Extended, Short-Term Neural Prediction Methodology, for European Electricity Production by Type

Miljana Lj. Milić*, Jelena B. Milojković and Andrija Z. Petrušić

Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, miljana.milic@elfak.ni.ac.rs; jelena.milojkovic@elfak.ni.ac.rs; andrija.petrusic@elfak.ni.ac.rs

Abstract: Accurate forecasting in the electrical energy supply sector is essential for cost savings, enhancing power system reliability, decisions on development, expansion, modification or reduction of facilities. It plays a vital role in the long-term development of the electric power industry. Electricity cannot be stored in large quantities, it is difficult to transfer, and requires continual production-consumption balance. The stochastic behavior of electricity consumption makes it challenging to anticipate. It is affected by a number of variables: climate, economy, population increase, pandemic breakout, etc. In this research we conduct experiments with different neural network forecasting topologies and establish the methodology that will most accurately anticipate the trend of the electricity production for various types of sources such as: wind, oil, coal, nuclear power plants, and bioenergy. An approach that incorporates Time Delay Neural Networks is proposed to reduce mistakes and improve forecasting confidence. It is shown that this strategy may significantly increase the forecasting accuracy of the individual networks regardless of their topologies, which improves the applicability of the method. The performance and efficiency of models are assessed using the appropriate performance criteria. Additional forecasting experiments, including ARIMA and Extreme Learning Machine Modeling, have been carried out to quantitatively compare the accuracy of the proposed technique with alternative state-of-the-art forecasting methodologies.

Keywords: Accuracy; Artificial Neural Networks; Forecasting; Time series analysis

1 Introduction

Energy is described as a capability measure, of an object to create motion, force, work, a change in shape or form, etc. It cannot be created or eliminated. It can only be transformed from one form to another. Energy as a physical quantity can have six basic forms: mechanical, chemical, electrical, nuclear, thermal, radiant [1]. Some additional forms of energy could be found in the literature, such as electrochemical, acoustic, electromagnetic etc. [2].

To produce electrical energy today, many traditional sources such as oil or coal-based power plants are still widely used. They all extensively pollute the planet. Human society is currently confronting and will continue to confront difficult challenges such as resource depletion and global climate change. If the international goal of reduced CO_2 emissions is to be met, shifting away from old "dirty" energy sources like coal and oil is crucial. To meet these challenges the exploration of renewable energy sources should be prioritized [3] [4]. Clean energy as sufficient, environmentally friendly, and widely available, has become an essential subject of study in the field of new energy resources [5]. However, certain renewable energy sources can be irregular and stochastic, which can have a negative impact on the cost management and power system, particularly when renewable and clean energy is integrated into traditional grid systems. Thus, increasing the precision and validity of energy production forecast can mitigate those problems, and improve power system management efficiency [6]. These all are the primary motivations for researchers in many areas of science and technology.

Solar energy, for example, as one of the renewable energy sources that neither pollutes the air nor affects climate change, has received the most attention in recent years, for generating power using photovoltaic technology [7]. It is crucial to identify different power system parameters in order to determine maximal power that can be collected from various sources.

The EU's strategy on climate change and energy is focused on energy transition. In 2008, the EU member states have adopted the legislation document referred to as "2020 Climate and Energy Package" (20-20-20), that contains three directives for the member states: reduction of gross energy consumption by 20%, 20% of gross energy production from renewable energy sources (RES), and 20% reduction of the greenhouse gas emissions [8]. The "2050 Long Term Strategy" for building a climate-neutral society, which set the target of reducing greenhouse gas (GHG) emissions by 80-95% by 2050, was added to this agenda in 2009. All of this is done to help achieving the 2-degree Celsius global warming reduction target, which was set in 2009 in Copenhagen during the 15th Conference of the Parties, also known as the "United Nations Climate Change COP 15". The power industry, which can quickly decrease emissions through the adoption of low-carbon technology and better energy efficiency, will benefit from these reductions [9].

From 2014 to 2018, the European Parliament crafted EU regulations, establishing mandatory targets for the "2030 Climate and Energy policy Framework." By 2030, EU Member States aim to:

1) Cut greenhouse gas emissions by 40% compared to 1990
2) Boost energy efficiency by 32.5% compared to a baseline
3) Raise renewable energy share in final consumption from 20% to 32%

In November 2018, the European Commission introduced a political vision for achieving a Net Zero economy by 2050, as well as the analytical foundations for the creation of an EU Long Term Strategy for energy and climate policy [10].

Due to the COVID-19 impact, EU Emissions Trading Scheme, GHG pollution fell by 12% in 2020, contributing to a 41% cumulative decline from the 2005 baseline. The 2020 achievement, reflecting a trend predating the pandemic, follows a 9% emission decrease from 2018 to 2019, indicating a shift from coal to less emission-intensive gas installations [11].

Within the 2030 framework, the EU mandated regulations, that needed to be adjusted based on identified trends. In 2016, the EU proposed a law aiming for 27% energy efficiency by 2030. However, by 2020, they adjusted the target to a 32.5% improvement over the 2007 baseline. A similar strategy increased the renewable energy participation objective from 27% to 32% [12].

In every level of the power system the generated power must be strongly corelated and balanced with the received power [13] [14]. Decision makers need advanced tools to track trends over time. This enables them to assess deviations from the set target and adjust course, either reaching the goal ahead of schedule or allowing for target revisions, as was done in 2020 for energy efficiency and the share of renewable energy sources.

Electricity production data represent an irregular time series that require non-linear methodologies for their processing and forecasting. Understanding every detail of an electricity production process isn't necessary for designing its regression model [15]. A multilayer neural network with backpropagation training can efficiently capture hidden patterns, providing superior predictive and statistical power compared to traditional models. Power systems are highly dynamic and nonlinear which is posing control challenges [16]. Similar to solving human cognition issues, building an efficient neural model from noisy data involves focusing on pertinent information and utilization of a small, highly relevant portion of the dataset [17]. Many different non-linear neural networks-based systems that analyze power systems behavior are lately being reported in the literature. In [18], a BI-LSTM algorithm proposes an accurate power prediction model for large-scale PV plants, specifically reliable for 1-hour large step ahead prediction. In [19], statistical time series models for forecasting half-daily values of global sun irradiance with a 3-day horizon are compared, revealing neural networks as the best approach. Authors in [20] introduce a framework for predicting energy consumption/distribution that uses LSTM, CNN, and Auto Encoders and is capable of managing uneven time series lengths with high prediction accuracy.

Two main methods for multi-step-ahead prediction are recursive and direct algorithms. In the recursive technique, the forecast from a one-step-ahead model is used for future predictions, accumulating errors over subsequent horizons [21] [22]. The direct technique treats the multi-step-ahead problem as a multi-output challenge [23] [24], allowing neural networks with multiple neurons in the output layer to represent prediction horizons. Despite challenges like highly chaotic time series and missing data, non-linear filters and neural networks have been employed to address these issues [25]. However, such approaches increase system complexity.

In this paper, a neural forecasting system has been proposed and expanded to anticipate the annual production of electrical energy at the European continent from seven distinct energy sources. This model should serve decision-makers, in preparing plans, strategies, transitions, reductions and increases within the energy policies of the European governments, and in understanding the size of the risk of their actions, as well as in taking adequate precautions to avoid major damages and losses. It will be demonstrated that the developed forecasting system can accurately model the trend of time series representing seven different types of electrical energy sources: oil, coal, wind, solar, nuclear, hydro and bioenergy.

Acknowledging limitations in current methods, we propose a robust modeling approach using Time Delay Neural Network (TDNN) topologies, known for their efficiency in short-term time series predictions [26] [27]. Employing TDNN topology as the foundational element for our extended forecasting system, we demonstrate that simple neural blocks with few neurons in one hidden layer can accurately forecast short time series within larger datasets. Illustrated with seven datasets of historical European electrical energy production, our approach showcases high accuracy and dependability for one-step-ahead and long-term predictions. This individual prediction extension exhibits notable accuracy and reliability, measured against well-known forecasting performance assessment metrics. Additionally, by tailoring our methodology to the peculiarities of energy production forecasting, we aim to provide a more precise and context-aware prediction process. Results show that our extended short-term ANN-based forecasting model achieves comparable accuracy to the latest state-of-the-art methodologies.

Neural network-based prediction models for electricity production find diverse applications in real-world scenarios. In renewable energy forecasting, they predict output from sources like solar and wind, aiding utilities in managing supply fluctuations and optimizing renewable energy use [28] [29]. For load forecasting, neural networks assist in planning resource allocation and maintenance scheduling, thus avoiding unnecessary costs [30] [31]. They're also integrated into power plant systems to predict equipment failures, reducing downtime and enhancing operational efficiency [32]. Traders optimize strategies using neural networks, and in smart grids, they predict consumption patterns, optimize distribution, and manage resilience [33-36]. In microgrid management, they predict local demand, ensuring reliable power supply in remote areas [37]. Industries use neural networks for energy efficiency by predicting consumption patterns [9] [38]. Finally, neural networks manage distributed resources in decentralized energy systems [39]. These applications showcase their role in improving decision-making, operational efficiency, and various aspects of electricity production and distribution. The specific use cases vary based on individual utility or organizational goals and needs.

The remainder of this paper is organized as follows: Section 2 analyzes European electrical energy production profiles, detailing time series datasets and the suggested neural network-based forecasting approach. Section 3 applies the methodology to seven datasets representing diverse electrical energy production types, comparing results with alternative methodologies like Extreme Learning Machine and ARIMA. Section 4 provides conclusions and includes several suggestions for further research.

2 Data and Methodology

2.1 Dataset

Seven different electricity production types and corresponding time series data will be the subject of this study and will be used for the development of the neural forecasting models. They are: coal, hydro, nuclear, oil, wind, solar and bioenergy-based types of electrical energy production. According to available data they represent the large majority of all sources of electrical energy in Europe [40] [41]. The residual part of the energy represents the energy that is generated from waves, geothermal sources, ocean, tidal energy etc.

The electrical energy production trends for different sources are shown in Fig. 1 [42]. They illustrate a slow and long-term mitigation from non-eco-friendly to eco-friendly types of energy sources at the European continent for 35 years. To evaluate a specific dataset, three groups of brief informative coefficients known as descriptive statistics are used [43]. They measure central tendency of the dataset, its variability, and its distribution. All datasets are analyzed for descriptive statistics, and the results are shown in Table 1. Obtained coefficients indicate that datasets express considerable variance and non-periodicity at the yearly level and are good candidates for forecasting methodology validation.

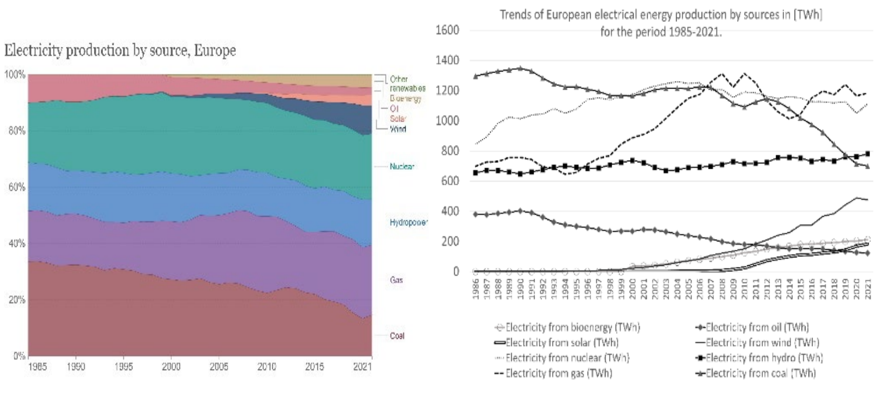


Figure 1
a) European electrical energy production profile over years [42], and b) Illustration of European electrical energy production time series by sources

Table 1
Descriptive statistics of data

Descrip. statistics measure	Electr. from bioenergy [TWh]	Electr. from oil [TWh]	Electr. from solar [TWh]	Electr. from wind [TWh]	Electr. from nuclear [TWh]	Electr. from hydro [TWh]	Electr. from coal [TWh]
Number of scores	22	36	32	44	57	57	36
Mean	125.46	331.62	41.64	103.87	792.73	644.44	1141.2
Median	129.25	256.36	1.99	18.24	1049.63	674.75	1186.9
25th percentile	68.53	161.18	0.06	0.38	248.02	621.34	1096.4
75th percentile	183.48	307.07	96.58	174.84	1158.07	716.21	1224.6
Interquartile range	114.95	145.89	96.52	174.46	910.05	94.85	128.19
Minimum	34.52	122.56	0.01	0.003	21.54	381.45	698.77
Maximum	211.05	3379.24	182.53	487.69	1258.56	787.68	1346.9
Range	176.53	3256.68	182.52	487.69	1237.02	406.23	648.13
Variance	3651.79	280254.5	3535.7	21949.2	214846.7	11111.7	28317
Std. deviation	60.43	529.39	59.46	148.15	463.51	105.41	168.28
Skew	-0.11	5.28	1.01	1.30	-0.65	-0.97	-1.23
Kurtosis	-1.57	27.51	-0.54	0.38	-1.36	-0.12	0.77

2.2 Methodology

The methodology proposed in this research is based on the application of artificial neural network forecasting using historical data and the modified ensemble learning concept [44]. In general, artificial neural networks perform complex nonlinear input-output transformations by imitating brain functions, which consist of over 86 billion neurons that receive environmental inputs, separate and recombine the most significant ones, and reason about the requirements and actions of the organism. ANNs consist of layers of neurons. Because of the non-linear nature of the activation function between its layers of processing components, they can be employed to solve a variety of problems. A neural network's performance is greatly influenced by the choice of activation function. Each epoch's error is determined by comparing the computed output of each input with the predicted output. A neural network optimization problem should be defined, with the primary objective of reducing error during the network training [45] [46].

The "wisdom of the crowd" suggests that a large group with average expertise can provide accurate forecasts, balancing out noise better than individual experts. This concept is applied to neural network-based AI through ensemble learning in ANN forecasting, where the combined outcome of multiple models often surpasses the accuracy of any single member for various tasks. Additionally, certain innovative models combine system dynamics with Bayesian degrees of truth to model both system behavior and the plausibility of each outcome [47].

The aim is to build an ANN forecasting model for estimating future power generation based on historical data. Despite various adjustments, individual models (referred to as "weak learners") lack satisfactory accuracy. However, these weak models can be valuable when combined into an ensemble. Applying the same data to each model for new predictions exploits their different individual learning approaches, effectively extending short-term learning. This combination balances out limitations, enhancing overall performance.

A time series is a collection of numerical data gathered successively over time. Short time series, characterized by frequency, volatility, and limited trend information, cannot result in statistically accurate predicting due to their small sample sizes (N). This study focuses on such time series and their forecasts, employing a technique [27] proven effective for handling short irregular series.

Creating a neural network architecture for time series prediction requires determining the number of layers and nodes in each layer. These features are often developed by experimenting with the existing data and there is no theoretical foundation for this. With the right number of hidden layer nodes and training dataset, neural networks with one hidden layer can simulate every nonlinear function. A Time Delay Neural Network block [48] is the neural network topology cell employed in this investigation. An extended neural network is built using this fundamental TDNN cell - TDNNC.

We employ TDNNC form of the ANN, where discretized input signals are time-shifted using delay elements before being sent to the neurons in the input layer. Typically, the topology requires a neural network with a single input and delaying components at the input layer as shown in Fig. 2.

Figure 3 describes the algorithm for the extended short-term prediction of time series. Main phases of the procedure include data preparation, networks construction, accuracy assessment, selecting the most accurate topologies. The data set for training a single neuron block requires a window of nine consecutive samples of the time series. For a dataset of N samples and prediction of the N^{th} moment, $(N+1)-9$ individual sets of nine successive samples are required. Successive mini sets are shifted by one moment in time. Corresponding network gives a prediction of the tenth step using nine historical samples. Each TDNNC is trained with a new shifted dataset, enabling one-step ahead forecasting in an arbitrarily large time series. Using simple, small neural structures (3 to 10 neurons in the hidden layer) easily trained with short sequences, we achieve independent prediction steps, avoiding error accumulations and constructing a modular forecasting system for electricity production.

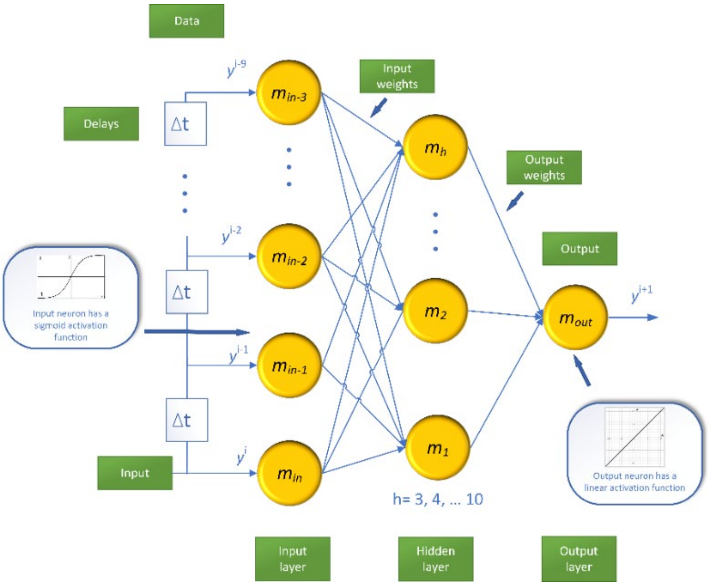


Figure 2
Basic neural network structure

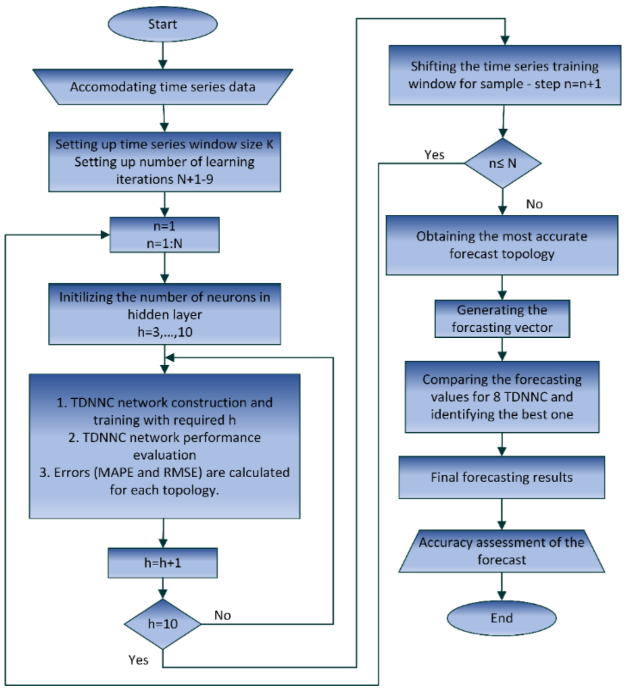


Figure 3
Extended short term neural prediction algorithm

The following formula describes the transformation function for one-step-ahead prediction:

$$y(i+1) = f(t_{i+1}) = \hat{y}(i+1) + \varepsilon \tag{1}$$

Here, $y(i+1)$ represents value of the initial i.e., training time series sample, $f(t_{i+1})$ stands for the transformation function of the next time instance, target forecasting of the next time instance is denoted with $\hat{y}(i+1)$, while ε is the forecasting error. More details about this block can be found in [49-51].

After testing with various delay block counts, due to most accurate results, it was chosen to employ nine prior observations for neural network training and prediction, along with nine neurons in the input layer. Neurons in the output layer have linear activation functions, and since performing one-step-ahead prediction is the main objective, just one output of the network and one neuron in the output layer are needed. During the ANN training, a variant of the steepest-descent minimization algorithm is employed.

Modelling and training of the individual TDNNC block can also be considered as an optimization problem. A step-by-step approach of this process is illustrated in the Algorithm 1 [52].

Algorithm 1: Training of the individual TDNNC block Initialize the network weights

1. In timestep t, deliver nine previous observations as an input vector $\boldsymbol{y}_t$ (y_{t-1}, y_{t-2}, ... y_{t-9}) to the input layer of the network:

$$\boldsymbol{y}_t = [y_{t-9}\, y_{t-8}\, y_{t-7}\, \dots\, y_{t-1}]^T$$

2. For each neuron in the hidden layer, compute the activations as:

$$a_i = S\left(\sum_{j=1}^{9} w_{ji}^{(in)}\, y_{t-i} + \theta_i^{(in)}\right)$$

or in a vector form:

$$a_i = S\left(\mathbf{w}_i^{(in)} \cdot \mathbf{y}_t + \theta_i^{(in)}\right)$$

Here:

- a_i is the activation of the i-th neuron in the hidden layer, while the vector of hidden neuron activations is:

$$\mathbf{a} = [a_1\, a_2\, \dots\, a_h]^T$$

- $w_{ji}^{(in)}$ is the weight connecting the j-th input and i -th hidden neuron,
- $\theta_i^{(in)}$ is the activation threshold value of each hidden neuron,
- $\mathrm{S(x)} = \frac{1}{1+\mathrm{e}^{-\mathrm{x}}} = \frac{\mathrm{e}^{\mathrm{x}}}{\mathrm{e}^{\mathrm{x}}+1}$ is the sigmoid activation function.

3. Compute the predicted observation ($\hat{y}$) at the network output:

$$\hat{y} = l\left(\sum_{i=1}^{h} w_i^{(out)}\, a_i + \theta^{(out)}\right)$$

or in a vector form

$$\hat{y} = l\left(\mathbf{w}^{(out)} \cdot \mathbf{a} + \theta^{(out)}\right)$$

Here, h is the number of neurons in the hidden layer,

- $w_i^{(out)}$ is the weight connecting the i-th hidden and the output neuron,
- $\theta^{(out)}$ is the activation threshold value of the output neuron,
- $l(x)$ is the linear activation function with arbitrary coefficients.

4. Calculate the prediction error as:

$$e_t = y_t - \hat{y}$$

5. According to the steepest descent method, for the next timestep $(t + 1)$ the weights should be updated by adding a value proportional to the gradient of the error function:

$$w_{ij_{t+1}} = w_{ij_t} + \eta \frac{\partial \varepsilon_t}{\partial w_{ij_t}}$$

Here:

- η is the "learning rate" i.e., the correction factor used to ensure the learning convergence,
- $\varepsilon(e_t)$ is the error function subject to the optimization process
- $\frac{\partial \varepsilon_t}{\partial w_{ij_t}}$ is the step of the steepest descent which can be calculated using the backpropagation algorithm

6. Shift the input values by one timestep and repeat the prediction procedure from step 2.
7. Repeat the above steps until the early stopping error condition was achieved in order to avoid overfitting.

2.3 Prediction Accuracy Measures

Root-Mean-Square-Error, or RMSE is one of the most often used metrics when training regression or time series models to measure the accuracy of the model's anticipated values vs. the actual or observed values. It illustrates how far away from the line of best fit the obtained data are. It can be a determination factor when selecting the best performing forecasting model from a group of models trained with the same dataset. It can be calculated as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}} \tag{2}$$

Here, N stands for the number of observations, and $\hat{y}_i$ and y_i represent the expected and the obtained value of the forecast, respectively.

Mean Absolute Percentage Error, or MAPE, is another performance indicator for the forecasting system that will be employed in this analysis. By averaging the absolute percentage errors of each value in a dataset, it demonstrates how accurate the predicted numbers were in comparison to the actual values. MAPE often allows for more efficient analysis of large datasets. The following equation can be used for determination of MAPE:

$$MAPE = \frac{100\%}{N} \sum_{i=1}^{N} \left| \frac{\hat{y}_i - y_i}{\hat{y}_i} \right| \tag{3}$$

Coefficient of determination or R^2 can have a value between 0 and 1. It indicates how accurately a forecasting model predicts the result. It serves as a criterion for the quality of fit. A model's forecast is better when it has a higher R^2 value. The formula for calculating R^2 is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2} \tag{4}$$

where $\bar{y}$ is calculated as the mean of the observed sample set.

3 Results and Discussion

As mentioned above, we proposed and developed an extended short-term prediction ANN model for electricity production forecasting. For reliable and accurate analysis of the results, we employed well-known error measuring methodologies. Seven different sources of electrical power generation were predicted for the European market: wind, solar, nuclear, hydro, oil, coal, and bioenergy. These predictions represent one-step-ahead annual forecasting values.

Different types of electricity production time series sizes ranged from 22 to 57 samples, depending on the "novelty" of the type and its initial introduction to the power production systems. In order to develop the suggested forecasting methodology, it is assumed that the values of the observed variable from the immediate past have the largest influence to the future prediction (nine previous samples are used to forecast the value in the tenth time-step). All data series were appropriately accommodated. They are also appropriately sequenced to train individual neural modules. A Sliding-window Time Series Cross-validation methodology was applied here, where entire dataset is divided into (N+1)-9 sets for iterative training and prediction of the 10th time step. First nine samples of the entire dataset are used for the prediction of the 10^{th} time step. After each iteration, the oldest sample is discarded, and the newest sample is added to perform the next neural network training and forecasting. This approach is applied here for two reasons: dataset size is huge, and the older observations are obsolete. This makes the length of each training set of nine samples (90% of the set used for training), while the 10^{th} sample is used to compare it to the obtained forecast of the 10^{th} sample (10% of the set used for testing), in order to check its accuracy against the expected value. Also, to obtain each one-step-ahead forecast for one type of electricity production, eight different neural modules were constructed and trained where number of the neurons in their hidden layers was altered from 3 to 10. After filling up the entire forecasting matrix, we have conducted the calculation of the performance measures for each type of the structure in order to select the most accurate ones. This is shown in Table 2.

The most accurate network, considering RMSE value of 3.28% was obtained for the coil data, while considering R^2, with the value of 0.989, the most accurate network was obtained for the wind energy data. At the other hand, the simplest network consists of four neurons in the hidden layer and predicts the bioenergy production, while the most complex neural structures consist of ten neurons in the hidden layer and extrapolate oil, wind and solar energy production data. The final forecast results obtained from a forecasting trend created by successive individual sliding window predictions were systematized in Fig. 4 a, c, e, g, i, k, m, along with their corresponding correct values. Relative errors of the forecasts on the observed prediction intervals are shown in Fig. 4 b, d, f, h, j, l, n. It can be seen that the recommended forecasting method performed quite well, keeping in mind that different sets of data have unique properties for each electricity type.

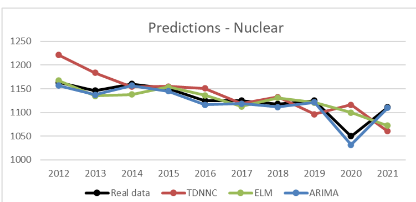
Table 2
Accuracy Measures of the forecasts

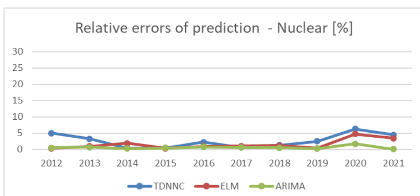
Most accurate neural structures	RMSE	MAPE	R^2
Nuclear - 8 hidden neurons	36.13	2.8%	0.944
Hydro - 5 hidden neurons	18.63	2.1%	0.531
Coil - 7 hidden neurons	3.28	2.5%	0.969
Oil - 10 hidden neurons	13.62	5.8%	0.978
Wind - 10 hidden neurons	17.00	5.6%	0.989
Solar - 10 hidden neurons	7.13	4.1%	0.935
Bioenergy -4 hidden neurons	4.86	2.2%	0.976

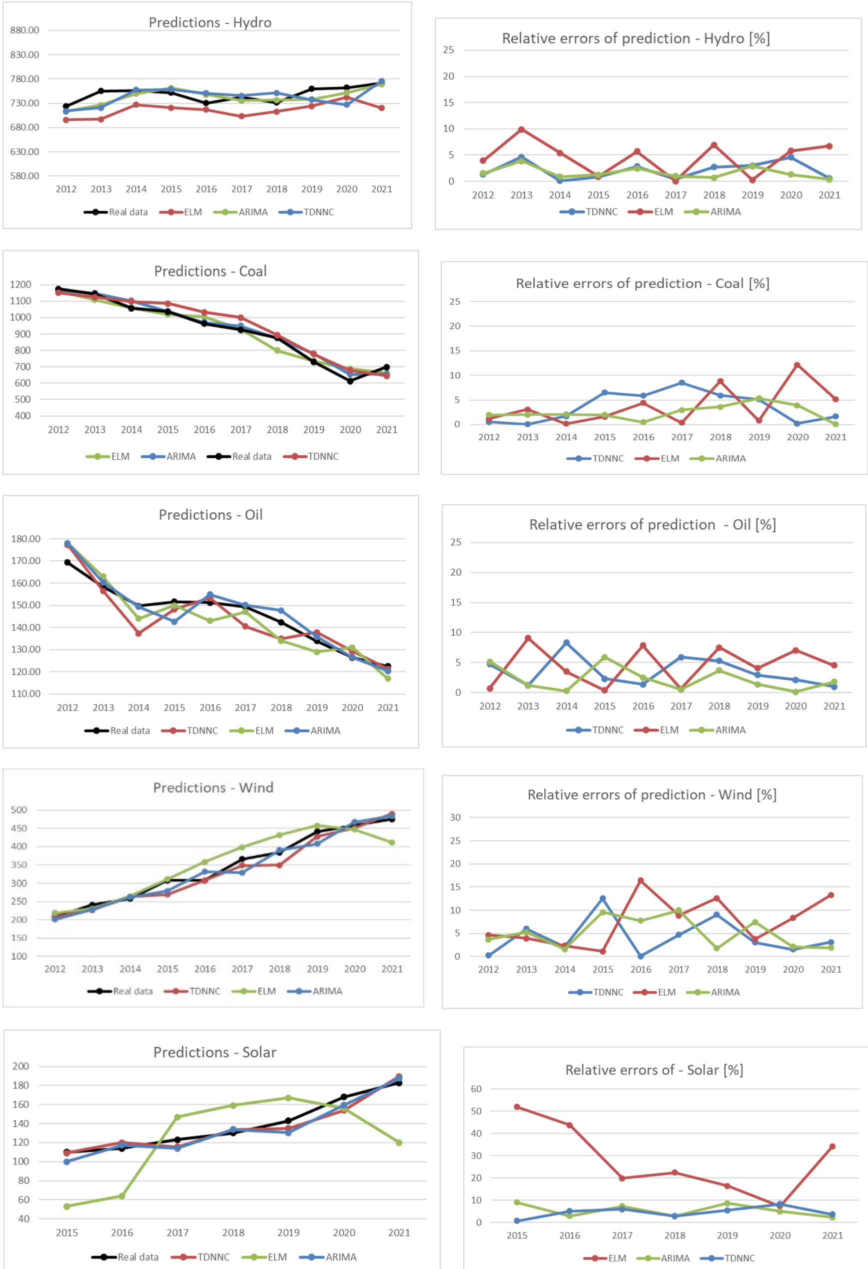
As was already said, we have only analyzed one parameter related to power generation during the entire research. It is the volume of electrical energy production at the European continent for seven different electrical energy source types. Since each form of electrical energy source encounters a unique set of impacts and various sorts of elements have an impact on its annual production, these datasets vary in different manners and are quite uncorrelated. As a result, we conclude that the chosen datasets may be utilized to evaluate the effectiveness of the proposed approach.

As shown in Table 2, the developed system obtained good prediction accuracy, which may be an indicator of the overfitting in the neural network training process. Overfitting can occur in the following situations [52]:

- Training data size is too small and does not contain enough data samples to accurately represent all possible input data values.
- The training data contains large amount of irrelevant information called noisy data.
- Overfitting due to prolonged training on a single sample set.
- Overfitting caused by high model complexity, where the model learns the noise within the training data.






Predictions - Hydro
880.00
830.00
780.00
730.00
680.00
630.00
580.00
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
Real data ELM ARIMA TDNNC
Relative errors of prediction - Hydro [%]
25
20
15
10
5
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
TDNNC ELM ARIMA
Predictions - Coal
1200
1100
1000
900
800
700
600
500
400
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
ELM ARIMA Real data TDNNC
Relative errors of prediction - Coal [%]
25
20
15
10
5
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
TDNNC ELM ARIMA
Predictions - Oil
180.00
170.00
160.00
150.00
140.00
130.00
120.00
110.00
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
Real data TDNNC ELM ARIMA
Relative errors of prediction - Oil [%]
25
20
15
10
5
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
TDNNC ELM ARIMA
Predictions - Wind
500
450
400
350
300
250
200
150
100
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
Real data TDNNC ELM ARIMA
Relative errors of prediction - Wind [%]
30
25
20
15
10
5
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021
TDNNC ELM ARIMA
Predictions - Solar
200
180
160
140
120
100
80
60
40
2015 2016 2017 2018 2019 2020 2021
Real data TDNNC ELM ARIMA
Relative errors of - Solar [%]
60
50
40
30
20
10
0
2015 2016 2017 2018 2019 2020 2021
ELM ARIMA TDNNC

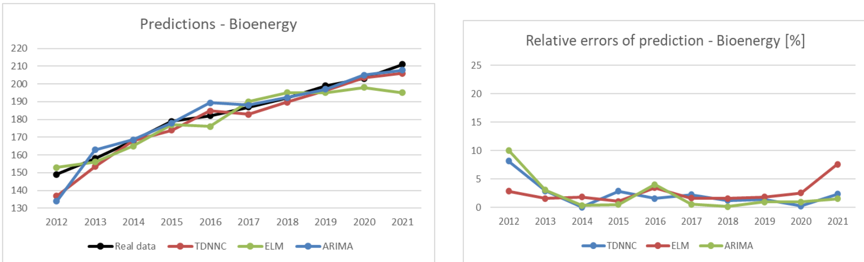


Figure 4
Prediction of different types of electrical energy production for the period 2012-2021 (2015-2021), and their relative prediction errors for three different methodologies:
a) Nuclear-based electrical energy and b) relative error of Nuclear energy prediction
c) Hydro-based electrical energy and d) relative error of Coal energy prediction
e) Coil-based electrical energy and f) relative error of Coal energy prediction
g) Oil-based electrical energy and h) relative error of Oil energy prediction
i) Wind-based electrical energy and j) relative error of Wind energy prediction
k) Solr-based electrical energy and l) relative error of Solar energy prediction
m) Bio-based electrical energy and n) relative error of Bio energy prediction

High accuracy of our predictions is not caused by overfitting since our data sets contain enough data, they were carefully prepared to avoid the presence of a large amount of irrelevant information, models were trained according to early stopping criteria while each neural network had at most ten neurons in the hidden layer.

Graphs in Fig. 5 illustrate the training progress of a typical neural network with 10 neurons in a hidden layer. Notably, despite its complexity, the network rapidly diminishes forecasting errors, showing a significant reduction even by the 5th iteration. This rapid error reduction highlights the efficiency of the training process, demonstrating the network's ability to learn and adapt quickly, ultimately achieving a highly accurate forecasting performance. Early stopping criteria here is achieved after 17 iterations with the prediction error for the particular step of 4.99269E-17.

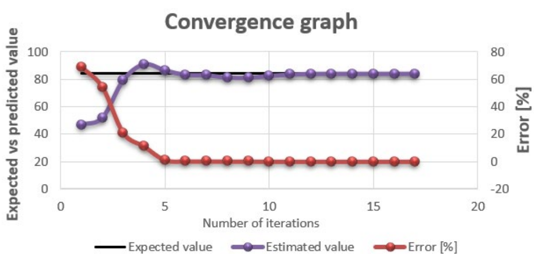


Figure 5
Learning speed of a typical TDNNC (10 hidden neurons) (Wind – based electrical energy prediction)

In order to assess the accuracy of the proposed methodology, two sets of alternative forecasting models have been developed, trained and verified for seven available datasets. These methodologies include a neural network based Extreme learning machine (ELM) forecasting [53] [54] and a traditional statistical Autoregressive integrated moving average (ARIMA) time series modelling [27] [55] [56]. Beside the TDNN forecasting results for the observed prediction intervals, Fig. 4 a, c, e, g, i, k, also shows results of the alternative forecasting methodologies. Corresponding relative errors of the European energy forecasts are shown in Fig. 4 b, d, f, h, j, l, n. Performance measures for all three methodologies are systematized in Table 3, and illustrated in Fig. 6. It should be emphasized that ARIMA models for the particular intervals were built using the entire available datasets. That is the reason for their better performance measures. However, these models do make mistakes in time series modelling even on the data that was used for the trend modelling. On the other hand, ELM approach could not always catch the trends of analyzed time series due to its simpler structure (10 neurons in the hidden layer). They would perform much better on larger datasets.

Table 3

Systematization of Prediction accuracy measures for TDNN, ELM, and ARIMA

Elec. energy type	**MAPE**			**RMSE [%]**			**R^2**		
	TDNN	ELM	ARIMA	TDNN	ELM	ARIMA	TDNN	ELM	ARIMA
Nuclear	2.68	1.54	**0.61**	29.87	22.60	**8.11**	0.64	0.48	**0.99**
Hydro	2.09	4.55	**1.61**	18.63	42.16	**14.43**	**0.53**	0.20	0.39
Coal	3.62	3.77	**2.43**	**4.28**	40.29	24.24	0.97	0.95	**0.99**
Oil	3.52	4.60	**2.25**	6.31	7.94	**4.59**	0.84	0.78	**0.93**
Wind	**4.24**	7.53	5.09	**17.00**	34.72	20.72	**0.97**	0.87	0.96
Solar	**4.64**	27.98	5.49	**7.62**	41.15	8.05	0.92	0.24	**0.94**
Bio	2.31	2.54	**2.24**	**5.08**	6.20	5.69	**0.97**	0.92	0.93

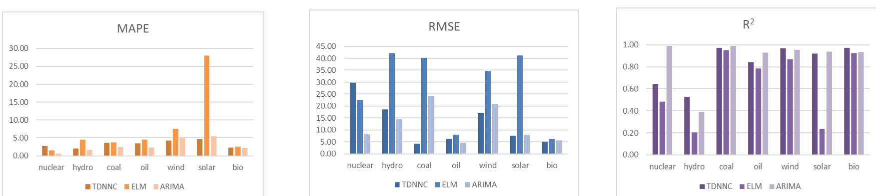


Figure 6

Prediction accuracy measures for TDNN, Extreme learning machine, and ARIMA long term energy forecasting methodologies

Results obtained with the proposed extended short-term prediction approach have various benefits for predicting the yearly power generation. This approach can find its purpose in business logic for balancing responsible parties regarding their strategy planning for participating in financial instruments markets (derivatives markets). For instance, they can more precisely calculate the "future amount of energy" they need to buy or sell at the beginning of the year, or on a quarterly basis.

It the same manner, the regulators can determine optimal level of subsidies for upcoming RES projects via feed-in tariff or market premiums, depending on how much energy can be anticipated from existing power plants based on the prediction model. The observed variables were quite prone to disturbances and a wide range of effects. Our system handled this issue well.

With our forecasting system, we have estimated the distribution of different types of electricity for Europe for the year 2022. This is shown in Fig. 7a. After acquiring the statistical measures for these parameters, relative errors of these out of sample forecasts for three methodologies could be evaluated. This is shown in Table 4, and in Fig. 7b.

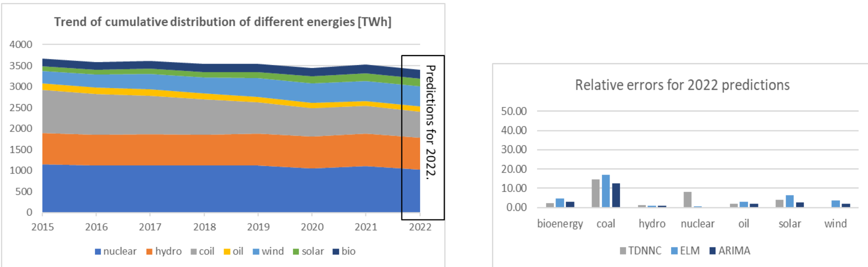


Figure 7

a) Trend of different energy types production distribution for the European continent. Values for the year 2022., represent the predictions achieved by the proposed forecasting system.

b) Relative errors of energy production predictions for 2022., for different methodologies

Table 4

Systematization of relative prediction errors for energy forecasting in 2022, for suggested TDNN, Extreme learning machine, and ARIMA forecasting methodologies

Electrical energy type	**Relative errors [%]**		
	TDNN	ELM	ARIMA
Bioenergy	**2.11**	4.76	3.03
Coal	14.53	17.04	**12.37**
Hydro	1.29	1.00	**0.96**
Nuclear	8.04	0.62	**0.10**
Oil	1.93	2.90	**1.79**
Solar	4.07	6.28	**2.65**
Wind	**0.12**	3.55	1.85

In summary, for small datasets, ARIMA models are usually computationally more efficient than a simple one-hidden-layer neural network. The choice between them depends on the specific characteristics of the data and the modeling requirements. If the time series patterns are simple and well-captured by ARIMA, it might be a more practical choice in terms of computational efficiency. However, if the data has complex patterns that ARIMA struggles to model, a small neural network like TDNN should be considered despite the higher computational cost.

We have also investigated some other solutions to the prediction problem related to electrical power production and consumption that are accessible in the current literature in order to assess the efficiency of the suggested technique. Table 5 summarizes these findings for seven different types of electrical energy sources. We have examined the accuracy of these approaches and the provided forecast results for different areas in the world.

Table 5
Systematization of the reported accuracies for latest state-of-the-art power production forecasting methods for various energy source types

Parameter	Methodology	Performance meas.		Ref.
		RMSE	MAPE [%]	
Solar energy output power from PV	Multivariate neural network ensemble framework	-	3.1	[57]
Wind speed prediction	Wavelet transformation and recurrent neural networks	1.21	0.93	[58]
Wind speed forecasting	Convolutional support vector machine	0.39	42.85	[59]
Oil and gas well production	ARIMA-LSTM model	11.584	0.144	[60]
Nuclear energy consumption	An optimized structure-adaptative grey model	5.889	2.094	[61]
Monthly peak energy demand in India	Fb Prophet models	4.23	3.3	[62]
Bioenergy power generation	Grey compositional data model	-	3.7	[63]
Biomass-based energy potential	Artificial intelligence and geographic information forecasting systems	-	-	[64]

Although we are aware that the datasets and methodologies, we have examined vary from one solution to the next, the most crucial performance metrics of our methodology, such as MAPE, are comparable to the most recent state-of-the-art forecasting power production methodologies.

Conclusions

This research proposes a unique way for extending the short-term prediction horizon, for various ANN-based forecasting systems. Any time-series dataset, can be “sliced” in windows that shift in time, and be used to train simple neural structures to perform one-step-ahead prediction. In our study we applied this modular neural topology to seven different electrical energy production sources, for Europe. Such modular neural structures were developed and assessed for their accuracy.

One may draw the conclusion that the results of case studies involving seven different sources of electrical energy demonstrate the ability of such individual predictions to anticipate the direction of future changes with high accuracy and reliability. The forecast's precision and reliability even surpassed some of the latest state-of-the-art forecasting methodologies. We may anticipate the method's applicability being broadened to other AI forecasting approaches, as well as other types of ANN topologies. Our future research will concentrate on further error reductions and a discussion concerning the smallest number of neurons in the ANN layers required to achieve the objective.

Acknowledgement

This research was partially funded by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia.

References

[1] L. Grigsby. Electric Power Generation, Transmission, and Distribution, 3rd Edition, Boca Raton, Florida, CRC Press, 2018

[2] A. J. Dangerman, and H. J. Schellnhuber. Energy systems transformation, *Proceedings of the National Academy of Sciences*, 110(7): E549-E558, 2013

[3] Z. Liu, P. Jiang, L. Zhang, X. Niu. A combined forecasting model for time series: Application to short-term wind speed forecasting, *Applied Energy* 259: 114-137, 2020

[4] C. Tian, Y. Hao, J. Hu. A novel wind speed forecasting system based on hybrid data preprocessing and multi-objective optimization, *Applied Energy* 231: 301-319, 2018

[5] C. Zuluaga, M. Álvarez, and E. Giraldo. Short-term wind speed prediction based on robust Kalman filtering: an experimental comparison, *Applied Energy* 156: 321-330, 2015

[6] R. Li, Y. Jin. A wind speed interval prediction system based on multi-objective optimization for machine learning method, *Applied Energy*, 228: 2207-2220, 2018

[7] N. Son, and L. Vinh. Parameter Estimation of Photovoltaic Model, Using Balancing Composite Motion Optimization, *Acta Polytechnica Hugarica*, *19*(11): 27-46, 2022

[8] J. I. Pena, and R. Rodriguez. Are EU's Climate and Energy Package 20-20-20 targets achievable and compatible? Evidence from the impact of renewables on electricity prices, *Energy*, 183: 477-486, 2019

[9] Energy and Climate Policies beyond 2020 in Europe (Online, accessed on 30 Jan. 2023), Available https://ens.dk/sites/ens.dk/files/Globalcooperation/eu_energy_and_climate_policy_overview.pdf

[10] European Energy Transition 2030: The Big Picture, (Online, accessed on 30 Jan. 2023), Available: https://www.agora-energiewende.de/fileadmin/Projekte/2019/EU_Big_Picture/153_EU-Big-Pic_WEB.pdf

[11] Trends and projections in Europe 2021, Report, EEA Report No 13/2021, Available: https://www.eea.europa.eu/publications/trends-and-projections-in-europe-2021/download, (accessed on 30 January 2023)

[12] Implementation of the Energy Efficiency Directive (2012/27/EU): Energy Efficiency Obligation Schemes, (Online), Available: https://www.europarl.europa.eu/RegData/etudes/STUD/2016/579327/EPRS_STU(2016)579327_EN.pdf, (accessed on 30 January 2023)

[13] G. Cerne, D. Dovzan, and I. Skrjand. Short-Term Load Forecasting by Separating Daily Profiles and Using a Single Fuzzy Model Across the Entire Domain, *IEEE Trans. on Industrial Electronics*, 65(9): 7406-7415, 2018

[14] P. Kapler. Forecasting of residential power consumer load profiles using a rype-2 fuzzy inference system, *Acta Polytechnica Hungarica,* 19(9), 2022

[15] R. Precup, G. Duca, S. Travin, and I. Zinicovscaia. Processing, neural network-based modeling of biomonitoring studies data and validation on republic of Moldova data. *Proc. of the Romanian academy series A*, 23(4): 403-410, 2022

[16] R. R. Hedrea, and E. Petriu. Evolving fuzzy models of shape memory alloy wire actuators, *Science and Technology*, 24(4): 353-365, 2021

[17] C. Pozna, and R. E. Precup. Aspects concerning the observation process modelling in the framework of cognition processes, *Acta Polytechnica Hungarica*, 9(1): 203-223, 2012

[18] H. Sharadga, S. Hajimirza, and R. S. Balog. Time series forecasting of solar power generation for large-scale photovoltaic plants, *Renewable Energy*, 150(C): 797-807, 2019

[19] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, and M. Cony. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning, *Solar Energy*, 84(10):1772-1781, 2010

[20] R. Rick, and L. Berton. Energy forecasting model based on CNN-LSTM-AE for many time series with unequal lengths, *Engineering Applications of Artificial Intelligence*, 113, 2022

[21] L. Zhang, W.-D. Zhou, P.-C. Chang, J.-W. Yang and F.-Z. Li, Iterated time series prediction with multiple support vector regression models, *Neurocomputing*, 99: 411-422, 2013

[22] A. Mellit, A. Pavan, and V. Lughi. Deep learning neural networks for short-term photovoltaic power forecasting, *Renewable Energy*, 172: 276-288, 2021

[23] A. Sorjamaa, J. Hao, N. Reyhani, Y. Ji and A. Lendasse. Methodology for long-term prediction of time series, *Neurocomp.*, 70(16): 2861-2869, 2007

[24] S. Taieb, A. Sorjamaa and G. Bontempi. Multiple-output modeling for multi-step-ahead time series forecasting, *Neurocomp.* 73(10): 1950-1957, 2010

[25] X. Wu, Y. Wang, J. Mao, Z. Du and C. Li. Multi-step prediction of time series with random missing data, *Appl. Math. Model.*, 3(14): 3512-3522, 2014

[26] J. Milojković. Prediction in Electronics Using Artificial Neural Networks Based on Limited Information, PhD Thesis, University of Niš, Faculty of Electronic Engineering, Niš, 2010

[27] M. Milić, J. Milojković, I. Marković, and P. Nikolić. Concurrent, Performance-Based Methodology for Increasing the Accuracy and Certainty of Short-Term Neural Prediction Systems, *Comp. Intell. and Neurosci*, 1687-5265, 2019

[28] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng. A review of deep learning for renewable energy forecasting, *Energy Conversion and Manag*, 198, 2019

[29] C. Goncalves, P. Pinson, and R. Bessa, Towards data markets in renewable energy forecasting, *IEEE Trans. on Sust. Energy*, 12(1): 533-542, 2020

[30] I. Nti, M. Teimeh, O. Nyarko-Boateng, and A. Adekoya. Electricity load forecasting: a systematic review, *Journal of Electrical Systems and Information Technology*, 7(1): 1-19, 2020

[31] L. Zhang, J. Wen, Y. Li, J. Chen, Y. Ye, Y. Fu, and W. Livingood. A review of machine learning in building load prediction, *Applied Energy*, 285, 2021

[32] A. Betti, E. Crisostomi, G. Paolinelli, A. Piazzi, F. Ruffini, and M. Tucci. Condition monitoring and predictive maintenance methodologies for hydropower plants equipment, *Renewable Energy*, 171, 246-253, 2021

[33] H. Kim, J. Lee, S. Bahrami, and V. Wong. Direct energy trading of microgrids in distribution energy market, *IEEE Trans. on Power Systems*, 35(1): 639-651, 2019

[34] E. Soto, L. Bosman, E. Wollega, and W. Leon-Salas. Peer-to-peer energy trading: A review of the literature, *Applied Energy*, 283, 2021

[35] A. Rosato, M. Panella, R. Araneo, and A. Andreotti. A neural network based prediction system of distributed generation for the management of microgrids, *IEEE Trans. on Industry Applications*, *55*(6): 7092-7102, 2019

[36] F. Mohammad, and Y. C. Kim. Energy load forecasting model based on deep neural networks for smart grids, *International Journal of System Assurance Engineering and Management*, 11, 824-834, 2020

[37] Y. Du, and F. Li, Intelligent multi-microgrid energy management based on deep neural network and model-free reinforcement learning, *IEEE Transactions on Smart Grid*, 11(2), 1066-1076, 2019

[38] D. Narciso, and F. Martins. Application of machine learning tools for energy efficiency in industry: A review. *Energy Reports*, 6, 1181-1199, 2020

[39] S. Teng, M. Touš, W. Leong, B. How, H. Lam, and V. Máša. Recent advances on industrial data-driven energy savings: Digital twins and infrastructures, *Renewable and Sustainable Energy Reviews*, 135, 2021

[40] D. Medved, L. Beňa, and R. Tailor. Energy Storage System Utilization, in a Distribution Power System, *Acta Polytechnica Hungarica*, 20(11), 2023

[41] B. Shaheen, A. Abu Hanieh, and I. Németh. Fault detection of a wind turbine's gearbox, based on power curve modeling and an on-line statistical change detection algorithm, *Acta Polytechnica Hugarica,* 2021

[42] Electricity Production by Source. Available online (accessed on 10th November 2023): https://ourworldindata.org/grapher/electricity-prod-source-stacked,

[43] J. Jones, and J. Goldring, Exploratory and Descriptive Statistics; SAGE Publications Ltd.: London, UK, 2022

[44] Z. Zografski. A novel machine learning algorithm and its use in modeling and simulation of dynamical systems, in Proc. of 5th Annual European Computer Conference, COMPEURO '91, Germany, pp. 860-864, 1991

[45] J. Milojković, and V. Litovski. Dynamic Short-Term Forecasting of Electricity Load Using Feed-Forward ANN, *Eng. Intell. Syst. Electr. Eng. Commun*, 17(1):39–48, 2009

[46] E. Hedrea, R. Precup, R. Roman, and E. Petriu. Tensor product-based model transformation approach to tower crane systems modeling, *Asian Journal of Control*, 23(3): 1313-1323, 2021

[47] C. Pozna, R. Precup, J. Tar, I. Škrjanc, and S. Preitl. New results in modelling derived from Bayesian filtering, *Knowledge-Based Systems*, 23(2): 182-194, 2010

[48] P. Zhang, E. Patuwo, and M. Hu. Forecasting with artificial neural networks, *International Journal of Forecasting*, 14(1):35-62, 1998

[49] J. Milojković, S. Bojanić, O. Nieto Taladriz, and V. Litovski. One-Week- and One-Month-Ahead Prediction of Suburban Electricity Load, in *Proc. of Small Systems Simulation Symposium*, Niš, Serbia, pp. 104-111, Feb. 2016

[50] J. Milojković, D. Topisirović, M. Milić, and M. Stanojević. Short term local road traffic forecast using feed-forward and recurrent ANN, *Facta Universitates, Working and Living Environ. Protection*, 13(1):1-12, 2016

[51] J. Milojković, M. Milić, and V. Litovski. ANN model for one day ahead Covid-19 prediction, in *Proceedings of the Conference IcETRAN*, Novi Pazar, Serbia, pp. 296-300, 2022

[52] D. Rumelhart, G. Hinton, and R. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088): 533-536, 1986

[53] M. Jeremić, M. Gocić, J. Milojković, M. Milić. A deep learning approach for hydrological time-series prediction with ELM model, *Proc. of the 9th Small Systems Simulation Symposium, Niš,* 61-65, 2022

[54] K. Mohammadi, S. Shamshirband, S. Motamedi, D. Petković, R. Hashim, and M. Gocić. Extreme learning machine based prediction of daily dew point temperature, *Computers and Electronics in Agriculture*, 117, 214-225, 2015

[55] T. Naylor, T. Seaks, and D. Wichern. Box-jenkins methods: an alternative to econometric models, *Int. Statistical Review/Revue Internationale De Statistique*, 40(2): 123-137, 1972

[56] H. Hejase and A. Assi. MATLAB-assisted regression modeling of mean daily global solar radiation in Al-Ain, UAE, in *Eng. Education and Research Using MATLAB*, IntechOpen, London, UK, 2011

[57] M. Raza, N. Mithulananthan, and A. Summerfield Solar output power forecast using an ensemble framework with neural predictors and Bayesian adaptive combination, *Solar Energy*, 166:226-241, 2018

[58] H. Liu, X. Mi, and Y. Li. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short-term memory neural network and Elman neural network, *Energy conversion and management*, 156:498-514, 2018

[59] X. Mi, H. Liu, and Y. Li. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine, *Energy conversion and management*, 180:196-205, 2019

[60] D. Fan, H. Sun, J. Yao, K. Zhang, X. Yan, and Z. Sun. Well production forecasting based on ARIMA-LSTM model considering manual operations, *Energy*, 220, 2021

[61] S. Ding, Z. Tao, H. Zhang, and Y. Li. Forecasting nuclear energy consumption in China and America: An optimized structure-adaptative grey model, *Energy*, 239, part A, 2022

[62] S. Chaturvedi, E. Rajasekar, S. Natarajan, and N. McCullen. A comparative assessment of SARIMA, LSTM RNN and Fb Prophet models to forecast total and peak monthly energy demand for India, *Energy Policy*, 168, 2022

[63] K. Zhang, K. Yin, and W. Yang. Predicting bioenergy power generation structure using a newly developed grey compositional data model: A case study in China, *Renewable Energy*, 198:695-711, 2022

[64] A. Senocak, and H. Goren. Forecasting the biomass-based energy potential using artificial intelligence and geographic information systems: A case study, *Science and Technology, an International Journal*, 26, 2021