energies

MDPI

Article

# Forecasting Electricity Demand in Turkey Using Optimization and Machine Learning Algorithms

**Mustafa Saglam [1,*], Catalina Spataru [1] and Omer Ali Karaman [2]**

[1] Energy Institute, Bartlett School Environment, Energy and Resources, University College London, London WC1E 6BT, UK; c.spataru@ucl.ac.uk

[2] Department of Electronic and Automation, Vocational School, Batman University, Batman 72100, Turkey; omerali.karaman@batman.edu.tr

* Correspondence: mustafa.saglam@ucl.ac.uk

**Abstract:** Medium Neural Networks (MNN), Whale Optimization Algorithm (WAO), and Support Vector Machine (SVM) methods are frequently used in the literature for estimating electricity demand. The objective of this study was to make an estimation of the electricity demand for Turkey's mainland with the use of mixed methods of MNN, WAO, and SVM. Imports, exports, gross domestic product (GDP), and population data are used based on input data from 1980 to 2019 for mainland Turkey, and the electricity demands up to 2040 are forecasted as an output value. The performance of methods was analyzed using statistical error metrics Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R-squared, and Mean Square Error (MSE). The correlation matrix was utilized to demonstrate the relationship between the actual data and calculated values and the relationship between dependent and independent variables. The *p*-value and confidence interval analysis of statistical methods was performed to determine which method was more effective. It was observed that the minimum RMSE, MSE, and MAE statistical errors are $5.325 \times 10^{-14}$, $28.35 \times 10^{-28}$, and $2.5 \times 10^{-14}$, respectively. The MNN methods showed the strongest correlation between electricity demand forecasting and real data among all the applications tested.

**Keywords:** medium neural networks; whale optimization algorithm; support vector machine; electricity demand forecast; machine learning; error metrics; multi regression equations; Turkey

check for updates

**Citation:** Saglam, M.; Spataru, C.; Karaman, O.A. Forecasting Electricity Demand in Turkey Using Optimization and Machine Learning Algorithms. *Energies* **2023**, *16*, 4499. https://doi.org/10.3390/en16114499

Academic Editor: Valentina E. Balas

Received: 24 April 2023
Revised: 30 May 2023
Accepted: 31 May 2023
Published: 2 June 2023



## 1. Introduction

Governments, researchers, and corporations are taking a serious interest in electricity consumption, production, and supply due to their crucial role in livelihoods and global economic development. Conventionally, electricity is generated using primary energy sources such as fossil fuels, nuclear, and renewable energy. Along with an increasing population, urbanization, the growth of industry, and technological advancement, electricity demand is also rising [1]. A power supply units' excessive number is activated when load estimates are higher than electricity demands, causing an excessive amount of electricity to be used and giving extra reserves. On the other hand, lower load projections can force the system to operate within limits, leading to insufficient supply [2]. Nevertheless, load and demand projections serve as the foundation for several energy market choices, enabling electricity system markets to be planned and administered in an effective, transparent, and dependable manner and to meet the sector's needs [3]. The demand for energy is impacted by various factors. Import, export, and GDP are significant impacts on estimating energy demand. Population not only affects the overall energy demand but also has an influence on the way energy is utilized and the occupancy per person [4]. According to Wang et al.'s [5] threshold regression model, as economies mature, the size of the working population has a weaker influence on energy demand. Electricity demand is more affected by urbanization than primary energy. Demographic and land urbanization's effects on energy demand were estimated by Yu et al. [6], while Liu et al. [7] analyzed the long-term monthly electricity energy demand forecast by examining the relationships among climate, socioeconomics (GDP, populations) and electricity consumption variables. Mutschler et al. [8] also explored the complex interplay between population growth, technology adoption, climate change, and energy demand to plan future energy systems policies.

Over the past 20 years, there has been a wealth of specialized literature about demand and load forecasting. There are different types of energy demand models, including static or dynamic, univariate or multivariate, and techniques that use time series or hybrid models. This literature seeks to categorize and analyze studies applicable and relevant to Turkey. The goal is to identify which methods have been considered, the input variables selection, and the arrangement of the parameter values [9].

This systematic review considered the following factors:

- To evaluate the effectiveness of different methods, an analysis of Key Performance Indicators (KPIs) is used to assess prediction accuracy. In this context, commonly used measures such as MAE in the literature result in the omission of crucial quality factors such as the highest forecast error and the distribution of error. By avoiding the mutual counteraction negative and positive errors in the prediction, RMSE, and MAE evaluate, respectively, how closely the anticipated value difference resembles the true value. While MAPE emphasizes the accuracy of the forecasting methodologies, MSE illustrates the difference between the actual data and the anticipated value. When different data sets are utilized, MAPE aids in examining how well the estimating methods function.
- The model hyper-parameters fine-tuning, data pre-processing methods, the validation and training data set selection, and the outcomes graphical display.
- The findings and precision verification of the large dataset collected for the mainland.

This research covers a mix of methods used to estimate electricity demand, including artificial neural networks (ANN), auto-regression, evolutionary algorithms, linear multi-variable regression, metaheuristic algorithms, and fuzzy logic techniques, which can be applicable to Turkey. Table 1 provides an overview of the research conducted on predicting electricity demand.

**Table 1.** Electricity Energy Consumption and Demand Forecasting Research from the Literature.

| Method | Forecasting for | Variables | Author |
|---|---|---|---|
| ANN, Gaussian regression, k-nearest neighbors, LR, random forest, and SVM | electricity supply and demand | system hourly demand, renewable generation sources | Cebekhulu et al. [10] |
| MLR, ANN, and PSO | island electricity demand | import, car numbers, passenger (tourist) numbers, export | Saglam et al. [11] |
| ANN- Generic Algorithm for power grid management | daily energy consumption | day data | Baba [12] |
| RT, GBT, RF, ANN, LSTM, and SVR | solar and wind energy oversupply in power system | biomass/geothermal units, output power of thermal power plants, load demands, power imports, nuclear units, wind turbines, solar farms, large hydro units, and WSPC | Shams et al. [13] |
| Bilateral long short-term memory (BILSTM), CNN, GWO, Time Series Prediction | Short-term electricity demand forecast | Buildings' electricity consumption times series data | Sekhar and Dahiya [14] |

**Table 1.** *Cont.*

| Method | Forecasting for | Variables | Author |
| --- | --- | --- | --- |
| SVM and ANN | electricity consumption | The population, inflation rate, GDP and unemployment rate | Sen et al. [15] |
| deep learning, SVM, and ANN | transportation energy demand | year, population, GDP, vehicle kilometer | Agbulut [16] |
| grey prediction model and SVM method | seasonal electricity generation | Eurostat database | Sahin et al. [17] |
| RNN | energy demand | past energy usage values | Tun et al. [18] |
| RNN, ANN, and adaptive network-based fuzzy inference system | electricity demand | historical electricity data | Ramsami and King [19] |
| SVR and PSO-ARIMA-ANN | long term electricity demand and peak load | energy and load data | Kazemzadeh et al. [20] |
| ANN, CNN, and compare with traditional ANN-ARIMA | energy demand | hours, week of the year, holidays, day of the week | Real et al. [21] |
| ANN, SVR, and RNN | electricity demand | electricity consumption dataset | Bedi and Toshniwal [22] |
| MLP optimization and ANN | energy demand for India, ustralia, China, the USA and France | Financial development, energy price, industrialization, FDI, economic growth, population, urbanization, | Bannor and Acheampong [23] |
| ANN and RNN | electrical energy demand | population, GDP, temperature, energy consumption | Abdulsalam and Babatundea [24] |

Kazemzadeh et al. [20] developed a support vector regression (SVR)-based prediction algorithm approach. The input samples dimension and the SVR technique parameters were both optimized using the particle swarm optimization (PSO) method. For yearly peak load and long-term total electrical energy demand, a hybrid forecasting method was considered to reduce forecasting error. The combination of ANN, autoregressive integrated moving average (ARIMA), and suggested SVR methods served as the foundation for the proposed hybrid method, which was used to estimate total electricity demand and annual peak load. It has been observed that the proposed hybrid methods and PSO-SVR give more precise results than ANN and ARIMA methods. The PSO-SVR method can estimate electricity and load demand with small errors without any sensitivity to seasonal patterns in the initial time series. On the other hand, the artificial bee colony (ABC) algorithm was utilized by Hao et al. [25] to demonstrate and use a unique ensemble forecasting model for predicting electricity consumption. As independent input variables, several historical time variables, including consumer price index, GDP, industrial structure, urbanization rate, technical innovation, and population, were employed. To train the model, China's primary electricity demand data was used, and it was determined that the suggested ensemble model delivers more precise results in forecasting hypothesis and accuracy testing than both the benchmark forecasting model and the basic mean ensemble estimation model.

Real et al. [21] have investigated deep learning application methods to predict energy demand. The authors suggest combining an ANN with a convolutional neural network (CNN) hybrid architecture. French energy demand forecast was trained and given context using Action de Recherche Petite Echelle Grande Echelle (ARPEGE), estimating meteorological data. The outcome demonstrates that this method outperforms the standard subscription-based service provided by Réseau de Transport d'Electricité (RTE, a French transmission system operator). Additionally, the proposed approach performs best when results are contrasted with other options, such as ARIMA and conventional ANN models. Using a deep learning-based system, Bedi and Toshniwal [22] calculated the demand for power by taking into account long-term historical dependencies. Based on all months' worth of electricity usage data, cluster analysis is used to create data segments based on the season. To classify load trends, it was necessary to have a thorough understanding of the metadata that belonged to each cluster. It was determined that the suggested method

(D-FED) outperforms ANN, SVM, and recurrent neural network (RNN) regression models and can be implemented to estimate electricity demand efficiently.

Using a hybrid model built on the ARIMA and least-square support vector machine (LSSVM), Kaytez [26] shows how to estimate Turkey's electricity consumption. Results from this suggested approach were evaluated against official prediction data, a single ARIMA, the body of literature, and a multiple linear regression approach. The results demonstrate that the model responds better to some unexpected reactions in the time series. Di Leo et al. [27] applied regression analysis (RA) to forecast trends in energy consumption in end-use industries. The suggested method was used to describe the positions of the long-term electricity demand by statistically characterizing the relationships between independent factors such as GDP, population, transportation, business, and residential energy demands. Traditional statistical tests have been used to assess and validate the non-linear and linear regression models' effectiveness for energy demand forecasting. The outputs demonstrated a close and logical correlation between transport and residential electricity demand with GDP and population, while the commercial energy demand is correlated with GDP.

To estimate the Rodrigues and Mauritius Islands' peak monthly electricity demand, Ramsami and King [19] used three techniques: the adaptive network-based fuzzy inference system (ANFIS), data handling group method, and ANN (recurrent and feedforward). Nine error measures were used in conjunction with the suggested models. The optimal value for seven error metrics was created using the proposed model with grid segmentation. As a result, it proved to be better suited for determining the peak electrical demand. For calculating peak demand, an ANFIS model with grid segmentation is more effective. Angelopoulos et al. [28] published the long-term predictions for electricity demand in Greece and also made use of the connection between time series and effective multiple criteria. Greece's value estimate model is examined using training-related data collected from 1999 to 2013. When determining the annual total net electricity consumption, the suggested approach was used for an interconnected power system in Greece during the following test period between 2014–2016. The results of the suggested model reveal that, after increasing the efficiency of electrical energy, taking into account the country's general weather conditions and economic growth as measured by the national GDP, those factors have the most impact on electricity demand. The regression models exhibit superior performance compared to the least squares MLR model regarding predictive dependability, with the minimum MAPE outcome being 0.74%.

The impact of European nations' power generation during the lockdown period was examined by Sahin et al. [17], who also reconstructed energy generation in these nations. The total monthly electricity generation from renewable and non-renewable sources in the UK, France, Germany, Turkey, and Spain was examined and compared from January 2017 to September 2020. Machine learning (ML) techniques and grey prediction (GP) models for seasonal periods were employed to predict future trends. Hou et al. [29] investigated the electricity output effects, GHG emissions, and consumption on climate change. They forecast electricity demand using an optimized ANN. The ANN approach was enhanced using the Improved Pathfinder algorithm. For estimating the demand for electricity, a more sensitive model with lower error numbers was produced by the ANN method's optimization technique.

Baba [12] compared and evaluated the effectiveness of three different strategies to forecast the daily energy consumption of the nearby industrial area. A probabilistic approach called the Multiple Model Particle Filter (MMPL) method was suggested. Then, ANNs with one and two hidden layers were developed and examined. Pegalajar et al. [30] used data from the Spanish Electricity Network between 2007–2019 to predict electricity demand. RNN, multilayer perceptron, linear regression, regression trees (RT), gradient boosting regression (GBR), and random forests (RF) were used among the six estimate models. These experiments show positive results in all scenarios, with a worst-case improvement of 12% and a best-case improvement of 37% over predictions made by the Spanish Electric Grid.

By utilizing load profiles, Bendaoud et al. [31] explored a load forecasting approach (LPs). The hourly temperature profiles used to incorporate seasonal data fluctuations and demand changes have been used to analyze Algeria's power consumption. For LP-based forecasting, daily, weekly, and annual LP propagation were employed. Mid–short-term load forecasting models have been developed using a variety of artificial intelligence (AI) techniques.

By utilizing ensemble ML techniques, Porteiro et al. [32] have produced a model for anticipating the electricity demand for residential and commercial buildings. In this research, computational intelligence models were created to estimate the demand for power one day in advance. An ensemble technique was also used to develop a day-ahead forecasting model. Standardization, addressing missing values, and outlier removal were the three processes in the pre-processing of the data. The real data sets for Burgos Industrial Park, Uruguay's total energy demand, and Montevideo's distribution substation electricity demand have been chosen for examination. The proposed models were assessed using common performance indicators. The primary findings demonstrate that the best day-ahead proposed model has a MAPE of 5.17% on total consumption data, 9.09% on substation value, and finally, 2.55% on industrial data based on Extra Trees Regressor. Gokceada Island's electricity demand was estimated using ANN, PSO, and MLR methods. Input values for car numbers, exports, imports, and passenger numbers related to tourism were based on the period from 2014 to 2019 [11]. The results obtained from these methods were analyzed using statistical errors such as RMSE, $R^2$, MAE, and MSE. The methods were analyzed by examining their confidence intervals. The correlation matrixes are utilized to indicate the association between the method outputs and actual values, as well as the relationship between dependent and independent parameters. Input parameters are separated into subsets, multi-regression equations related to these parameters, and $p$-value performances and $R^2$ were demonstrated. The results showed that the ANN method had the widest confidence interval of 95% among the techniques used, and the statistical error metrics had the strongest correlation with the actual data and electricity demand output for the ANN method [11].

A new approach to power load forecasting is introduced by Lu et al. [33], which involves utilizing an SVR model in combination with WOA that incorporates elite and chaotic opposition-based learning (ECWOA) to enhance forecasting results. The results of experiments indicate that incorporating electricity price information results in higher forecasting accuracy. To enhance the accuracy of $CO_2$ emissions and energy demand predictions in the transportation industry, Javanmard et al. [34] implement a mixed method that combines a mathematical model with multiple objectives with ML algorithms that use data-driven approaches. An estimation of the energy consumption demand in China is conducted by Rao et al. [35]. Firstly, a two-stage model based on least absolute shrinkage and selection operator-random forest (Lasso-RF) is introduced to determine the factors affecting energy demand. Secondly, the SVR-compositional data second exponential smoothing (SVR-CDSES) model is developed to predict the demand for primary electricity, oil, natural gas, and coal. The results indicate that primary electricity will experience significant growth at an annual rate of 8.05%, reflecting a growing focus on clean energy.

Li et al. [36] propose a hybrid approach that utilizes Manta ray foraging optimization to optimize the parameters of SVM for short-term load forecasting. To assess the precision of this approach, five other optimizers, the Satin Bowerbird optimizer, Tug of War optimization, Fruit-fly optimization, Moth Flame optimization, and Slime Mould algorithm, are utilized to compare the proposed method's superiority. Huang et al. [37] developed a transformer-based model for estimating energy consumption in an actual university library and compared it to a baseline model, SVR. To account for the various factors that can influence the development and computation of building electricity energy models, advanced ML models driven by the inherent electricity consumption patterns are proposed by Huang et al. [38]. In this study, the performance of three ML algorithms, LSTM, SVR, and XGBoost, are evaluated using one-year datasets with sub-hourly temporal granularity to determine the most accurate predictor. Additionally, the performance and robustness of the ML model, assessed by the coefficient of variation (CV), are compared across XGBoost and LSTM trained with the same datasets that include data size, temporal granularity, and building type attributes.

The literature contains various case studies on electricity demand forecasting. Velasquez et al. [39] utilized regression with seasonality to predict the long-term electricity demand in Brazil. Pallonetto et al. [40] compared the performance of SVM and LSTM to forecast commercial buildings' hourly electricity demand in Ireland and found that LSTM outperformed SVM. May et al. [41] utilized ANN and ANFIS to forecast Mexico's electricity market hourly electricity demand and observed that ANN outperformed ANFIS. Meanwhile, Niu et al. [42] suggested a hybrid model to forecast the aggregated four-grid electricity load in China for quarter-hour periods. Their approach involved population-based metaheuristic algorithms and integrating a signal decomposition tool with an ML model. Luzia et al. [43] conducted a study to assess the suitable application of ANN, ARIMA combined with Wavelet Transform, or Fourier Transform for predicting electricity demand at various time horizons and frequencies. The findings indicate that when considering both time frequencies, ANN proves to be a superior method for short-term predictions. Işık et al. [44] employed DL techniques to predict the electricity demands of select Fortune 500 companies in Turkey. The study focused on LSTM and MLP techniques, which have demonstrated effectiveness in previous research. Additionally, for the first time in electricity demand forecasting, the Multiple Seasonal-Trend Decomposition using Loess (MSTL) technique was utilized, and the results of MSTL outperformed better.

Albuquerque et al. [45] employed regularized Lasso Lars and RF models to predict Brazil's electricity consumption. Energy forecasting model based on a DL approach demonstrated by Rick and Berton [46] that combined CNN, LSTM, and auto-encoder (AE) for time series with unequal lengths. Maaouane et al. [47] utilized ANN modelling to estimate and predict energy demand in the transport sector of developing countries. Recently, Chaturvedi et al. [48] conducted a comparison of several time-series models, such as SARIMA, and LSTMRNN models, for the purpose of predicting India's peak and total monthly energy demand. Unlike the ANN model, the SVR model is designed to avoid local overfitting and minimization. In this study, the electricity demand of Turkey's mainland was projected through the year 2040 using MNN, SVM, and WOA. Both error measures (RMSE, $R^2$, MSE, and MAE) and confidence interval and $p$-value analysis statistical techniques were used to compare the prediction performances of the different methods. The correlation matrix was employed to demonstrate the association between the observed value and method-predicted values, as well as the relationship between the independent parameters (import, GDP, population, export) and the dependent data (electricity consumption). The outputs of correlation matrices have revealed which variables influence the result and by how much. Subsets of multiple regression equations for the input variables (import, export, population, and GDP) were developed. The parameters affecting the output with R-squared and $p$-value performances were provided and compared in the resulting equations. A statistical procedure known as the confidence interval analysis of the methodologies was also carried out. Overall, the electrical energy demand forecasting performance of MNN and SVM among ML methods and WOA chosen as optimization methods were compared using error metrics (RMSE, MSE, MAE), correlation matrices, and multi-regression equations. In addition, $p$-value and confidence interval analysis of statistical methods was performed to determine which method was more effective. The contributions of this study can be summarized in the following five points:

1. In this study, multi-objective forecasting models were created using various traditional ML methods and a new optimization method, WOA, to improve forecasting accuracy. In the Turkey case study, forecast performances were verified with error metrics by using inter-year data in electrical energy demand forecasting. The predicted results provided reliable and informative references for annual energy demand for the coming decades.
2. The effect of independent inputs used for electrical energy demand forecasting on forecast output has been investigated with MLR subsets and different combinations.
3. Statistical performance error metrics are included to effectively improve forecast accuracy and demonstrate the effectiveness of the method used.
4. It includes the technical analysis of determining the optimal parameters of methods by means of input-output correlation matrices. Thus, it is determined how much the independent variables affect the dependent variable.
5. The effective electricity demand estimation made in this study prevents extra reserves and limited operation of the system.

The rest of this paper is organized as follows. Section 2 demonstrates the data sources, pre-processing, and exploration. MNN, SVM, WOA, and Error Metrics are given in Section 3 under Materials and Methods. The Analysis and Results part under Section 4 presents electricity demand forecasting results, error metrics, multi-regression equations, and correlation matrix by using the proposed methods. The results discussion, novelty of this study, limitations, and future works are given in the Discussion section.

## 2. Exploration, Pre-Processing, and Data Sources

The study was carried out on the mainland's electrification system to analyze which socio-economic variables could affect electricity demand growth. First, yearly electricity demand data for the period from 1980 to 2019 was collected from the Turkish Electricity Transmission Corporation to understand the historical trend [49]. The following phase of the data investigation aimed to determine the variables that had the most significant influence on predicting electricity demand. The model was constructed using five input variables that encompass socio-economic indicators and electricity consumption measured in gigawatt-hours (GWh). Population data is collected from the Turkish Statistical Institute (TSI) [50]. Import, export, and GDP data are obtained from the World Bank Open Database. Import, population, export, and GDP data are used as inputs, while electricity demand is selected as an output.

To develop and train the models, data pre-processing was a crucial initial step. Initially, all the variables were consolidated and organized into a single Excel file in a suitable format. This organized file was subsequently imported into MATLAB R2021b version. Afterward, the complete data set was partitioned into a test set (20%), training set (70%), and validation set (10%) while maintaining the data in chronological order. Specifically, data from 1980 to 2007 were assigned as the training set, data between 2008–2015 were designated as the test set, and values from 2016 to 2019 were assigned as the validation set on a yearly basis. GDP is an economic indicator. The fact that this value is developing, stagnant, or receding also affects electricity estimation. A developing country will industrialize, and as it industrializes, both income and electricity needs will increase. New constructions mean new consumption points. In other words, as the population increases, energy consumption is also likely to increase [51]. Previous studies in the literature have shown that exports and imports typically have positive effects on electricity consumption [52].

## 3. Materials and Methods

The historical energy consumption data was utilized for training, testing, and validating forecasting methods such as MNN, SVM, and WOA, as shown in Figure 1 of the flowchart of the method. By incorporating demographic input data from the past, the electricity demand was predicted until the year 2040.

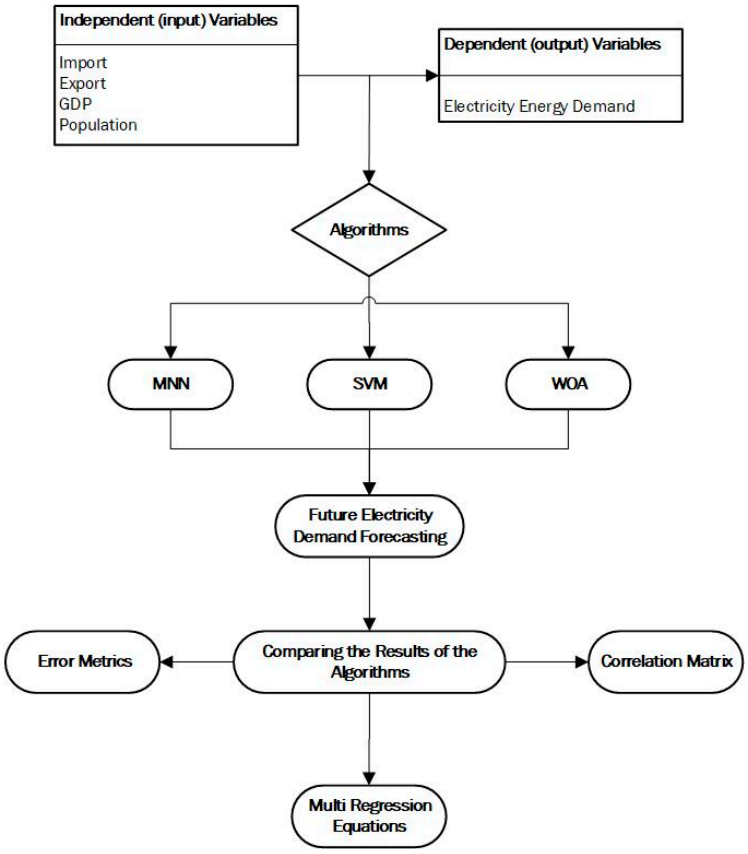


**Figure 1.** ML modelling framework proposed to estimate electricity demand.

According to statistical analysis results, the best-forecasted method output will be used for the next step as future electricity demand data. To compare the performance of the methods, error metrics such as MAE, $R^2$, MSE, and RMSE were used along with statistical methods such as confidence interval and *p*-value analysis.

The correlation matrix was employed to demonstrate the correlation between the real data and the predicted data derived from the methods, as well as the relationship between the independent (GDP, import, population, and export) and the dependent variable (electricity consumption) for Turkey. The correlation matrices have demonstrated the extent to which variables affect the magnitude of the output. The input parameters, GDP, population, export, and import, were split into subsets, with multiple regression equations calculated for each parameter. The equations that were generated demonstrated the extent to which each parameter impacted the output, and their $R^2$ and *p*-value performances were compared and presented. Furthermore, a statistical method known as confidence interval analysis was employed to evaluate the reliability of the methods used.

There are three distinct scenarios (low, base, and high) for each method. These scenarios are created by considering the relative differences among the input data's historical values employed in the proposed methods. Thus, diverse scenario ratios and input values are utilized for the mainland. For the population, the official scenarios of the Turkish Statistical Institute were considered. In addition, for the future estimation of economic data, percentages were designed by considering previous studies in the literature, International Monetary Fund (IMF), and World Bank scenarios [53–55]. Table 2 demonstrates the mainland's assumption and scenarios to estimate future electricity demand.

**Table 2.** Mainland scenarios assumptions and input values.

| Input Variables | Mainland | | |
|---|---|---|---|
| | Low Scenario | Base Scenario | High Scenario |
| Import | 1% | 2% | 3% |
| Export | 3% | 5% | 6% |
| GDP | 3% | 4.5% | 6% |
| Population | 1% | 2% | 3% |

### *3.1. Medium Neural Networks (MNN)*

MNNs are computer-based systems that have been created to imitate the functions of the human brain, such as learning, comprehension, and the discovery of new information through experience [56]. Inspired by the human brain structure, MNN performs information analysis by learning, relating, and generalizing over data. MNNs consist of layers connected to each other in parallel [57]. In this study, the neural networks consist of three layers, including the input, hidden, and output layers. The function of the network is the connections between these layers. By adjusting the weight values that the layers are connected to each other, the network is trained to perform a certain function. Thus, an output is produced in response to an input in the network.

With MNN, predictions can be made in multiple time intervals, including short-, medium-, and long-term, in electricity demand forecasting [58]. Outputs are transmitted via synapse connections. Synaptic link weights are expressed in numerical values. The architecture of the MNN model is depicted in Figure 2. The connection between all layers in the ANN, which consists of 3 components, the input, hidden, and output layer, is provided by weights. When inputs from the input layer are transmitted to the hidden layer, they are multiplied by the link weights between the hidden layer and the input layer. The inputs to the neurons in the hidden layer are summed. Then, this expression is multiplied by the link weights between the hidden layer and the output layer and transmitted to the output layer. During the learning process, the weight values of the connections are determined [59].

In this study, the electricity demand forecasting model for Turkey is modelled using a feedforward multilayer perceptron neural network. Layers and neurons arranged in a feedforward way are evaluated in multilayer perceptron neural networks. In this structure, the input layer of the neurons collects the information outside the system, while the output layer calculates the values according to the input values [60].

Within this research, the backpropagation method with gradient descent is employed to train the model. This method, also known as backpropagation, begins by comparing the target value with the output value. Then, moving back through the input, the network adjusts the neuron weights to minimize the error, which is the difference between the target and output data. The Levenberg–Marquardt algorithm is utilized in the backpropagation process as a gradient-based algorithm.

The backpropagation method determines the output by finding the optimal combination of weights that minimizes the error function. To ensure continuity and differentiability of the error function, an interneuron transfer function is employed. In feedforward multilayer networks, linear transfer functions are utilized in the output neuron layer [61].

In a neural network, each neuron, except for the neurons in the input layer, is connected to another neuron in the subsequent layer through weights. The neuron calculates a weighted sum of all the neuron values in the preceding layer. This situation is added to the weighted sum transfer function, and the output value is calculated. The neuron output is calculated using the following Equations (1) and (2) [11].

$$a_j = \sum_{i=0}^{n} W_{ji} X_i \tag{1}$$

$$y_i = f(a_i) \tag{2}$$

where $a$ is the weighted sum of the inputs, $j$ is the neuron number, $f$ represents the transfer function, $w$ is the weight, $i$ is the input number, and $y$ resembles the output value. The sigmoid transfer function is determined as in Equation (3) [11,62]:

$$f(a_j) = y_j = \frac{1}{1 + e^{-a_j}} \tag{3}$$

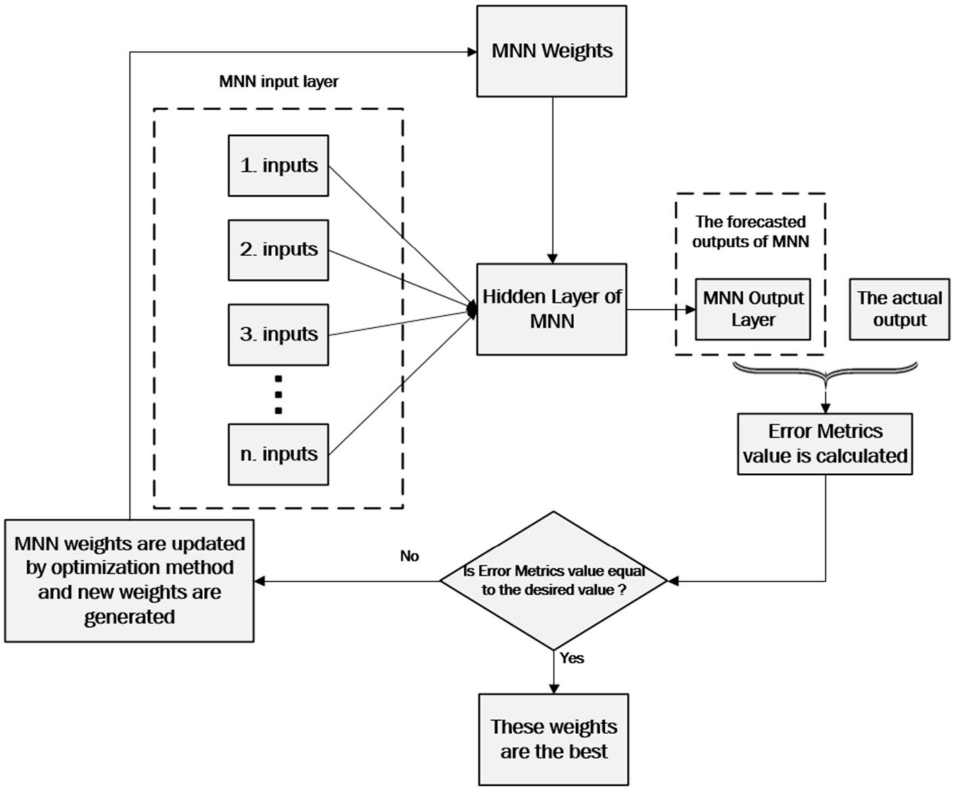


**Figure 2.** MNN working flowchart.

To simulate, train, and design the system, MATLAB R2021b–Neural Network Toolbox was utilized. By employing this software, an MNN model was trained, constructed, and evaluated using 40 distinct raw data samples obtained from official sources. The input data for the MNN model consists of four parameters, import, GDP, population, and export, while the output parameter comprises electricity consumption, as depicted in Figure 3.

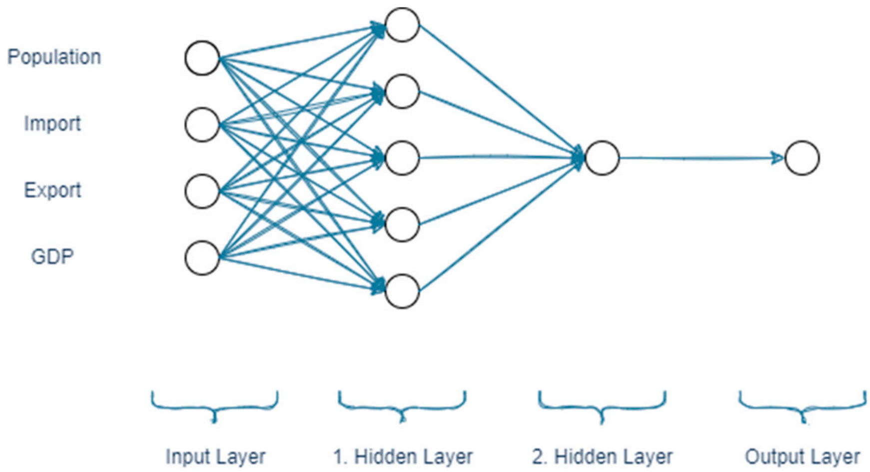


**Figure 3.** MNN model structure used in this study.

The MNN model architecture comprises an input layer with four neurons, an output layer with one neuron, and each hidden layer containing five to one neuron. The collected data from the experimental setup are split into two sections, namely, testing and training. The design, training, and simulation of the MNN model were performed using MATLAB-Neural Network Toolbox. Levenberg–Marquardt algorithm used the MNN training algorithm in this study.

### 3.2. Support Vector Machine

In this research, the SVM algorithm, which is a non-parametric supervised classification method, was utilized. The SVM was originally designed for the classification of two-class linear data but was insufficiently developed for non-linear and multi-class data and applied to engineering applications [63,64].

The approximate function in the SVM algorithm is given in Equation (4) [63].

$$f(x) = \omega \cdot \varphi(x) + b \tag{4}$$

where $\varphi(x)$ denotes the higher dimensional feature space transformed from the input vector $x$, and $\omega$ and $b$ denote the weight vector and bias term, respectively. The risk function obtained by minimizing $b$ and $\omega$ values is given in Equation (5) [64].

$$R(C) = C\sum_{i}^{N} L_{\varepsilon}(f(x_i), y_i)\frac{1}{2}\|\omega\|^2 \tag{5}$$

where $\frac{1}{2}\|\omega\|^2$ term represents the regulation term of SVM, and $C$ represents the compensator parameter, which indicates the error rate in the optimization.

The most important difference between the Vapnik linear function and the classical regression functions is shown in the Novell loss function $L_{\varepsilon}$ in Equation (6) [63]. Two variables ($\xi$ and $\xi$*) with positive values are defined to avoid unexpected outliers. Lagrange multipliers ($a$, $a$*) are added to solve the optimization problems.

$$\sum_{i}^{N} L_{\varepsilon}(f(x_i), y_i) = \begin{cases} |f(x_i) - y_i| \leq \varepsilon = 0 \\ |f(x) - y| - \varepsilon \neq 0 \end{cases} \tag{6}$$

After calculating the Lagrange multipliers, Equation (4) can be written as Equation (7) [63].

$$f\left(x, a, a^*\right) = \sum_{i=1}^{N} \left(a_i - a_i^*\right) K(x_i, x_j) + b \tag{7}$$

where $K(x_i, x_j)$ is called the kernel function and is illustrated in Equation (8) [64].

$$K(x_i, x_j) = \varphi(x_i)\varphi(x_j) \tag{8}$$

After these regulations, the main function of SVM is [64]

$$f(x) = \left\{\sum_{i=1}^{N} a_i K(x_i, x)\right\} - b \tag{9}$$

Here, $a_i$ and $b$ are the parameters of the SVM, $K$ is the kernel function, $N$ is the number of training data, $x$ is the independent vector, and $x_i$ is the vectors used in the training process. The selection of an appropriate function is important in terms of obtaining more accurate results from the data to be used. In this study, the linear SVM function given in Equation (10) was determined [64].

$$K(x_i - x_j) = x_i \cdot x_j \tag{10}$$

### 3.3. Whale Optimization Algorithm

WOA is a swarm-based metaheuristic algorithm inspired by the hunting behaviour of humpback whales by Mirjalili and Lewis [65]. Humpback whales, which feed on small fish near the surface, hunt these creatures with the bubble net method shown in Figure 4. With the water bubble method, humpback whales create a narrowing circle and spiral bubbles while rising to the water's surface, and they collect their prey together and reduce the target [65].

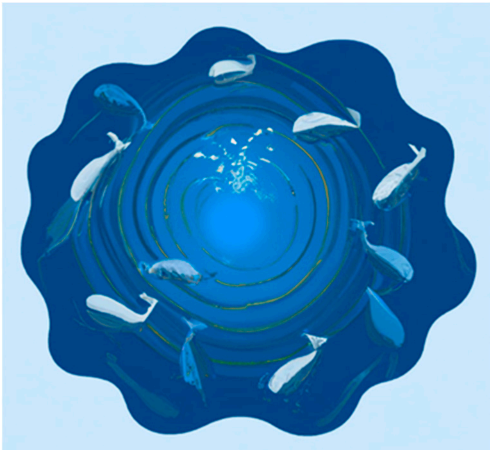

**Figure 4.** Humpback whales feeding behavior.

WOA is used in classification, image processing, network, and other engineering problems [66]. In WOA, each solution is represented by a whale (agent), while the prey represents the assumed best solution to the problem. The algorithm starts with the number of whales determined by the user, and the positions of these whales are updated according to the position of the best whale found so far, and the search for the best solution to the problem continues until the number of iterations determined by the user is reached (Figure 5). The algorithm is mathematically modelled in three parts, namely, prey wrapping, bubble net attack, and prey search [67].

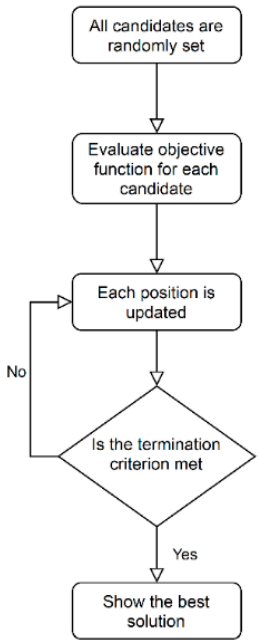


**Figure 5.** WOA working flowchart.

In this study, WOA is applied to forecast electricity demand. The recommended model applies the errors as a WOA objective function to measure its solutions in the training section. Firstly, WOA initializes the position vector and score of the leader and the search whales' positions. Secondly, optimization returns the search whales that go out of the search space boundaries and compute each search whale's objective function. The leader is updated if the current objective function is not at the desired value. This situation continues until the maximum iteration is reached. Ultimately, the best solution is obtained [68]. The number of whales, location coordinates, and launch parameters were randomly generated. To make predictions with WOA, the number of search agents was determined as 40, and the Maximum Number of Iterations was selected as 1000.

### 3.3.1. Encircling Prey

Once the search algorithm has identified the search agent with the most optimal solution discovered thus far, the positions of the remaining search agents are adjusted based on that of the best search agent. Thus, the prey that represents the best solution is surrounded. This behaviour is mathematically expressed in Equations (11) and (12) [67].

$$\vec{D} = \left| \vec{C} \cdot \overrightarrow{X_{Best}}(t) - \vec{X}(t) \right| \tag{11}$$

$$\vec{X}(t+1) = \overrightarrow{X_{Best}}(t) - \vec{A} \cdot \vec{D} \tag{12}$$

Here, $\vec{D}$ represents the direct distance between the food and the whale, $\overrightarrow{X_{Best}}$ represents the best humpback whale position so far, $t$ represents the number of iterations available, $\vec{X}(t)$ represents the whale position, while $\vec{A}$ and $\vec{C}$ are the specific coefficients shown below in Equations (13) and (14) [67].

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{13}$$

$$\vec{C} = 2\vec{r} \tag{14}$$

Additionally, $\vec{r}$ denotes a random vector with values in the range [0, 1] and $\vec{a}$ denotes a vector that decreases from 2 to 0 over the iterations.

### 3.3.2. Bubble-Net Attacking Method

During a bubble net attack, agents can perform either the constricting motion or the spiral motion around the prey with equal probability, as shown in Figure 5. While the narrowing of the circle around the prey is provided by Equation (10), the spiral movement is provided by Equation (15).

$$\vec{X}(t+1) = \vec{D'} \cdot e^{bl} \cdot \cos(2\pi l) + \overrightarrow{X_{Best}}(t) \tag{15}$$

Here, $\vec{D'}$ shows the difference between the search agent and the best agent obtained so far [67].

### 3.3.3. Search for Prey

During the bubble net attack stage, each whale updates its position towards the best whale, for which a linearly decreasing parameter ($A$) is considered. Instead of relying on the best-known point, the update of new locations for prey search agents is based on a search agent selected at random. Thanks to this move, the algorithm avoids the risk of getting stuck with the best local solutions and performs a global search. Equations (16) and (17) express the hunting movement mathematically [67].

$$\vec{D} = \left|\vec{C}\cdot\overrightarrow{X_{rand}} - \vec{X}\right| \tag{16}$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}} - \vec{A}\cdot\vec{D} \tag{17}$$

In WOA, local search is provided by encircling the prey, and holistic search is provided by the act of searching for prey. The type of search is decided according to the value of the vector $\vec{A}$.

### 3.4. Error Metrics

In this paper, several statistical criteria were used to assess the accuracy of the MNN-SVM-WOA model's predictions. These criteria included commonly used error metrics such as RMSE, MAE, MSE, MAPE, and Pearson's correlation coefficient (R). The MAE and RMSE were used to evaluate the difference between the predicted and true values without considering positive or negative errors or their mutual counteraction. The MSE indicated the discrepancy between the actual and estimated values, while the MAPE measured the precision of the forecasting methods, particularly when diverse data sets were used. The goal is to achieve low RMSE, MAE, and MAPE data. Additionally, the R value represented the correlation between the estimated and actual data [69].

$R^2$ is a statistical parameter used to determine the extent to which changes in the independent variable can explain changes in the dependent variable, and its value ranges from 0 to 1. If the $R^2$ value is close to 1, it indicates that the regression line fits well, implying that the changes in the dependent parameter are mostly due to changes in the independent variable. Equations (18)–(21) provide the formulas for $R^2$, RMSE, MSE, and MAE [10,11,70–74].

$$R^2 = \frac{\left(\sum_{i=1}^{N}(x_i^* - \overline{x_i^*})(x_i - \overline{x_i})\right)^2}{\sum_{i=1}^{N}(x_i^* - \overline{x_i^*})^2 \sum_{i=1}^{N}(x_i - \overline{x_i})^2} \tag{18}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i^* - x_i)^2} \tag{19}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - x_i^*\right)^2 \tag{20}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|x_i - x_i^*\right| \tag{21}$$

Note: $x_i, x_i^*, N, \overline{x_i}, \overline{x_i^*}$ denote the estimated data, real value, sample size, mean predicted data, and mean actual data, respectively [11].

## 4. Analysis and Results

### 4.1. Electricity Demand Forecasting

Turkey's electricity needs are high, and a significant part of this need is imported. With the increase in population and the introduction of new technologies into human life in a changing world, the need for energy increases even more. In terms of sustainable development, it is necessary for Turkey to meet the energy needs of its growing population in 2023 and beyond, to a large extent, by its own means. For this purpose, it is important to calculate the primary energy need in the coming years and to investigate how this need can be met.

The R regression values for the training data set, validation data set, and testing data set in the estimation of electrical energy consumption by MNN are 0.99661, 0.99903, and 0.99697, respectively, as shown in Figure 6. The R overall regression value was calculated at 0.99448; this result indicates that the MNN has very high reliability in forecasting electricity consumption. The actual data were very similar to MNN's estimated results.

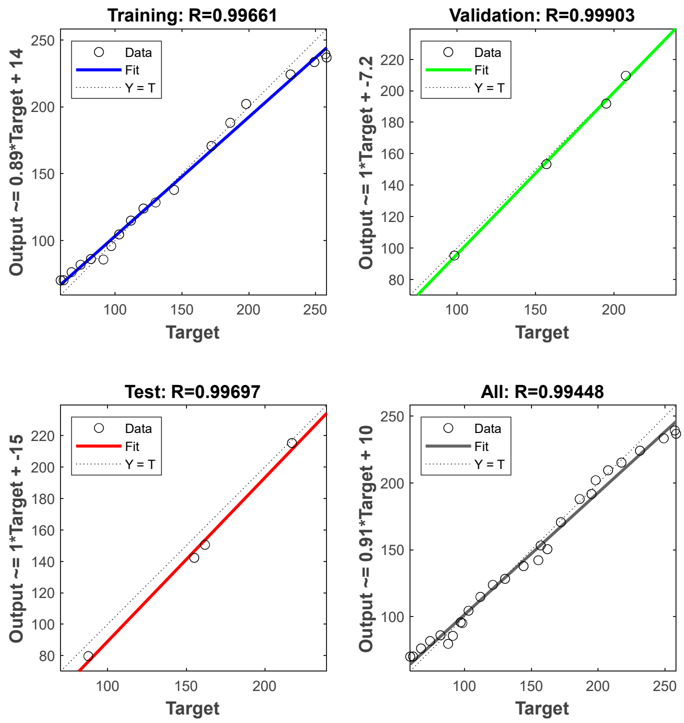


**Figure 6.** MNN Method Relationship between estimated values and actual data for Turkey's mainland electricity consumption.

The actual electricity demand data from the mainland between 1980 to 2019 and the predicted values from the three methods' outputs (MNN, SVR, and WOA) are presented (Figure 7). Upon examining the graph, it can be observed that the MNN method closely aligns with the actual value.

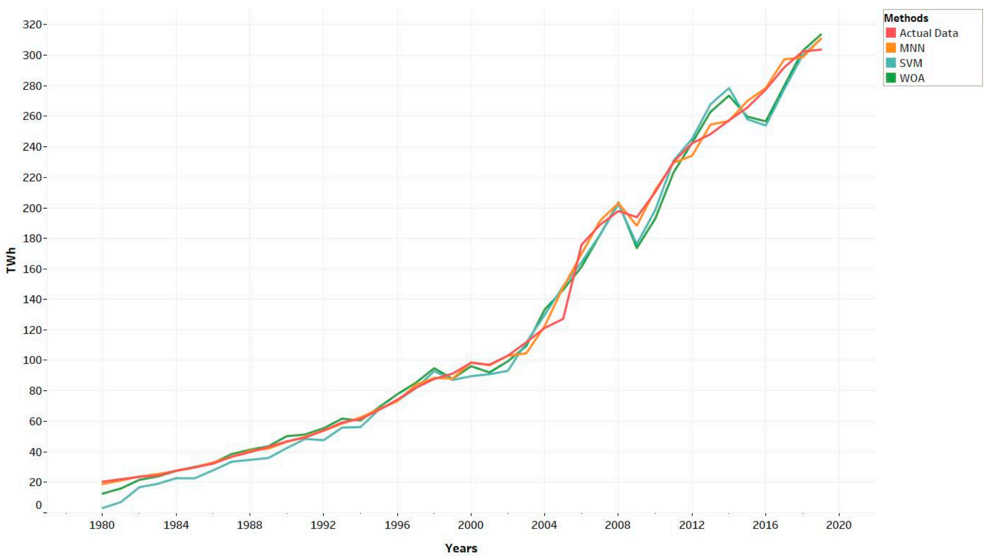


**Figure 7.** Mainland electricity demand predicted with MNN, SVR, and WOA vs. actual demand.

Figure 8 presents a graph that predicts electricity energy using three different methods (MNN, SVR, and WOA) for three different scenarios (low, base, and high) until 2040. It was observed that there were only slight variations in the results produced by the SVM and WOA methods. On the other hand, the MNN method showed significant changes across all three scenarios.

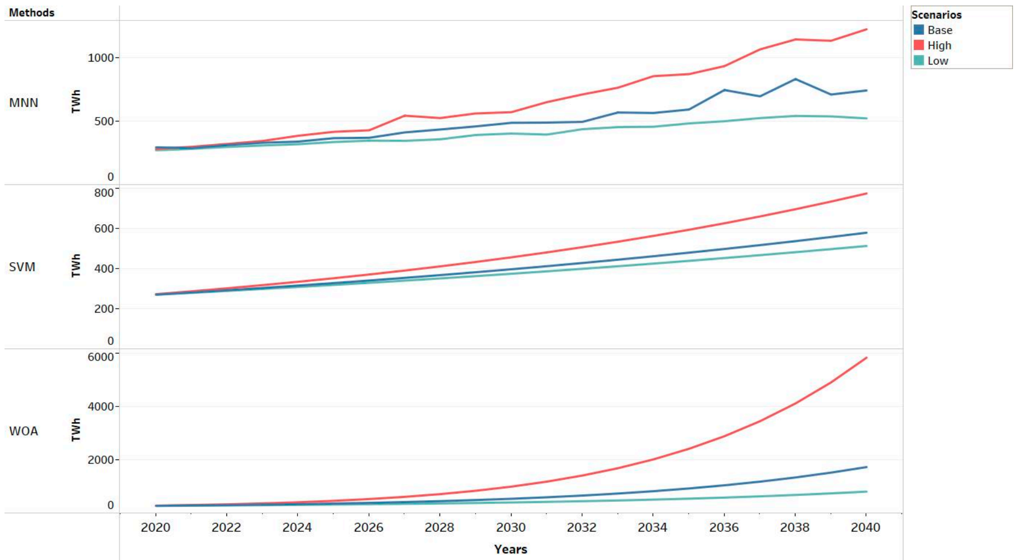


**Figure 8.** Electricity demand forecasting results using MNN, WOA, and SVM.

### 4.2. Error Metrics

The mean values of training time, MAE, MSE, RMSE, and $R^2$ were drawn by using the SVM, MNN, and WOA models demonstrated in Table 3. It is also observed that the MSE, $R^2$, RMSE, and MAE standard deviation values achieved via MNN have lower values than WOA and SVM models.

**Table 3.** Mainland statistical analysis results.

| Methods | | Mainland |
|---|---|---|
| $R^2$ | SVM | 0.9978 |
| | WOA | 0.9966 |
| | MNN | 0.9984 |
| RMSE | SVM | 3.4335 |
| | WOA | 2.9873 |
| | MNN | $5.325 \times 10^{-14}$ |
| MSE | SVM | 11.78 |
| | WOA | 8.923 |
| | MNN | $28.35 \times 10^{-28}$ |
| MAE | SVM | 2.9982 |
| | WOA | 2.3276 |
| | MNN | $2.5 \times 10^{-14}$ |

A perspective of the analysis of RMSE, $R^2$, and MSE for the training model is that the values of RMSE are always positive values, and the units match with the system response. Moreover, $R^2$ is constantly between 1 and 0 on the mainland. The comparison between the trained model and the investigation model is constant and equal for the training response. In addition, there are no defined negative values for $R^2$. Therefore, when the values of $R^2$

become negative, the model is classified as worse than the constant model. Consequently, the MSE is equal to the square of the RMSE and is always defined as positive for all circumstances. Additionally, the MAE is always a positive value, similar to the RMSE; however, it is less sensitive to deviations.

### 4.3. Multi Regression Equations

The importance of expressing experimental data as a mathematical equation can be applied more broadly [75]. This section of the paper indicates the study findings that evaluate the accuracy of estimating F using four input parameters: import, export, GDP, and population (represented as $a$, $b$, $c$, and $d$, respectively). The four parameters were also divided into various subsets (such as $a$, $b$, $c$ or $b$, $c$) to assess their individual performances, which were measured using $R^2$ and $p$-value metrics.

Table 4 presents the performance of subsets of parameters, their related equations, and the corresponding $R^2$ and $p$-values. The first-row regression equation, which includes all four parameters ($a$, $b$, $c$, and $d$), shows the maximum $R^2$ value (0.995). This indicates that the equation strongly represents the relationship between the input parameters and F. Conversely, the equations that exclude $c$ and $d$ variables in Equation (11) exhibit low generalization abilities due to their low $R^2$ performances in Table 4. However, the inclusion of the $d$ coefficient, which exhibits the strongest correlation among Equations (1)–(7), leads to an increase in the $R^2$ performance. Although the $R^2$ performances of Equations (1)–(7) are similar, it was observed that parameter $a$ has the lowest generalization ability and thus has a low correlation value.

**Table 4.** $R^2$ performance, regression equations, and multiple subset parameters.

| Eq No | Parameters | Multi Regression Equations | $R^2$ | $p$-Value |
|---|---|---|---|---|
| 1 | a, b, c, d | $F = -49.914 + 0.089284 * a + 0.48065 * b + 0.15825 * c + 0.78607 * d$ | 0.995 | $3.53 \times 10^{-39}$ |
| 2 | b, c, d | $F = -47.462 + 0.61379 * b + 0.14806 * c + 0.78086 * d$ | 0.994 | $0.03 \times 10^{-40}$ |
| 3 | a, c, d | $F = -64.02 + 0.3626 * a + 0.21054 * c + 0.08503 * d$ | 0.993 | $4.07 \times 10^{-39}$ |
| 4 | c, d | $F = -102.91 + 0.32548 * c + 1.1863 * d$ | 0.98 | $4.41 \times 10^{-32}$ |
| 5 | a, b, d | $F = -82.719 - 0.12819 * a + 1.0247 * b + 2.0838 * d$ | 0.99 | $5.48 \times 10^{-36}$ |
| 6 | b, d | $F = -90.488 + 0.8576 * b + 2.2348 * d$ | 0.99 | $1.96 \times 10^{-37}$ |
| 7 | a, d | $F = -166.68 + 0.5579 * a + 3.763 * d$ | 0.98 | $3.74 \times 10^{-32}$ |
| 8 | a, b, c | $F = -16.317 + 0.086601 * a + 0.49675 * b + 0.1895 * c$ | 0.994 | $4.1 \times 10^{-40}$ |
| 9 | b, c | $F = -14.155 + 0.62581 * b + 0.17942 * c$ | 0.994 | $9.28 \times 10^{-42}$ |
| 10 | a, c | $F = -28.091 + 0.36962 * a + 0.24634 * c$ | 0.993 | $3.85 \times 10^{-40}$ |
| 11 | a, b | $F = 23.942 - 0.30766 * a + 1.5105 * b$ | 0.984 | $5.42 \times 10^{-34}$ |

The closer the $p$-value gets to smaller values, the larger the statistically significant difference. The $p$-value ranges from 0.01 to 0.05; it can be said that there is a significant difference. If the $p$-value ranges from 0.001 to 0.01, there is a high level of significant difference. If the $p$-value is less than 0.001, there is a very high level of statistically significant difference [76,77]. If the $p$-values in Table 4 are less than 0.001, there, therefore, is a very high level of statistically significant difference. The equations in the second and ninth rows have the lowest $p$-values.

### 4.4. Correlation Matrix

A correlation matrix is a valuable tool for analyzing the relationship between several variables in a dataset. It takes into account both the signs and the matrix size of the correlation coefficients. The correlation coefficient evaluates how strong the linear correlation between two variables is, and it ranges from −1 to 1. A value close to 1 indicates a robust positive correlation, while a value near −1 implies a robust negative correlation. If the correlation coefficient is near 0, it shows that there is no linear correlation between the two variables [78].

In this study, two nonidentical correlated matrices are generated. The first matrix illustrated the correlation between input and output data which are utilized in MNN, SVM, and WOA methods. The second matrix displays the association between the outcomes of the methods and the real electricity consumption.

Table 5 displays the input variables correlation matrix (import, population, GDP, and export) and the output variable (electricity consumption). Based on the outcomes, it is evident that there is a strong positive linear relationship between electricity consumption and export (0.991). Furthermore, the input variables of import and export also exhibit a strong positive linear relationship (0.9895).

**Table 5.** Correlation Matrix Between Independent and Dependent Variables.

| Variables | Import | Export | GDP | Population | Electricity Consumption |
|---|---|---|---|---|---|
| Import | 1 | 0.9895 | 0.946 | 0.9232 | 0.9742 |
| Export | 0.9895 | 1 | 0.9727 | 0.9478 | 0.991 |
| GDP | 0.946 | 0.9727 | 1 | 0.9684 | 0.9892 |
| Population | 0.9232 | 0.9478 | 0.9684 | 1 | 0.9669 |
| Electricity Consumption | 0.9742 | 0.991 | 0.9892 | 0.9669 | 1 |

The results of Table 6 presented the correlation between the proposed methods, which are SVM, WOA, MNN, and the real data outcomes. The correlation of real data with MNN, SVM, and WOA methods was observed at 0.9988, 0.9952, and 0.9957, respectively. Additionally, the maximum correlation observed between MNN and actual values equals 0.9988.

**Table 6.** Methods Correlation Matrix.

| Methods | Actual Data | MNN | SVM | WOA |
|---|---|---|---|---|
| Actual Data | 1 | 0.9988 | 0.9952 | 0.9957 |
| MNN | 0.9988 | 1 | 0.996 | 0.9967 |
| SVM | 0.9952 | 0.996 | 1 | 0.9994 |
| WOA | 0.9957 | 0.9967 | 0.9994 | 1 |

The impact of export, GDP, population, and import on actual electricity consumption is analyzed using the Ordinary Least Squared (OLS) method. The results reveal that all the independent variables have $p$-values of less than 0.05, which means they are statistically significant at a 5% level. The import coefficient is positive, indicating a positive relationship between import and real consumption. Specifically, an increase in import leads to a rise in actual consumption by 0.72 kWh while holding all other variables constant. A one-unit increase in import results in a 2.24 kWh increase in actual electricity consumption while holding all other variables constant. On the other hand, export has a negative effect on actual consumption, indicating that a 1% increase in export is associated with roughly a 0.77 decrease in actual consumption.

The F statistics are used to determine if the model is significant as a whole, and a $p$-value of less than 0.05 indicates that the model is statistically significant at a 5% level (Table 7). Another way to interpret the results is through the R-squared adjusted, which indicates the percentage change in actual consumption that can be explained by population, GDP, export, and import. The adjusted R-squared is 0.99, meaning that the variables included in the model can explain 99% of the variation in actual consumption.

**Table 7.** Independent Variables Statistical Values According to Methods.

| Methods | Variables | Coefficient | 95% Confidence Internal | | t | $p > \|t\|$ |
|---|---|---|---|---|---|---|
| Real | import | 0.72 | 0.053 | 1.387 | 2.24 | 0.035 |
| | export | −0.771 | −1.438 | −0.103 | −2.4 | 0.025 |
| | GDP | 9.925 | 6.044 | 13.807 | 5.3 | 0 |
| | population | 5.198 | 4.135 | 6.26 | 10.15 | 0 |
| MNN | import | 0.913 | 0.054 | 1.772 | 2.18 | 0.036 |
| | export | −0.693 | −1.254 | −0.133 | −2.1 | 0.026 |
| | GDP | 9.698 | 6.088 | 13.29 | 5.39 | 0 |
| | population | 5.432 | 4.368 | 6.497 | 10.04 | 0 |
| SVM | import | 0.72 | 0.298 | 1.142 | 3.54 | 0.002 |
| | export | −0.77 | −1.193 | −0.348 | −3.78 | 0.001 |
| | GDP | 9.931 | 7.472 | 12.39 | 8.38 | 0 |
| | population | 5.197 | 4.524 | 5.871 | 16.02 | 0 |
| WOA | import | 0.72 | 0.389 | 1.051 | 4.52 | 0 |
| | export | −0.771 | −1.102 | −0.44 | −4.83 | 0 |
| | GDP | 9.912 | 7.986 | 11.838 | 10.67 | 0 |
| | population | 5.202 | 4.674 | 5.729 | 20.46 | 0 |

Table 7 presents the confidence intervals, t-values, and coefficient values for each variable in the different methods used. In the 95% confidence interval analysis, it can be observed that the MNN method shows a positive effect of import, GDP, and population on the dependent variable, while exports have a negative effect. This negative trend of exports is consistent across all methods.

## 5. Discussion

Forecasting electricity demand is crucial for the efficient planning of capacity and the establishment of electricity networks with minimal expenses. To achieve accurate predictions, policymakers must assess various alternative methods and determine which one will yield the most advantageous results.

This paper utilized MNN, SVM, and WOA models to estimate the electricity demand in Turkey by implementing three distinct scenarios. The architecture of the MNN model comprises an input layer, a hidden layer, and an output layer. The input parameters of the three methods are classified into two primary categories: demographic and economic. The output layer produces the forecasted electricity demand. The model performance is evaluated using various indices, including RMSE, MSE, MAE, and $R^2$. The training process and evaluation algorithm are presented and analyzed, along with the limits of the 95% confidence intervals. Four independent variables, which are export, population, GDP, and import, are identified as the possible electricity demand predictors from 1980 to 2019. Equations for the mainland were derived using data spanning from 1980 to 2019. These equations were subsequently utilized to estimate Turkey's future electricity demand under various scenarios.

Electricity consumption between 1980 and 2019 is obtained from Turkish Electricity Transmission Corporation. The population data is collected from the TSI. The World Bank Open Dataset was utilized to obtain values for imports, GDP, and export. These variables were then subjected to stepwise regression to state which of them best predicted the dependent variable.

To determine the significance of the correlation coefficient, statistical techniques were employed, and the results were compared. The SVM, WOA, and MNN methods' correlation coefficient ranges were established using a 95% confidence interval. Based on the statistical analysis, it was determined that the MNN method is more reliable and consistent than the others, with better results at a 95% confidence level. Nevertheless, all the proposed method frameworks demonstrate promise as models that can assist engineers and policymakers in developing energy-related budgets more effectively. The MNN method's adaptability renders it suitable for identifying optimal solutions regarding forecasting future trends in electricity demand. Moreover, the high regression values obtained from the models demonstrate that MNN is an effective tool for electricity demand prediction. Specifically, the R regression values for the testing, training, and validation of datasets in the prediction of electrical energy consumption using MNN were 0.99661, 0.99903, and 0.99697, respectively. The overall R regression value was calculated as 0.99448, which indicates that MNN is highly reliable when estimating electricity consumption. In addition, the $R^2$ values for MNN, SVM, and WOA in the prediction of electricity consumption were shown as 0.9984, 0.9978, and 0.9966, respectively. These results indicate that MNN is highly reliable when estimating electricity consumption. The RMSE, MSE, and MAE values for the MNN method are $5.325 \times 10^{-14}$, $28.35 \times 10^{-28}$, $2.5 \times 10^{-14}$, respectively. These results show that the electrical energy estimation performance of MNN is successful.

The official findings closely matched the predicted results of MNN, which is crucial for the advancement of productive and practical electrification systems policy planning. These findings demonstrate that the proposed model can be employed effectively and actively for Turkey's long-term electricity demand forecasting. The obtained results can also serve as a guide for future electricity system network designs. In addition, statistical methods were utilized to determine the confidence intervals of the algorithms used to estimate electricity demand, and the findings were compared. The proposed model's performance was assessed using various metrics such as RMSE, MSE, MAE, $R^2$, the independent and dependent variables correlation matrix across the different methods, and the multiple regression equations correlation. The metrics for measuring errors provided a clear indication of how accurate and precise the estimation techniques were.

Typically, a single hidden layer is utilized for developing the MNN architecture; however, the determination of MNN's architecture, including hidden layers and neurons, is frequently based on trial and error in most articles. The accuracy of the recognition process and training speed is influenced by the number of neurons in the hidden layer. In future studies, the number of neurons in the hidden layer, as well as the number of intermediate layers, can be modified to assess their performances in forecasting electricity demand. The $R^2$ and $R^2$ adj values indicate that the proposed regression model fits very well. The T-statistic for the historical electricity demand is over 2, indicating that it is statistically significant. The electricity demand and past data both have significant *p*-values, indicating that they are valuable additions to the model as they serve as significant regressors.

The correlation matrix depicts the relationship between input values (import, population, GDP, and export) and output values (electricity consumption). A strong linear correlation was observed between electricity consumption and export (0.991). In addition, imports and export show a strong linear relationship (0.9895). Import, export, GDP, and population affect the real data positively and significantly. Their positive effect continues with the different dependent variables used interchangeably. The *p*-values in Table 4 are less than 0.001, and therefore there is a very high level of statistically significant difference. The equations in the second and ninth rows have the lowest *p*-values.

The Levenberg–Marquardt technique is an effective learning approach that outperformed other learning methods. Though Levenberg–Marquardt is the quickest training method, it is only employed for small- and medium-sized networks. While WOA is a sufficiently efficient algorithm, it does have limitations in terms of exploration ability, slow solution generation, and susceptibility to becoming trapped in local solutions. In addition, the issue of inadequate convergence accuracy and speed is present in the original WOA algorithm. There are some general shortcomings in the optimization process, including a tendency to get stuck in local optima and a low level of convergence accuracy. Additionally, when the data exceeds a certain number, the prediction performance of SVM decreases. In future studies, hybrid and machine learning approaches will be incorporated, and various optimization techniques will be developed to assess the accuracy of the forecasting models.

**Author Contributions:** Conceptualization, M.S. and C.S.; methodology, M.S., C.S. and O.A.K.; software, M.S.; validation, M.S.; formal analysis, M.S.; data curation, M.S.; writing—original draft preparation, M.S., C.S. and O.A.K.; writing—review and editing, M.S., C.S. and O.A.K.; visualization, M.S.; supervision, C.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Publicly available data sets were analyzed in this study.

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

## References


1. Nie, R.X.; Tian, Z.P.; Long, R.Y.; Dong, W. Forecasting household electricity demand with hybrid machine learning-based methods: Effects of residents' psychological preferences and calendar variables. *Expert Syst. Appl.* **2022**, *206*, 117854. [CrossRef]
2. Gunay, M.E. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. *Energy Policy* **2016**, *90*, 92–101. [CrossRef]
3. Sultana, N.; Hossain, S.M.; Almuhaini, S.H.; Düştegör, D. Bayesian Optimization Algorithm-Based Statistical and Machine Learning Approaches for Forecasting Short-Term Electricity Demand. *Energies* **2022**, *15*, 3425. [CrossRef]
4. Beşkirli, A.; Beşkirli, M.; Haklı, H.; Uğuz, H. Comparing Energy Demand Estimation Using Artificial Algae Algorithm: The Case of Turkey. *J. Clean Energy Technol.* **2018**, *6*, 349–352. [CrossRef]
5. Wang, T.; Wen, Y.Y.; Lin, B.Q. Energy consumption and the influencing factors in China: A nonlinear perspective. *J. Clean Prod.* **2020**, *249*, 119375. [CrossRef]
6. Yu, Y.T.; Zhang, N.; Kim, J.D. Impact of urbanization on energy demand: An empirical study of the Yangtze River economic belt in China. *Energy Policy* **2020**, *139*, 111354. [CrossRef]
7. Liu, S.; Zeng, A.; Lau, K.; Ren, C.; Chan, P.W.; Ng, E. Predicting long-term monthly electricity demand under future climatic and socioeconomic changes using data-driven methods: A case study of Hong Kong. *Sustain. Cities Soc.* **2021**, *70*, 102936. [CrossRef]
8. Mutschler, R.; Rüdisüli, M.; Heer, P.; Eggimann, S. Benchmarking cooling and heating energy demands considering climate change, population growth and cooling device uptake. *Appl. Energy* **2021**, *288*, 116636. [CrossRef]
9. Román-Portabales, A.; López-Nores, M.; Pazos-Arias, J.J. Systematic review of electricity demand forecast using ann-based machine learning algorithms. *Sensors* **2021**, *21*, 4544. [CrossRef]
10. Cebekhulu, E.; Onumanyi, A.J.; Isaac, S.J. Performance Analysis of Machine Learning Algorithms for Energy Demand–Supply Prediction in Smart Grids. *Sustainability* **2022**, *14*, 2546. [CrossRef]
11. Saglam, M.; Spataru, C.; Karaman, O.A. Electricity Demand Forecasting with Use of Artificial Intelligence: The Case of Gokceada Island. *Energies* **2022**, *15*, 5950. [CrossRef]
12. Baba, A. Advanced AI-based techniques to predict daily energy consumption: A case study. *Expert Syst. Appl.* **2021**, *184*, 115508. [CrossRef]
13. Shams, M.H.; Niaz, H.; Hashemi, B.; Liu, J.J.; Siano, P.; Anvari-Moghaddam, A. Artificial intelligence-based prediction and analysis of the oversupply of wind and solar energy in power systems. *Energy Convers. Manag.* **2021**, *250*, 114892. [CrossRef]
14. Sekhar, C.; Dahiya, R. Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand. *Energy* **2023**, *268*, 126660. [CrossRef]
15. Sen, D.; Tunç, K.M.; Günay, M.E. Forecasting electricity consumption of OECD countries: A global machine learning modeling approach. *Util. Policy* **2021**, *70*, 101222. [CrossRef]
16. Ağbulut, Ü. Forecasting of transportation-related energy demand and $CO_2$ emissions in Turkey with different machine learning algorithms. *Sustain. Prod. Consum.* **2022**, *29*, 141–157. [CrossRef]
17. Şahin, U.; Ballı, S.; Chen, Y. 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. [CrossRef]
18. Tun, Y.L.; Thar, K.; Thwal, C.M.; Hong, C.S. Federated Learning based Energy Demand Prediction with Clustered Aggregation. In Proceedings of the Conference: 2021 IEEE International Conference on Big Data and Smart Computing, Jeju Island, Republic of Korea, 17–20 January 2021; pp. 164–167.
19. Ramsami, P.; King, R.T.A. Neural Network Frameworks for Electricity Forecasting in Mauritius and Rodrigues Islands. In Proceedings of the 2021 IEEE PES/IAS PowerAfrica, Nairobi, Kenya, 23–27 August 2021; pp. 1–5.
20. Kazemzadeh, M.-R.; Amjadian, A.; Amraee, T. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy* **2020**, *204*, 117948. [CrossRef]
21. del Real, A.J.; Dorado, F.; Durán, J. Energy Demand Forecasting Using Deep Learning: Applications for the French Grid. *Energies* **2020**, *13*, 2242. [CrossRef]
22. Bedi, J.; Toshniwal, D. Deep learning framework to forecast electricity demand. *Appl. Energy* **2019**, *238*, 1312–1326. [CrossRef]
23. Bannor, E.; Acheampong, A.O. Deploying artificial neural networks for modeling energy demand: International evidence. *Int. J. Energy Sect. Manag.* **2020**, *14*, 285–315. [CrossRef]
24. Abdulsalam, K.A.; Babatunde, O.M. Electrical energy demand forecasting model using artificial neural network: A case study of Lagos State Nigeria. *Int. J. Data Netw. Sci.* **2019**, *3*, 305–322. [CrossRef]

25. Hao, J.; Sun, X.; Feng, Q. A Novel Ensemble Approach for the Forecasting of Energy Demand Based on the Artificial Bee Colony Algorithm. *Energies* **2020**, *13*, 550. [CrossRef]
26. Kaytez, F. 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. [CrossRef]
27. Di Leo, S.; Caramuta, P.; Curci, P.; Cosmi, C. Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models. *Energy* **2020**, *196*, 117058. [CrossRef]
28. Angelopoulos, D.; Siskos, Y.; Psarras, J. Disaggregating time series on multiple criteria for robust forecasting: The case of long-term electricity demand in Greece. *Eur. J. Oper. Res.* **2019**, *275*, 252–265. [CrossRef]
29. Hou, R.; Li, S.; Wu, M.; Ren, G.; Gao, W.; Khayatnezhad, M. 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. [CrossRef]
30. Pegalajar, M.; Ruíz, L.G.B.; Cuéllar, M.P.; Rueda, R. Analysis and enhanced prediction of the Spanish Electricity Network through Big Data and Machine Learning techniques. *Int. J. Approx. Reason.* **2021**, *133*, 48–59. [CrossRef]
31. Bendaoud, N.M.M.; Farah, N.; Ben Ahmed, S. Applying load profiles propagation to machine learning based electrical energy forecasting. *Electr. Power Syst. Res.* **2022**, *203*, 107635. [CrossRef]
32. Porteiro, R.; Hernández-Callejo, L.; Nesmachnow, S. Electricity demand forecasting in industrial and residential facilities using ensemble machine learning. *Rev. Fac. De Ing.* **2022**, *102*, 9–25. [CrossRef]
33. Lu, Y.; Wang, G. A load forecasting model based on support vector regression with whale optimization algorithm. *Multimed. Tools Appl.* **2023**, *82*, 9939–9959. [CrossRef]
34. Emami Javanmard, M.; Tang, Y.; Wang, Z.; Tontiwachwuthikul, P. Forecast Energy Demand, $CO_2$ Emissions and Energy Resource Impacts for the Transportation Sector. *Appl. Energy* **2023**, *338*, 120830. [CrossRef]
35. Rao, C.; Zhang, Y.; Wen, J.; Xiao, X.; Goh, M. Energy Demand Forecasting in China: A Support Vector Regression-Compositional Data Second Exponential Smoothing Model. *Energy* **2023**, *263*, 125955. [CrossRef]
36. Li, S.; Kong, X.; Yue, L.; Liu, C.; Khan, M.A.; Yang, Z.; Zhang, H. Short-Term Electrical Load Forecasting Using Hybrid Model of Manta Ray Foraging Optimization and Support Vector Regression. *J. Clean. Prod.* **2023**, *388*, 135856. [CrossRef]
37. Huang, J.; Kaewunruen, S. Forecasting Energy Consumption of a Public Building Using Transformer and Support Vector Regression. *Energies* **2023**, *16*, 966. [CrossRef]
38. Huang, J.; Algahtani, M.; Kaewunruen, S. 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. [CrossRef]
39. Velasquez, C.E.; Zocatelli, M.; Estanislau, F.B.; Castro, V.F. Analysis of time series models for Brazilian electricity demand forecasting. *Energy* **2022**, *247*, 123483. [CrossRef]
40. Pallonetto, F.; Jin, C.; Mangina, E. Forecast electricity demand in commercial building with machine learning models to enable demand response programs. *Energy AI* **2022**, *7*, 100121. [CrossRef]
41. May, E.C.; Bassam, A.; Ricalde, L.J.; Soberanis, M.E.; Oubram, O.; Tzuc, O.M.; Alanis, A.Y.; Livas-García, A. 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.
42. Niu, W.J.; Feng, Z.K.; Li, S.S.; Wu, H.J.; Wang, J.Y. 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. [CrossRef]
43. Luzia, R.; Rubio, L.; Velasquez, C.E. Sensitivity analysis for forecasting Brazilian electricity demand using artificial neural networks and hybrid models based on Autoregressive Integrated Moving Average. *Energy* **2023**, *274*, 127365. [CrossRef]
44. Işık, G.; Öğüt, H.; Mutlu, M. Deep learning based electricity demand forecasting to minimize the cost of energy imbalance: A real case application with some fortune 500 companies in Türkiye. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105664. [CrossRef]
45. Albuquerque, P.C.; Cajueiro, D.O.; Rossi, M.D.C. Machine learning models for forecasting power electricity consumption using a high dimensional dataset. *Expert. Syst. Appl.* **2022**, *187*, 115917. [CrossRef]
46. Rick, R.; Berton, L. Energy forecasting model based on CNN-LSTM-AE for many time series with unequal lengths. *Eng. Appl. Artif. Intell.* **2022**, *113*, 104998. [CrossRef]
47. Maaouane, M.; Chennaif, M.; Zouggar, S.; Krajaci'c, G.; Duic, N.; Zahboune, H.; ElMiad, A.K. Using neural network modelling for estimation and forecasting of transport sector energy demand in developing countries. *Energy Convers Manag.* **2022**, *258*, 115556. [CrossRef]
48. Chaturvedi, S.; Rajasekar, E.; Natarajan, S.; McCullen, N. A comparative assessment of SARIMA, LSTM RNN and Fb Prophet models to forecast total and peak monthly energy demand for India. *Energy Pol.* **2022**, *168*, 113097. [CrossRef]
49. Turkish Electricity Transmission Corporation. Available online: https://www.teias.gov.tr/en-US/interconnections (accessed on 23 May 2022).
50. Turkish Statistical Institute. Available online: https://data.tuik.gov.tr/Kategori/GetKategori?p=nufus-ve-demografi-109&dil=1 (accessed on 18 September 2021).
51. Binici, M. Turkey's Energy Consumption Forecast by Using Mathematical Modeling. Master's Thesis, Sivas Cumhuriyet University, Sivas, Turkey, 2019.

52. Aghay Kaboli, S.H.; Selvaraj, J.; Rahim, N.A. Long-term electric energy consumption forecasting via artificial cooperative search algorithm. *Energy* **2016**, *115*, 857–871. [CrossRef]
53. International Monetary Fund (IMF). Available online: https://www.imf.org/en/Data (accessed on 19 June 2022).
54. World Bank. Available online: https://data.worldbank.org/?intcid=ecr_hp_BeltD_en_ext (accessed on 2 May 2022).
55. Aslan, M.; Beşkirli, M. Realization of Turkey's Energy Demand Forecast with the Improved Arithmetic Optimization Algorithm. *Energy Rep.* **2022**, *8*, 18–32. [CrossRef]
56. Efekemo, E.; Saturday, E.G.; Ofodu, J.C. Electricity Demand Forecasting: A Review. *Educ. Res. IJMCER* **2022**, *4*, 279–301.
57. Alhendi, A.A.; Al-Sumaiti, A.S.; Elmay, F.K.; Wescaot, J.; Kavousi-Fard, A.; Heydarian-Forushani, E.; Alhelou, H.H. Artificial intelligence for water–energy nexus demand forecasting: A review. *Int. J. Low-Carbon Technol.* **2022**, *17*, 730–744. [CrossRef]
58. Moustris, K.; Kavadias, K.A.; Zafirakis, D.; Kaldellis, J.K. Medium, short and very short-term prognosis of load demand for the Greek Island of Tilos using artificial neural networks and human thermal comfort-discomfort bio-meteorological data. *Renew Energy* **2020**, *147*, 100–109. [CrossRef]
59. Birecikli, B.; Karaman, O.A.; Çelebi, S.B.; Turgut, A. Failure load prediction of adhesively bonded GFRP composite joints using artificial neural networks. *J. Mech. Sci. Technol.* **2020**, *34*, 4631–4640. [CrossRef]
60. Karaman, O.A.; Tanyıldızı Ağır, T.; Arsel, İ. Estimation of Solar Radiation Using Modern Methods. *Alex. Eng. J.* **2021**, *60*, 2447–2455. [CrossRef]
61. Akbal, B. Hybrid GSA-ANN Methods to Forecast Sheath Current of High Voltage Underground Cable Lines. *J. Comput.* **2018**, *13*, 417–425. [CrossRef]
62. Onaolapo, A.K.; Carpanen, R.P.; Dorrell, D.G.; Ojo, E.E. A Comparative Assessment of Conventional and Artificial Neural Networks Methods for Electricity Outage Forecasting. *Energies* **2022**, *15*, 511. [CrossRef]
63. Korkut, F.G. Prediction of Interactions between SARS-CoV-2 Protein and Human Protein Using Machine Learning Methods. Master's Thesis, Yalova University, Yalova, Turkey, 2022.
64. Dinsever, L.D.; Gümüş, V.; Şimşek, O.; Avşaroğlu, Y.; Kuş, M. Adıyaman ve Gaziantep İstasyonlarındaki Güneşlenme Şiddetinin Destek Vektör Makineleri ile Tahmini. *Bilecik Şeyh Edebali Üniversitesi Fen Bilim. Derg.* **2021**, *8*, 753–769. [CrossRef]
65. Mirjalili, S.; Lewis, A. The Whale Optimization Algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]
66. Gharehchopogh, F.S.; Gholizadeh, H. A Comprehensive Survey: Whale Optimization Algorithm and Its Applications. *Swarm Evol. Comput.* **2019**, *48*, 1–24. [CrossRef]
67. Aydın, Z.E. Missing Data Imputation Methods Based on Whale Optimization Algorithm and Random Subspaces. Ph.D. Thesis, Eskisehir Technical University, Eskisehir, Turkey, 2021.
68. Shang, Z.; He, Z.; Song, Y.; Li, L.; Chen, Y. A Novel Combined Model for Short-Term Electric Load Forecasting Based on Whale Optimization Algorithm. *Neural Process. Lett.* **2020**, *52*, 1207–1232. [CrossRef]
69. Zhang, W.; Zhang, L.; Wang, J.; Niu, X. Hybrid system based on a multi-objective optimization and kernel approximation for multi-scale wind speed forecasting. *Appl. Energy* **2020**, *277*, 115561. [CrossRef]
70. Houimli, R.; Zmami, M.; Ben-Salha, O. Short-term electric load forecasting in Tunisia using artificial neural networks. *Energy Syst. Berl. Period.* **2020**, *11*, 357–375. [CrossRef]
71. Shah, I.; Jan, F.; Ali, S. Functional Data Approach for Short-Term Electricity Demand Forecasting. *Math. Probl. Eng.* **2022**, *2022*, 6709779. [CrossRef]
72. Soyler, I.; Izgi, E. Electricity Demand Forecasting of Hospital Buildings in Istanbul. *Sustainability* **2022**, *14*, 8187. [CrossRef]
73. Moradzadeh, A.; Moayyed, H.; Zare, K.; Mohammadi-Ivatloo, B. Short-term electricity demand forecasting via variational autoencoders and batch training-based bidirectional long short-term memory. *Sustain. Energy Technol. Assess.* **2022**, *52*, 102209. [CrossRef]
74. Aponte, O.; McConky, K.T. Forecasting an electricity demand threshold to proactively trigger cost saving demand response actions. *Energy Build.* **2022**, *268*, 112221. [CrossRef]
75. Brown, S.H. Multiple linear regression analysis: A matrix approach with MATLAB. *Ala. J. Math.* **2009**, *34*, 1–3.
76. Kaya, N.; Guler, H. Examining the Solution-focused Thinking and Hope Levels of Turkish Midwifery Students During Distance Education: A Cross-sectional Study. *Mod. Care J.* **2023**, *20*, e131139. [CrossRef]
77. Khelouf, N.; Haoud, K.; Meziani, S.; Fizir, M.; Ghomari, F.N.; Khaled, M.B.; Kadi, N. Effect of infant's gender and lactation period on biochemical and energy breast milk composition of lactating mothers from Algeria. *J. Food Compos. Anal.* **2023**, *115*, 104889. [CrossRef]
78. Steiger, J.H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **1980**, *87*, 245. [CrossRef]