MDPI

Article

Power Factor Modelling and Prediction at the Hot Rolling Mills' Power Supply Using Machine Learning Algorithms

Manuela Panoiu * , **Caius Panoiu** and **Petru Ivascanu**

Department of Electrical Engineering and Industrial Informatics, University Polytechnica Timisoara, 300006 Timișoara, Romania; caius.panoiu@fih.upt.ro (C.P.); petru.ivascanu@student.upt.ro (P.I.)
* Correspondence: manuela.panoiu@fih.upt.ro

Abstract: The power supply is crucial in the present day due to the negative impacts of poor power quality on the electric grid. In this research, we employed deep learning methods to investigate the power factor, which is a significant indicator of power quality. A multi-step forecast was developed for the power factor in the power supply installation of a hot rolling mill, extending beyond the horizontal line. This was conducted using data obtained from the respective electrical supply system. The forecast was developed via hybrid RNN (recurrent neural networks) incorporating LSTM (long short-term memory) and GRU (gated recurrent unit) layers. This research utilized hybrid recurrent neural network designs with deep learning methods to build several power factor models. These layers have advantages for time series forecasting. After conducting time series forecasting, qualitative indicators of the prediction were identified, including the sMAPE (Symmetric Mean Absolute Percentage Error) and regression coefficient. In this paper, the authors examined the quality of applied models and forecasts utilizing these indicators, both in the short term and long term.

Keywords: rolling mill; power factor; prediction; deep learning; RNN; LSTM

MSC: 68T07

1. Introduction

The transition from semi-automation to knowledge automation, process intelligence, and production information is transforming metallurgy, especially steel rolling processes. Rolling is crucial to steel production because it affects material properties and product quality. The rolling process is nonlinear and unbalanced, presenting numerous challenges. High-speed continuous rolling mills make monitoring process information, modelling behaviour characteristics, and implementing high-speed operational controls extremely difficult. In rolling mills, maintaining high power quality is critical for guaranteeing reliable equipment operation, consistent product quality, accurate process control, energy efficiency, reduced downtime, and compliance with standards and regulations. Poor power quality can cause equipment failures, safety hazards, and increased operating costs, highlighting the importance of a stable and reliable power supply in metal production [1].

Power quality is currently a critical issue, particularly in plants with high power consumers that are occasionally unbalanced [2]. These consumers can be found in metallurgical plants like electric arc furnaces (EAF) [3,4], ladle furnaces [5], and rolling mills [1]. These high power consumers use a lot of electricity and can have an impact on power quality via reactive power, low power factor, and electric current harmonics [6–8]. The largest consumers of electric power in steel plants are EAF [5,9,10], but rolling mills also have significant power and produce harmonic currents [11,12], overcurrents that cause reactive power and low power factor [13,14] or flicker [12].

Power quality affects hot rolling performance and lamination product quality. Voltage and frequency affect hot rolling mills. Significant variations may compromise equipment

operation and cause production losses or damage [15,16]. A consistent voltage and frequency are essential for rolling efficiency and precision. Power quality has a substantial impact on the efficiency of hot rolling operations and the quality of laminated products. Low power quality might result in frequent device failures and operational interruptions. Fluctuations in voltage, such as sags and surges, can lead to the failure of critical control systems, resulting in interruptions to the rolling process [15,17]. This impacts production and might result in inconsistent thickness or defects in the surface of the rolled products caused by unexpected stops and accelerations.

Harmonic disturbances in the electrical grid may damage electronic and electrical equipment [18]. Harmonics can cause overheating in frequency or lamination control problems in hot rolling mills.

Power losses or outages can induce unexpected shutdowns, causing equipment failure. Process continuity may require interruption management [15,16,18].

Overvoltage may affect the rolling process. The quality of power can affect the process energy efficiency [17]. Overvoltage can damage drives, motors, and control systems immediately or progressively. Effects may result in failure of equipment, causing cost repairs or replacements and unexpected interruption [19]. Motor insulation breakdown resulting from overvoltage may cause short or open circuits, interrupting rolling motion. Control systems that control roll speed, pressure, and alignment influence rolling precision. High voltage may interfere with these control systems, resulting in imprecise rolling process control. This could lead to rolling defects in the product including thickness or surface irregularities, decreasing quality. A quality energy source can minimize operational expenses and improve energy efficiency [5,12].

Rolling process control systems may be affected by power grid oscillations. A reliable energy source is needed for precise process control.

Power quality affects product quality. Rolling instabilities can cause product defects, unequal dimensions, and other quality issues. Power quality issues might cause device malfunctions. If the voltage at the point of common coupling (PCC) drops for over 50 milliseconds, it will cause the variable frequency drive to shut down and interrupt the process [19]. This results in emergency situations, faults, and the under-release of products [20]. Rolling mills' variable frequency drives are susceptible to voltage drops. Voltage drops have significant repercussions for production. Malfunctioning of the electric drive in the technical installation causes interruptions in the technological process, resulting in equipment breakdowns, rejects, and the underproduction of products [18].

Monitoring and managing equipment power quality are crucial for hot sheet operation and product quality. Harmonic filters and surge protection devices can improve electrical quality and reduce disturbances in the rolling process and completed product [12,16,18].

Harmonic currents are caused in large part by the widespread use of power converters, which are found in many electrical and electronic systems, including household appliances. While the power of household appliances is relatively low [21,22], harmonic currents cause significant power losses in industrial electrical installations due to the very high power [15–18]. As a result, many papers in specialized journals present research on the presence of harmonic currents, load unbalances, or overcurrents. In recent years, there has been a significant amount of research conducted on power quality and harmonic currents, focusing on the utilization of artificial intelligence and machine learning techniques [22–30]. Our research also focuses on utilizing deep learning and recurrent neural networks (RNN) for predicting the power factor in hot rolling mils factories. We developed an algorithm employing deep learning methods to forecast the power factor in a hot rolling mill's power supply system. The method used is a hybrid RNN model that includes LSTM and GRU layers. The forecasts' effectiveness was evaluated using qualitative indicators such as sMAPE and regression coefficients, highlighting the accuracy of the models in predicting both short-term and long-term results.

We developed an algorithm using deep learning techniques to predict the power factor in the power supply system of a hot rolling machine. The approach involves a hybrid RNN

model of LSTM and GRU layers. The effectiveness of the forecasts has been evaluated by qualitative measures including sMAPE and regression coefficients, demonstrating the accuracy of the models in forecasting short-term and long-term predictions.

The paper's contributions are summarized as follows:

- Measurements were conducted in the electrical equipment powering a hot rolling mill. Measurements provided datasets for voltages, currents, active and reactive powers, as well as active and reactive energy;
- Deep machine learning algorithms were developed to train a hybrid recurrent neural network for forecasting power factor in a hot rolling mill's power supply system.
- The datasets from the measurements were utilized to train the hybrid RNN with various parameter values;
- The power factor forecasting results were analyzed using quantitative metrics such as RMSE, *MAEMAE*, and R-squared.

The paper is organized as follows: The first section provides an overview of the challenges associated with power quality in hot rolling mill power supplies. Section 2: The literature review summarizes some studies, primarily from the metallurgical industry, which address power quality issues; additionally, this section presents related works that investigate and address energy quality improvements using artificial intelligence-based solutions. Section 3: Materials and Methods describes the authors' proposed method for power factor forecasting. Section 4: Results describes the power factor forecasting results. Section 5: Discussions presents the obtained results with various prediction parameter configurations, while Section 6: Conclusions provides research conclusions.

2. Literature Review

In this section, some research that aims to study the impact of industrial power plants, such as those used in rolling mills, on power quality are briefly presented.

The design and analysis of an expanding steel plant's power supply system are presented in paper [13]. The plant has, in addition to the EAF, a hot strip rolling mill system. It has been noticed that many large loadings in the steel plant, such as the hot strip rolling mill and induction motor starting, may cause serious voltage fluctuation problems. The conclusion is that if proper capacitor banks and SVC are installed, the voltage fluctuation caused by hot strip rolling operation can be kept within 5% and the power factor above 0.95 [13].

Paper [6] focuses on harmonic and electromagnetic interference and power factor at a rolling mill factory and concludes that the 5th, 7th, 11th, and 13th harmonics are the primary sources of power pollution.

The articles [8,31] consider the use of a static synchronous compensator (STATCOM) for dynamic reactive power management of a hot rolling mill plant, as well as active harmonic mitigation. Paper [8] proposes a strategy for managing reactive power flow, reducing active power loss, and controlling the voltage at the coupling point, as well as reducing harmonics and reducing unscheduled shutdowns in the event of a trip in the passive filtering system.

The focus of paper [32] is on high-power frequency converters with active rectifiers from steel plants. Voltage dips of 15–30% lasting 150–300 ms cause rolling mill main electric drives to shut down. It was demonstrated that the frequency converter with active rectifiers could operate during voltage surges caused by switching high harmonic filters and voltage dips caused by switching the furnace transformer.

The study in [33] centers on power factor improvement, particularly in rolling mills with variable loads, using thyristor switch modules combined with a series-connected detuned reactor and capacitor. It details the design and implementation of a thyristor-based automatic power factor correction unit for three-phase industrial circuits designed to achieve near-unity power factor.

In [30], the focus is on applying intelligent techniques to optimize metal rolling control, highlighting the process's complexity due to its multi-scale, multi-variable, nonlinear

nature and the challenges posed by high-speed operations. The study concludes with an emphasis on the increasing trend toward more sustainable, intelligent, and efficient control systems in metal rolling.

Research in [34] introduces an advanced simulation model for analyzing voltage quality in an industrial power supply system and rolling mill electric drives equipped with active rectifiers. This model, developed in Matlab-Simulink, accounts for complex resonances in the network's frequency response. Its accuracy is validated by comparing simulation outcomes with operational data from a metallurgical plant. The model serves as a tool for designing and enhancing power systems and electric drives' electromagnetic compatibility.

Hot strip mills' abnormal operation during filtering system switching is simulated in [20]. Various steel milling parameters predict the key electrical variables' evolution. Voltage stability and distortion at the point of common coupling (PCC), reactive power flows upstream, harmonic current prediction, and new filter and rolling stand operating conditions under contingencies are examined. The methodology provides the operator with expected electrical variables to help determine if the hot rolling mill can handle unexpected filter bank switching states. This forecast may prevent unscheduled stops, reducing production and costs.

In paper [16], a typical industrial load represented by an electric arc furnace and a rolling mill is investigated for load model construction and mechanism analysis using actual measured voltage and current waveforms. The simulation model results are compared to the measured waveforms, and the harmonic content of the model current is investigated in order to demonstrate the model's dependability and accuracy.

Recent grid connection circuit advancements for rolling mills' main AC REDs (AC regenerative electric drives) are investigated in [16]. Matlab/Simulink simulations calculated THD factors up to the 60th harmonic. Researchers and engineers may develop and ensure the electromagnetic compatibility of nonlinear consumers in similar circuits using the results. Comparisons of powerful AC RED 6-, 12-, and 18-pulse connection circuits with three-level AFE's PPWM and SHE algorithms are shown. The results assist in selecting the optimal connection circuit and algorithm.

The preceding articles in this section illustrate the various issues that can occur in the power supply to hot rolling mills. The following are several research studies that employ intelligent solutions to investigate and address enhancements in energy quality.

In paper [23], the deep learning method LSTM predicts future harmonics in a power grid. Voltage harmonics were predicted using LSTM, a deep learning method. The deep learning algorithm effectively identified harmonic features, resulting in accurate forecasting and classification.

Total harmonic distortion (THD) is an important measure for evaluating power quality; however, predicting THD is difficult. This issue is addressed in the work [24] by developing a harmonic characteristics detection experiment and employing an artificial intelligence algorithm. The simulation results reveal that the suggested technique outperforms BP and GRNN (generalized regression neural network) in prediction accuracy, reaching 95.48%.

The research [25] presents a fuzzy approach to estimating voltage and current total harmonic distortions (THD) and assessing their power quality effects.

The fuzzy approach uses average THD indices to diagnose power quality using well-known standards. The proposed system was tested in a lab for power quality disturbances caused by nonlinear loads.

In research [26], an artificial neural network (ANN) system using location-specific data is used to estimate solar PV inverter harmonic distortions. A simple power system is modelled and simulated for various scenarios to train the ANN system and enhance prediction. The approach computed harmonic components with a maximum inaccuracy of 10% and a median of 5.4%.

The research [27] presents a thorough assessment of the progress developed in using DL for forecasting PQ indices time series, revealing that this field is still developing. For this

scenario, an LSTM network is proposed to predict the steady state of PQ indices time series, which assesses the current distortion at the point of common coupling (PCC) of a residence.

Deep machine learning-based algorithms and novel data augmentation are used to forecast flicker, voltage dip, harmonics, and interharmonics from highly time-varying electric arc furnace (EAF) currents and voltage in [28]. The prediction aims to reduce response and reaction time delays in electric arc furnaces (EAF)-specific active power filters (APFs). Three strategies existed. A low-pass Butterworth filter and linear finite impulse response (FIR) or long short-term memory network (LSTM) are utilized in two of them. In the third method, a deep CNN and LSTM network filter and predict simultaneously. The Butterworth and linear prediction, Butterworth and LSTM, and CNN and LSTM approaches yield 2.06%, 0.31%, and 0.99% dqo (direct-quadrature-zero) component prediction errors for a 40 ms prediction horizon.

Machine learning approaches were also employed for predicting THD in the research [29]. The model was developed with the ANN GMDH (group method of data handling) technique.

In [35], four forecasting models were investigated with 3x3 SOM maps to predict power quality parameters.

A method for predicting power factor variations in three-phase electrical power networks is presented in study [36], which makes use of machine learning techniques. Three main linear regression algorithms were used: ordinary least square (OLS), polynomial (Poly) and random forest (RF). Evaluation metrics such as mean absolute error (*MAE*), mean standard error (MSE), root mean square error (*RMSE*), and R squared were calculated.

The forecast of energy efficiency (EE), power factor (PF), and carbon emission was the main focus of the paper described in the reference [37]. Multi-gene genetic programming (MGGP), least square–support vector machine (LS-SVM), and fuzzy logic are the three soft computing approaches that are utilized in this study to model the energy efficiency (EE), power factor (PF), and carbon emission for a machine tool. Other techniques include fuzzy logic. The performance of the models was evaluated based on a number of statistical measures, including the coefficient of determination (R^2), the mean absolute error, the root mean square error, the mean square error, the sum of squared error, and the relative percentage error. The comparative performance evaluation of the models showed that the LS-SVM consistently outperforms the MGGP and fuzzy logic.

3. Materials and Methods

In a metallurgical factory, measurements were performed in the power supply installation of a hot rolling mill. Europrofiles' laminator has a designed capacity of 400,000 tons per year. It was equipped with continuous lamination equipment consisting of 12 stands organized in tandem. These measurements were important for determining the quality of the power. There are several power substations located in the station where the measurements were taken, and each one of them is connected to the national power grid by means of a transformer rated at 110 kV and 6 kV. The measurements were taken in a substation that supplies a hot rolling line with electricity.

Figure 1 depicts an electrical power supply system. The measurements were taken at Lines 1 and 2 of ST1, the power supply station that supplies the rolling mill. The experimental measurements were performed using a three-phase power analyzer C.A. 8334. The power analyzer was used to record the RMS and THD values of electric currents and voltages, as well as the phase, total active and reactive power, and power factor. The parameter values are recorded at the rate of one complete set of parameters per second, with a one-hour recording period. The power analyzer's sampling frequency is 12.8 KHz/channel, allowing for an analysis of 50 orders of current and voltage harmonics.

The RMS values of the phase and line voltages, as well as the voltage's THD, were also recorded. These recordings were made both with and without the SVC compensator turned on.

Figure 2 displays the active and reactive power values measured throughout the specified time period. The presence of high reactive power values, as anticipated due to the absence of SVC, is evident. This is also represented in the power factor values depicted in Figure 3. The low power factor is a drawback due to the diminished energy efficiency it entails. Figure 4 displays an overall active energy consumption and reactive energy use. It is seen that the reactive energy demonstrates significantly high values, which is also evident in the correspondingly low values of the power factor.
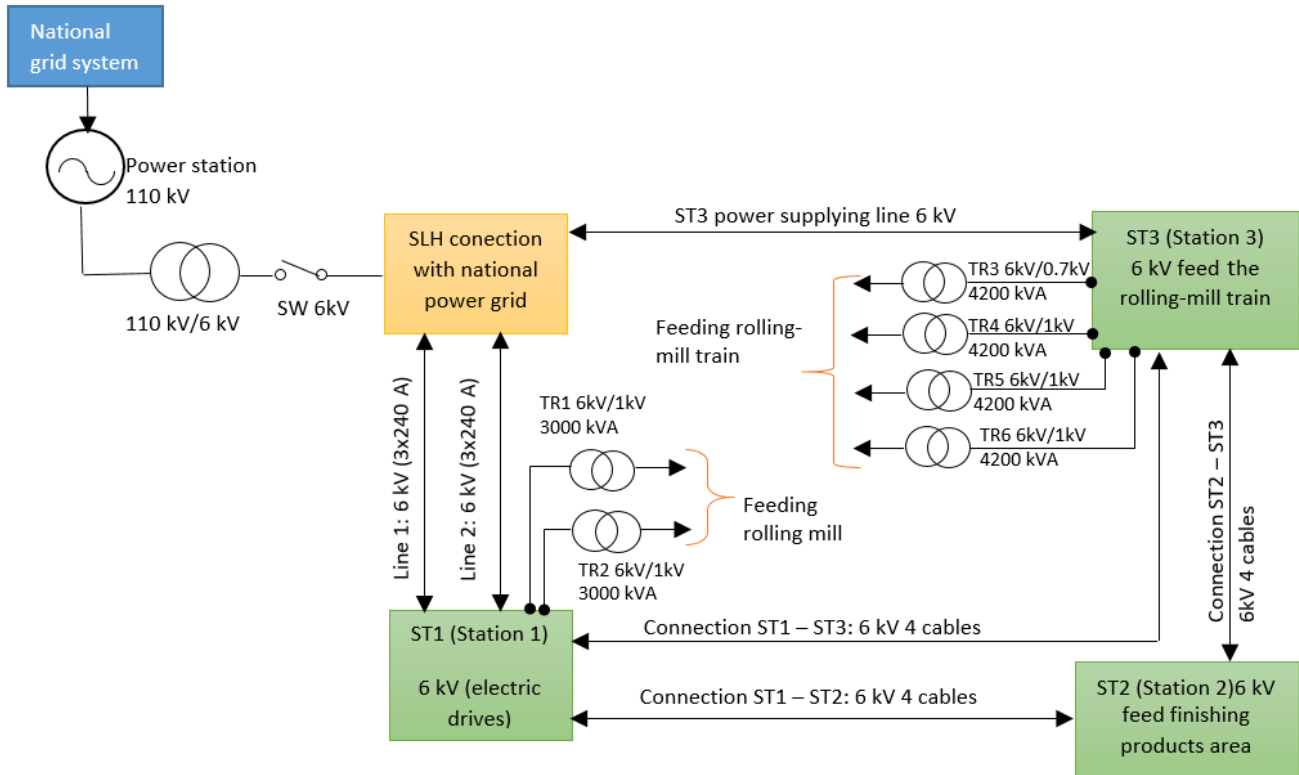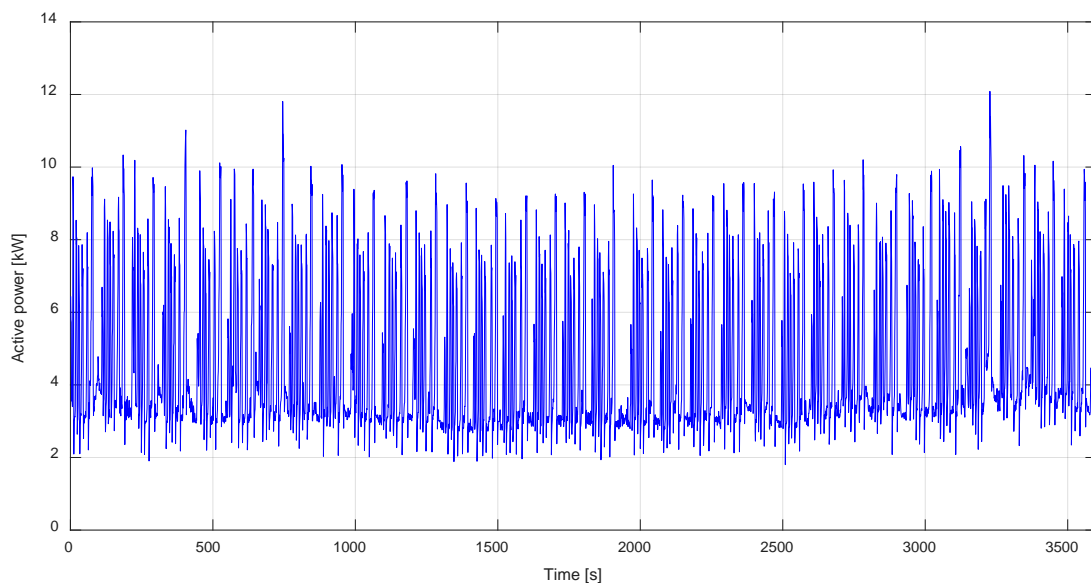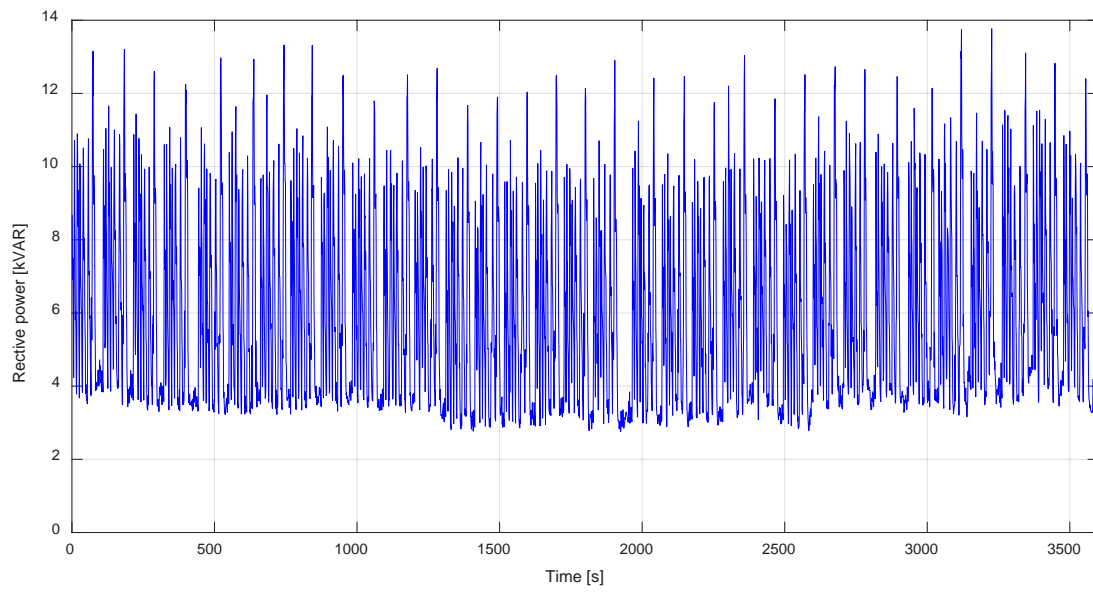

Figure 1. The power-supplying system of the hot rolling mill.

(**a**) Active power

Figure 2. *Cont.*

(**b**) Reactive power

Figure 2. The variation in the active (**a**) and reactive power (**b**) in the substation of the hot rolling mill for one hour.
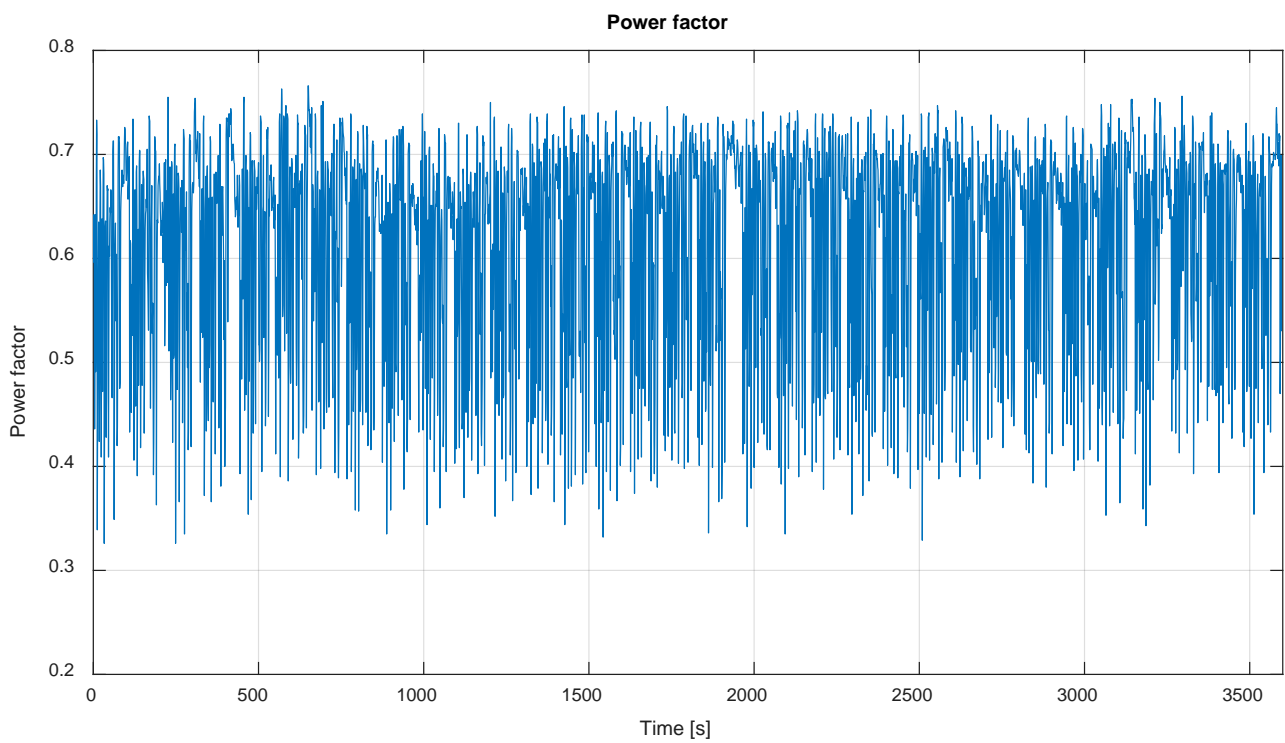

Figure 3. The variation in the power factor in the substation of the hot rolling mill for one hour.
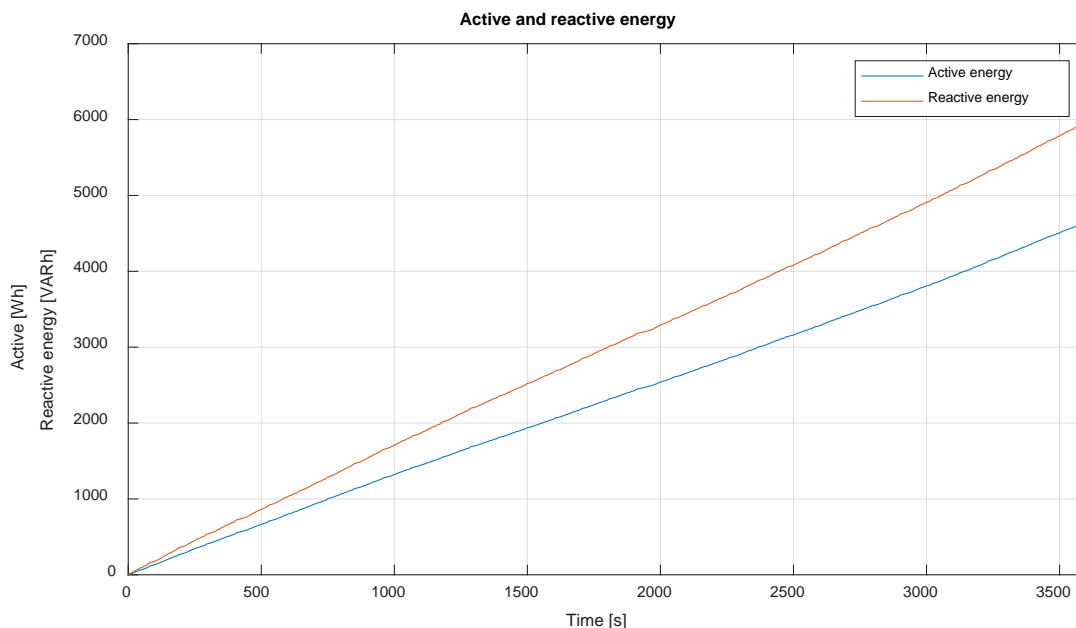
Figure 4. The variation in the active and reactive energy in the substation of the hot rolling mill for one hour.

4. Deep Learning Model for Power Factor Prediction

Time series prediction has important theoretical implications as well as various technical applications. Forecasting is a procedure that determines future values of time series data based on past data and can be utilized in applications where estimation is not feasible [36–47].

The research presented in our paper addresses the use of deep learning in the study of energy efficiency in a metallurgical plant. The use of deep learning methods in the energy efficiency analysis of hot rolling mills in metallurgical plants can bring numerous benefits by providing detailed insights and process optimization. Using deep learning models can allow for forecasting the energy consumption of hot rolling mills. By analyzing historical data, these models can provide accurate energy demand forecasts based on different variables, such as product type, rolling speed or temperature. The real-time prediction of power factor dips allows the system to immediately activate power factor correction devices like capacitor banks to maintain an appropriate power factor. Factories can use power factor forecasts to adjust equipment and operations in real time by integrating the deep learning model with automated control systems. This could include adjusting motor speeds, aligning production schedules with optimal power factors, or dynamically managing HVAC systems to preserve energy. Real-time operational modifications can reduce energy costs.

By analyzing data in real time, models can recommend adjustments of rolling speed according to production requirements and operating conditions. This can help reduce energy consumption and improve process efficiency. Implementing advanced control systems based on deep learning technologies can help optimize operating parameters, ensuring efficient energy consumption during the lamination process. Using deep learning algorithms can help optimize the use of electricity, thereby reducing network losses and ensuring efficient energy distribution within the metallurgical plant.

The implementation of these technologies requires efficient data collection and processing, access to sensors and appropriate monitoring systems, as well as the ability to integrate these technologies into the existing hot sheet control and operation system. However, the benefits can include significant energy savings, reduced operating costs, and improved process sustainability in metallurgical plants.

Various topologies of artificial neural networks that employ deep learning methods can be utilized for time series forecasting. Traditional neural networks are unsuitable for

sequential or time series predictions. In time series data, current observations depend on past observations, making them not independent. Since they cannot store historical data, traditional neural networks treat each observation as independent. Basically, they do not remember the past.

This study employed hybrid artificial neural network (ANN) models that incorporate deep learning methods to carry out a forecasting analysis. The prediction models utilized in this study incorporated RNN hybrid architectures, which also included LTSM and GRU layers.

Although both convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are applicable to time series data, they possess distinct advantages and are better adapted for distinct problem domains. When it comes to analyzing spatial relationships in data, such as video or image data, CNNs are ideally suited. Conversely, RNNs exhibit exceptional suitability when it comes to the examination of temporal associations within data, including time series data [38]. They can be employed to discern recurring patterns and characteristics in the data, such as seasonality or shifts in trends, which can subsequently be utilized for the purpose of data classification or prediction. In brief, RNNs are utilized to analyze temporal relationships in data, whereas CNNs are employed to analyze spatial relationships. In the context of time series data analysis, the suitability of a CNN or an RNN for the given problem may vary.

4.1. Recurrent Neural Networks

Cell feedback loops give RNNs memory. This is the main difference between RNNs and conventional neural networks. Feed-forward neural networks only carry information between layers, but the feedback loop passes data within a layer [38,39].

Time series forecasting is frequently performed with recurrent neural networks (RNNs), although they have limitations beyond the training horizon. The main issues are as follows [39–41]:

- Vanishing and exploding gradients: RNNs cannot capture long-term dependencies in time series data due to vanishing or exploding gradients during training. RNNs may struggle to forecast long-term occurrences due to this restriction.
- Short-term memory: Standard RNN architectures forget prior time steps, especially for extended sequences. Short-term memory can generate inadequate predictions for time points beyond the training horizon.
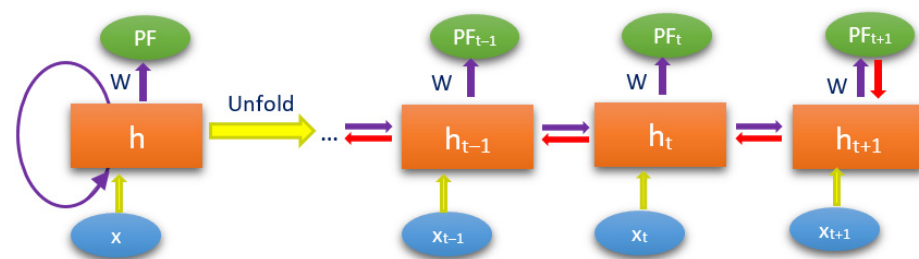
Figure 5 shows a typical RNN.

Figure 5. The typical structure of an RNN.

In step t, the network calculates h_t using input x_t and the hidden internal state of the previous step, h_{t-1}. The RNN's hidden memory is now linked to the output layer PF_t. As seen in Figure 5, this is like having numerous input–output architecture copies with connected hidden layers. The RNN cell can be basic with one activation function. Each step of a basic RNN updates the hidden state.

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_{t-1} + b_h), \tag{1}$$

Additionally, if input data x_t have dimension n, and hidden state h_t has dimension m, weight matrices W_{hh} and W_{ih} have dimensions $m \times m$ and $m \times n$, respectively. The hidden state h_t can be utilized to calculate the output, specifically the *PF* at step t.

$$PF_t = W_{h0}h_t, \tag{2}$$

Figure 5 depicts a schematic of this approach.

4.2. Long Short-Term Memory

The LSTM adds a cell state to the RNN, similar to the hidden state, which is transferred from cell to cell. Unlike the hidden state, no matrix multiplication occurs; gates add or remove cell state information. LSTMs typically have four gates covering three steps: forget, input/update, and output [38,39].

4.2.1. Forget Step

The initial step in an LSTM neural network cell is to select whether to maintain or forget the previous time step's information. Equation for the forget gate is as follows (Equation (3)):

$$f_t = \sigma\left(W_{hf}h_{t-1} + W_{xf}x_t\right), \tag{3}$$

- The symbol σ represents the sigmoid function, which is applied element-wise.
- The x_t represents the input given at the current step.
- The W_{xf} represents the weight that is linked to the input.
- The term h_{t-1} represents the hiding state from the preceding step.
- The weight matrix associated with the hidden state is denoted as W_{hf}.
- Based on the previous Equation (3), if f_t equals 0, it leads to the outcome of forgetting everything, whereas if f_t equals 1, it results in not forgetting anything.

4.2.2. Input Step

The significance of the newly transmitted information is quantified by the input gate. The equation for the input gate is as follows (Equation (4)):

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1}), \tag{4}$$

As a result of the sigmoid function being applied, the input value at step t will fall within the range of 0 to 1.

4.2.3. Update Step

The new information that has to be transmitted to the cell state at step t is determined by the hidden state at step $t-1$ and the input x at step t. Consequently, the value given by Equation (5) will be integrated into the cell state.

$$n_t = \tanh(W_{hn}h_{t-1} + W_{xn}x_t), \tag{5}$$

As a result of the tanh function, new information will have a value between -1 and 1. Nevertheless, the n_t will not be instantly incorporated into the cell state. The updated equation is displayed in (6).

$$c_t = f_t \cdot c_{t-1} + i_t \cdot n_t, \tag{6}$$

In this context, c_{t-1} represents the cell state at the current step, whereas the remaining values correspond to the previously determined values.

4.2.4. Output Step

The output is determined by Equation (7). In our paper, we labeled the output as o.

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1}), \tag{7}$$

To determine the current hidden state, we will use the PF_t and tanh of the updated cell state, as illustrated in (8).

$$h_t = PF_t \cdot \tanh c_t, \tag{8}$$

The hidden state is determined by the combination of the long-term memory (c_t) and current output.

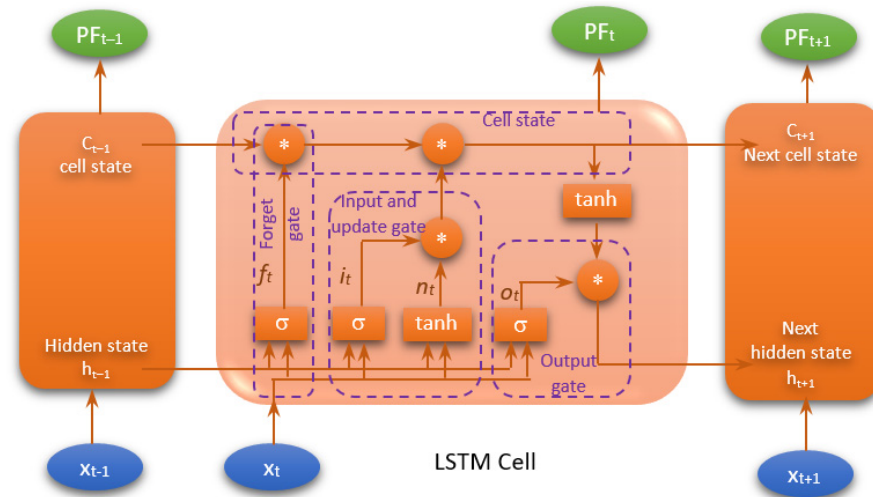Figure 6 displays a basic illustration of an LSTM network.

Figure 6. LSTM network.

4.3. Gated Recurrent Unit (GRU)

A GRU cell demonstrates similarities to both LSTM and RNN cells [39,40].

GRUs have similarities to long short-term memory (LSTM) in many different respects. GRU, like LSTM, controls the movement of information via gates. GRU layers have a more straightforward architecture and provide some enhancement over LSTM. GRU, in contrast to LSTM, lacks a distinct cell state (c_t). It contains only a hidden state. Due to their more straightforward architecture, GRUs require less time to train.

GRU receives an input x_t and the hidden state h_{t-1} from the previous step $t-1$ at each step t. It afterward returns a new hidden state, Ht, which is transmitted to the subsequent step $t + 1$.

4.3.1. Reset Gate (Short-Term Memory)

How much of the hidden state to forget is determined by the reset gate. It accepts the previous hidden state and the current input and outputs a vector of values between 0 and 1 that controls how much the previous hidden state is "reset" at the current time step. The reset gate stores the network's hidden state—short-term memory. The reset gate Equation is shown in (9) with values between 0 and 1.

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1}), \tag{9}$$

4.3.2. Update Gate (Long Term Memory)

The update gate selects the extent of the candidate activation vector to add to the hidden state. It uses the earlier hidden state and present input and outputs a vector of integers between 0 and 1 that sets the candidate activation vector's inclusion into the new hidden state. Equation (10) presents the updated value.

$$u_t = \sigma(W_{xu}x_t + W_{hu}h_{t-1}), \tag{10}$$

To determine the GRU's hidden state H_t, it requires the candidate activation vector. Using the current input and the reset gate, the candidate activation vector is a modified

version of the previous hidden state. Using a tanh activation function, it outputs −1 to 1. Equation (11) shows the candidate activation vector.

$$\hat{h}_t = \tanh\left(W_{xg}x_t + (r_t \cdot h_{t-1})W_{hg}\right), \tag{11}$$

The input and hidden state from step $t - 1$ multiplied by the reset gate output r_t provided Equation (10). This information is given to tanh to calculate the candidate activation vector. All information from the previous hidden state h_{t-1} is taken into account if r_t is 1. Similarly, if r_t is 0, the previous hidden state is ignored.

4.3.3. Hidden State

Instead of employing a separate gate like LSTM, GRU uses a single update gate to handle both h_{t-1} and candidate state information to calculate h_t. Relation (12) illustrates this, with the two components displayed in green and blue.

$$h_t = u_t \cdot h_{t-1} + (1 - u_t) \cdot \hat{h}_t, \tag{12}$$

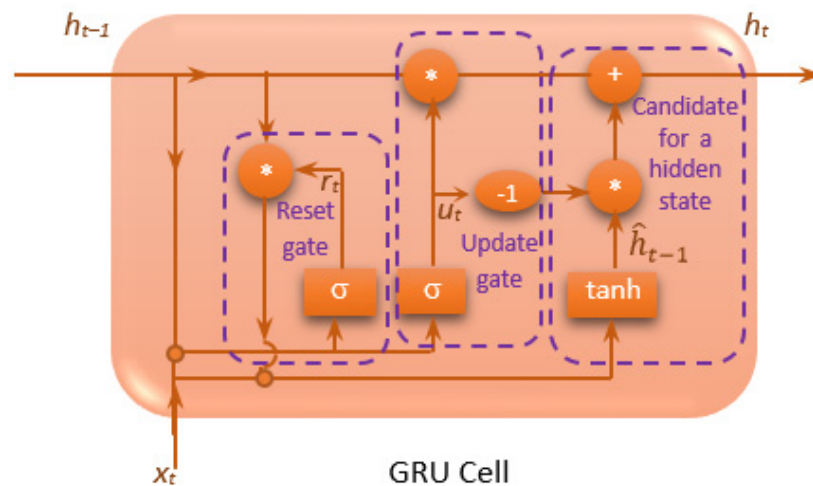Figure 7 shows the structure of a GRU cell.

Figure 7. The structure of a GRU cell.

Deep neural networks that combine LSTM or GRU architectures have been proposed for multiple applications. A hybrid model uses different statistical or machine learning methods to improve predicting accuracy and robustness. Various deep neural architectures for time series forecasting aim to achieve the same goal [40]. For the purpose of achieving a more accurate long-term prediction beyond the horizon, in this article, a version of a recurrent artificial neural network that incorporates both LSTM and GRU layers was developed.

5. Results

The power factor is one of the characteristics that has a considerable impact on energy efficiency. Its values should be very close to one. The power factor in hot rolling mill power supplies is the ratio of real power (active power) to apparent power in the electrical system. It is a critical parameter for determining power efficiency and how effectively the hot rolling process converts electrical energy into productive activity. As a result, we considered that this important parameter of power quality could benefit from being studied using intelligent deep learning approaches. As a result, the data collected from the measurements described in Section 3 were utilized to train different kinds of artificial neural networks, yielding valuable and relevant information on the variation in the power factor during the rolling mill process.

To generate a useful prediction that extends beyond the horizon, we built hybrid RNN models using deep learning. Prediction beyond the horizon in time series forecasting is the process of predicting future time points that are beyond the most recent observed data point. In other words, it entails forecasting values in time series that exceed the existing historical data. When working with time series data, forecasting models are frequently used to predict future values based on historical patterns and trends. In this context, the term "horizon" refers to the time period in the future for which predictions are intended [41,42].

Short-term forecasting predicts values for the next few points, while long-term forecasting covers a longer period. Forecasting beyond the horizon involves predicting a time beyond the last observed data point [41,42]. Predicting further into the future is difficult due to uncertainty. Longer-term forecasts are susceptible to more unknown factors, and small starting errors can accumulate.

Time series forecasting approaches like statistical models and machine learning algorithms can predict beyond the observed horizon. Long-term forecasting requires adequate models and consideration of accuracy-computational complexity compromises. Additionally, characteristics, model assumptions, and historical data quality are critical to making accurate long-term predictions.

The research employed Matlab and Deep Learning Toolbox to train a neural network. Thus, various power factor models were obtained by training a hybrid RNN with the measured power factor values. On the basis of these models, power factor forecasting was performed.

To obtain the most accurate information on the variation in the power factor, data recorded over a half an hour period were used. This data collection was important in order to obtain as much information about the variance of the power factor as feasible.

We used the train function to train the model, and the parameters used for training were as follows:

- Model parameters defined included the following:
  - Training and testing data ratio: 0.5 to 0.9;
  - Lag (number of prior samples): 2 to 350;
  - Forecasting horizon duration beyond the current time frame: 2 to 400.

- Parameters utilized by the training function, train the following:
  - Adjust MiniBatchSize from 16 to 128;
  - MaxEpochs from 300 to 1000;
  - Learningrate from 0.0001 to 0.1;
  - Solver "adam";
  - Trainfcn "trainbr".

The method used for predicting the power factor comprises the subsequent stages, as illustrated in the flowchart depicted in Figure 8. The steps are detailed below:

Step 1: the first step involves the establishment of an architecture for artificial neural networks (ANN).

Step 2: The data were processed and made ready. In order to achieve better results, the signal followed a process of filtration. The filtered signal is depicted in Figure 9. Subsequently, data standardization was employed. By employing this data standardization technique, a more effective process convergence may be guaranteed.

Step 3 involves determining the training parameters, such as the learning rate, maximum number of epochs, and training function.

Step 4: Divide the data into training and testing subsequences. Partitioning the dataset into separate training and testing sets facilitates the assessment of the forecasting model's performance.

Step 5: Preparing the deferred sequence. The time series can be delayed based on the desired extent of historical observation, as in Figure 10. For the purpose of forecasting using lagged time series data, it is necessary to generate features from past observations in

order to predict future values. In order to produce lagged features, the time series values are shifted by a specified quantity of time steps.
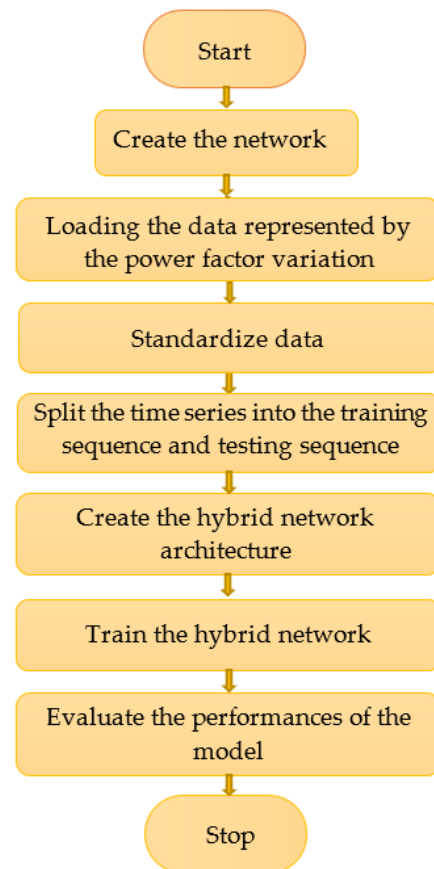

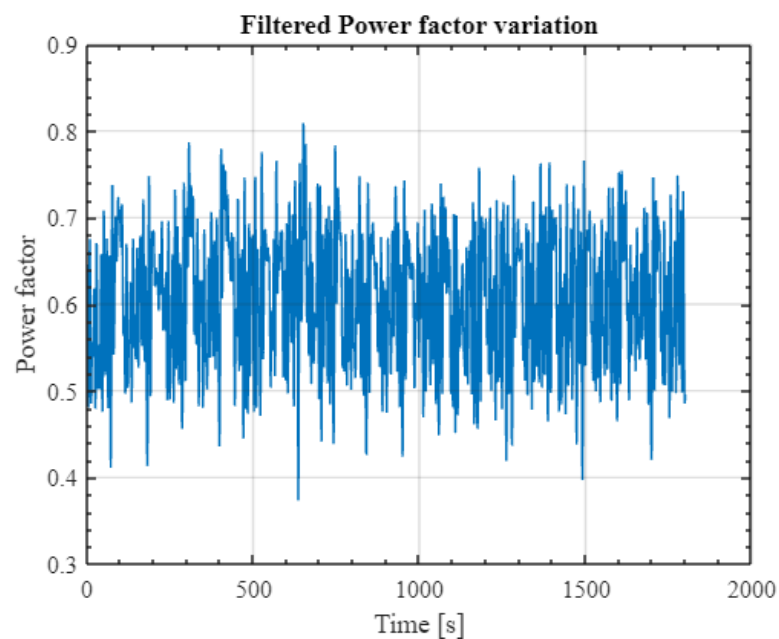**Figure 8.** The flowchart of modelling and prediction for the power factor (PF).

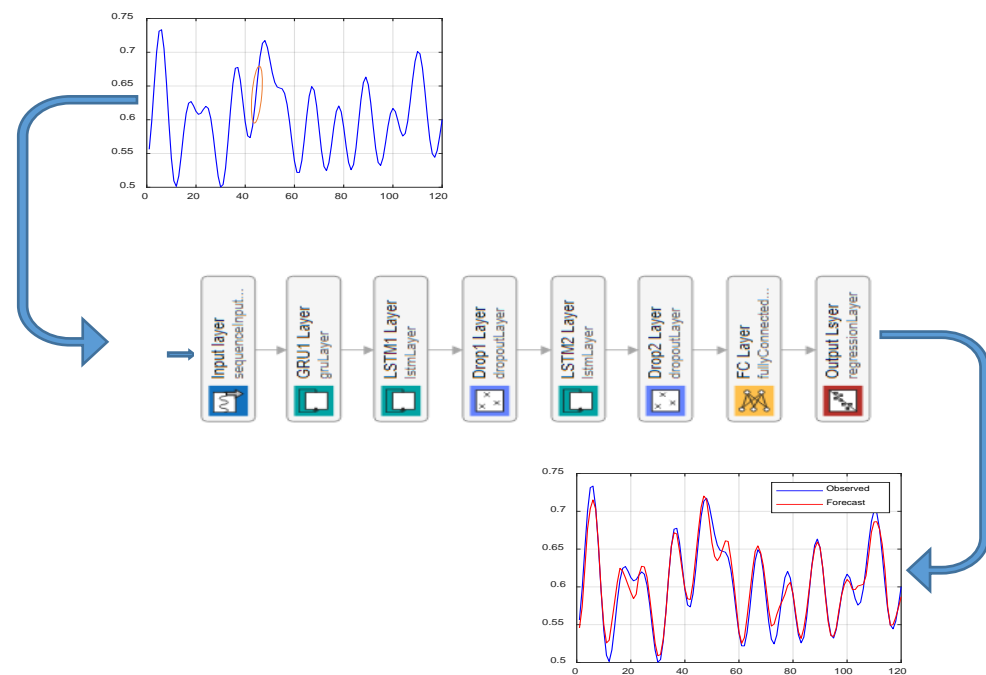**Figure 9.** The filtered signal for power factor.

Figure 10. The method of prediction for the power factor (PF).

In our paper, we performed a multistep prediction beyond the horizon. When making a forecast, it is rare to predict merely the next element in the series ($t + 1$). Instead, the most typical goal is to predict the entire future interval ($t + 1, ..., t + n$) or a distant point in time $t + n$. Several ways exist for making this type of prediction. Because the value $PF(n - 1)$ is necessary to predict $PF(n)$, yet $PF(n - 1)$ is unknown, a recursive approach is used, with each subsequent prediction relying on the preceding one. This approach is called recursive forecasting or recursive multi-step forecasting. Figure 11 illustrates that.
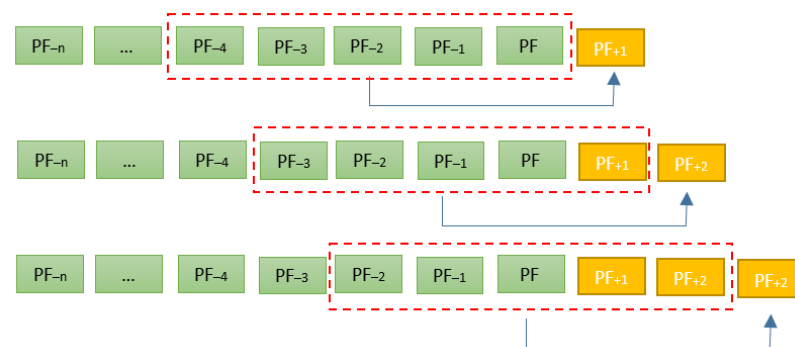
Figure 11. Multi-step prediction of the PF tests.

In the paper, multiple employing diverse models of artificial neural networks. The results of the prediction are presented using a hybrid RNN model, the architecture of which is depicted in Figure 10. The main parameters utilized in the training process are as follows: The term "Lag" refers to the number of prior samples taken into consideration. The term "ratio" denotes the percentage of data used for training. The term "horizon" refers to the forecasting period beyond the current time frame. The variable MaxEpoch defines the upper limit for the number of training epochs, whereas the variable learningrate denotes the rate at which the model learns.

The parameters utilized in training for the model depicted in Figure 12 are as follows:

Lag = 200;
Ratio = 0.8;
Horizon = 350;

MaxEpoch = 600;
Learningrate = 0.007.

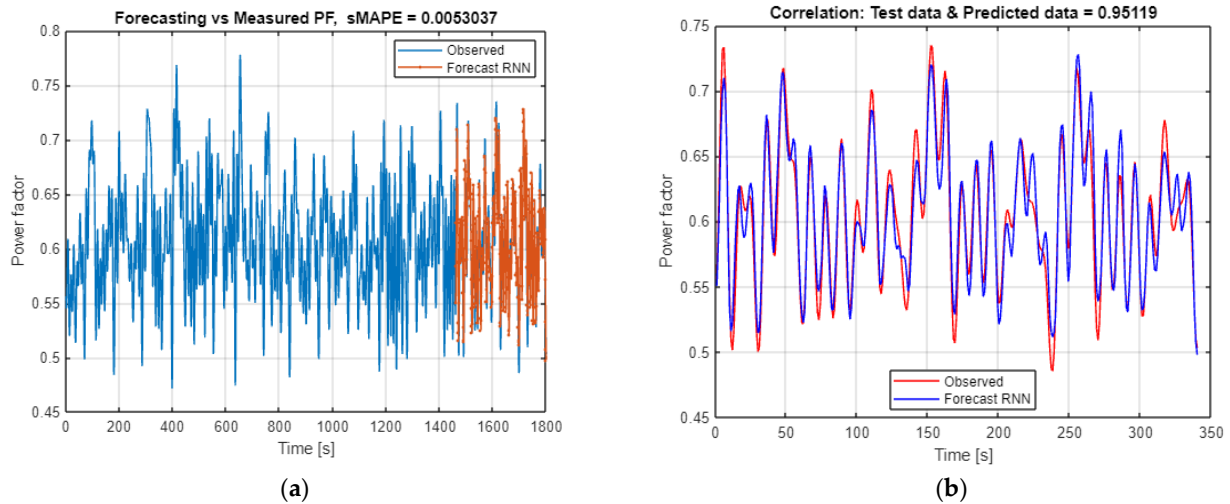(**a**)   (**b**)

Figure 12. The prediction of the PF based on 1800 samples, lag = 200, horizon = 350. (**a**) The predicted values and the measured values for PF; (**b**) correlation between test data and predicted data.

The network expects a series of "Lag" values to forecast the cases for the subsequent PF (power factor) value by utilizing a rollback window. The trained network preserves the memory of the training time sequence and anticipates a new sequence in order to make predictions one step ahead. Following the prediction, it is necessary to unstandardize the data in order to scale them to generic values for accurate interpretation. In order to assess the quality of the fitting, a basic correlation analysis is performed between the actual and predicted values.

The network anticipates a series of Lag values to forecast the power factor (PF) value for the upcoming step using a rollback window. The trained network retains a "memory" of the training time sequence and anticipates one step ahead of a new sequence in its predictions. The trained neural network maintains its "memory" intact, which indicates that it recalls the sequence of data from the time of training. The net specifically retains the last Lag time step; consequently, it anticipates a new sequence in order to make a prediction that is one step ahead. For the model presented previously, the predicted PF values are illustrated in Figure 12a, and Figure 12b shows the correlation between the forecasted values and the testing data. The regression coefficient and *RMSE* value are shown in Figure 13.
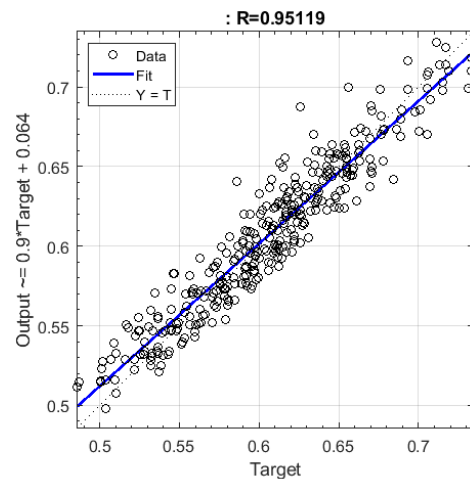

Figure 13. The correlation coefficient and the *RMSE*, for prediction based on 1800 samples, lag = 200, horizon = 350.

It is possible to use this hybrid RNN to generate a new network and transfer its experience and memory to it. This neural network will retrieve the most recent sequence provided in the training data. Nevertheless, in order to proceed with the prediction to the subsequent stage, the network must first generate a new sequence in accordance with the previous prediction. By employing this, one can generate a sequence prediction that is proportional to the number of steps in the testing data (horizon). To maintain a constant number of features, the preceding predicted value is placed at the top of the sequence while the list moves downward. The predicted series obtained from the testing data (complete testing series) and the new prediction created from a sequence generated with the previous forecasted value are illustrated in Figure 14, while Figure 15 shows the predicted values beyond the horizon.

(**a**)   (**b**)
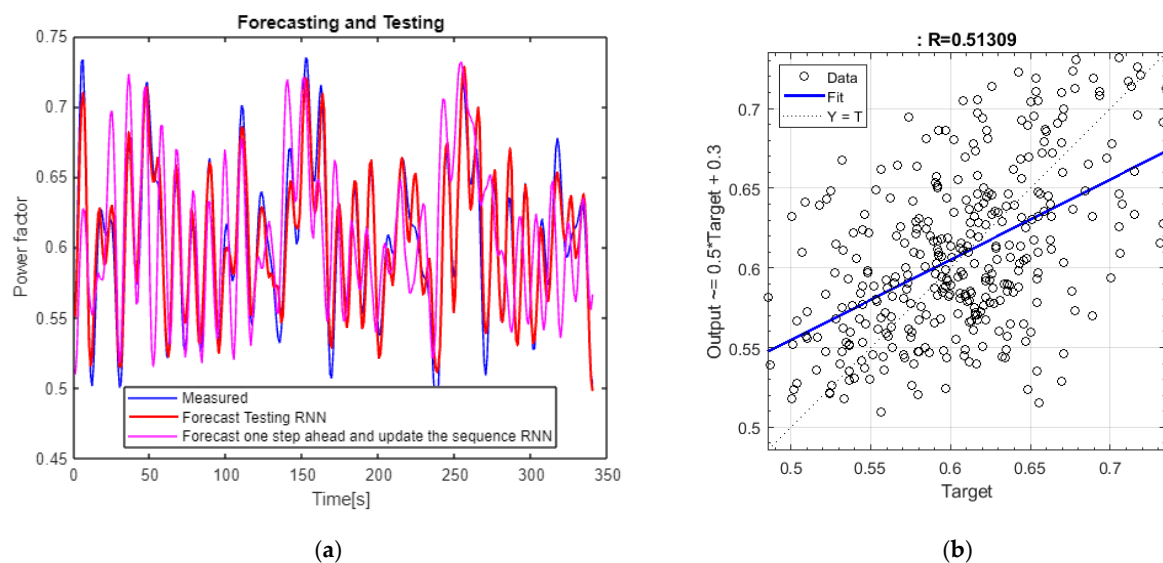
Figure 14. The prediction results based on 1800 samples, lag = 200, horizon = 350. (**a**) The one-step prediction and also multistep prediction of the PF; (**b**) correlation between test data and predicted data for multi-step prediction.
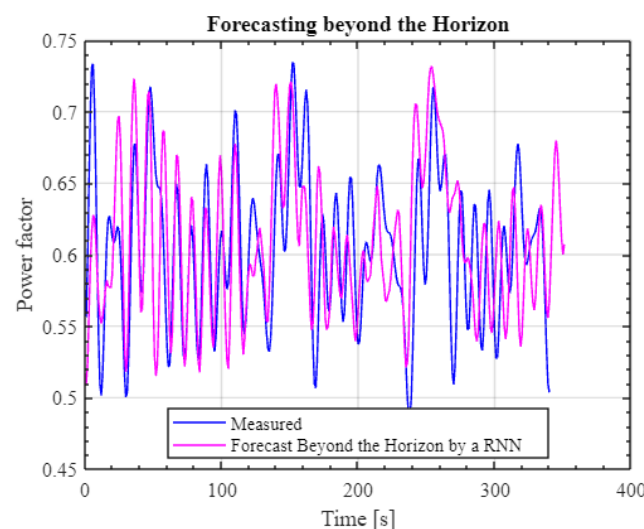
Figure 15. The results of the forecasting beyond the horizon based on 1800 samples, lag = 200, horizon = 350.

5.1. Evaluation Metrics

The accuracy and performance of time series forecasting models can be assessed using several indicators [43–46]. The type of time series data and forecasting goals influence the

evaluation indicators. Below are several frequently employed evaluation measures for time series prediction that we have utilized in our research.

5.1.1. Mean Absolute Error (*MAE*)

This computes the mean absolute deviation between the expected (*PF*) and actual values $\hat{PF}$ and is presented in (13).

$$MAE\left(PF_i, \hat{PF_i}\right) = \frac{1}{n} \sum_{i=1}^{n} \left| PF_i - \hat{PF_i} \right| \tag{13}$$

5.1.2. Root Mean Square Error

This is generated from *MSE*, which calculates the average of the squared differences between predicted and measured values. This emphasizes greater errors over *MAE*. *RMSE* is the square root of *MSE*, which provides a more interpretable measure of error, being displayed in Equation (14).

$$RMSE\left(PF_i, \hat{PF_i}\right) = \sqrt{\frac{1}{n}\left(PF_i - \hat{PF_i}\right)^2} \tag{14}$$

5.1.3. Mean Absolute Percentage Error

This determines the percentage difference between the predicted and measured values, which is useful in quantifying errors as a percentage in relation to the measured values. Equation (15) represents this indicator.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{PF_i - \hat{PF_i}}{PF_i} \right| \tag{15}$$

5.1.4. Symmetric Mean Absolute Percentage Error (*SMAPE*)

The calculation involves determining the symmetric absolute percentage difference by taking into account the average of the absolute percentage error for each observation. Equation (16) represents this indicator.

$$SMAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{2\left| PF_i - \hat{PF_i} \right|}{\left| PF_i \right| + \left| \hat{PF_i} \right|} \tag{16}$$

5.1.5. R-Squared—The Coefficient of Determination

In time series forecasting, the R^2 coefficient, also known as the coefficient of determination, evaluates a forecasting model's quality of fit. However, there are significant complexities and considerations when using R^2 with time series data. The R^2 for time series forecasting is derived as the squared correlation coefficient (r^2) between the measured values (PF_i) and the forecasted values ($\hat{PF_i}$), as shown in Equation (17).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(PF_i - \hat{PF_i}\right)^2}{\sum_{i=1}^{n}\left(PF_i - \overline{PF_i}\right)^2} \tag{17}$$

These metrics are the ones that are utilized the most frequently in order to evaluate various aspects of performance when time series forecasting models are being evaluated. A full knowledge of the model's accuracy, precision, and overall goodness of fit may be obtained by taking into account the mean absolute error (*MAE*), root mean squared error (*RMSE*), Symmetric Mean Absolute Percentage Error (*SMAPE*), and R-squared, respectively, together. An assessment of a prediction model that considers only one indicator is insufficient. It is essential to assess these indicators collectively. Analyzing these parameters collectively has several benefits. The *MAE* and *SMAPE* metrics offer insights into the model's accuracy, whilst *RMSE* provides information regarding precision. While it is

important to have high precision, it is essential to maintain accuracy and not compromise it in the pursuit of precision. *MAE* and SMAPE are more resistant to outliers than *RMSE*. If the time series data contains outliers, *MAE* and SMAPE may offer a more balanced assessment. *MAE*, *RMSE*, and SMAPE are easier to understand, but R-squared indicates how well the model explains the variance. However, R-squared may be a less reliable indicator in time series forecasting; consequently, it is best used in conjunction with other metrics. A full evaluation of a time series forecasting model requires analyzing multiple indicators, each of which provides a distinct view of different elements of model performance.

We forecast the power factor by employing several artificial neural network structures for the models. We initially created a model using a traditional neural network, specifically a Multilayer Perceptron (MLP) model. This artificial neural network includes multiple hidden layers. After studying the results of a model with three layers, we reached the following conclusions: the model has been very well trained, with a regression coefficient close to one. The sMape parameter is around 4.4×10^{-5}, with an *RMSE* of 1.54×10^{-5}, suggesting a highly successful training phase. The prediction is inaccurate due to the parameters having very low values, for example, $R = 0.01$, which suggests a low-quality forecast. Various scenarios were examined with different numbers of layers. Nevertheless, the anticipated accuracy consistently performed poorly.

We also investigated employing the ResNet50 network, a 50-layer deep convolutional neural network, on the same dataset. This network comprises numerous layers and is expected to generate accurate forecasts. Training has shown significant positive results with an R-squared value of 0.992, sMape of 0.012, and *RMSE* of 0.035 during the training phase. The forecasting results show good values with *RMSE* of 0.1122, *MAE* of 0.09, and R squared of 0.36. However, the architecture of these networks results in inefficient performance due to extremely expanded training times.

We consider that the hybrid structure of ANN-RNN used in our research provides good results for both short-term and long-term prediction.

6. Discussion

A study was conducted to assess the performance of the developed models, focusing on the primary parameters of the models and their predictive capabilities. As a result, LAG values ranging from 100 to 320 and horizon parameter values between 100 and 400 were implemented. Furthermore, we examined the impact of the ratio parameter, which was employed to calculate the proportion of the dataset that was allocated for training and testing purposes. The outcomes of some of the models' predictions are displayed in Table 1. The analyses are conducted using the aforementioned indicators to determine the one-step advance prediction and beyond-horizon prediction, respectively.

Table 1. The prediction performances.

Model Parameters		One Step Ahead				Prediction Beyond Horizon		
	Horizon	sMape	RMSE	MAE	R	RMSE	MAE	R
	160	0.006489	0.019683	0.015457	0.963825	0.092392	0.075237	0.288532
	200	0.006390	0.018737	0.015263	0.96760	0.090654	0.073941	0.324873
	240	0.006673	0.019687	0.016026	0.964244	0.095409	0.077594	0.150331
Lag = 160	280	0.007124	0.021192	0.017132	0.958140	0.091739	0.074751	0.026225
	320	0.007168	0.020631	0.017140	0.961027	0.095048	0.077251	0.160760
	360	0.006929	0.020721	0.016595	0.960881	0.092077	0.074983	0.145100
	400	0.006608	0.019332	0.015952	0.96523	0.090610	0.073909	0.217332
	160	0.008471	0.025325	0.020117	0.938433	0.092560	0.075202	0.020544
	200	0.008661	0.025332	0.020491	0.937462	0.093857	0.076196	0.033937
	240	0.008647	0.024865	0.020516	0.941373	0.091105	0.074174	0.125700
Lag = 200	280	0.008842	0.025672	0.020965	0.936302	0.089814	0.073141	0.321996
	320	0.009467	0.027266	0.022588	0.927733	0.091985	0.074688	0.038184
	360	0.008297	0.025073	0.019685	0.939253	0.090619	0.073766	0.017390
	400	0.008282	0.024109	0.019534	0.945042	0.090007	0.073272	0.279986

Table 1. *Cont.*

Model Parameters	Horizon	One Step Ahead				Prediction Beyond Horizon		
		sMape	*RMSE*	*MAE*	*R*	*RMSE*	*MAE*	*R*
Lag = 240	160	0.009775	0.028738	0.022959	0.931563	0.099912	0.081576	0.294496
	200	0.010913	0.032659	0.025910	0.913753	0.095454	0.078326	0.359388
	240	0.009967	0.029112	0.023464	0.930081	0.095710	0.078482	0.392096
	280	0.009816	0.028729	0.023296	0.931552	0.099546	0.081309	0.351608
	320	0.010403	0.030344	0.024536	0.925773	0.099880	0.081498	0.354081
	360	0.009641	0.029050	0.022620	0.930357	0.100956	0.082392	0.376369
	400	0.010696	0.030860	0.025142	0.921431	0.101547	0.082809	0.363546
Lag = 280	160	0.010918	0.032299	0.026064	0.895651	0.098910	0.080598	0.566225
	200	0.011695	0.034596	0.028074	0.879655	0.100376	0.081652	0.570846
	240	0.011535	0.033365	0.027670	0.886827	0.099517	0.081136	0.647117
	280	0.011746	0.034617	0.028210	0.880539	0.102660	0.083644	0.620837
	320	0.011261	0.033107	0.027099	0.889766	0.096725	0.078849	0.572887
	360	0.011719	0.034218	0.028231	0.883458	0.099361	0.081198	0.522786
	400	0.010868	0.032184	0.025959	0.895913	0.102770	0.083815	0.586158
Lag = 320	160	0.013265	0.037889	0.031732	0.880825	0.101182	0.083082	0.109687
	200	0.011788	0.038033	0.028702	0.876656	0.101793	0.083475	0.125691
	240	0.010896	0.035366	0.026485	0.895311	0.098212	0.080433	0.080207
	280	0.011746	0.034617	0.028210	0.880539	0.102660	0.083644	0.620837
	320	0.011202	0.033706	0.027039	0.906087	0.095732	0.078700	0.185491
	360	0.011950	0.036144	0.028806	0.889896	0.098534	0.081027	0.221544
	400	0.012625	0.037413	0.030036	0.885345	0.103011	0.084218	0.282126

When analyzing the data presented in Table 1, it is obvious that the forecasting indicators vary in accordance with the model's parameters. Regarding one-step ahead prediction, it was observed that all the parameters have outstanding values. The root mean square error (*RMSE*) values are less than 0.034, and the mean absolute error (*MAE*) values are less than 0.028. Furthermore, the root mean square error (*RMSE*) values exceed the mean absolute error (MEA) values for all instances outlined in Table 1, as anticipated. The R-squared value is near to 1, with a minimum of 0.87, indicating a strong correlation between the measured data and the data utilized in the test. Therefore, we may conclude that the one-step-ahead prediction is quite accurate.

Regarding to the prediction beyond the horizon, good results have been achieved for some of the model parameters. The indicators *MAE* and *RMSE* typically have values that are generally less than 0.1 (*RMSE*) and less than 0.08 (*MAE*), respectively. However, the R-squared metric shows lower values in comparison to the one-step-ahead forecast. R-squared, also known as the coefficient of determination, is a statistical metric that quantifies the proportion of the variability in the dependent variable that can be explained by the independent variables in a regression model. Although R-squared is frequently employed as a statistic in regression analysis, its interpretation and application may vary when applied to time series forecasting. A moderate R-squared does not necessarily indicate an inadequate forecasting performance. Instead, it shows that the model is able to explain a portion of the variability in the data, but there is still a significant amount of unexplained variability remaining. While R-squared can provide insight into how well the model explains data variability, it is important to take into account other measures (such as *MAE* and *RMSE*) for a more thorough evaluation of predicting effectiveness. An average root mean square error (*RMSE*) of 0.1 signifies a minimal divergence between the predicted and observed power factor values. Considering the typical range of power factor values (0.3 to 0.6), a root mean square error (*RMSE*) of 0.1 indicates that the model's predictions approximate the actual values to a good level.

In addition to what is shown in Table 1, we carried out more experiments to further comprehend the efficiency of the one-step-ahead and beyond-the-horizon forecast. These experiments are presented as variation graphs of SMAPE, *RMSE*, *MAE*, and R-square, as shown in Figure 16a, b, c, and d, respectively, Figure 17a, b, and c.
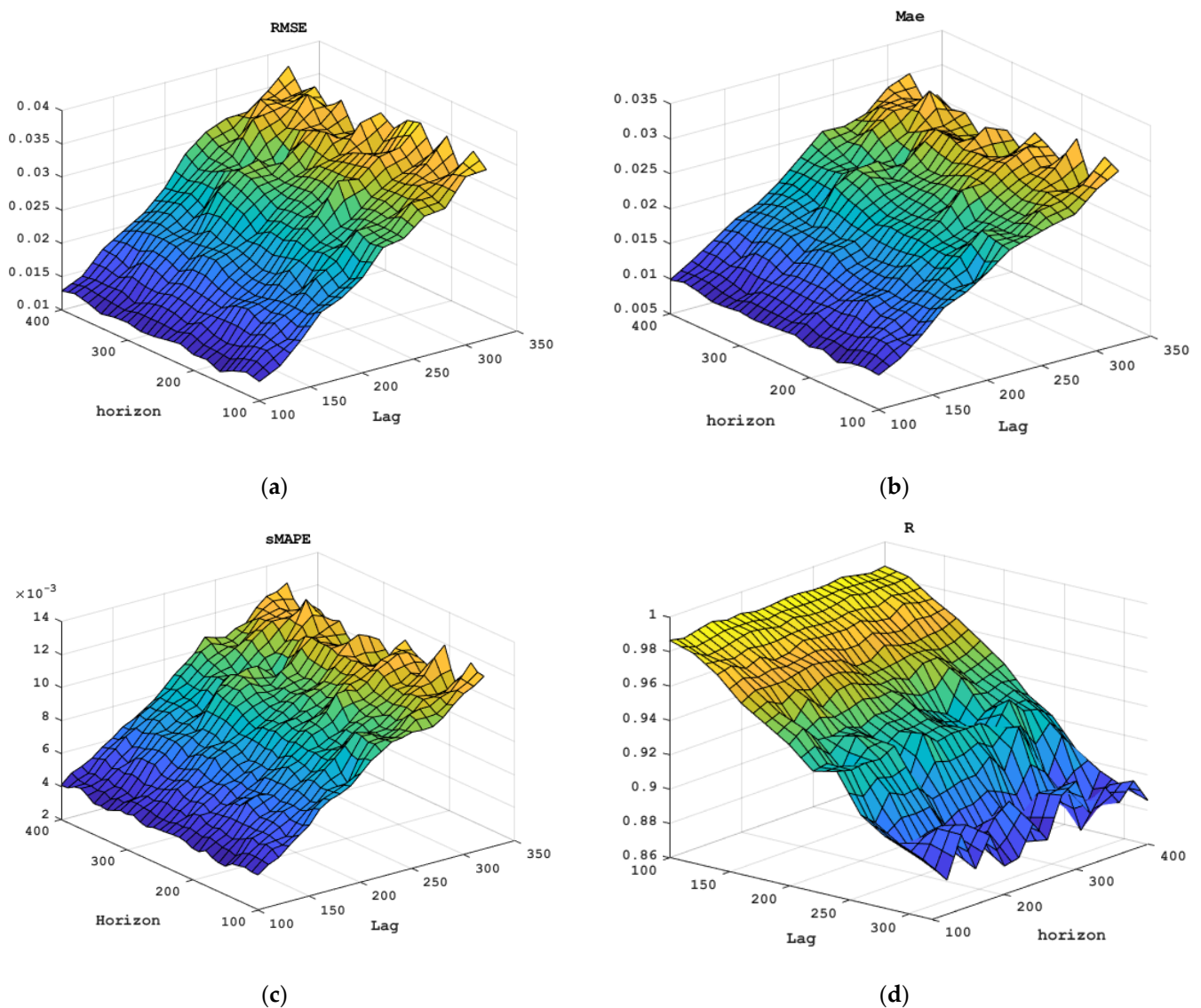
Figure 16. The metrics variation for one-step-ahead prediction. (**a**) The *RMSE* coefficient; (**b**) the *MAE* coefficient; (**c**) the sMape coefficient; (**d**) the R-squared coefficient.

The graphs illustrate that the experiments were conducted by manipulating two parameters of the analyzed models: Lag and Horizon. The LAG value ranged from 100 to 320, whereas the Horizon parameter ranged from 100 to 400. The efficiency of the forecast varies to some extent based on these two characteristics. Therefore, it is evident that the one-step-ahead prediction yields the lowest *MAE*, *RMSE*, and *sMAPE* values when the LAG parameter is set to lower values. Furthermore, the R-squared indicator shows the highest values when the LAG parameter has been configured to lower values. Regarding the impact of the horizon parameter, it has no effect on the efficiency of the one-step ahead forecast, as anticipated. We evaluated performance using Lag and Horizon values ranging from 100 to 320 and 100 to 400, respectively. We included a summary in Table 1 to observe the variation trend of parameters that determine the forecast performance. The graphs illustrate the performance metrics for LAG and Horizon values ranging from 10 to 10, while the table displays these values from 40 to 40. We studied the data from Figures 16 and 17 and found that the best result is reached when Horizon = 240 and Lag = 180. We highlighted these in Figure 17a,b.
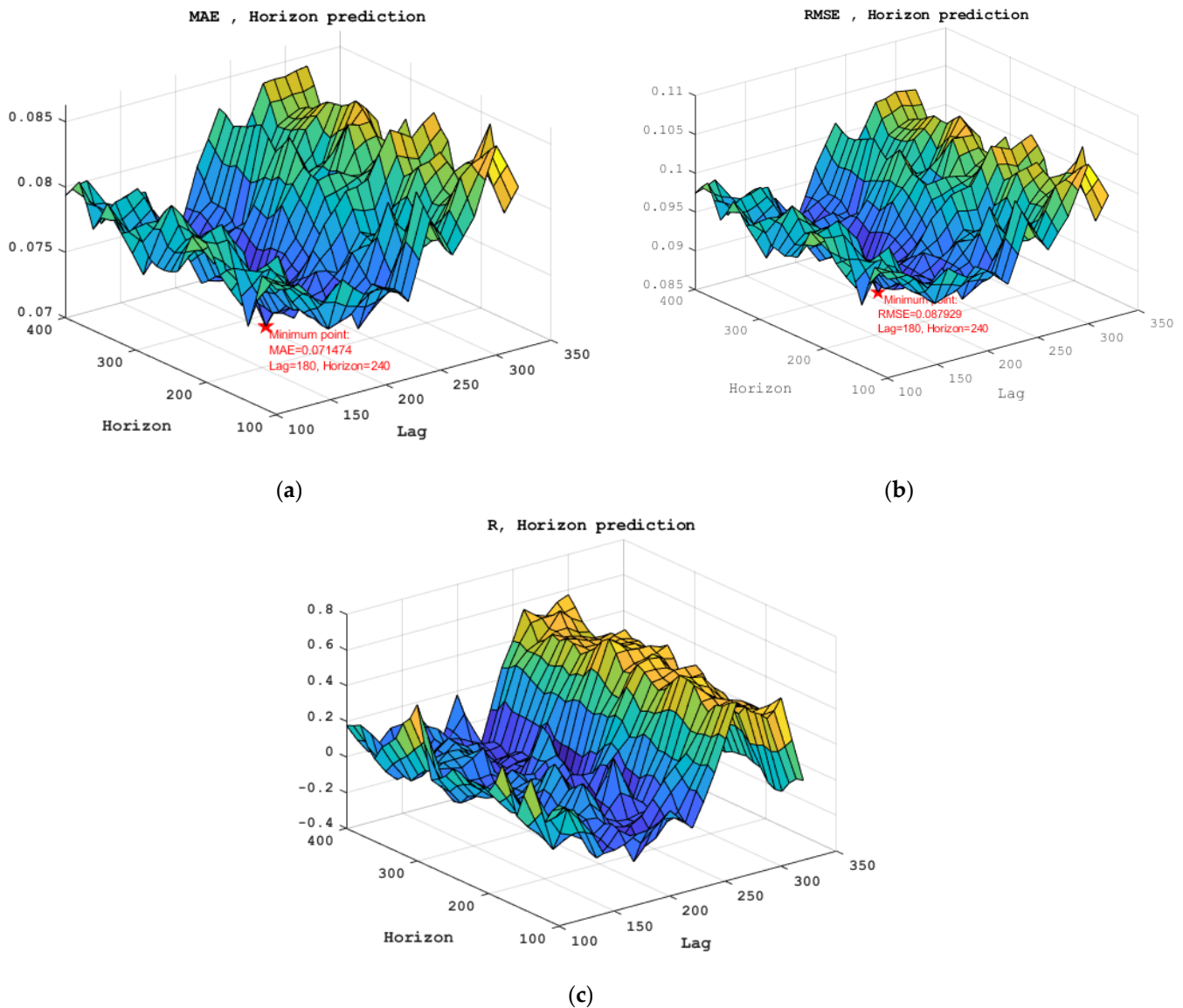
Figure 17. The metrics variation for prediction beyond the horizon. (**a**) The *RMSE* coefficient; (**b**) the *MAE* coefficient; (**c**) the R-squared coefficient.

We compared our research with previous research that performed time series prediction on different datasets found in the literature. Forecasting of power quality parameters is performed in reference [35] utilizing SOM Maps with a KNN Algorithm. This study analyzed and resulted in RMSE values for power factor ranging from 0.011 to 0.33, which were comparable to our results. Article [45] presents the authors' results on time series forecasting using techniques like CNN ANN and RNN ANN on a variety of datasets from various applications, similar to the methods we utilized. Datasets that are comparable to ours are those related to electricity, showing mean squared error (MSE) values between 0.129 and 0.197 and mean absolute error (*MAE*) values between 0.222 and 0.290.

Reference [36] presents the research on forecasting power factors in electrical power systems using machine learning methods, such as supervised, unsupervised, and reinforcement algorithms. The *MAE* values varied between 0.099 and 0.135, whereas the *RMSE* values ranged from 0.029 to 0.175.

On the other hand, when looking at the long-term forecast, which is the most significant aspect, we notice that there are some differences. At low values of the LAG parameter, the *MAE* and *RMSE* indicators have low values; we could say that they are very good;

nevertheless, the R-squared parameter also has low values, and in fact, it has values that are far too low. On the other hand, if we take into account these parameters in their whole, as is required, we can see from the graphs and from Table 1 that the best results (in their entirety) are obtained for values that fall between the range of 240 to 300. Regarding this particular interval, the forecast can be considered as being accurate. Because of this, it is possible to make an accurate prediction of the power factor values, even beyond the horizon, up to 400 samples in advance. We analyze the performance of the forecasting in relation to the model parameters, and we came to the conclusion that the best prediction performance is obtained for LAG in 170 to 210 and horizon 100 to 300. Considering training parameters, we have observed that performance is significantly influenced by the proportion of data utilized for training and testing. The accuracy of the prediction slightly decreases beyond the horizon for values between 0.6 and 0.7, but the overall prediction quality remains high. Values greater than 0.85 lead to a slight decrease in prediction accuracy while also increasing the training time. The prediction quality significantly reduces beyond these intervals.

7. Conclusions

After evaluating a deep learning model for PF forecasting, we determined that its accuracy is exceptionally high. Specifically, for long-term predictions beyond the horizon, the model achieved outstanding results for the root mean square error (*RMSE*) ranging from 0.088 to 0.105 and the mean absolute error (*MAE*) ranging from 0.071 to 0.086. When considering one step ahead, the prediction is considerably enhanced, with *RMSE* values ranging from 0.004 to 0.013 and *MAE* values ranging from 0.009 to 0.032. We have compared our work with other similar research studies that have conducted time series prediction on various datasets presented in the literature. Power quality parameters forecasting is conducted in reference [35] using SOM Maps with a KNN Algorithm. This study analyzed and obtained *RMSE* values for power factors ranging from 0.011 to 0.33, which were comparable to those obtained by us. Article [45] discusses the authors' results on time series forecasting utilizing methods such as CNN ANN and RNN ANN on various datasets from diverse applications, similar to the methods we implemented. The datasets most similar to ours are likely those related to electricity, with reported mean squared error (MSE) values ranging from 0.129 to 0.197 and mean absolute error (*MAE*) values ranging from 0.222 to 0.290. The authors in reference [36] present an investigation on power factor prediction in electrical power systems utilizing machine learning techniques, including supervised, unsupervised, and reinforcement algorithms. They obtained *MAE* values ranging from 0.099 to 0.135 and *RMSE* values ranging from 0.029 to 0.175. Predicting the power factor in metallurgical factories, particularly where rolling mills and electric arc furnaces are present, can offer several advantages. Accurate power factor forecasts can assist in optimizing energy usage. Comprehending the power factor enables suppliers to effectively control their electrical systems, resulting in lower energy costs and enhanced total energy efficiency. Via the prediction of power factor values, companies may develop ways to enhance power factor correction and prevent penalties linked to low power factor. A low power factor shows that the factory is not utilizing electricity efficiently, resulting in increased demand charges from utilities [36]. By forecasting these periods, measures can be taken to increase the power factor, such as using power factor correction equipment, minimizing wasted energy and increasing overall efficiency. This can result in cost savings by reducing consumption charges and enhancing the utilization of electrical infrastructure. Power factor forecasts can assist in capacity planning operations. Industries can proactively predict power factor fluctuations to effectively strategize for future electrical capacity requirements, thus ensuring that the existing infrastructure is capable of handling the expected demands. To implement a power factor forecasting strategy, a factory may monitor and analyze electrical system data and develop power factor correction installations. Being proactive in managing energy may contribute to significant cost savings. Load management can be optimized by investigating and forecasting power factor trends. This allows operators to properly schedule high power factor loads during peak demand periods in order to maintain energy usage while

maintaining a higher power factor level. Deep learning models have advantages over traditional forecasting models like Multilayer Perceptrons (MLP), neuro-fuzzy systems, and NARX models because they can effectively manage complex, high-dimensional data and recognize complex patterns. Deep learning models are proficient at handling extensive amounts of complex data and autonomously identifying relevant characteristics without requiring manual selection of features. Due to RNNs and LSTM networks, they can identify and learn from complex data patterns [42]. As the dataset expands, these models become more accurate and flexible for different forecasting tasks [45]. Deep learning models can generalize well to new, unknown data, making them robust for forecasting in dynamic situations. We also investigated the power factor prediction using ANN MLP models, although the accuracy was not as high as with deep learning models.

As a continuation of our research, we plan to extend the forecasting to include other elements that can also have an effect on the power quality (PQ), such as total harmonic distortion (THD), reactive power, and so on. We will take into account all of these parameters as a whole and use various hybrid deep learning models to forecast the PQ parameters.

Author Contributions: Conceptualization M.P. and P.I.; methodology, P.I.; software, M.P.; validation, C.P.; investigation, P.I. and C.P.; writing—original draft preparation, M.P.; writing—review and editing, C.P.; visualization, P.I.; supervision, M.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data came from the metallurgical plant and is under privacy restriction.

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

Abbreviations

ANN	Artificial Neural Networks
APF	Active Power Filter
CNN	Convolutional Neural Network
DL	Deep Learning
EAF	Electric Arc Furnace
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
PCC	Point of Common Coupling
PF	Power Factor
PWM	Pulse Width Modulation
PQ	Power Quality
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
sMAPE	Symmetric Mean Absolute Percentage Error
SVC	Static Var Compensator
THD	Total Harmonic Distortion

References

1. Orcajo, G.A.; Rodriguez, D.J.; Cano, J.M.; Norniella, J.G.; Ardura, G.P.; Llera, T.R.; Cifrian, R.D. Dynamic estimation of electrical demand in hot rolling mills. In Proceedings of the 2015 IEEE Industry Applications Society Annual Meeting, Addison, TX, USA, 18–22 October 2015; pp. 1–9. [CrossRef]
2. Ardura, P.; Alonso, G.; Cano, J.M.; Norniella, J.G.; Pedrayes, F.; Briz, F.; Cabanas, M.F.; Melero, M.G.; Rojas, C.H.; Suárez, J.R.G. Power Quality Analysis and Improvements in a Hot Rolling Mill using a STATCOM. In Proceedings of the International Conference on Renewable Energies and Power Quality (ICREPQ'14), Cordoba, Spain, 8–10 April 2014.
3. Nikolaev, A.A.; Anokhin, V.V.; Lozhkin, I.A. Estimation of accuracy of chosen SVC power for steel-making arc furnace. In Proceedings of the 2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Chelyabinsk, Russia, 19–20 May 2016; pp. 1–5. [CrossRef]

4. Panoiu, M.; Panoiu, C.; Sora, I.; Osaci, M. Using a model based on linearization of the current-voltage characteristic for electric arc simulation. In Proceedings of the 16th IASTED International Conference on Applied Simulation and Modelling, Palma de Majorca, Spain, 29–31 August 2007.

5. Liu, Y.-J.; Hung, J.-P. On real-time simulation for power quality assessment of a power system with multi steel plants. In Proceedings of the 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), Florence, Italy, 7–10 June 2016; pp. 1–5. [CrossRef]

6. Liu, B.; Liu, C.; Liu, J.; Zheng, X.; Xu, G.; Wang, B. Harmonic Control Measures in Yili Rolling Mill. In Proceedings of the 2018 China International Conference on Electricity Distribution (CICED), Tianjin, China, 17–19 September 2018; pp. 571–575. [CrossRef]

7. Ferreira, D.M.; Jota, P.R.S.; Souza, C.P. Harmonics in electrical industrial systems: A case study in a steel manufacturing facility. In Proceedings of the 11th International Conference on Electrical Power Quality and Utilisation, Lisbon, Portugal, 17–19 October 2011; pp. 1–6. [CrossRef]

8. Alonso Orcajo, G.A.; Diez, J.R.; Cano, J.M.; Norniella, J.G.; González, J.F.; Rojas, C.H.; Ardura, P.; Cifrián, D. Enhancement of Power Quality in an Actual Hot Rolling Mill Plant Through a STATCOM. *IEEE Trans. Ind. Appl.* **2020**, *56*, 3238–3249. [CrossRef]

9. Ghiormez, L.; Panoiu, M.; Panoiu, C.; Rob, R. Electric Arc Model in PSCAD—EMTDC as embedded component and the dependency of the desired Active Power. In Proceedings of the 2016 IEEE 14th International Conference On Industrial Informatics (INDIN), Poitiers, France, 19–21 July 2016; pp. 351–356.

10. Panoiu, M.; Panoiu, C.; Sora, I.; Osaci, M. Simulation results on the reactive power compensation process on electric arc furnace using PSCAD-EMTDC. *Int. J. Model. Identif. Control.* **2007**, *2*, 250–257. [CrossRef]

11. Orcajo, G.A.; Diez, J.R.; Cano, J.M.; Norniella, J.G.; González, J.F.P.; Rojas, C.H. Coordinated Management of Electrical Energy in a Steelworks and a Wind Farm. *IEEE Trans. Ind. Appl.* **2022**, *58*, 5488–5502. [CrossRef]

12. IAbdulveleev, R.; Khramshin, T.R.; Kornilov, G.P. Novel Hybrid Cascade H-Bridge Active Power Filter with Star Configuration for Nonlinear Powerful Industrial Loads. In Proceedings of the 2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Moscow, Russia, 15–18 May 2018; pp. 1–7. [CrossRef]

13. Hsu, C.-T.; Chen, C.-S.; Lin, C.-H. Electric Power System Analysis and Design of an Expanding Steel Cogeneration Plant. *IEEE Trans. Ind. Appl.* **2011**, *47*, 1527–1535. [CrossRef]

14. Nikolaev, A.; Maklakov, A.; Bulanov, M.; Gilemov, I.; Denisevich, A.; Afanasev, M. Current Electromagnetic Compatibility Problems of High-Power Industrial Electric Drives with Active Front-End Rectifiers Connected to a 6–35 kV Power Grid: A Comprehensive Overview. *Energies* **2023**, *16*, 293. [CrossRef]

15. Nikolaev, A.A.; Bulanov, M.V.; Gilemov, I.G.; Maklakov, A.S. Improving the Electromagnetic Compatibility of Hot Rolling Mill Electric Drives at CJSC "MMK Metalurji with Supply Network 34.5 kV. In Proceedings of the 2023 International Ural Conference on Electrical Power Engineering, Magnitogorsk, Russia, 29 September–1 October 2023; pp. 455–460.

16. Maklakov, A.S.; Jing, T.; Nikolaev, A.A.; Gasiyarov, V.R. Grid Connection Circuits for Powerful Regenerative Electric Drives of Rolling Mills: Review. *Energies* **2022**, *15*, 8608. [CrossRef]

17. Orcajo, G.A.; Gonzalez, F.P.; Cano, J.M.; Norniella, J.G.; Rojas, C.H.; Rodríguez Diez, J. Voltage Sag Ride-Through in a Joint Installation of a Hot Rolling Mill Plant and a Wind Far. *IEEE Trans. Ind. Appl.* **2023**, *59*, 5190–5200. [CrossRef]

18. Nikolaev, A.A.; Bulanov, M.V.; Antropova, L.I. Ways to Ensure Electromagnetic Compatibility of Powerful Frequency Converters in Internal Power Supply Systems of Industrial Enterprises in the Presence of Resonance Phenomena. In Proceedings of the 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russia, 25–29 March 2019; pp. 1–6. [CrossRef]

19. Orcajo, G.A.; Ardura, P.; Rodríguez, J.; Cano, J.M.; Norniella, J.G.; Llera, R.; Cifrián, D. Overcurrent Protection Response of a Hot Rolling Mill Filtering System: Analysis of the Process Conditions. *IEEE Trans. Ind. Appl.* **2017**, *53*, 2596–2607. [CrossRef]

20. Nowak, S.; Kocman, S. Unfavourable Reactive Power in a Rolling Mill. In Proceedings of the 19th International Conference on Renewable Energies and Power Quality (ICREPQ'21), Almeria, Spain, 28–30 July 2021.

21. Apse-Apsitis, P.; Krievs, O.; Avotins, A. Impact of Household PV Generation on the Voltage Quality in 0.4 kV Electric Grid—Case Study. *Energies* **2023**, *16*, 2554. [CrossRef]

22. Yan, K.; Wang, X.; Du, Y.; Jin, N.; Huang, H.; Zhou, H. Multi-Step Short-Term Power Consumption Forecasting with a Hybrid Deep Learning Strategy. *Energies* **2018**, *11*, 3089. [CrossRef]

23. EKuyunani, M.; Hasan, A.N.; Shongwe, T. Improving voltage harmonics forecasting at a wind farm using deep learning techniques. In Proceedings of the 2021 IEEE 30th International Symposium on Industrial Electronics (ISIE), Kyoto, Japan, 20–23 June 2021; pp. 1–6. [CrossRef]

24. Yang, J.; Ma, H.; Dou, J.; Guo, R. Harmonic Characteristics Data-Driven THD Prediction Method for LEDs Using MEA-GRNN and Improved-AdaBoost Algorithm. *IEEE Access* **2021**, *9*, 31297–31308. [CrossRef]

25. DNolasco, H.S.; Alves, D.K.; Costa, F.B.; Palmeira, E.S.; Ribeiro, R.L.A.; Nunes, E.A.F. Application of Fuzzy Systems in Power Quality: Diagnosis of Total Harmonic Distortions. In Proceedings of the 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–6. [CrossRef]

26. Alghamdi, T.A.H.; Abdusalam, O.T.E.; Anayi, F.; Packianather, M. An artificial neural network based harmonic distortions estimator for grid-connected power converter-based applications. *Ain Shams Eng. J.* **2023**, *14*, 101916. [CrossRef]

27. Garrido-Zafra, J.; Gil-de-Castro, A.; Calleja-Madueño, A.; Linan-Reyes, M.; Moreno-Garcia, I.; Moreno-Munoz, A. LSTM-based Network for Current Harmonic Distortion Time Series Forecasting. In Proceedings of the 2023 IEEE International Conference on Environment and Electrical Engineering and 2023 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), Madrid, Spain, 6–9 June 2023; pp. 1–6. [CrossRef]
28. Balouji, E.; Salor, Ö.; McKelvey, T. Deep Learning Based Predictive Compensation of Flicker, Voltage Dips, Harmonics and Interharmonics in Electric Arc Furnaces. *IEEE Trans. Ind. Appl.* **2022**, *58*, 4214–4224. [CrossRef]
29. Panoiu, M.; Panoiu, C.; Mezinescu, S.; Militaru, G.; Baciu, I. Machine Learning Techniques Applied to the Harmonic Analysis of Railway Power Supply. *Mathematics* **2023**, *11*, 1381. [CrossRef]
30. Hu, Z.; Wei, Z.; Sun, H.; Yang, J.; Wei, L. Optimization of Metal Rolling Control Using Soft Computing Approaches: A Review. *Arch. Comput. Methods Eng.* **2021**, *28*, 405–421. [CrossRef]
31. Orcajo, G.A.; Cano, J.M.; Norniella, J.G.; Pedrayes, F.; Rojas, C.H.; Josué, R.D.; Ardura, P.; Diego, C.R. Power Quality Improvement in a Hot Rolling Mill Plant Using a Cascaded H-Bridge STATCOM. In Proceedings of the 2019 IEEE Industry Applications Society Annual Meeting, Baltimore, MD, USA, 29 September–3 October 2019; pp. 1–9. [CrossRef]
32. Nikolaev, A.A.; Denisevich, A.S.; Antropova, L.I. Sustainability of High-Power Frequency Converters with Active Rectifiers Connected in Parallel with "EAF-SVC Complex. In Proceedings of the 2019 IEEE Russian Workshop on Power Engineering and Automation of Metallurgy Industry: Research & Practice (PEAMI), Magnitogorsk, Russia, 4–5 October 2019; pp. 127–133. [CrossRef]
33. Shende, A.G.; Khubalkar, S.W.; Vaidya, P. Hardware Implementation of Automatic Power Factor Correction Unit For Industry. *J. Phys. Conf. Ser.* **2021**, *2089*, 012032. [CrossRef]
34. Meshcheryakov, V.N.; Evseev, A.M.; Didenko, E.E. Joint Control of Looper Electric Drive of Finishing Mill Group and Active Energy Filter. In Proceedings of the 2018 International Russian Automation Conference (RusAutoCon), Sochi, Russia, 9–16 September 2018; pp. 1–4. [CrossRef]
35. Jahan, I.; Mohamed, F.; Blazek, V.; Prokop, L.; Misak, S.; Snasel, V. Power Quality Parameters Forecasting Based on SOM Maps with KNN Algorithm and Decision Tree. In Proceedings of the 2023 23rd International Scientific Conference on Electric Power Engineering (EPE), Brno, Czech Republic, 24–26 May 2023; pp. 1–6. [CrossRef]
36. Gámez Medina, J.M.; de la Torre y Ramos, J.; López Monteagudo, F.E.; Ríos Rodríguez, L.d.C.; Esparza, D.; Rivas, J.M.; Ruvalcaba Arredondo, L.; Romero Moyano, A.A. Power Factor Prediction in Three Phase Electrical Power Systems Using Machine Learning. *Sustainability* **2022**, *14*, 9113. [CrossRef]
37. Pawanr, S.; Garg, G.K.; Routroy, S. Prediction of energy efficiency, power factor and associated carbon emissions of machine tools using soft computing techniques. *Int. J. Interact. Des. Manuf.* **2023**, *17*, 1165–1183. [CrossRef]
38. Available online: https://towardsdatascience.com/a-brief-introduction-to-recurrent-neural-networks-638f64a61ff4 (accessed on 11 December 2023).
39. Samanta, I.S.; Panda, S.; Rout, P.K.; Bajaj, M.; Piecha, M.; Blazek, V.; Prokop, L. A Comprehensive Review of Deep-Learning Applications to Power Quality Analysis. *Energies* **2023**, *16*, 4406. [CrossRef]
40. Casolaro, A.; Capone, V.; Iannuzzo, G.; Camastra, F. Deep Learning for Time Series Forecasting: Advances and Open Problems. *Information* **2023**, *14*, 598. [CrossRef]
41. Wang, K.; Zhang, J.; Li, X.; Zhang, Y. Long-Term Power Load Forecasting Using LSTM-Informer with Ensemble Learning. *Electronics* **2023**, *12*, 2175. [CrossRef]
42. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. *Energies* **2018**, *11*, 1636. [CrossRef]
43. Iordan, A.-E. An Optimized LSTM Neural Network for Accurate Estimation of Software Development Effort. *Mathematics* **2024**, *12*, 200. [CrossRef]
44. Azevedo Costa, M. Forecasting Hierarchical Time Series in Power Generation. *Energies* **2020**, *13*, 3722. [CrossRef]
45. Westergaard, G.; Erden, U.; Mateo, O.A.; Lampo, S.M.; Akinci, T.C.; Topsakal, O. Time Series Forecasting Utilizing Automated Machine Learning (AutoML): A Comparative Analysis Study on Diverse Datasets. *Information* **2024**, *15*, 39. [CrossRef]
46. Chis, V.; Barbulescu, C.; Kilyeni, S.; Dzitac, S. ANN based Short-Term Load Curve Forecasting. *Int. J. Comput. Commun. Control.* **2018**, *13*, 938–955. [CrossRef]
47. Wang, Z.; Jia, L.; Ren, C. Attention-Bidirectional LSTM Based Short Term Power Load Forecasting. In Proceedings of the 2021 Power System and Green Energy Conference (PSGEC), Shanghai, China, 20–22 August 2021; pp. 171–175. [CrossRef]