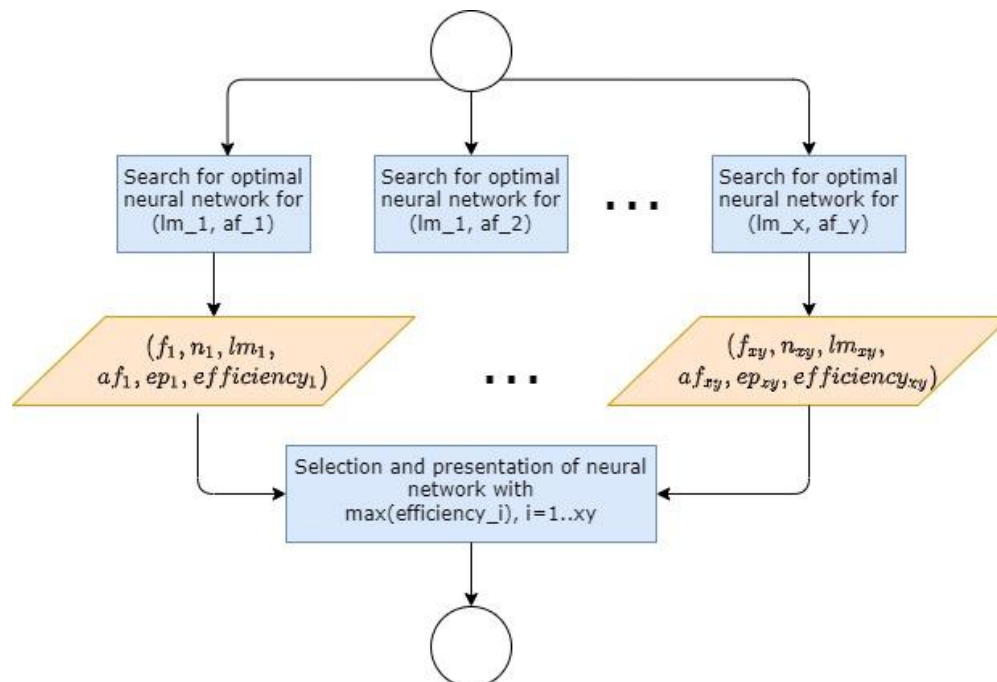

Figure 4. A variant of a parallel algorithm for constructing an optimal neural network for forecasting.

Table 1. Used activation functions.

Function
Hyperbolic tangent sigmoid transfer function
Log-sigmoid transfer function
Hard-limit transfer function
Symmetric hard-limit transfer function
Competitive transfer function
Elliot symmetric sigmoid transfer function
Elliot 2 symmetric sigmoid transfer function
Inverse transfer function
Positive linear transfer function
Radial basis transfer function
Normalized radial basis transfer function
Saturating linear transfer function
Symmetric saturating linear transfer function
Soft max transfer function
Triangular basis transfer function

Table 2. Used training methods.

Method
Levenberg–Marquardt
Bayesian Regularization
Scaled Conjugate Gradient
BFGS Quasi-Newton
Resilient Backpropagation
Conjugate Gradient with Powell/Beale Restarts
Fletcher-Powell Conjugate Gradient
Polak–Ribiére Conjugate Gradient
One Step Secant
Variable Learning Rate Backpropagation
Levenberg–Marquardt
Bayesian Regularization

Finding the optimal solution faster is associated with:

- Indication of parameters for changing the number of neurons. Reducing the limits of variation in the number of neurons will reduce the total number of iterations in the algorithm and therefore will accelerate the achievement of the final result—the creation of an optimal neural network. In some cases, the expected number of neurons can be justified mathematically [40,41].

- Arrangement of the used training methods and activation functions. Different activation functions are suitable for different tasks. To solve problems requiring the application of mathematical logic, we included the hard-limit transfer function and its variant—the Symmetric hard-limit transfer function. These functions provide an interrupted binary signal along the axon of neurons and are suitable in case there is a need to solve problems requiring binary logical thinking. On the other hand, to forecast the future development of continuous processes, we used a large number of efficient and convenient transfer functions. Priority in the order of use is given to those with a sigmoid character such as hyperbolic tangent sigmoid transfer function and log-sigmoid transfer function, useful in a very wide range of problems [42].

- Necessity to change the number of learning epochs. Sometimes, to reach the optimal neural network it is only necessary to change the number of training epochs. Increasing the iterations slows down the neural network learning process, but often even a slight change in this parameter leads to a surprising overcoming of a small plateau of neural network errors and leads to a sharp improvement in the final approximation.

The iterative algorithm for automated construction of artificial neural networks (Figure 5) has the following main steps:

(1) Preparation of input data:

- Tensor data—input data for the factors, which are usually in the form of a one-dimensional—$\{x_i\}_{i=1...k}$—or two-dimensional array—$\{x_{i,\,j}\}_{i=1...k}^{j=1...n_i}$.

- Number of forecasted results—c—which is 1 for single-point forecasts or a larger integer for multi-step forecasts.

- Desired efficiency—*efficiency*—of the trained neural network.

(2) Preparation of the parameters on which iteration is performed:

- List of training methods $lms = \{m_i\}$, where i varies from *1* to the number of methods (Table 2). Depending on the task, to achieve the desired result faster, it is possible to arrange the methods in the list according to the expected efficiency, and some of them may even be excluded if they are considered inappropriate.

- List of activation functions $afs = \{af_j\}$, where j varies from *1* to the number of functions (Table 1). Here, too, the functions can be arranged at the discretion of the appropriateness of their use in the specific task.

- Minimal and maximal number of neurons—*min_n u max_n,* as well as a step by which neurons change—*step_n*. The current number of neurons we denote by n. For more elementary tasks, the number of neurons may start from *1* (*min_n* = 1) and the step by which their number increases is also *1*. The maximum number limits the possible iterations related to the number of neurons.

- The epochs *epochs* change from *min_ep* to *max_ep* with a step *step_ep*. Values we have experimented with are $min_{ep} = 1000$, $max_{ep} = 5000$, $step\_ep = 1000$, where usually 3–4 iterations are enough to assess whether the change of epochs affects the efficiency of the trained neural network.

- By iterating over the number of neurons, training methods and activation functions, a neural network with their current values is created—the ordered triple (n, lm, af) and the input data. The nesting of the loops for the specific task is a matter of judgment, which determines the sequence of the parameters change. In the experiments, we chose to increase the number of neurons in the outermost loop, as we wanted to find a neural network with the lowest number of neurons. Training methods and activation functions change in inner loops.

(3) After the creation of the current neural network, it is trained, tested, and validated.

(4) The desired efficiency is compared to the efficiency of the current neural network, and then:

- If the neural network meets the condition, its data is saved and the task is completed;
- Otherwise, attempts are made to increase the efficiency of the neural network by increasing the number of learning epochs. The information about the most efficient neural network found (with the smallest error) is saved, and it can be current or obtained in a previous iteration.

(5) The end result of the algorithm is a trained neural network having the closest possible to the specified efficiency, as well as parameters for its architecture and training—*efficiency*, number of neurons *n*, training method *lm*, activation function *af*, and epochs *ep*.

The use of the "brute force" method, by traversing all possible values of the iterative parameters and finding the neural network with the optimal ratio "number of neurons—efficiency" is not the most rational approach. Tracking changes in the specified ratio can lead to the creation of heuristic variants of the algorithm by automated changing of the order of change of the iterative parameters.

The availability of sufficient computing power predisposes to the use of different parallel algorithms for such automated search for an optimal neural network (Figure 5) in which, for example, all combinations of training methods and activation functions can be started in parallel processes—(lm_i, af_j), where *i* and *j* are changed from *1* to the corresponding number of training methods and activation functions. After the completion of the individual processes, the efficiencies of all neural networks created during their implementation are compared and the most appropriate one is selected.
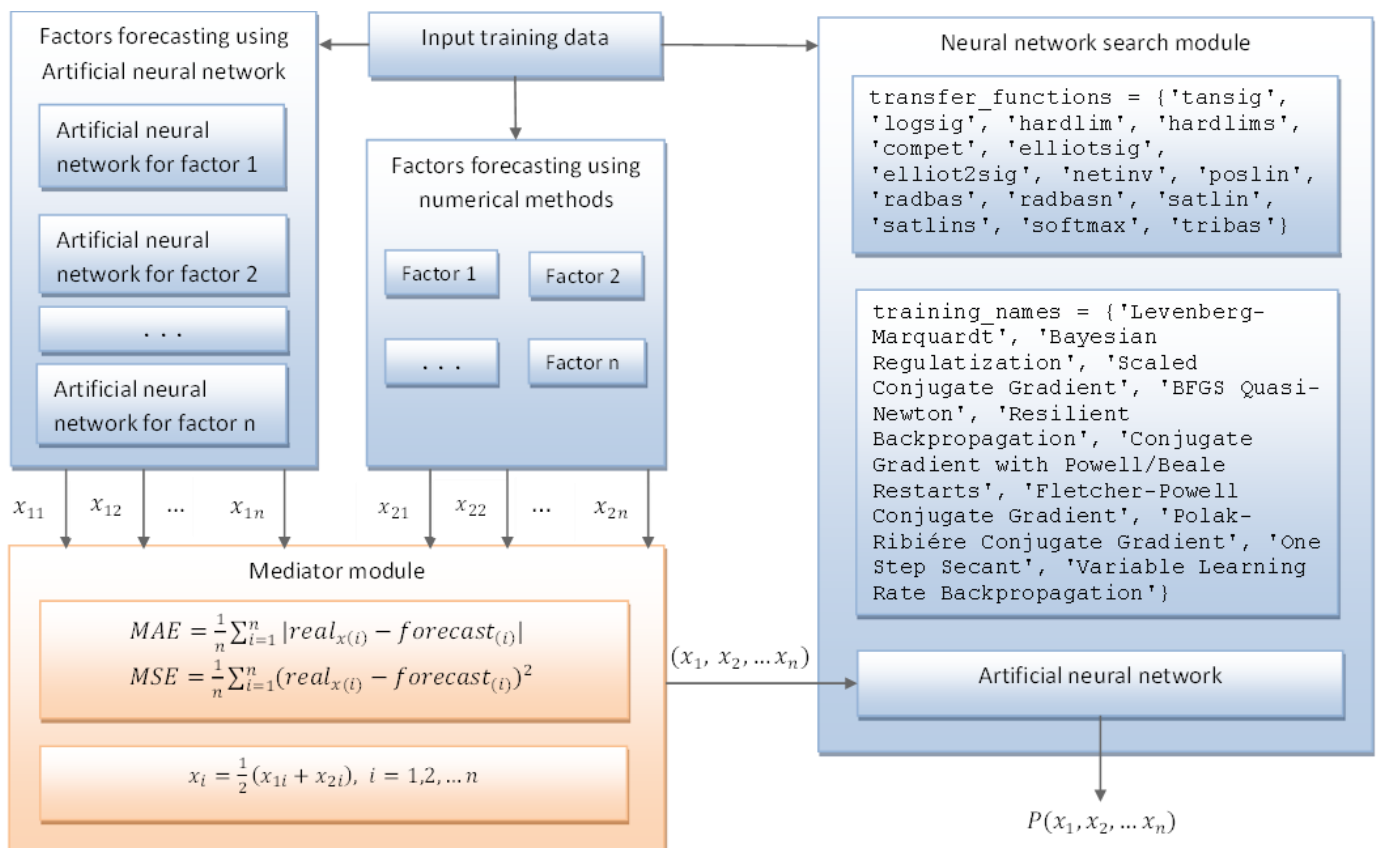

Figure 5. General scheme of forecasting processes.

4. Results and Discussion

A prototype was developed for approbation of the presented model. The program MatLab was used, with which the described basic modules and functionalities were implemented. The prototype was tested to solve prognostic tasks in the field of energy.

4.1. Setup of the Experiment

The problem of forecasting the demand and respectively—the consumption of electricity is extremely important for the planning and management of the national energy system of every country. Accurate forecasting of probable electrical loads is an important prerequisite for effective planning of production capacity, proper maintenance of the transmission and distribution network, planning of future exports or imports of electricity, the behavior and direction of energy flows both in the country and in related international networks.

The developed prototype was used to solve three tasks for forecasting electricity consumption in the National Power System of Bulgaria. The targets subject to forecasting were:

(1) Total final consumption in the national power system;
(2) Electricity consumption in the industry, the public sector, and services;
(3) Electricity consumption in households.

The forecasts were made by taking into account their dependencies on the following socio-economic factors:

(1) Gross domestic product (GDP);
(2) Energy intensity (EI);
(3) Population;
(4) Income of the population;
(5) Energy efficiency;
(6) Price of electricity.

These factors have been identified as significant in studies carried out by the Bulgarian Academy of Sciences (BAS) and Risk Management Lab [10,11]. A subsequent step would be to build a module for the automated study of correlations between the factors with the aim of minimizing their number and optimizing post-processing.

In order to forecast the target values for a specific year, we first made forecasts for the factors for the respective year.

In the process of work, data from official sources such as the National Statistical Institute [42], Information System INFOSTAT [43], Electricity System Operator [44], The World Bank Group [45], and Eurostat [46] were used. All available data on factors and target values for 17 years were used for the study.

The factor forecasting module uses different models of the time series trend. A substructure involving the use of neural networks to predict these factors provides for the possibility of further processing of the input data. Some of the data are submitted to the neural networks normalized, which is implemented by multiplying them by a specific coefficient. This reduces the radius vector of the input data and facilitates the training of this type of artificial intelligence. The general scheme showing the joint operation of the modules is shown in Figure 6:

(1) Modules for factor forecasting, respectively through time series and neural networks;
(2) Error estimation module (mediator module) for the various methods and parameters;
(3) Consumption forecasting module.

4.2. Single-Point Forecasts for the Factors

The use of the prototype for forecasting electricity consumption in single-point (annual) factor forecasting showed different results for the effectiveness of different types of trends in time series (Figure 6).

When forecasting GDP, population, and average annual income, maximum efficiency was achieved with a linear trend of the time series. The forecasting of the energy intensity, the price of the electricity for the household, and the price of the electricity for the industry

achieved good results respectively in logarithmic, quadratic, and hyperbolic trends of the time series. Models for the most efficient neural networks, forecasting the same factors, are presented in Figure 7.
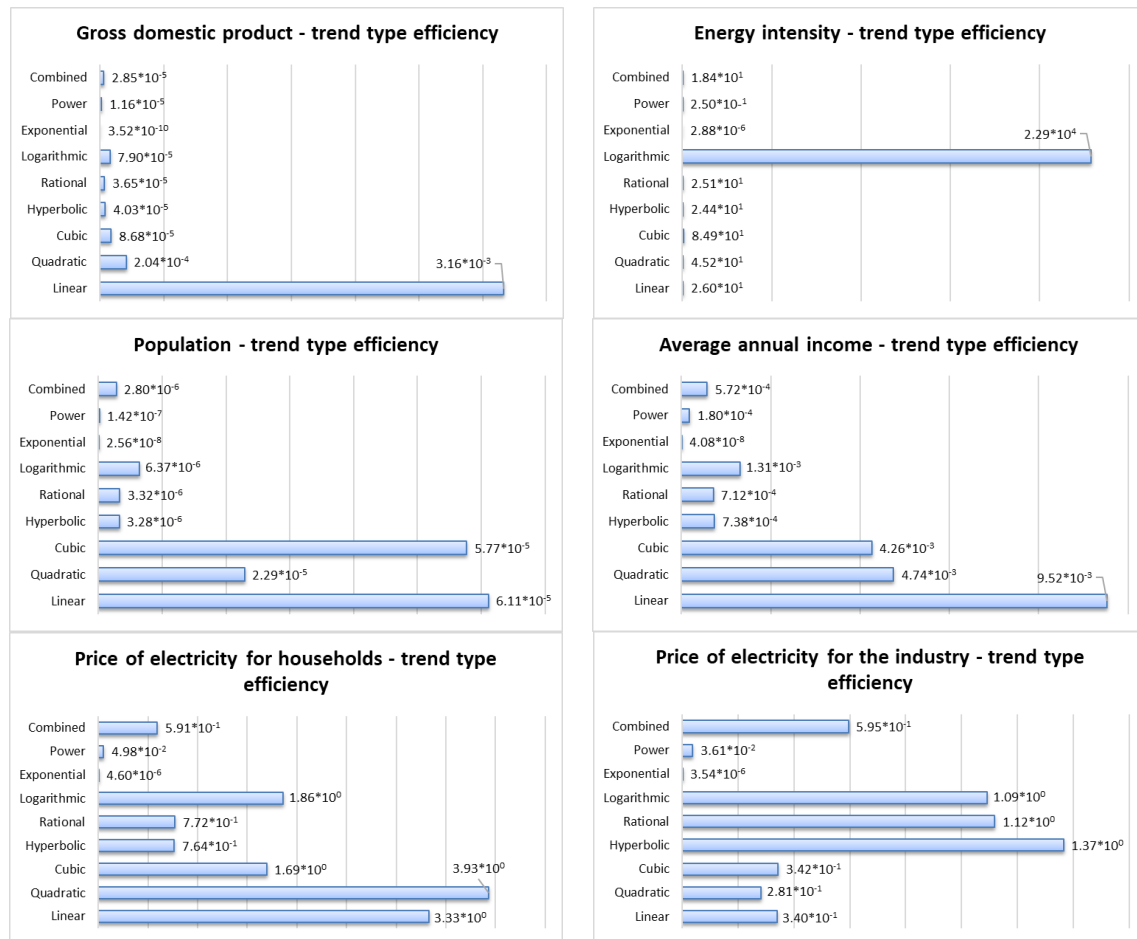

Figure 6. Comparison of efficiency in different types of time series trend.

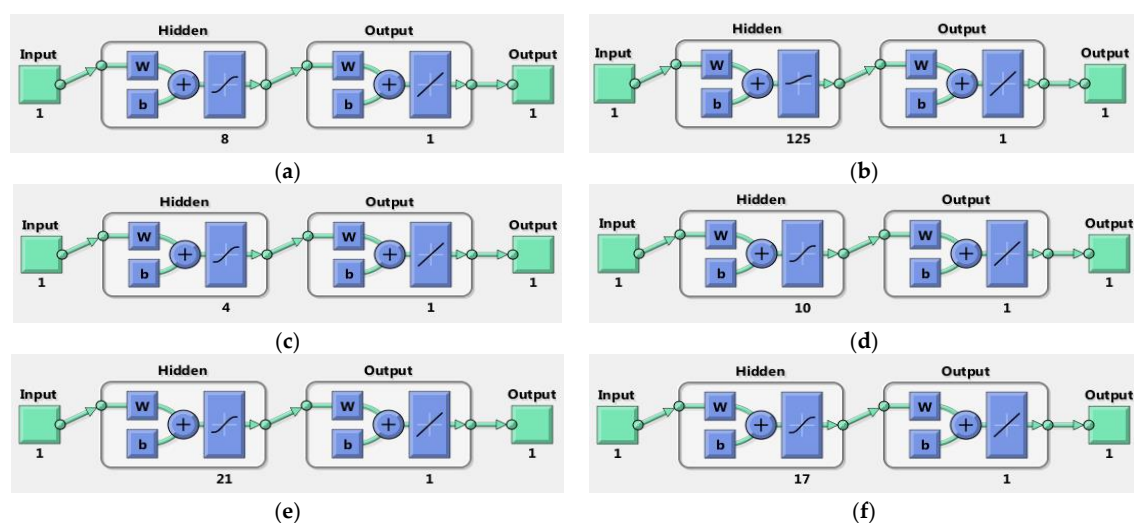**Figure 7.** Forecasting the behavior of the factors influencing the electricity consumption in NEES through separate artificial neural networks. (**a**) Gross domestic product, (**b**) energy intensity of the economy (Net_Int), (**c**) population, (**d**) average annual income, (**e**) price of electricity for households, (**f**) price of electricity for the industry.

During neural network training, all factors were considered as functions of time in order to enable the comparison of the results of forecasting in time series. The type of training that proved to be most effective for current tasks is the Lavenberg–Marquardt algorithm. The most appropriate activation function of neurons in the hidden layers for Energy intensity of the economy neural network is the logarithmic sigmoid function:

$$g(x) = \sigma(x) = \frac{1}{1 - e^{-x}},$$

and for every other it is the hyperbolic tangent:

$$g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

The activation function of the output neuron in all neural networks is linear:

$$g(x) = x.$$

The comparative analysis between the efficiency of the found neural networks in a one-year forecast and the type of trend in time series, leading to minimal error, is presented in Figure 8.
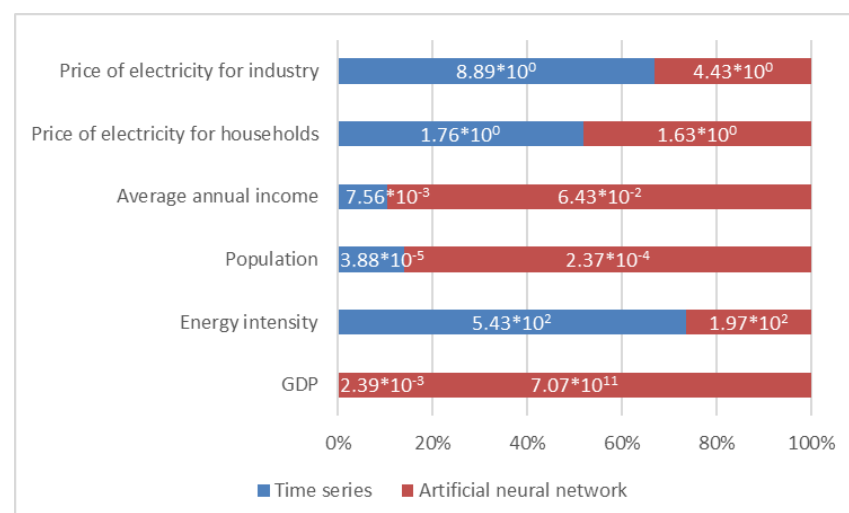

Figure 8. Comparison of efficiency between the most effective time series trend and the results of the forecasts of artificial neural networks.

The use of the presented prototype with time series and artificial neural networks showed a significant advantage in favor of neural networks in the case of forecasting GDP, population, and average annual income. In the other three cases, single-point forecast of energy intensity, price of electricity for households and its price for the economy sector, the module for comparison of errors showed some advantage in the efficiency of time series. In these cases, the type of trend was logarithmic, quadratic, and hyperbolic, respectively. Based on them, forecasts were made for the development of the values of the factors presented in Table 3.

Table 3. Single-point forecasts for the factors for the last year of the available dataset.

Factor	Predicted Value	The Most Effective Method	SMAPE (%)
GDP	102,954.73	Neural network	0.15
Energy intensity of the economy	0.42	Time series with logarithmic trend	3.65
Population	7,006,425.85	Neural network	0.65
Average annual income	6196.5	Neural network	1.57
Price of electricity for households	22.56	Time series with quadratic trend	4.98
Price of electricity for the industry	21.77	Time series with hyperbolic trend	4.8

4.3. Multifactor Single-Point Forecasts of Target Values

In the second part of the conducted experiments, the influence of the considered factors on the target values in the National Power System of Bulgaria was studied. Appropriate optimal neural structures were created for their forecast (Figure 9):

(1) Total final consumption in the National Power System—the neural network named Net_Nees, with 54 neurons in the hidden layer;
(2) Electricity consumption in industry, public sector and services—Net_Industry neural network with 35 neurons in the hidden layer;
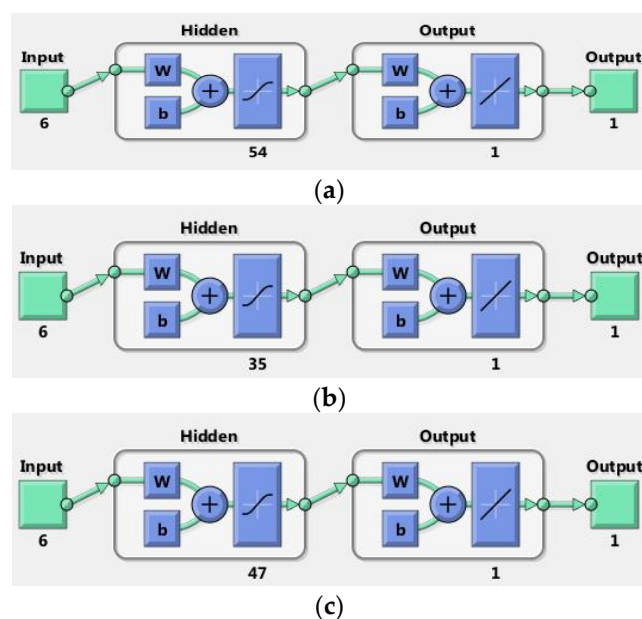(3) Electricity consumption on households—Net_Household neural network with 47 neurons in the hidden layer.

(a)

(b)

(c)

Figure 9. Neural networks for consumption forecasting. (**a**) Total final consumption in the NES (Net_Nees), (**b**) electricity consumption in industry, public sector, and services (Net_Industry), and (**c**) electricity consumption in households (Net_Households).

Activation functions of all neural networks found by the software system were hyperbolic tangent, and of the output neuron, linear.

The deviations of the forecast results from the actual data for single-point forecasts (for one-year consumption) were relatively small (Table 4) as the largest deviations in forecasting consumption in the industry are about 41 thousand tons of oil (toe) equivalent, i.e., about 1.525%. The deviation in forecasting consumption in the entire energy system was 0.0739%, and household consumption was 0.0302%.

Table 4. Forecasting of consumption with artificial neural networks, in thousand toe.

Neural Network	Actual Value	Predicted Value	Absolute Error	SMAPE (%)
Net_Nees	9737.9	9745.1	7.2	0.04
Net_Industry	2721.3	2679.8	41.48	0.77
Net_Households	2318.7	2318.0	0.7	0.02

Considering complex values as time series did not provide good results. The errors obtained with the best approximations of time series were many times worse than the corresponding neural networks (Table 5).

Table 5. The comparative table of errors in the use of these neural networks and time series.

Consumption	Method Type	Method	Absolute Error (Thousand Toe)	SMAPE (%)
Total final consumption in the National Power System	Time series	Cubic trend	199.33	1.01
	Neural network	Net_Nees	7.2	0.04
Electricity consumption in industry, public sector and services	Time series	Combined trend	111.16	2
	Neural network	Net_Industry	41.48	0.77
Electricity consumption in households	Time series	Linear trend	24.47	0.52
	Neural network	Net_House holds	0.7	0.02

The analysis of the weights of the neural networks can show us the influence of the individual factors on the predicted value of the target variable.

Let us denote by $w_{i,p}$ the weights by which the *i*-th factor is transmitted to the neuron *p* from the hidden layer of the already trained neural network, as $i, \, p \in N$. We choose the value as a criterion for the significance of factor *i*:

$$INPF_i = \left| \sum_{p=1}^{r} w_{i,p} \right|,$$

where *r* is the number of neurons in the hidden layer of the neural network.

The study showed that the largest impact on the total electricity consumption in the national network belongs to the gross domestic product of the country with $INPF_1 = 11.6$, and the least to the average annual income per capita with $INPF_4 = 1.06$. For consumption in the industry sector, things are similar. The most significant factor was GDP with $INPF_1 = 5.01$, and the one with the least importance was the price of electricity for the household—$INPF_6 = 2.3$. The most important factors in the energy consumption of households are population ($INPF_3 = 13.82$) and average annual income—$INPF_4 = 6.16$.

4.4. Multi-Step Forecasts

Using the created neural structures for factor forecasting, we created forecasts for a period of 7 years. Their accuracy can be determined over time (Table 6). Similarly, we created 7-year forecasts for the studied multifactorial values (Table 7). The forecast results showed that household electricity consumption, as well as total final consumption, will gradually increase, while electricity consumption in industry, the public sector, and services will decrease.

Table 6. Seven-year forecasting of factors.

Factor	2021	2022	2023	2024	2025	2026	2027
GDP (Million levs)	102,932	106,693	110,455	114,217	117,978	121,740	125,501
Energy intensity of the economy	0.400	0.395	0.390	0.385	0.380	0.376	0.371
Population	6,804,441	6,746,144	6,687,847	6,629,550	6,571,253	6,512,956	6,454,659
Average annual income	6671	6798	6942	7013	7105	7241	7339
Price of electricity for industry (in stotinki per kWh and without VAT)	26.79	27.62	28.46	29.29	30.12	30.96	31.79
Price of electricity for households (in stotinki per kWh and without VAT)	19.22	19.35	19.47	19.58	19.69	19.79	19.89

Table 7. Seven-year forecast for electricity consumption in the National Energy System, made through artificial neural networks.

Consumption	2021	2022	2023	2024	2025	2026	2027
Total final consumption in the National Power System	13,832	13,856	13,858	13,858	13,858	13,858	13,858
Electricity consumption in industry, public sector and services	3103	3101	3100	3099	3099	3098	3098
Electricity consumption in households	2800	2800	2800	2800	2800	2800	2800

5. Conclusions

Forecasting is a complex task. The availability of a wide variety of mathematical and statistical methods, and artificial intelligence methods, combined with the pursuit of the most accurate forecasting, usually requires a lot of time and effort. The use of software tools to automate some of the activities greatly simplifies the work.

The article proposed a model for multifactor forecasting which automatically selects the best method for forecasting significant factors and then uses the data predicted by them to create a complex multifactor forecast. The developed model was successfully experimented with to make multiple forecasts for the energy consumption of households, industry, and total consumption—for a one-year and a seven-year period. By automating the forecasting process in an indicative way, we made it easier to make predictions with fewer errors than previously.

The presented model has wide applications in various subject areas. It can be used for air quality forecasting, demographic forecasting, forecasting in industry, etc.

In the future, the system can be expanded in several directions. An important part of its development is the integration of additional forecasting methods. The development of a module to evaluate the correlation between the individual factors and the target variable would also help to optimize the forecasting process. Automated generation of graphics and documentation for each step in the overall forecasting process would be beneficial to the end user of the system.

Author Contributions: Conceptualization, K.Y., E.H. and S.H.; data curation, K.Y. and S.C.; formal analysis, K.Y., E.H. and S.H.; funding acquisition, S.H. and S.C.; investigation, K.Y., E.H. and S.C.; methodology, K.Y., E.H. and S.H.; project administration, E.H.; software, K.Y. and S.C.; supervision, E.H.; validation, K.Y., E.H. and S.H.; visualization, K.Y. and E.H.; writing—original draft, K.Y., E.H. and S.H.; writing—review and editing, S.H. and S.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available through the public databases mentioned in the text.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. ReferencesHyndman, R.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Melbourne, Australia, 2021.
2. Slavova, I. *Forecasting and Planning*; New Bulgarian University: Sofia, Bulgaria, 2012.
3. Skiadas, C.H.; Bozeman, J. *Data Analysis and Applications 1: Clustering and Regression, Modeling-Estimating, Forecasting and Data Mining*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2019; Volume 2.
4. Brockwell, P.; Davis, R. *Introduction to Time Series and Forecasting*; Springer: Berlin/Heidelberg, Germany, 2016; ISBN 9783319298528.
5. Agung, I.G. *Advanced Time Series Data Analysis: Forecasting Using EViews*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2019; ISBN 978-1-119-50471-9.
6. Ostertagová, E.; Ostertag, O. Forecasting Using Simple Exponential Smoothing Method. *Acta Electrotech. Inform.* **2012**, *12*, 64–66. [CrossRef]
7. Chen, T.; Honda, K. *Fuzzy Collaborative Forecasting and Clustering: Methodology, System Architecture, and Applications*; Springer: Berlin/Heidelberg, Germany, 2019.
8. Lazzeri, F. *Machine Learning for Time Series Forecasting with Python*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2020; ISBN 1119682363.
9. Gridin, I. *Time Series Forecasting using Deep Learning: Combining PyTorch, RNN, TCN, and Deep Neural Network Models to Provide Production-Ready Prediction Solutions*; BPB Publications: Noida, India, 2021; ISBN 9391392571.
10. BAS Strategy. *Preparation of the National Strategy in the Field of Energy*; BAS: Sofia, Bulgaria, 2017.
11. Risk Management Lab. Forecast of the Electricity Balance of the Republic of Bulgaria 2025. 2014. Available online: https://www.bia-bg.com/uploads/files/events/Energy_balans_BG2025.pdf (accessed on 15 July 2022).
12. Mohamed, Z.; Bodger, P. Forecasting electricity consumption in New Zealand using economic and demographic variables. *Energy* **2005**, *30*, 1833–1843.
13. Lee, Y.; Gaik, T.; Yee, C.H. Forecasting Electricity Consumption Using Time Series Model. *Int. J. Eng. Technol.* **2018**, *7*, 218–223.
14. Sun, T.; Zhang, T.; Teng, Y.; Chen, Z.; Fang, J. Monthly Electricity Consumption Forecasting Method Based onX12 and STL Decomposition Model in an Integrated Energy System. *Math. Probl. Eng.* **2019**, *2019*, 9012543.
15. Chou, J.S.; Tran, D.S. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. *Energy* **2018**, *165*, 709–726. [CrossRef]
16. Ding, S.; Hipel, K.; Dang, Y. Forecasting China's electricity consumption using a new grey prediction model. *Energy* **2018**, *149*, 314–328. [CrossRef]
17. Lee, Y.S.; Tong, L.T. Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Convers. Manag.* **2011**, *52*, 147–152. [CrossRef]
18. Huang, Y.; Wang, C.; Dang, H.; Lai, S. Evaluating performance of the DGM(2,1) model and its modified models. *Appl. Sci.* **2016**, *6*, 73. [CrossRef]
19. Chung, Y. Electricity consumption prediction using a neural-network-based grey forecasting approach. *J. Oper. Res. Soc.* **2017**, *68*, 1259–1264.
20. Chernykh, I.; Chechushkov, D.; Panikovskaya, T. Prediction of electrical loads in operational control of power systems based on neural network structures. *WIT Trans. Ecol. Environ.* **2014**, *190*, 109–117.
21. Khosravani, H.; Castilla, M.; Berenguel, M.; Ruano, A.; Ferreira, P. A Comparison of Energy Consumption Prediction Models Based on Neural Networks of a Bioclimatic Building. *Energies* **2016**, *9*, 57.
22. Yoo, S.; Myriam, H.A. Predicting residential electricity consumption using neural networks: A case study. *J. Phys. Conf. Ser.* **2018**, *1072*, 012005. [CrossRef]
23. Jahn, M. *Artificial Neural Network Regression Models: Predicting GDP Growth*; HWWI Research Paper, No. 185; Hamburg Institute of International Economics (HWWI): Hamburg, Germany, 2018.
24. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies* **2018**, *11*, 1636. [CrossRef]
25. Alamaniotis, M.; Bargiotas, D.; Tsoukalas, L.H. Towards smart energy systems: Application of kernel machine regression for medium term electricity load forecasting. *Springer Plus* **2016**, *5*, 58. [CrossRef]
26. Alamaniotis, M. Synergism of Deep Neural Network and ELM for Smart Very-Short-Term Load Forecasting. In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Novi Sad, Serbia, 10–12 October 2019; pp. 1–5. [CrossRef]
27. Chou, J.S.; Truong, D.N. Multistep energy consumption forecasting by metaheuristic optimization of time-series analysis and machine learning. *Energy Res.* **2021**, *45*, 4581–4612. [CrossRef]
28. Kong, W.; Dong, Z.Y.; Jia, Y.; Hill, D.J.; Xu, Y.; Zhang, Y. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Trans. Smart Grid* **2019**, *10*, 841–851. [CrossRef]

29. Moradzadeh, A.; Moayyed, H.; Zakeri, S.; Mohammadi-Ivatloo, B.; Aguiar, A.P. Deep Learning-Assisted Short-Term Load Forecasting for Sustainable Management of Energy in Microgrid. *Inventions* **2021**, *6*, 15. [CrossRef]
30. Khwaja, A.S.; Anpalagan, A.; Naeem, M.; Venkatesh, B. Joint bagged-boosted artificial neural networks: Using ensemble machine learning to improve short-term electricity load forecasting. *Electr. Power Syst. Res.* **2020**, *179*, 106080. [CrossRef]
31. Khwaja, A.S.; Zhang, X.; Anpalagan, A.; Venkatesh, B. Boosted neural networks for improved short-term electric load forecasting. *Electr. Power Syst. Res.* **2017**, *143*, 431–437. [CrossRef]
32. Cao, Z.; Wan, C.; Zhang, Z.; Li, F.; Song, Y. Hybrid Ensemble Deep Learning for Deterministic and Probabilistic Low-Voltage Load Forecasting. *IEEE Trans. Power Syst.* **2020**, *35*, 1881–1897. [CrossRef]
33. Karatasou, S.; Santamouris, M.; Geros, V. Modeling and predicting building's energy use with artificial neural networks: Methods and results. *Energy Build.* **2008**, *38*, 949–958. [CrossRef]
34. Cervená, M.; Schneider, M. Short-term forecasting of GDP with a DSGE model augmented by monthly indicators. *Int. J. Forecast.* **2014**, *30*, 498–516. [CrossRef]
35. Farsi, B.; Amayri, M.; Bouguila, N.; Eicker, U. On Short-Term Load Forecasting Using Machine Learning Techniques and a Novel Parallel Deep LSTM-CNN Approach. *IEEE Access* **2021**, *9*, 31191–31212. [CrossRef]
36. Chung, W.; Gu, Y.H.; Yoo, S.J. District heater load forecasting based on machine learning and parallel CNN-LSTM attention. *Energy* **2022**, *246*, 123350. [CrossRef]
37. Sajjad, M.; Khan, Z.A.; Ullah, A.; Hussain, T.; Ullah, W.; Lee, M.Y.; Baik, S.W. A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting. *IEEE Access* **2020**, *8*, 143759–143768. [CrossRef]
38. Alamo, D.H.; Medina, R.N.; Ruano, S.D.; García, S.S.; Moustris, K.P.; Kavadias, K.K.; Zafirakis, D.; Tzanes, G.; Zafeiraki, E.; Spyropoulos, G.; et al. An Advanced Forecasting System for the Optimum Energy Management of Island Microgrids. *Energy Procedia* **2019**, *159*, 111–116. [CrossRef]
39. Botchkarev, A. A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. *Interdiscip. J. Inf. Knowl. Manag.* **2019**, *14*, 45–76. [CrossRef]
40. Sheela, K.; Deepa, S.N. Review on Methods to Fix Number of Hidden Neurons in Neural Networks. *Math. Probl. Eng.* **2013**, *2013*, 425740. [CrossRef]
41. Lörke, A.; Schneider, F.; Heck, J.; Nitter, P. Cybenko's Theorem and the Capability of a Neural Network as Function Approximator. 2019. Available online: https://www.mathematik.uni-wuerzburg.de/fileadmin/10040900/2019/Seminar__Artificial_Neural_Network__24_9__.pdf (accessed on 15 July 2022).
42. National Statistical Institute. Available online: https://www.nsi.bg/ (accessed on 15 July 2022).
43. INFOSTAT Information System. Available online: https://infostat.nsi.bg/infostat/pages/external/login.jsf (accessed on 15 July 2022).
44. Electricity System Operator. Available online: http://www.eso.bg/ (accessed on 15 July 2022).
45. The World Bank Group. Available online: https://www.worldbank.org (accessed on 15 July 2022).
46. Eurostat. Your Key to European Statistic. Available online: https://ec.europa.eu/eurostat (accessed on 15 July 2022).