Electricity Consumption Forecasting Based on a Bidirectional Long-Short-Term Memory Artificial Neural Network

Dana-Mihaela Petrosanu,1,*, and Alexandru Pirjan2,  

1 Department of Mathematics-Informatics, Faculty of Applied Sciences, University Politehnica of Bucharest, 060042 Bucharest, Romania  
2 Department of Informatics, Statistics and Mathematics, Romanian-American University, 012101 Bucharest, Romania; alex@pirjan.com

Abstract: The accurate forecasting of the hourly month-ahead electricity consumption represents a very important aspect for non-household electricity consumers and system operators, and at the same time represents a key factor in what regards energy efficiency and achieving sustainable economic, business, and management operations. In this context, we have devised, developed, and validated within the paper an hourly month ahead electricity consumption forecasting method. This method is based on a bidirectional long-short-term memory (BiLSTM) artificial neural network (ANN) enhanced with a multiple simultaneously decreasing delays approach coupled with function fitting neural networks (FITNETs). The developed method targets the hourly month-ahead total electricity consumption at the level of a commercial center-type consumer and for the hourly month ahead consumption of its refrigerator storage room. The developed approach offers excellent forecasting results, highlighted by the validation stage’s results along with the registered performance metrics, namely 0.0495 for the root mean square error (RMSE) performance metric for the total hourly month-ahead electricity consumption and 0.0284 for the refrigerator storage room. We aimed for and managed to attain an hourly month-ahead consumed electricity prediction without experiencing a significant drop in the forecasting accuracy that usually tends to occur after the first two weeks, therefore achieving a reliable method that satisfies the contractor’s needs, being able to enhance his/her activity from the economic, business, and management perspectives. Even if the devised, developed, and validated forecasting solution for the hourly consumption targets a commercial center-type consumer, based on its accuracy, this solution can also represent a useful tool for other non-household electricity consumers due to its generalization capability.

Keywords: long term electricity consumption forecasting; artificial neural networks (ANNs); bidirectional long-short-term memory (BiLSTM) networks; function fitting neural networks (FITNETs); large commercial center-type consumers

1. Introduction

1.1. Motivation

According to the International Energy Agency, between 1971 and 2018, the world total final energy consumption multiplied by 2.3, growing from 4243 "millions of tons of oil equivalent (Mtoe)” in 1971 to 9938 Mtoe in 2018 [1]. While the share of energy consumption has been almost constant over the years in most sectors, energy consumption in non-residential sectors has increased over the course of time, from 76% in 1971 to 79% in 2018. The non-residential sector, comprising the consumption arising from industry, transport, commercial activities, services rendered in the public interest, agricultural activities, forest-related activities and unspecified other sectors, represents the largest sector, which accounts for the largest percentage within the world total final energy consumption. Considering the mentioned indicators, one is able to deduce that attaining efficiency with regard to
energy within non-household sectors impacts in a considerable manner the worldwide efforts to save fuel, to reduce pollutants, to achieve sustainable economic, business, and management operations, and to enhance the quality of life.

Consequently, in order to attain worldwide energy efficiency, a factor of major importance consists in providing an accurate prediction of the consumers’ electricity consumption belonging to the large non-household category. Moreover, such an accurate forecast helps in optimizing the electricity consumption and influencing the consumption strategy, and facilitates the efforts made in view of attaining an appropriate environment and natural resource management. In this context, the accurate forecasting of the hourly month ahead electricity consumption is a very important aspect for non-household electricity consumers and system operators, and at the same time represents an essential factor in what regards energy efficiency.

In our country, Romania, the total energy consumption increased at a pace of 1.3% per year between 2014 and 2019, year in which it reached the value of 34 Mtoe. The non-household final consumers represent the largest sector, representing an amount of 74% of the total electricity consumption in 2019 (45% of the total electricity consumption is due to the industry, while 17% is represented by the services) [2].

Lately, in the scientific literature, the interest regarding the accurate forecast of the consumed electricity has been growing at an ever-accelerating pace, the researchers developing a wide range of applications, methods, and techniques in order to obtain forecasts of time series related to energy datasets. In this paper, we have devised, developed, and validated an electricity consumption forecasting method based on a bidirectional long-short-term memory (BiLSTM) artificial neural network (ANN) enhanced with a multiple simultaneously decreasing delays approach coupled with function fitting neural networks (FITNETs) for the hourly month-ahead total electricity consumption at the level of a commercial center-type consumer and for the hourly month-ahead consumption of its refrigerator storage room.

In order to emphasize the results related to the approached subject reported within the scientific literature and to contextualize the research developed within this paper, a literature review is performed in the following subsection, having as a main purpose the necessity to pinpoint an exact deficiency, an unsolved problem, and a gap that still exists in the current body of knowledge that is being addressed by the research presented within this paper.

Afterwards, the “Materials and Methods” section depicts the most important notions that have been applied in the development of the prediction method along with details concerning the stages and steps that represent the building blocks for our proposed approach.

Afterwards, the following section, “Results”, presents the results obtained over the course of the stages and steps of the prediction method developed based on a BiLSTM ANN enhanced with a multiple simultaneously decreasing delays approach coupled with FITNETs, for the month-ahead total electricity consumption at the level of the commercial center type-consumer and the month-ahead consumption of the refrigerator storage room. Firstly, we present the results regarding the developed BiLSTM ANN total hourly electricity consumption forecasting solution with multiple delays support, based on the adaptive moment estimation (ADAM), stochastic gradient descent with momentum (SGDM), and root mean square propagation (RMSProp) training algorithms. Secondly, we depict the results with regard to the FITNET ANN prediction approach for the refrigerator storage room, based on the Levenberg–Marquardt (LM), Bayesian regularization (BR), and scaled conjugate gradient (SCG) training algorithms. Lastly, we present the results registered during the validation process of the developed method, by analyzing the registered forecasting results.

Subsequently, in the fourth section, the “Discussion”, we analyze and interpret the forecasting results obtained through the developed approach, in contrast with other existing studies and approaches from the body of knowledge targeting resembling issues, highlighting also future research directions and limitations of our proposed approach.
In the “Conclusions” section, we emphasize the most important outcomes and findings of the paper, by interpreting our findings at a higher level of abstraction than in the “Discussion” section. We highlight and justify the fact that we have managed to address the target proposed within the “Introduction” section.

1.2. Literature Review

In the research paper [3], Moradzadeh et al. devised a hybrid method comprising of an adjusted support vector regression (SVR) and long short-term memory (LSTM) ANN, entitled SVR-LSTM, with a view to obtaining a short-term forecast of the electricity consumption of a microgrid located in Sub-Saharan Africa. For the input variables, the authors took into account different factors that have an effect on the microgrid’s load; factors like the commercial establishments, different family units consisting of houses and their occupants, while for the outputs they used electricity load profiles. The hybrid approach was benchmarked against its single constituent parts and similar approaches from the scientific literature, providing superior results in terms of the correlation coefficient (R), mean squared error (MSE), RMSE, mean absolute error (MAE), and mean absolute percentage error (MAPE) performance metrics.

Kim et al. propose in [4] a hybrid approach, entitled “(c, l)-LSTM + convolutional neural network (CNN)”, used for predicting the electricity demand in view of assuring an efficient management and appropriate operational activities for a smart grid. The authors’ proposed method involves two variates sequences, comprising pairs of key values, associated with contextual information, that are inserted into LSTM ANNs to obtain sets of features that are afterwards used by a CNN to obtain a forecast of the electricity demand. After assessing the performance of their method by comparing it with the autoregressive integrated moving average (ARIMA), the (c, l)-LSTM, and with a sequence-to-sequence (S2S) LSTM in terms of the MAPE and relative RMSE (RRMSE) performance metrics, the authors state that their method yielded better results in terms of accuracy.

Within the scientific article [5], Yaprakdal et al. propose a forecasting method based on deep recurrent neural networks (DRNNs), consisting of a BiLSTM ANN, entitled “DRNN Bi-LSTM”, in order to obtain an accurate estimation for the day-ahead electricity demand and the total photovoltaic electricity production for a microgrid equipped with both photovoltaic panels and diesel generators. The authors applied their method on a real production dataset. They compared their proposed method with other approaches from the literature, namely with a deep CNN (D-CNN) [6], a Copula model deep belief network (Copula-DBN) [7], a DRNN-LSTM [8], and with parallel convolutional recurrent neural networks (CNN-RNN) [9]. Finally, the authors concluded that their devised method offers superior results as depicted by the MAPE performance metric.

Recognizing the importance of attaining an accurate prediction of the day-ahead hourly electricity consumption in order to optimize and appropriately manage the energy related resources, in [10], Santra et al. propose an approach consisting of LSTM ANNs for solving the problems related to long-term data dependencies and genetic algorithm (GA) to get the most appropriate parameters for the LSTM ANNs. The authors trained the network using real electricity load and meteorological datasets, and afterwards, they tested their proposed method and compared it with the results provided by a LSTM ANN that does not have its parameters tuned by means of a GA, and concluded the superiority of their approach based on the MAPE performance metric that highlighted a higher accuracy level.

Acknowledging the importance that an accurate monthly electricity demand forecast exerts at both local and countrywide levels, along with the numerous challenges that must be overcome in attaining these goals when referring to the residential segment, in the scientific article [11], Son et al. put forward a prediction method based on LSTM ANNs. The authors explain this choice as being driven by the robustness and efficiency that LSTM ANNs exhibited during the course of time when it came to predicting time series. In training the model, the authors used a dataset comprising of 22 years of monthly electricity consumption from South Korea, while for evaluating the performance, the MAE, RMSE,
MAPE, post-error ratio (C), mean bias error (MBE), and the unpaired peak accuracy (UPA) performance metrics were computed and analyzed. After having benchmarked the proposed method against the SVR, the backpropagation ANN, the ARIMA, and the multiple linear regression (MLR), the authors state that they registered very good results, the LSTM approach having proved to be superior in all of the analyzed cases.

Considering the increase in the number of smart metering equipment and infrastructure throughout the world, an increase that has led to huge volumes of useful data related to electricity consumption and disaggregated loads, in [12], Alonso et al. introduce a processing and forecasting methodology for multiple time series datasets recorded by means of smart-metering devices. In contrast to classical and univariate methods, the authors put forward a RNN model with LSTM for the purpose of identifying different patterns of electricity consumption stemming from different individual electricity consumers and households, with the goal of achieving day-ahead accurate predictions while consuming very low computational resources. After benchmarking and computing the MSE along with the MAE performance metrics, the authors conclude the efficiency of their devised LSTM model in forecasting electricity consumption time series. In future work, the authors aim to achieve a larger prediction time frame, namely a week-ahead forecast, to experiment with various LSTM architectures and different experimental conditions, from diverse datasets to dissimilar geographic regions.

In the scientific article [13], taking into account that attaining a long-term forecast of the electricity consumption is of paramount importance from both the monetary and time resources perspectives, Almazrouee et al. perform a comparative analysis between the Prophet and Holt–Winters long-term forecasting models in terms of prediction accuracy and efficiency. When conducting their study, the authors used real electricity consumption datasets ranging between the years 2010 and 2020 from the powerplants located in Kuwait in order to predict the electricity consumption for the following 10 years, up to the year 2030. After benchmarking their method, the authors observed that the Prophet model yielded forecasts with a higher level of accuracy, as highlighted by the RMSE, the MAPE, coefficient of determination ($R^2$), MAE, and coefficient of variation of RMSE (CVRMSE) performance indicators. In addition, the authors inserted Gaussian white noise of different intensities in order to assess the robustness of the two analyzed models and concluded that the Prophet model is more resilient to noise than the Holt–Winters one.

In their paper [14], Shao et al. put forward a hybrid model in order to provide “multiple forecasts by combining CNNs and LSTM” ANNs. The hybrid model was developed in order to parallelly process the dataset, therefore improving the robustness of the conventional CNN–LSTM approach. Afterwards, the features extracted using the above-mentioned methods were used along with six statistical variables, namely “mean, max, min, standard deviation (Sd), skewness (Skew), and kurtosis (Kurt)”, and the developed model was validated using three real datasets coming from three American companies. The authors compare their model with others, based on deep learning (DL), such as DNN-LSTM [15], nonpooling convolutional neural network LSTM (NPCNN-LSTM) [16], LSTM-LSTM [17], CNN-LSTM [18] and conclude that their developed approach has proven its robustness and efficiency when forecasting the electricity consumption for different time horizons, ranging from very short term to long term.

In the context of the worldwide ever-increasing electricity demand, Mir et al. present an insight of the forecasting approaches designed in the scientific literature in order to estimate future electricity demands in the cases of countries having low and middle incomes [19]. The devised study takes into account different prediction methodologies from the body of knowledge, targeting various time horizons, and remark that in the cases when forecasting for long [20] or medium time [21] intervals, the models based on time series have been mainly used, while for short term intervals [22], the techniques have made use of artificial intelligence (AI). The authors remark that in each country the electricity demand is influenced by a series of determinants, such as the population of the respective country, the gross domestic product (GDP), the local weather conditions, and the
consumption for different moments and intervals of time (as local consumption habits). In this paper, the authors also present a comparative analysis of the results retrieved from the scientific literature with the ones from their country, Pakistan. The authors conclude their developed analysis of the scientific literature by identifying a series of research gaps in the case of Pakistan that will be worth narrowing in the future, such as the limited number of studies regarding short-term load forecasting (STLF) and the lack of medium-term load forecasting (MTLF) studies.

In the article [23], Marulanda et al. develop a method with a view to obtaining long-term scenarios with regard to the forecasting of the produced wind energy, considering spatial and temporal dependencies between energy markets from different areas. After having assessed the dependencies using a bottom-up strategy and real production datasets from France, Spain, and Portugal, the authors concluded that the ordered quantile transformation provided the most appropriate normalization of the time series when compared to the logit Yeo–Johnson and Box–Cox ones. In what concerns the data division process, the authors remarked that segmenting the datasets monthly led to better results, while the comparison with the pinball loss function and with the Winkler score metrics highlighted the superiority of their approach.

Being aware of the importance of electricity load forecasting in view of planning efficient operation of supply networks, within the scientific article [24], Sloyali implemented a series of machine learning methods, namely ANN, MLR, adaptive neuro-fuzzy inference system (ANFIS), and support vector machine (SVM) with the purpose of estimating energy demand, and discussed criteria related to its generation in Cyprus. The author remarked that the electricity load depends on external parameters such as weather, humidity, and solar irradiation, but also on national factors, obtaining therefore the parameters used as inputs for the machine learning algorithms. Afterwards, based on datasets representing the electricity usage in the years 2016 and 2017, the author developed simulations in the purpose of evaluating the algorithms’ performance and analyzing the short- and long-term forecasting. The study concludes that in both cases the implemented approaches offer a better performance than the rest of the analyzed ML-based ones, due to the fact that they provide more accurate forecasting results.

In light of the extreme importance and numerous benefits that a long-term accurate forecast of the electricity consumption can bring to energy providers and dispatchers, Elkamel et al. propose in [25] a model based on the MLR along with three CNNs for predicting electricity consumption for the state of Florida. The authors developed and compared the performance of univariant, multichannel, and multilayer ANNs along with the MLR. The MSE, RMSE, MAE, and MAPE performance metrics revealed the superiority of the multichannel and multilayer CNNs in contrast to the model based on the MLR, the accuracy of the CNN approach having increased along with the size of the dataset. The authors concluded that the multichannel CNN provided the best results, being the most suitable option of the analyzed ones for fulfilling the objective of attaining a long-term electricity forecast in the case of Florida.

Within the research paper [26], Son et al. bring forward an electricity consumption prediction method developed based on deep neural network (DNN) and LSTM ANNs covering the medium- and long-term forecasting time frame. The performance metrics, consisting of the MAE, RMSE, MAPE, $R^2$, and the CVRMSE computed according to the definition from [27], convey the superiority of the DNN approach in contrast to the long short-term one. The authors stated that in their future work they intend to address the existing limitations pertaining to the lack of weather data and seasonality.

In a recent paper, Ribeiro et al. started their research by highlighting the importance of benefitting from accurate energy consumption forecasts in the case of companies that are energy intensive consumers [28]. In addition, the authors highlight the importance of implementing DL techniques in energy demand forecasting. In this purpose, the authors compare two benchmarks, namely the ARIMA and a technique that exists on site, against three DL approaches, namely RNN, LSTM, and gated recurrent unit (GRU), and two
approaches based on the SVR and random forest models. The comparison was conducted using a short-term time horizon for the obtained forecasts, within which a dataset retrieved from a Brazilian company was used. The developed comparison showed that the approach already implemented at the case site provided the worst results, while the ones provided by the GRU model surpassed all the other approaches that were tested within this study.

Considering from both the integration and operational costs point of views how beneficial is for power grid operators to be able to obtain their own power generation forecasts, in [29], Kim et al. develop a photovoltaic electricity production forecasting method for a one-hour-ahead, an approach that combines a BiLSTM in two steps with an ANN that uses the exponential moving average (EMA) for pre-processing the datasets. After benchmarking and computing the RMSE, MAPE, and the $R^2$, the authors concluded that the combined method is superior to others.

In order to better highlight the results presented above, related to the approached subject reported within the scientific literature, we have synthetized them within a summarization table (Table 1).

| Reference | Publication Year | Achieved Forecasting Timeframe | Forecasting Purpose | Approach | Comparison with Other Methods | Performance Metrics |
|-----------|------------------|--------------------------------|---------------------|----------|-------------------------------|---------------------|
| [3]       | 2020             | short-term                    | electricity consumption | SVR-LSTM | SVR, LSTM                      | R, MSE, RMSE, MAE, MAPE |
| [4]       | 2019             | long-term                     | electricity demand    | LSTM-CNN | ARIMA, (c, l)-LSTM, S2S LSTM   | MAPE, RRMSE         |
| [5]       | 2020             | short-term                    | electricity demand    | DRNN Bi-LSTM | D-CNN, Copula-DBN, DRNN-LSTM, CNN-RNN | MAPE |
| [10]      | 2019             | short-term                    | electricity consumption | LSTM, GA | LSTM                          | MAPE |
| [11]      | 2020             | long-term                     | electricity demand    | LSTM     | backpropagation ANN, ARIMA, MLR | MAE, RMSE, MAPE, C, MBE, UPA |
| [12]      | 2020             | short-term                    | electricity consumption | RNN-LSTM | other state-of-the-art approaches, not explicitly stated | MSE, MAE |
| [13]      | 2020             | long-term                     | electricity consumption | Prophet, Holt-Winters | Prophet against Holt-Winters | RMSE, MAPE, $R^2$, MAE, CVRMSE |
| [14]      | 2020             | very short-/medium-/long-term | electricity consumption | CNN-LSTM | DNN-LSTM, NPCNN-LSTM, LSTM-LSTM, CNN-LSTM | RMSE, MAE, MAE |
| [19]      | 2020             | short-/medium-/long-term      | electricity demand    | models based on time series; AI models | N/A (review paper) | N/A (review paper) |
| [23]      | 2020             | long-term                     | produced wind energy  | seasonal ARIMA | two approaches for seasonal ARIMA | “a bottom-up approach” |
| [24]      | 2020             | short-/long-term              | electricity load      | ANN, MLR, ANFIS, SVM | ML methods | prediction error, RMSE |
| [25]      | 2020             | long-term                     | electricity consumption | MLR, CNNs | univariant, multichannel, and multihead ANNs | MSE, RMSE, MAE, MAE |
Table 1. Cont.

| Reference | Publication Year | Achieved Forecasting Timeframe | Forecasting Purpose                     | Approach                        | Comparison with Other Methods                                      | Performance Metrics   |
|-----------|------------------|--------------------------------|----------------------------------------|--------------------------------|---------------------------------------------------------------------|-----------------------|
|           |                  |                                | electricity consumption                | DNN, LSTM                       | DNN against LSTM                                                    | MAE, RMSE, MAPE, R², CVRMSE |
| [26]      | 2020             | medium-/long-term              |                                        |                                |                                                                     |                       |
|           |                  |                                | energy consumption                    | ARIMA, a technique existing at the case site, RNN, LSTM, GRU, SVR, random forest | ARIMA, a technique existing at the case site, RNN, LSTM, GRU, SVR, Random Forest |                       |
| [28]      | 2020             | short-term                     | photovoltaic electricity production   | BD-LSTM, EMA                    | BD-LSTM, EMA, the combined approach                                 | RMSE, MAPE, MAE       |
| [29]      | 2020             | short-term                     |                                        |                                |                                                                     |                       |

Through the analysis of the current state of knowledge presented above, one can identify a true necessity, a gap in the current body of knowledge with regard to achieving an accurate long-term hourly forecast for the month-ahead consumed electricity in the case of large commercial center-type electricity consumers without experiencing a significant drop in the forecasting accuracy that usually tends to occur after the first two weeks. Therefore, by pursuing this path, one achieves a reliable, accurate forecasting method that satisfies the contractor’s needs, helping him/her achieve sustainable economic, business, and management commercial operations. In addition to these, another gap insufficiently addressed within the scientific literature in the case of large commercial center-type electricity consumers, is represented by a forecasting method for the hourly month-ahead total electricity consumption at the level of a commercial center-type consumer and for an individual appliance that helps the contractor by warning him/her regarding an abnormal growth trend of the consumed electricity of the respective appliance, therefore signaling him/her the fact that it is not functioning properly and needs maintenance operations.

In the following, we have synthesized the main contributions brought by the proposed forecasting method, highlighting the aspects related to narrowing and filling the gap identified within the scientific literature.

1.3. Contributions of the Paper

In order to fill the gap identified within the body of knowledge, in this paper we have devised, developed, and validated an hourly month-ahead electricity consumption forecasting method based on a BiLSTM ANN enhanced with a multiple simultaneously decreasing delays approach coupled with FITNETs, for the hourly month-ahead total electricity consumption at the level of a commercial center-type consumer and for the hourly month-ahead consumption of its refrigerator storage room. The highlights of the main contributions brought by our developed approach are the following:

- Within the paper, a forecasting method for the hourly month-ahead total electricity consumption at the level of a commercial center-type consumer and for the hourly month ahead consumption of its refrigerator storage room is devised, developed, and validated in a real production environment, overcoming in this way the numerous challenges that the forecasting of the electricity consumption poses to the researchers.
- The proposed method is based on obtaining the total hourly electricity consumption forecasting solution with multiple delays support by developing first a series of BiLSTM ANNs based on the ADAM, SGDM, and RMSProp training algorithms. Afterwards, this forecast is further processed by using a series of FITNET ANN forecasting solutions for the hourly consumption of the refrigerator storage room based on the LM, BR, and SCG training algorithms.
The developed forecasting method encompasses the advantages of BiLSTM ANNs enhanced with a multiple simultaneously decreasing delays approach coupled with FITNETs, the main advantage of the BiLSTM approach being their capacity to learn long-term dependencies in both directions, by passing through data twice: backward and forward; while the main advantage of the FITNETs consists in their capacity to forecast accurately in the context of a fast-training process. The developed approach offers excellent forecasting results, highlighted by the validation stage’s results along with the registered performance metrics. For example, in the case of the BiLSTM ANNs, the best registered values for RMSE are: 0.0307 during the training process, 0.0327 for the integral set, and 0.0495 during the testing process. When forecasting the month-ahead electricity consumption of the refrigerator storage room using the best of the developed FITNET ANNs, we obtained an excellent value of the RMSE performance metric during the forecasting process, namely 0.0284.

One of the most important aspects related to the forecasting accuracy is represented by the fact that we managed to attain an hourly month-ahead consumed electricity prediction without experiencing a significant drop in the forecasting accuracy that usually tends to occur after the first two weeks, therefore achieving a reliable method that satisfies the contractor’s needs, being able to help him/her attain energy efficiency and achieve sustainable economic, business, and management operations.

In all the analyzed cases, the training times registered very good values (namely, in the case of the ANNs that provided the best forecasting accuracy, 1462.415 s for the FITNET ANN and 0.64699 s for the BiLSTM ANN), without requiring investments in an expensive hardware configuration. This represents a very important advantage for the moment when the devised solution is put into operation and the ANNs need subsequent re-training steps, because over the course of time new datasets are registered and provided to the developed neural networks as inputs.

In the “Discussion” and “Conclusions” sections of the paper, we present more details regarding the added value of this research. In the following, the “Materials and Methods” section is presented.

2. Materials and Methods

The main purpose behind our devised research methodology has been the need to design, develop, and validate an efficient forecasting method that is able to provide large commercial center-type electricity consumers with as accurate as possible month-ahead hourly electricity consumption projections. We have acknowledged from the very beginning the fact that the provided forecasts must cover the whole hourly month-ahead time frame and at the same time be as reliable as possible in what concerns the prediction accuracy for tackling the contractor’s real needs and therefore to be able to improve the overall economic, management, and business aspects of his/her activity.

Considering that we have tackled previous challenges related to electricity forecasting over the course of our previous studies and research grants, the first logical step was to try and tap into our already devised materials and methods in this field and assess their suitability for the current purpose. Therefore, as in one of our research group’s previous studies [30], the profile of the electricity consumer was of commercial center-type and the forecasting horizon consisted of hourly month-ahead predictions, we first tried to investigate if the employed methods from that study, namely an ANNs approach based on the non-linear autoregressive (NAR) and the non-linear autoregressive with exogenous inputs (NARX) models could provide satisfactory results for our current situation.

The method based on the NAR model yielded results with insufficient accuracy, especially considering that in our current situation the contractor needed not only an hourly forecast for the overall activity, but an accurate forecast for the refrigerator storage room as well. When trying to apply the NARX approach from the previous work, we were hindered by the unavailable outdoor temperature forecasts that the contractor was not inclined to purchase in the long run, making this choice for a temperature exogenous
dataset unfeasible. However, building a time stamp dataset and using it as an exogenous variable for the NARX approach improved the accuracy in contrast with the NAR model when it came to hourly forecasts for the week-ahead, but was unsatisfactory for the hourly month-ahead. At this point, we decided to explore other approaches that yielded improved results while preserving the findings regarding the potential benefits of an exogenous timestamp dataset.

Another previous work of one of the authors of the current paper targeted the forecasting of the consumed electricity for a whole month on an hourly resolution, in the case of an industrial consumer [31]. Even if the profiles of the two electricity consumers differ, we investigated the approach consisting of using the ANNs based on the NARX model for obtaining first a forecast of the total daily aggregated electricity consumption that is being processed further using LSTM ANNs for refining the forecasted aggregated daily electricity consumption for a whole month on an hourly resolution. The obtained results were superior when compared to [30], and if an hourly two-weeks ahead forecast had sufficed for the contractor, the approach would have been an acceptable solution. In addition, the refrigerator storage room had to undergo the same method applied separately all over again leading to exceedingly high training times for the ANNs. These findings, although not satisfactory for our purpose, led us to the decision of investigating further the potential benefits that LSTM ANNs could bring in devising our forecasting method, considering their ability to learn dependencies in the long-term.

As we wanted to attain a high degree of accuracy for all of the weeks comprising the hourly month-ahead forecasts, after having experimented with regular LSTM ANNs that provided very good results for the first two weeks, only to register a drop afterwards in the forecasting accuracy, especially for the last week of the month-ahead forecast horizon, we decided to traverse the input data in both directions. Therefore, we developed and implemented within the second stage of the forecasting method a BiLSTM ANN enhanced with multiple simultaneously decreasing delays, as depicted below when presenting in detail the stages and steps of the proposed forecasting method. As we wanted to save considerable processing time without affecting the forecasting accuracy, we tried to avoid re-applying the whole process once again for the refrigerator storage room and tried to disaggregate its hourly month-ahead electricity consumption from the total hourly month-ahead electricity consumption forecasted by the BiLSTM enhanced with multiple simultaneously decreasing delays.

We considered the fact that in a previous work, along with a part of our research group, we managed to achieve an electricity disaggregation method based on FITNETs, a method that has provided a high degree of accuracy when it came to forecasting the disaggregated electricity loads [32]. This led us to the decision of coupling the BiLSTM enhanced with multiple simultaneously decreasing delays approach with a FITNETs one used in the purpose of disaggregating the hourly month-ahead electricity consumption of the refrigerator storage room out of the total hourly month-ahead forecasted electricity consumption.

The devised forecasting method was developed using the following software configuration: the Windows 10 Education operating system, version 1803, OS build 17134.523 (produced by Microsoft Corporation from Redmond town, Washington, United States of America); the MATLAB development environment software version R2020(a) (produced by MathWorks, Inc. from Natick town, MA, USA) and an entry to mid-level hardware configuration from the computational point of view. The advantage of developing the forecasting method using an inexpensive hardware configuration consists of reduced costs for the current operator and other potential ones that will not have to buy expensive hardware equipment for the frequent retraining operations of the ANNs. The research developed and presented within this paper emerged as a consequence of the request of our contractor, a commercial center-type consumer from Romania. The electricity datasets registered and given by the contractor comprise the total hourly consumption at the commercial center-type level and the hourly consumption of the refrigerator storage room for the period January–December 2019.
Cold-Rite, one of the most renowned companies in the refrigeration industry from Sydney, Australia that provides “commercial and industrial refrigeration solutions, including walk in fridges, commercial water chillers, humidity-controlled rooms and more”, published a list of the most important symptoms signaling the fact that a cool room does not function properly and it needs maintenance, otherwise risking serious side effects such as a skyrocket in the electricity consumption and even a potential spoilage of the cooling system [33].

According to the specialists in refrigeration systems from this company, the most important five signs that a cool room needs maintenance are: condensation (consisting in “droplets of water on the inside of the cool room”), ice (representing “spot ice forming on the freezer door”), buzzing sounds (when refrigeration units are “excessively noisy”), spoiling food (in the case when perishable items in the fridge spoil before their expiration date), and most of all, higher energy bills. This final sign usually reflects the fact that some of the used appliances are inefficient and might need servicing. Therefore, at the request of the contractor, in order to be able to warn him/her regarding an abnormal growth trend of the consumed electricity of the refrigerator storage room, we targeted in our study obtaining an as accurate as possible forecasting approach for the month-ahead hourly consumption of the refrigerator storage room along with the main purpose of the study, the month-ahead hourly total electricity consumption at the level of the commercial center-type consumer, based on which the contractor can choose his/her most suitable billing plan.

In the following, there are presented details concerning the stages and steps that represent the constructing elements of our developed forecasting method, along with a series of details regarding the main concepts used in developing our forecasting approach, such as BiLSTM neural networks and FITNETs.

2.1. Stage I: Acquiring and Preprocessing the Datasets

The first stage of the forecasting method has five steps over the course of which the datasets were acquired and preprocessed. Throughout the first step of this stage, two datasets for the whole year were acquired from the commercial center-type consumer: a first dataset comprising the previously recorded total consumed electricity per hour (THED) and a second dataset representing the previously recorded electricity consumption per hour of the refrigerator storage room (RHED), each of the datasets consisting in 8760 records representing the hourly consumption for the year 2019, measured in MWh. The datasets were provided by the contractor who retrieved them from his own metering system.

We designed the first stage’s second step to search for abnormal or missing values that might occur within the two above-mentioned datasets (due to the malfunctioning of the metering device or due to recording errors) for obtaining the final preprocessed previously recorded consumed electricity per hour datasets (entitled PTHED for the historical total hourly consumed electricity and PRHED for the refrigerator storage room). However, even if when we devised our forecasting approach, we remarked that in this particular case, the two datasets given by the contractor did not contain such type of values, we designed this step to make sure that the devised forecasting method can be applied for a wider range of commercial center-type consumer cases, similar to the one for which we conducted the current study and therefore to be able to generalize the devised forecasting approach. In order to solve the issues regarding the potential missing or abnormal values, we designed this step to apply a gap-filling approach that has proven its effectiveness when applied in previous studies, some of them being conducted within our research team [30–32,34], the method being based on a linear interpolation approach.

Over the course of the first stage’s third step of the method, each of the two preprocessed datasets obtained during the previous step was divided into two hourly data subsets. Therefore, the PTHED dataset was divided into a training subset (PTHEDT) and a validation subset (PTHEDV), while the PRHED dataset was divided into a training subset (PRHEDT) and a validation subset (PRHEDV). Both of the PTHEDT and PRHEDT subsets
consisted of 8016 samples corresponding to the first 11 months of the year 2019, while each of the PTHEDV and PRHEDV subsets consisted of 744 samples corresponding to the last month of the year 2019. The PTHEDT and PRHEDT subsets were used to develop, train, and validate the prediction method, while the PTHEDV and PRHEDV subsets were used in the purpose of validating the forecasting approach, through a comparison of the real values for the last month of the year, stored in the validation subsets PTHEDV and PRHEDV, with the forecasted ones for the total hourly consumed electricity and for the refrigerator storage room.

The Supplementary Materials file contains detailed information about the datasets.

2.2. Stage II: Developing the BiLSTM ANN Total Hourly Electricity Consumption Forecasting Solution with Multiple Delays Support

One of the main research directions when dealing with the analysis of time series consists in forecasting the behavior of the respective series in the future, with various time horizons ranging from short- to medium- and long-term ones, depending on the specificity of the analyzed problem and on the implemented forecasting approach.

Over the course of time, in order to obtain an accurate forecasting in time series problems, researchers have developed various techniques, within which the DL and machine-based algorithms represent useful and efficient approaches. Comparing these methods to the ones based on regression techniques, one can remark their irrefutable advantages highlighted by the accuracy of the obtained results. For example, the LSTM models, particular cases of recurrent neural networks (RNNs), also entitled feedback-based models [35], provide improved prediction results compared to the ones obtained when using ARIMA models, due to the fact that LSTM models benefit from an enhanced memory, capable of learning a larger amount of input data due to specific gates that they incorporate for this purpose [31,34].

The LSTM ANNs have been implemented by many scientists within their studies [36–41], due to their main advantage that they offer when compared to the feed forward ANNs, consisting in the LSTM’s architecture that comprises a sequence input layer allowing the input of the time series or sequences in the ANN, along with a certain LSTM layer, conceived in order to store and be able to retrieve patterns over long time periods. The LSTMs incorporate certain loops enabling the prolonged existence of information that is passed inside the LSTM ANN from one step to another. In this type of ANN, the input data are crossed only once (from the input to the output) and the network computes the forecasting based on the past and current data.

With the purpose of obtaining an additional improvement of the forecasting accuracy, an enhanced architecture of LSTMs was proposed, namely the BiLSTM one, which enables additional training steps through which the input data are used for training in both directions, by passing through data twice: backward and forward [35,42,43]. In fact, the BiLSTMs models are based on the aggregation of input information coming from the previous and upcoming steps related to a certain time step in LSTM models and therefore, in this kind of model, in any moment of time the information from past and future is preserved [44].

One particular detail that differs when comparing the LSTM and BiLSTM approaches is the fact that the equilibrium is reached at a slower pace in the case of BiLSTM models when compared to the LSTMs. In fact, when referring to the BiLSTMs, two LSTM networks are applied to input data, as follows: firstly, the input sequence is processed through the first LSTM network, in a so-called “forward layer”, while secondly, the reversed input sequence is input into another LSTM model, therefore obtaining the “backward layer” [35]. A comparison between the LSTM’s [36] and BiLSTM’s [45] architectures is presented in Figure 1 below, adapted from [46].
The BiLSTM ties the two hidden layers having reverse directions in order to compute the same output, which takes into account both the past and future states at the same time. In many situations, the simultaneous access to these two states helps improve the model’s performance. The main principle on which the architecture of a BiLSTM is based, consists in splitting each neuron into two directions: one corresponds to the future and reflects the forward states, while the second corresponds to the past, highlighting the backward states [42].

In what concerns the training process of the BiLSTMs, this is achieved by using similar algorithms to the ones used in the case of regular LSTMs, for example the ADAM, SGDM, and RMSProp training algorithms [32]. Nevertheless, when the training process is based on the back-propagation technique, a series of supplementary processes are required due to the fact that the input and output layers updates cannot take place simultaneously. In the BiLSTM’s case, the training steps consist in [42]:

- Within the forward pass, both the forward and backward states are passed first and afterwards, the output neurons are passed.
- Within the backward pass, firstly the output neurons are passed, while the forward and backward states are passed subsequently.
- Afterwards, when the two passes into the two opposite directions have been performed, the network updates the weights.

Over the course of time, researchers have developed various applications of the BiLSTM models, such as: dependency parsing [47], protein structure prediction [43,48], translations [49], entity extraction [50], handwritten recognition [51], and speech recognition [52,53].

In order to take advantage of the numerous benefits of the BiLSTM ANNs, we decided to make them an integral part of developing the second stage of our forecasting approach, consisting in six steps. Over the course of the second stage’s first step, the training subset PTHEDT was retrieved and normalized by processing it as to have a mean of 0 and a variance of 1, corresponding to a standard normal distribution [54].

Subsequently, we constructed, in the second stage’s second step, an array of delays $D = \{x_1, \ldots, x_d\}$ and the BiLSTM inputs based on the number and values of the delays, therefore obtaining a set of d training inputs (BiLSTM-I) for the BiLSTM ANN total hourly electricity consumption month-ahead forecasting solution enhanced with a multiple simultaneously decreasing delays approach. According to our devised approach, when considering a certain $x_k \in D, k \in \{1, \ldots, d\}$ as a delay, the developed method takes into
consideration as delays all the values within the set \( \{x_k, x_{k-1}, \ldots, 2, 1\} \) and uses them as simultaneous input sequences.

Throughout the course of the steps 3, 4, and 5, using the set of delays \( D \) and the BiLSTM inputs, a set of BiLSTM ANNs was developed, trained using the ADAM, SGDM, and RMSProp algorithms. In the case of each training algorithm, we developed 50 BiLSTM ANNs with the purpose of obtaining the most accurate forecast of the electricity consumption per hour, covering a timeframe of an entire month, by developing networks with various architectures comprising a number of hidden units \( n \in \{100, 200, \ldots, 1000\} \), the delays varying within the set \( D = \{1, 6, 12, 18, 24\} \). In addition to the values contained in the set \( D \), we also tested supplementary intermediate values of the delays, but we saved only the most eloquent ones in order to highlight the variation and especially the impact of the values of the delays. Following the experimental tests, a noticeable improvement in performance accuracy was found when increasing the delay value with a minimum of 6 units. However, in order to depict the impact that the delay has on the forecasting accuracy, we also studied, for benchmarking purposes, the case where the delay has the minimum value, namely when \( d = 1 \).

When developing an ANN-based forecasting solution, and particularly in the cases of our developed BiLSTM and FITNET neural networks, one has to take into consideration the fact that when the devised solution is implemented in a real production environment, the ANNs need follow-up steps for retraining, because as time passes, the volume of data increases and has to be submitted to the developed neural networks as inputs. In this context, the training time is of great importance in assessing the performance of a certain forecasting solution based on ANNs and therefore, for both types of developed ANNs (BiLSTM and FITNET), we registered and compared the training times.

2.3. Stage III. Elaborating the FITNET ANN Prediction Approach for the Electricity Consumption of the Refrigerator Storage Room

The ANNs were developed starting from the natural, biological networks existing in the brains of human beings or animals, with the purpose of learning how to perform certain activities starting from pre-existing examples, such as pattern classification, social network, image recognition, speech recognition, and medical diagnosis [55,56].

Being modeled as to resemble the biological neurons of a natural brain, an ANN comprises a group of interconnected nodes (the neurons) that are working in parallel and are connected through weights, having the purpose of transmitting the signal from one neuron to another. After receiving a signal, each artificial neuron processes and passes it to an interconnected node. In most of the ANN models, the neurons compute the outputs through a function that depends on the input data and is nonlinear. These neurons are characterized by their weights that get adjusted over the course of a learning process. The artificial neurons are positioned on different layers defined as to correspond to specific tasks that the neurons have to execute on the received inputs. All the signals pass through the ANN’s layers, starting with the input layer and going further to the output one.

The first developed ANN model is represented by the feed-forward neural networks, case in which the information flows forward, in a unidirectional way. In the scientific literature, one of the most frequently used feed-forward neural networks model consists in function fitting ANNs (FITNET ANNs), useful in the cases when one needs to fit a relationship that connects the inputs and the outputs [57].

The learning process of the ANNs is based on different training algorithms selected by taking into consideration the specific tasks that should be performed within each problem. When dealing with the training algorithms, one of the most frequently approached problems within the scientific community concerns the determination of the best training algorithm for a certain ANN. The accuracy and the training time represent the most frequently used criteria based on which the classification of the algorithms is done, while the best training algorithm is different from case to case, depending on the specificity of the problem and of the training set [34]. In the forecasting method’s third stage, we elaborated
a series of FITNETs based on the most frequently used training techniques: the LM \cite{58,59}, the BR \cite{60,61}, and the SCG \cite{62,63} algorithms.

After we developed and analyzed numerous possible options to devise an accurate forecasting solution for the month-ahead hourly refrigerator storage room consumption, by testing different types of ANNs architectures, we noticed that this type of networks can accurately predict, in the context of a fast training process, the month-ahead hourly refrigerator storage room consumption when they are provided as inputs the training subset PTHEDT with their associated timestamps and as an output the PRHEDT dataset.

Consequently, the third stage of the devised approach targets within its seven steps the month-ahead forecasting of the refrigerator storage room consumption. In the third stage’s first step, we associated a timestamp dataset (TSJN) to the PTHEDT dataset. Each of the 8016 entries of the timestamp dataset contains a number ranging from 1 to 24 representing the current hour; a number ranging from 1 to 7 representing the current day of the week (the first day of the week is considered to be Monday); a number between 1 and the maximum number of days of the current month (which can be 28, 30, or 31); and a number ranging from 1 to 11 representing the current month, namely one of the first eleven ones of the year 2019. Over the course of steps 2, 3, and 4, we elaborated a sequence of FITNET forecasting ANNs for the hourly consumption of the refrigerator storage room trained using the LM, BR, and SCG algorithms. In elaborating the ANNs, we used as inputs the TSJN and PTHEDT datasets and as output the PRHEDT dataset. In view of identifying the FITNET ANN that offers the highest performance level, we benchmarked different values for the hidden layer’s size \(N\), by choosing \(N \in \{10, 20, \ldots, 100, 200, \ldots, 1000\}\).

Consequently, we obtained 57 ANNs, consisting in 19 networks per training algorithm. The dimension of the involved datasets, PTHEDT, TSJN and PRHEDT, was of 744 samples, corresponding to the hourly measurements during the 31 days of the month of December.

When developing the FITNETs trained using the LM or SCG algorithms, the PTHEDT dataset was divided into three subsets by using the 70%-15%-15% approach for the training, validation, and testing processes, composing each of these percentages by randomly chosen samples.

In the case of the BR algorithm, the input dataset was also divided into three subsets, but their allocation was slightly different from the cases of the other two training algorithms, because the BR training algorithm does not imply a validation step. Therefore, even if for this training algorithm the PTHEDT dataset was divided according to the same 70%-15%-15% approach, the subset that would have been allocated to the validation step remained unused but kept separately. This allocation strategy has the main advantage that it provides the same data volumes in the testing and training processes for the LM, BR, and SCG algorithms, and therefore creates the premise of attaining an appropriate comparison between the results obtained using the three different training algorithms. We ran a number of 10 training iterations for each of the analyzed cases, and selected the best network based on execution time and MSE performance metric for identifying and storing the network that provided the most accurate forecasting results. Finally, after the steps 2, 3, and 4, we obtained a total number of 57 FITNET ANNs for the hourly consumption of the refrigerator storage room, whose forecasting accuracy was compared afterwards, in the third stage’s fifth step. By analyzing the MSE, the values of \(R\), and the training times, we saved the FITNET ANN that offered the most accurate prediction for the hourly consumption of the refrigerator storage room (BestFITNET), while discarding the remaining networks.

2.4. Stage IV: Forecasting with an Hourly Resolution the Total Electricity Consumption and the Consumption of the Refrigerator Storage Room, and Validating the Results

Stage four of the devised prediction method had four steps. During the second stage’s last step, we obtained a BiLSTM ANNs total hourly electricity consumption forecasting solution with multiple delays support.

Over the course of the fourth stage’s first step, for each of the analyzed cases, with the purpose of identifying and saving the network that provided the most accurate forecasting, we ran a number of 10 training iterations and selected the best network based on execution
time and RMSE performance metric. In this way, for each of the three training algorithms, we saved 15 BiLSTM ANNs with a multiple simultaneously decreasing delays approach, therefore obtaining 45 neural networks. Subsequently, we forecasted the total hourly electricity consumption for the month of December (obtaining the FHD dataset) for each of the 45 ANNs and denormalized these datasets (obtaining a new dataset, FHDDN) in order to compare them with the PTHEDV validation subset by computing the RMSE performance metric. Therefore, we saved the BiLSTM ANN that offered the most accurate prediction (BestBiLSTM) based on the RMSE performance metric and the registered training times, while discarding the remaining networks.

Afterwards, over the course of the fourth stage’s second step of our devised approach, we associated a timestamp dataset (TSD) to the FHDDN dataset. Each entry of the timestamp dataset contains: a number ranging from 1 to 24 representing the current hour; a number ranging from 1 to 7 representing the current day of the week (the first day of the week is considered to be Monday); a number between 1 and 31 representing the day of the month; and the number 12 representing the current month, namely the month of December of the year 2019. Subsequently, during the fourth stage’s third step, by using the BestFITNET ANN, the denormalized FHDDN dataset, and the TSD timestamp dataset, we forecasted the hourly electricity consumption of the refrigerator storage room for the month of December (FHDR).

Over the course of the fourth stage’s last step, in order to obtain a validation of the forecasting results of the refrigerator storage room’s electricity consumption, we compared the FHDR and PRHEDV datasets and computed the RMSE performance metric.

The forecasting method designed, developed, validated, and presented in Section 2.1, Section 2.2, Section 2.3, Section 2.4 is synthesized in the following flowchart (Figure 2).

In the following, we provide the results recorded during the development of the experimental tests and their implications.
In the following, we provide the results recorded during the development of the experimental tests and their implications.
3. Results

In the following, we present the main experimental results registered during the stages and steps of the prediction method.

3.1. Results Regarding the Developed BiLSTM ANN Total Hourly Electricity Consumption Forecasting Solution with Multiple Delays Support

As we have mentioned before, within steps 3, 4, and 5 of the second stage of our forecasting approach, we developed the BiLSTM ANN forecasting solution with multiple delays support, a solution that was used afterwards, during the fourth stage’s first step, in order to forecast the total consumed electricity per hour for the last month of the year and to identify the best BiLSTM ANN total hourly electricity consumption forecasting solution with multiple decreasing delays support (BestBiLSTM) based on the RMSE performance metric and on the registered training times. Analyzing the developed BiLSTM ANNs, we remarked that regarding the number of hidden units, starting with $n = 300$, and considering the dimension of the minibatch size as 128, the performance reached a plateau so that any further increase in the number of neurons no longer brought a noticeable increase in the performance accuracy, but only a significant increase in the execution time (e.g., $2713.504 \text{ s} = 45.23 \text{ min}$ when considering the ADAM training algorithm with $n = 1000$ and $d = 1$, compared to $1259.877 \text{ s} = 20.99 \text{ min}$ registered in the case when considering the ADAM training algorithm with $n = 300$ and $d = 1$). Consequently, the developed method tests various values for the parameters of the BiLSTM ANNs and saves the network that provides the best results.

Concerning the number of hidden layers, we considered firstly a single layer and afterwards two layers, concluding that an increase in the number of layers did not bring further improvements to the forecasting accuracy, but a considerable penalty in the execution time (e.g., $12,249.532 \text{ s} = 204.158 \text{ min} = 3.402 \text{ h}$ when considering the ADAM training algorithm, with 300 hidden units in each of the two layers and a delay $d = 24$, compared to $1259.877 \text{ s} = 20.99 \text{ min}$ registered in the case when considering the ADAM training algorithm with $n = 300$, $d = 1$ and a single hidden layer). We synthetized the obtained results in Table 2, which contains the RMSE performance metric along with the times $t$ (measured in seconds) representing the duration of the training process, registered during the various experimental tests, making use of each training algorithms, considering the number of hidden units $n = 300$ and the delay parameter varying within the set $D = \{1, 6, 12, 18, 24\}$. Regarding the maximum value of the delay parameter, when we tried to increase it further, we noticed that the forecasting accuracy did not register significant improvements, while the training times increased to a great extent (for example, for all the training algorithms, when considering the maximum value of the delay parameter equal to 30, the training time doubled its value, while the value of the RMSE performance metric barely improved). We recall the fact that, according to the second step of the second stage of our devised methodology, when choosing a certain value within the set $D$, all the integers obtained by decreasing the initial value up to 1 are taken into consideration by the devised approach as delays and used as simultaneous input sequences, therefore achieving a multiple simultaneously decreasing delays approach.

Analyzing Table 2, we observed that the vast majority of the BiLSTM networks from this table offer an accurate forecast as depicted by the RMSE values and consequently are suitable to be used by the contractor in actual daily activities. Nevertheless, one also has to take into consideration the running times that significantly increase the higher the size of the hidden layer is, whereas the forecasting accuracy remains almost the same. Regarding the variation of the delay parameters, according to the results obtained during the experimental tests, the best results were the ones obtained when considering the maximum value of the delay equal to 24, meaning that, according to our developed approach, in the context of the decreasing delays support, the network took into consideration as delays all the values within the set \{24, 23, \ldots, 2, 1\} and used them as inputs.
Table 2. An overview of the results obtained when developing the BiLSTM artificial neural network (ANN) total hourly electricity consumption forecasting solution with multiple simultaneously delays support (t is measured in seconds).

|                      | ADAM Training Algorithm |                      |                      |                      |                      |                      |                      |
|----------------------|--------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                      | n/d                      | 1                    | 6...1                | 12...1               | 18...1               | 24...1               |                      |
|                      | RMSE                     | 0.0865               | 0.0595               | 0.0579               | 0.0583               | 0.0495               |                      |
|                      | t                        | 1259.877             | 1419.438             | 1445.786             | 1454.187             | 1462.415             |                      |
|                      | RMSE                     | 0.0789               | 0.0692               | 0.0558               | 0.0597               | 0.0504               |                      |
|                      | t                        | 1291.090             | 1291.941             | 1292.111             | 1337.825             | 1358.041             |                      |
|                      | SGDM Training Algorithm  |                      |                      |                      |                      |                      |                      |
|                      | n/d                      | 1                    | 6...1                | 12...1               | 18...1               | 24...1               |                      |
|                      | RMSE                     | 0.0787               | 0.0686               | 0.0545               | 0.0609               | 0.0526               |                      |
|                      | t                        | 965.800              | 966.605              | 992.054              | 1015.9377            | 1294.2159            |                      |

Analyzing the obtained results synthetized above in Table 2, registered by the 15 BiLSTM networks trained with the ADAM, SGDM, and RMSProp algorithms, it can be observed that the ANN trained based on the ADAM training algorithm provides the most accurate forecast as depicted by the RMSE value (0.0495) and the computational time of $1462.415 \text{s} = 24.373 \text{ min}$. According to the devised approach, we entitled this total hourly electricity consumption forecasting solution with multiple delays support, as BestBiLSTM, a network whose architecture is depicted in Figure 3.

![Figure 3. The architecture of the BestBiLSTM network.](image)

One can find the obtained BestBiLSTM network within the Supplementary Materials.

3.2. The Obtained Results in the Case of the FITNET ANNs Provided Forecasts for the Refrigerator Storage Room

Over the course of the third stage’s second, third, and fourth steps, we registered for each of the developed function fitting ANNs, the running time along with MSE and R computed for the training dataset. As we mentioned before, we used a hidden layer’s
size comprising \( N \) neurons, where \( N \in \{10, 20, \ldots, 100, 200, \ldots, 1000\} \). The results are summarized in Table 3.

**Table 3.** An overview of the results obtained when developing the function fitting neural network (FITNET) ANNs hourly electricity consumption forecasting solution for the refrigerator storage room (t is measured in seconds).

| N   | The Training Algorithm | The LM Training Algorithm | The BR Training Algorithm | The SCG Training Algorithm |
|-----|------------------------|----------------------------|----------------------------|----------------------------|
| 10  | MSE 0.00027402          | 0.00028113                 | 0.00028285                 | 0.00028285                 |
|     | R 0.99076               | 0.99169                    | 0.99143                    | 0.99143                    |
|     | t 1.9959                | 1.0384                     | 0.9029                     | 0.9029                     |
| 20  | MSE 0.00027298          | 0.00028325                 | 0.00028755                 | 0.00028755                 |
|     | R 0.99101               | 0.99205                    | 0.99043                    | 0.99043                    |
|     | t 0.64699               | 3.9916                     | 0.77644                    | 0.77644                    |
| 30  | MSE 0.00029745          | 0.00027939                 | 0.00029644                 | 0.00029644                 |
|     | R 0.99139               | 0.991118                   | 0.99073                    | 0.99073                    |
|     | t 1.2039                | 6.4184                     | 1.6761                     | 1.6761                     |
| 40  | MSE 0.00029059          | 0.00027873                 | 0.00041093                 | 0.00041093                 |
|     | R 0.99158               | 0.99119                    | 0.98805                    | 0.98805                    |
|     | t 0.84041               | 16.3761                    | 1.1379                     | 1.1379                     |
| 50  | MSE 0.0003061           | 0.00027842                 | 0.00052929                 | 0.00052929                 |
|     | R 0.99153               | 0.99127                    | 0.98904                    | 0.98904                    |
|     | t 0.89097               | 11.8194                    | 2.4604                     | 2.4604                     |
| 60  | MSE 0.0003061           | 0.000282                    | 0.00032761                 | 0.00032761                 |
|     | R 0.99092               | 0.99165                    | 0.98894                    | 0.98894                    |
|     | t 0.92051               | 42.9114                    | 3.1738                     | 3.1738                     |
| 70  | MSE 0.00029641          | 0.00027996                 | 0.00032691                 | 0.00032691                 |
|     | R 0.99074               | 0.99178                    | 0.99044                    | 0.99044                    |
|     | t 0.96072               | 26.3337                    | 3.0564                     | 3.0564                     |
| 80  | MSE 0.00028144          | 0.00027975                 | 0.00041828                 | 0.00041828                 |
|     | R 0.99157               | 0.99153                    | 0.98795                    | 0.98795                    |
|     | t 1.5911                | 42.3873                    | 2.1301                     | 2.1301                     |
| 90  | MSE 0.0002975           | 0.00027954                 | 0.00039089                 | 0.00039089                 |
|     | R 0.99073               | 0.99142                    | 0.98733                    | 0.98733                    |
|     | t 1.2714                | 85.5294                    | 2.4269                     | 2.4269                     |
| 100 | MSE 0.00030012          | 0.00028045                 | 0.00042968                 | 0.00042968                 |
|     | R 0.99129               | 0.99159                    | 0.98743                    | 0.98743                    |
|     | t 1.7473                | 111.3506                   | 2.7927                     | 2.7927                     |
| 200 | MSE 0.00028921          | 0.00027885                 | 0.00033792                 | 0.00033792                 |
|     | R 0.99095               | 0.99099                    | 0.98931                    | 0.98931                    |
|     | t 6.4525                | 662.0224                   | 8.1396                     | 8.1396                     |
| 300 | MSE 0.00030733          | 0.00027797                 | 0.00043193                 | 0.00043193                 |
|     | R 0.98974               | 0.99132                    | 0.98714                    | 0.98714                    |
|     | t 14.089                | 1488.41                    | 17.444                     | 17.444                     |
| 400 | MSE 0.00030762          | 0.00028059                 | 0.00052717                 | 0.00052717                 |
|     | R 0.99                    | 0.99131                    | 0.9846                     | 0.9846                     |
|     | t 24.6715               | 2693.6973                  | 25.6914                    | 25.6914                    |
| 500 | MSE 0.00033178          | 0.00028349                 | 0.00033315                 | 0.00033315                 |
|     | R 0.98871               | 0.99168                    | 0.98939                    | 0.98939                    |
|     | t 37.5447               | 4291.0526                  | 59.1329                    | 59.1329                    |
| 600 | MSE 0.00043466          | 0.00027972                 | 0.00035855                 | 0.00035855                 |
|     | R 0.98576               | 0.99176                    | 0.98962                    | 0.98962                    |
|     | t 65.1606               | 4481.8194                  | 85.3138                    | 85.3138                    |
Table 3. Cont.

| N  | The Training Algorithm | The LM Training Algorithm | The BR Training Algorithm | The SCG Training Algorithm |
|----|------------------------|----------------------------|----------------------------|----------------------------|
| 700| MSE 0.00043285         | 0.00028053                 | 0.00058308                 |                             |
|    | R 0.98621              | 0.99126                    | 0.98229                    |                             |
|    | t 93.8144              | 6084.2019                  | 68.2735                    |                             |
| 800| MSE 0.00042612         | 0.00027583                 | 0.00053732                 |                             |
|    | R 0.98746              | 0.99076                    | 0.98389                    |                             |
|    | t 121.1446             | 8205.7725                  | 106.8163                   |                             |
| 900| MSE 0.00045703         | 0.00027928                 | 0.0005471                  |                             |
|    | R 0.98545              | 0.99136                    | 0.98346                    |                             |
|    | t 156.5832             | 10453.9573                 | 132.4432                   |                             |
| 1000| MSE 0.00069004         | 0.00027433                 | 0.00073457                 |                             |
|    | R 0.98107              | 0.99139                    | 0.98049                    |                             |
|    | t 214.5622             | 13230.2844                 | 145.866                    |                             |

The registered results emphasize the fact that the developed FITNET ANNs provided a very good forecasting accuracy in all the cases, highlighted by the MSE values that were very small, ranging between 0.00027298 (registered for \( N = 20 \) and the LM training algorithm) and 0.00073457 (registered for \( N = 1000 \) and the SCG training algorithm); and the values of R that were around 1, ranging between 0.98049 (registered for \( N = 1000 \) and the SCG training algorithm) and 0.99205 (registered for \( N = 20 \) and the BR training algorithm). As regards the execution times, their values varied from 0.64699 s (registered for \( N = 20 \) and the LM training algorithm) and a maximum of 13,230.2844 s (registered for \( N = 1000 \) and the BR training algorithm).

The most accurate prediction results (as highlighted by the MSE and R performance metrics) were registered when the FITNET ANNs were trained using the LM and BR algorithms. However, even if when the FITNET ANNs were trained using the SCG training algorithm, the performance metrics were lower than in the other two cases, for this algorithm, the training process was the fastest. Consequently, if new datasets appear often and the FITNET ANN requires frequent retraining steps, the SCG represents a suitable solution due to its high computational speed and reduced memory requirements.

According to the devised methodology, in the third stage’s final step, by comparing the performance metrics of the 57 developed FITNETs for the hourly consumption of the refrigerator storage room synthetized in Table 3, we determined the network that offered the most accurate prediction hourly (BestFITNET), while discarding the remaining ones. The BestFITNET ANN was the one with a hidden layer’s size comprising \( N = 20 \) neurons, trained using the LM algorithm, case in which the values of the performance metrics were \( MSE = 0.00027298, \ R = 0.99101 \). One can find within the Supplementary Materials the BestFITNET network.

For the purpose of highlighting the accuracy provided by the BestFITNET ANN, we present and analyze in the following the performance plots of this network. Firstly, we computed and represented the three curves corresponding to the training, validation and testing processes (Figure 4a). Analyzing this plot, one can remark that the performance peak was attained at the 11th epoch, when the MSE had the value 0.00027298. Figure 4a highlights the fact that none of the three curves increased significantly before the others, they are almost overlapping. This remark confirms that we obtained an accurate prediction and that an overfitting process does not occur. Another important conclusion that can be drawn from the analysis of Figure 4a is the efficiency of the training process and the confirmation that the dataset was divided in an appropriate way. Analyzing the plots, one can remark that for any of the curves there is no increase after the convergence has taken place, which represents a strong argument in stating that the FITNET ANN approach represents a robust solution, providing an increased level of accuracy in the prediction.
of the refrigerator storage room’s electricity consumption for a whole month, with an hourly resolution.

(a) The most accurate validation result highlighted by the mean squared error (MSE).

(b) The histogram of error values.

Figure 4. Cont.
The error histogram related to the prediction of the refrigerator storage room’s electricity consumption for a whole month on an hourly resolution, through the means of the BestFITNET ANN underlines the very good prediction accuracy of this forecasting solution, as the recorded errors were generally within a narrow range, namely \((-0.0377, 0.03613)\), while the difference between the outputs of the network and its targets had the median value of \(-0.00078\) (Figure 4b).

Afterwards, by computing and plotting the BestFITNET ANN’s regressions between the outputs of the network and its targets, one can observe that all the values of R were higher than 0.99101, which is very close to 1. The above-mentioned value corresponds to the testing steps and, along with all the other values of R, highlights a high level of fitting registered by the BestFITNET ANN (Figure 4c).

In the following, we depict the results obtained during the validation process that has been developed over the course of the fourth stage of the devised approach, by forecasting the total electricity consumption and the consumption of the refrigerator storage room for a whole month, on an hourly resolution.

### 3.3. Results Registered within the Validation Process

Over the course of the fourth stage’s first step, for the purpose of validating the method, the forecasted values obtained using the BestBiLSTM ANN (the FHDDN dataset) were compared with the real ones (the PTHEDV dataset). For obtaining a proper evaluation of the forecasting accuracy, we evaluated the RMSE and we remarked that the registered values were very good, namely: 0.0307 during the training process, 0.0327 for the integral set, and 0.0495 during the testing process. Assessing and confirming the forecasting accuracy of the developed method represents a very important step in making the final decision regarding the suitability of this method to be used in actual daily activities.
The comparison described above was also made by representing on the same plot the total real electricity consumption per hour for a whole month (PTHEDV) and the corresponding forecasted dataset (FHDDN) (Figure 5a). The two curves are almost identical, consequently confirming the fact that the developed approach has provided an increased level of prediction’s accuracy. The forecasting accuracy has also been emphasized by computing and representing for each month ahead time moment, the difference \( \Delta \) between the corresponding values of the PTHEDV and FHDDN datasets (Figure 5b). By analyzing these two plots, one can observe that the differences were very small during the whole month ahead period and therefore one can confirm the high level of prediction’s accuracy during the entire targeted period.

![The real and forecasted total hourly month-ahead electricity consumption datasets](image1)

(a) The real and forecasted total hourly month-ahead electricity consumption datasets

![The differences \( \Delta \) between the real and forecasted total hourly month-ahead electricity consumption datasets](image2)

(b) The differences \( \Delta \) between the real and forecasted total hourly month-ahead electricity consumption datasets

**Figure 5.** The comparison between the real and forecasted total hourly month-ahead electricity consumption datasets.

Afterwards, over the course of the fourth stage’s last step, for the purpose of validating the method, the forecasted values of the electricity consumption for a whole month on an hourly resolution, obtained using the BestFITNET ANN (the FHDR dataset), were compared with the real ones (the PRHEDV dataset). In order to obtain a relevant evaluation...
of the prediction solution’s accuracy, we also computed the RMSE performance metric during the forecasting process and registered a very good value of 0.0284.

The above-mentioned comparison was realized by plotting the real electricity consumption dataset (PRHEDV) of the refrigerator storage room for a whole month, on an hourly resolution, and the corresponding forecasted (FHDR) dataset on the same chart (Figure 6a). By analyzing this plot, one can remark that the two curves are almost identical, having the same trend and passing through very close points, therefore confirming the excellent forecasting accuracy provided by the developed method. Another way of highlighting the forecasting accuracy through the means of a graphical representation consists in computing and representing for each month ahead time moment, the difference $\Delta'$ between the corresponding real and forecasted values of the electricity consumption of the refrigerator storage room (Figure 6b). Both of these plots emphasize that these differences are very low for the whole month-ahead period and therefore one can confirm the high level of prediction’s accuracy of the developed method for the entire targeted period.

![The real and forecasted hourly month-ahead consumption electricity datasets of the refrigerator storage room](image1)

(a) The real and forecasted hourly month-ahead consumption electricity datasets of the refrigerator storage room

![The differences $\Delta'$ between the real and forecasted hourly month-ahead consumption electricity datasets of the refrigerator storage room](image2)

(b) The differences $\Delta'$ between the real and forecasted hourly month-ahead consumption electricity datasets of the refrigerator storage room

**Figure 6.** The comparison between the real and forecasted hourly month-ahead consumption electricity datasets of the refrigerator storage room.
By analyzing the registered forecasting results, the computed performance metrics, the above-described performance plots, the comparisons between the forecasted electricity consumption datasets for a whole month on an hourly resolution, and the corresponding real datasets for both the total consumption and for the refrigerator storage room, one obtains a validation of the developed forecasting approach, consequently obtaining a useful and accurate forecasting tool for the electricity consumption per hour, covering a timeframe of an entire month at the level of the commercial center-type consumer and also for the refrigerator storage room. One can find within the Supplementary Materials file thorough details with regard to the prediction results, their validation, and the above-mentioned charts.

In the following, we present a discussion concerning the interpretation of the obtained forecasting results in contrast with other existing studies and approaches from the body of knowledge that target resembling issues.

4. Discussion

The evolution of modern-day computing hardware architectures has made it possible to significantly reduce the training times of artificial intelligence (AI) techniques down to a point where developing forecasting methods in a custom-tailored approach, by means of multiple AI algorithms tackling different parts of the same bigger challenge, has become an attainable approach. This makes it possible to develop not only more efficient forecasting methods in terms of performance accuracy contrasted with the necessary training times, but also makes it possible to traverse a long sequence of steps from within a process of testing numerous parameters for the different AI algorithms involved. In this way, one can be sure that after a few hours of computational work, one is able to obtain an optimum method having the capability of being frequently re-trained as new data is encountered and stored throughout the course of operation, therefore counter-acting the impact that data dynamics can pose to an ANN approach over a period of time.

Taking into account all of the above mentioned aspects, we devised and validated an efficient prediction method for a large commercial center-type electricity consumer, attaining an hourly month-ahead consumed electricity prediction without experiencing a significant drop in the forecasting accuracy that usually tends to occur after the first two weeks, therefore achieving a reliable method that satisfies the contractor’s needs, being able to help him/her to attain energy efficiency and achieve sustainable economic, business, and management operations.

In rapport with previous studies from the literature that target the very important aspect of electricity forecasting and aim to develop prediction methods with a high degree of accuracy, one can distinguish several aspects. When compared to the article of a part of our research group [30] that targets the same commercial center-type electricity consumer profile and the same forecast horizon of an hourly month-ahead, one can notice that the proposed methods differ significantly. The methods proposed in the referenced article consist in separate ANNs developed based on nonlinear autoregressive models, using for the NARX models a timestamp dataset and a meteorological one. The use of a meteorological dataset was not a viable option for us in the current situation as the contractor did not agree to have to invest money monthly to purchase long-term accurate forecasts from the Romanian National Meteorological Administration or other specialized institutes. By developing in the second stage of our proposed approach a BiLSTM ANN enhanced with multiple simultaneously decreasing delays, we were able to attenuate the drop in forecasting accuracy that registered after the first two weeks in the NARX case without the need to acquire meteorological exogenous variables. In addition, as we needed to address another prediction for the refrigerator storage room, we managed to avoid reapplying the BiLSTM approach all over again by developing and coupling in the third stage a FITNET ANN prediction approach for the refrigerator storage room, therefore achieving an accurate electricity disaggregation forecast.
In the paper [30], together with a member of our group of researchers, we managed to successfully develop and implement a hybrid LSTM-FITNET approach. The LSTM ANNs proved very useful in refining a weather forecast corresponding to a certain region at the level of individual wind turbines so as to obtain accurate input values for subsequent FITNET ANNs that provided an accurate forecast for the hourly day-ahead production and consumption of electricity in the case of a wind farm located on a hilly terrain. In the case of the large hypermarket chain consumer, when an hourly month-ahead time horizon was needed, a BiLSTM ANN enhanced with multiple simultaneously decreasing delays proved to be more useful than a LSTM approach for avoiding a steeper decline of the forecasting accuracy towards the end of the month. In contrast to the previous study, the FITNET ANNs proved useful for electricity disaggregation purposes, while being unsuitable for an initial forecasting of the total consumption per hour for a whole month.

Within the scientific article [35], Siami et al. performed a comparative analysis between a BiLSTM approach and a LSTM one, along with analyzing the implications that a larger training data volume related to financial stocks has in terms of forecasting accuracy when devising forecasting methods based on BiLSTMs models compared with LSTM ones. Instead of financial stock data, we used in our study data related to the consumption of electricity in the commerce sector. In addition, as we needed a higher degree of forecasting accuracy in order to cover the hourly month-ahead timeframe, we devised the BiLSTM ANN with multiple simultaneously decreasing delays support in contrast with the analyzed approach. The registered results of the authors are in accordance with our findings, namely that, when provided with sufficient training data, a BiLSTM ANN provides a more accurate forecast in contrast to unidirectional LSTM ANNs, benefiting from the fact that long term dependencies are learned and stored by passing through data twice, both backward and forward.

The goal of achieving an as accurate as possible electricity consumption or production forecasting is of extreme importance to all the involved parties, starting from electricity producers, dispatchers, policy regulators, up to the different types of electricity consumers. Therefore, in the scientific literature, one can find a great variety of papers that focus on achieving accurate electricity forecasts from the production and consumption perspectives, proposing different forecasting methods developed based on the electricity profiles of the involved beneficiaries and on the durations of the forecasting windows.

In order to attain these goals, scientists have used different methods and carried out numerous performance improvement techniques, such as: SVR [3,11,28], LSTM ANNs [3,4,10,11,26,28], RNN [5,28], BiLSTM ANN [5], Copula-DBN [7], DRNN-LSTM [8], CNN-RNN [9], the Prophet and Holt–Winters long-term forecasting models [13], CNNs [4,6,14,25], MLR [24,25], SVM [24], and ANFIS [24]. In contrast with these, within our article, we devised and developed a forecasting method for the hourly month-ahead electricity consumption based on a BiLSTM ANN enhanced with a multiple simultaneously decreasing delays approach coupled with FITNETs.

Analyzing the scientific literature, one can remark that when forecasting the electricity consumption, different forecasting time horizons are of interest, encompassing short- [3,14,19,22,24,26,28], medium- [19,21,26], and long-term timeframes [10,13,14,20,23–25,64], each of them bringing their own particular advantages in line with the actual requirements and business needs of the contractors. We targeted the hourly month-ahead electricity prediction, considering the numerous benefits that such a forecast brings to the large-electricity commercial center-type consumer, ranging from the negotiation and choosing of the most appropriate hourly billing tariffs and correct estimations for the month-ahead electricity consumption submitted to the dispatcher to proper decisional support in what concerns the return on investment in more energy efficient equipment and assessing expanding options.

In what concerns the registered results, it is clear that every method was customized in view of the particular situations it had to address and the different targeted purposes, therefore making use of distinctive datasets covering certain scenarios with their particular
characteristics. Consequently, one cannot attain an entirely eloquent comparison between the forecasting methods from the scientific literature. Nevertheless, it can be noticed that the forecasting methods yielded results with a high degree of accuracy which meets the contractors’ real needs, attaining the objectives of the respective studies.

Although the final form of our proposed forecasting method that we reached after having applied the described research methodology within which we tried to progressively improve the steps and methods with a view to reaching the final goal registered very good results in term of prediction accuracy, as the performance metrics, the analysis of the graphical performance plots, and the direct comparison between the real recorded and the forecasted datasets depict, our proposed method nonetheless has certain limits.

One of the most protruding limits comes from the need to have an initial sufficient amount of recorded data to facilitate the proper training of ANNs so as to obtain reliable, accurate prediction results. Consequently, as we are targeting a long-term forecast, the initial necessary historical amount of data needs to be of a minimum of 11 months, otherwise we have to accept an initial lower degree of forecasting accuracy that will increase as new data accumulate over time and the re-training process has access to them.

By analyzing the existing body of knowledge, one can observe that many studies have focused on developing a variety of forecasting methods targeting particular situations and case studies regarding the consumption of electricity in the case of large commercial center-type consumers. Achieving a comprehensive, all-encompassing comparison with a view to the particularities and the most important characteristics of the multitude of proposed methods is an extremely difficult endeavor, even an unreachable one, taking into account the vast number of records of studies and processes that have been carried out for different scenarios, targeting distinctive objectives, using different datasets.

Even so, when several case studies share similar features, a certain method can be adapted and applied from a specific case to another by means of a generalization procedure. In this regard, the large hypermarket chain for which we devised the approach constitutes a non-limitative case where our proposed method can be successfully applied. This statement is justified by the fact that the large hypermarket chain can be regarded as a representative example with regard to other similar electricity consumers, which are placed by the European Regulations within the same electricity consumption band, namely the category “Band-ID”, characterized by an annual electricity consumption which varies within the range [2000; 20,000] MWh [65].

Consequently, the proposed electricity forecasting method, consisting in a BiLSTM ANN enhanced with a multiple simultaneously decreasing delays approach coupled with FITNETs, a method designed, developed, and validated in accordance to the devised methodology detailed in the second section of the paper, can be put into practice efficiently, having a generalizing capability, being able to cover a wide range of scenarios that exhibit similar features to the studied one.

Moreover, we designed, developed, validated, and compiled the hourly month-ahead forecasting method using a state of the art programming platform, aimed precisely at scientists, in order to assure the possibility of putting it into operation in a large variety of situations that require efficient long term electricity forecasting tools, ensuring in view of the software quality perspective its functional suitability, reliability, usability, maintainability, and performance efficiency with respect to the processing time and necessary computational resources. All of these characteristics reinforce the potential of our proposed method of being widely applicable, also from the hardware and software point of views, as a result emphasizing once more that the proposed approach has the ability to meet both expressed and implicit requirements of diverse beneficiaries, therefore being able to deliver bona fide business value.

Currently, worldwide, in 2020, humanity has to deal with disruptive circumstances caused by the rapid spread of the SARS-CoV-2 pandemic and its terrifying side effects, a fact that highlights the intricate connection between people, companies, countries, and governments. This unprecedented situation in the contemporary history of humanity
has generated unexpected consequences at personal, local, and global levels, raising in front of the human beings questions about the vulnerability of individuals, companies, and local and global economies, in the context of job losses, lack of trust in authorities, and concerns regarding a general sense of uncertainty. Consequently, it is natural for any research tackling issues regarding future projections to discuss the impact of disruptive circumstances such as economic crises [66] or the one mentioned above.

In Romania, the virus was detected for the first time on 26 February 2020 [67] and in the time period that has elapsed since then, more and more effects have cascaded. A detailed analysis of the particular situation approached in our research showed that in the context of the pandemic, our contractor did not have to close or restructure his business (except for the implementation of social distancing measures and other measures to prevent or limit the effects of the pandemic). Even during the lockdown period, the large hypermarket chain for which we have designed, developed, and implemented the forecasting method has continued its daily activity according to its usual schedule.

The developed prediction method allows retraining with new datasets, related to the period after the outbreak of the coronavirus pandemic. Therefore, in order to be able to analyze the impact of this situation on the input datasets and on the forecasting results, we intend as future research to retrain the developed networks and revalidate the obtained results as soon as we have the new datasets corresponding to the disruptive circumstances. A first approach that we plan to experiment with in further research includes narrowing the retraining timeframe window, from once a month to twice a month, or, if deemed necessary, even once a week. In addition, we also intend to explore the usage of exogenous variables related to the disturbing pandemic event when developing the BiLSTM with multiple simultaneously decreasing delays support, if the retraining of the network will have proved to be insufficient, or even consider developing a separate AI stage that forecasts the necessary adjustments and apply them in subsequent processing stages.

Another future research direction consists in the assessment of the economic operators’ intentions towards using alternative solar energy resources using analytical tools [68], as government subsidies have become more prevalent during the pandemic in order to sustain economic operators in reducing their electricity expenditures, and the electricity produced by the commercial operators’ own renewable installations must be forecasted as well.

5. Conclusions

The aim of this paper, namely, to design and develop an electricity consumption prediction method for the consumed electricity per hour, for a whole month, in the case of an industrial consumer and for its refrigerator storage room, is a very important aspect for the non-household electricity consumers and for the system operators and at the same time represents a key-factor in what regards energy efficiency, achieving sustainable economic, business, and management operations.

In addition to this, an accurate forecast offers benefits in what regards the electricity consumption optimization, influencing the consumption strategy and facilitating the efforts in view of attaining an appropriate natural resource and environmental management. In addition, an accurate prediction method for the consumed electricity per hour, for a whole month, at the level of the large hypermarket chain helps the contractor in choosing the most suitable billing plan, offering him/her also the opportunity to negotiate in advance with the dispatch operator the most appropriate billing tariff.

Using the designed, developed, and validated electricity consumption forecasting method based on a BiLSTM ANN enhanced with a multiple simultaneously decreasing delays approach coupled with FITNETs, we managed to predict the consumed electricity per hour for a whole month in the case of a large hypermarket chain consumer and for its refrigerator storage room. The analysis of the registered forecasting results and the performance plots, along with the comparison of the real values for the last month of the year, with the forecasted ones for both the total consumption and for the refrigerator
storage room, the analysis of the devised approach in contrast to other existing studies and approaches from the scientific literature body of knowledge targeting resembling issues, all emphasize the advantages of our proposed method and its capacity to narrow the gap identified after having analyzed the current state of knowledge.

Although the devised, developed, and validated forecasting solution for the hourly consumption focuses on a commercial center-type consumer, this solution also represents a useful, accurate tool for other non-household electricity consumers, due to its generalization capability.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/2071-1050/13/1/104/s1: the Sheet “12-month dataset” of the “Dataset.xlsx” Excel file that comprises detailed information about the datasets; the BestBiLSTM ANN (ADAM_300_BILSTM_1_STRAT_24_DELAYS.mat file); the BestFITNET ANN (FITNET_TRAINLM_20_NEURONI.mat file); the Sheet “BiLSTM” of the “Dataset.xlsx” Excel file that comprises the forecasted results for the last month of the year for the total electricity consumption and the validation of the forecasting results; the Sheet “FITNET” of the “Dataset.xlsx” Excel file that comprises the forecasted results for the last month of the year for the refrigerator storage room and the validation of the forecasting results.

**Author Contributions:** Conceptualization, D.-M.P. and A.P.; methodology, D.-M.P. and A.P.; software, D.-M.P. and A.P.; validation, D.-M.P. and A.P.; formal analysis, D.-M.P. and A.P.; investigation, D.-M.P. and A.P.; resources, D.-M.P. and A.P.; data curation, D.-M.P. and A.P.; writing—original draft preparation, D.-M.P. and A.P.; writing—review and editing, D.-M.P. and A.P.; visualization, D.-M.P. and A.P.; supervision, D.-M.P.; project administration, D.-M.P. and A.P.; funding acquisition, D.-M.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** The article processing charge (APC) was discounted integrally by the Multidisciplinary Digital Publishing Institute (MDPI).

**Acknowledgments:** The authors would like to express their gratitude for the logistics support received from the Center of Research, Consultancy and Training in Economic Informatics and Information Technology RAU-INFORTIS of the Romanian-American University.

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

**References**

1. World Energy Balances Overview (2020 edition). Available online: https://webstore.iea.org/world-energy-balances-overview-2020-edition (accessed on 14 November 2020).

2. Romania Energy Market Report. Available online: https://www.enerdata.net/estore/country-profiles/romania.html (accessed on 14 November 2020).

3. Moradzadeh, A.; Zakert, S.; Shoaran, M.; Mohammadi-Ivatloo, B.; Mohammadi, F. Short-Term Load Forecasting of Microgrid via Hybrid Support Vector Regression and Long Short-Term Memory Algorithms. *Sustainability* 2020, 12, 7076. [CrossRef]

4. Kim, M.; Choi, W.; Jeon, Y.; Liu, L. A hybrid neural network model for power demand forecasting. *Energies* 2019, 12, 931. [CrossRef]

5. Yaprapkald, F.; Yilmaz, M.B.; Baysal, M.; Anvari-Moghaddam, A. A deep neural network-assisted approach to enhance short-term optimal operational scheduling of a microgrid. *Sustainability* 2020, 12, 1653. [CrossRef]

6. Khan, S.; Javid, N.; Chand, A.; Khan, A.B.M.; Rashid, F.; Afridi, I.U. Electricity Load Forecasting for Each Day of Week Using Deep CNN. In *Web, Artificial Intelligence and Network Applications*; Barolli, L., Takizawa, M., Xhafa, F., Enokido, T., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 1107–1119.

7. Ouyang, T.; He, Y.; Li, H.; Sun, Z.; Baek, S. Modeling and Forecasting Short-Term Power Load With Copula Model and Deep Belief Network. *IEEE Trans. Emerg. Top. Comput. Intell.* 2019, 3, 127–136. [CrossRef]

8. Wen, L.; Zhou, K.; Yang, S.; Lu, X. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. *Energy* 2019, 171, 1053–1065. [CrossRef]

9. He, W. Load Forecasting via Deep Neural Networks. In *Procedia Computer Science*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 308–314.

10. Santra, A.S.; Lin, J.L. Integrating long short-term memory and genetic algorithm for short-term load forecasting. *Energies* 2019, 12, 2040. [CrossRef]

11. Son, H.; Kim, C. A deep learning approach to forecasting monthly demand for residential-sector electricity. *Sustainability* 2020, 12, 3103. [CrossRef]

12. Alonso, A.M.; Nogales, F.J.; Ruiz, C. A single scalable LSTM model for short-term forecasting of massive electricity time series. *Energies* 2020, 13, 5328. [CrossRef]
13. Almazroue, A.I.; Almeshal, A.M.; Almutairi, A.S.; Alenezi, M.R.; Alhajeri, S.N. Long-term forecasting of electrical loads in Kuwait using prophet and holt-winters models. *Appl. Sci.* 2020, 10, 5627. [CrossRef]

14. Shao, X.; Kim, C.S.; Sontakke, P. Accurate deep model for electricity consumption forecasting using multi-channel and multi-scale feature fusion CNN-LSTM. *Energies* 2020, 13, 1881. [CrossRef]

15. Hu, Q.; Zhang, R.; Zhou, Y. Transfer learning for short-term wind speed prediction with deep neural networks. *Renew. Energy* 2016, 85, 83–95. [CrossRef]

16. Liu, S.; Ji, H.; Wang, M.C. Nonpooling convolutional neural network forecasting for seasonal time series with trends. *IEEE Trans. Neural Netw. Learn. Syst.* 2020, 31, 2879–2888. [CrossRef] [PubMed]

17. Kong, W.; Dong, Z.Y.; Jia, Y.; Hill, D.J.; Xu, Y.; Zhang, Y. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Trans. Smart Grid* 2019, 10, 841–851. [CrossRef]

18. Kim, T.Y.; Cho, S.B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* 2019, 182, 72–81. [CrossRef]

19. Mir, A.A.; Alghassab, M.; Ullah, K.; Khan, Z.A.; Lu, Y.; Imran, M. A review of electricity demand forecasting in low and middle income countries: The demand determinants and horizons. *Sustainability* 2020, 12, 5931. [CrossRef]

20. Perwez, U.; Sohail, A.; Hassan, S.F.; Zia, U. The long-term forecast of Pakistan’s electricity supply and demand: An application of long range energy alternatives planning. *Energy* 2015, 93, 2423–2435. [CrossRef]

21. Sigauke, C. Forecasting medium-term electricity demand in a South African electric power supply system. *J. Energy S. Afr.* 2017, 28, 54–67. [CrossRef]

22. Potapov, V.; Khamitov, R.; Makarov, V.; Gritsay, A.; Tyunkov, D.; Chervenchuk, I. Short-Term Forecast of Electricity Load for LLC Omsk Energy Retail Company Using Neural Network. In Proceedings of the 12th International Scientific and Technical Conference “Dynamics of Systems, Mechanisms and Machines”, Dynamics 2018, Omsk, Russia, 13–15 November 2018; IEEE Publisher: Piscataway, NJ, USA, 2019.

23. Marulanda, G.; Bello, A.; Cifuentes, J.; Reneses, J. Wind power long-term scenario generation considering spatial-temporal dependencies in coupled electricity markets. *Energies* 2020, 13, 3427. [CrossRef]

24. Solyali, D. A Comparative Analysis of Machine Learning Approaches for Short-/Long-Term Electricity Load Forecasting in Cyprus. *Sustainability* 2020, 12, 3612. [CrossRef]

25. Elkamel, M.; Schleider, L.; Pasiliao, E.L.; Diabat, A.; Zheng, Q.P. Long-term electricity demand prediction via socioeconomic factors-a machine learning approach with Florida as a case study. *Energies* 2020, 13, 3996. [CrossRef]

26. Son, N.; Yang, S.; Na, J. Deep Neural Network and Long Short-Term Memory for Electric Power Load Forecasting. *Appl. Sci.* 2020, 10, 6489. [CrossRef]

27. Hong, T.; Kim, J.; Koo, C. LCC and LCCO 2 analysis of green roofs in elementary schools with energy saving measures. *Energy Build.* 2012, 45, 229–239. [CrossRef]

28. Ribeiro, A.M.N.C.; do Carmo, P.R.X.; Rodrigues, I.R.; Sadok, D.; Lynn, T.; Endo, P.T. Short-Term Firm-Level Energy-Consumption Forecasting for Energy-Intensive Manufacturing: A Comparison of Machine Learning and Deep Learning Models. *Algorithms* 2020, 13, 274. [CrossRef]

29. Kim, Y.; Seo, K.; Harrington, R.J.; Lee, Y.; Kim, H.; Kim, S. High accuracy modeling for solar PV power generation using Noble BD-LSTM-based neural networks with EMA. *Appl. Sci.* 2020, 10, 7339. [CrossRef]

30. Pirjan, A.; Oprea, S.V.; Caruțăsu, G.; Petoșanu, D.M.; Bără, A.; Cociulescu, C. Devising hourly forecasting solutions regarding electricity consumption in the case of commercial center type consumers. *Energies* 2017, 10, 1727. [CrossRef]

31. Petoșanu, D.M. Designing, developing and validating a forecasting method for the month ahead hourly electricity consumption in the case of medium industrial consumers. *Processes* 2019, 7, 310. [CrossRef]

32. Oprea, S.-V.; Pirjan, A.; Caruțăsu, G.; Petoșanu, D.-M.; Bără, A.; Stânică, J.-L.; Cociulescu, C. Developing a Mixed Neural Network Approach to Forecast the Residential Electricity Consumption Based on Sensor Recorded Data. *Sensors* 2018, 18, 1443. [CrossRef]

33. Signs that Your Cool Room Needs Maintenance. Available online: https://www.cold-rite.com.au/post/signs-your-cool-room-needs-maintenance (accessed on 14 November 2020).

34. Pirjan, A.; Caruțăsu, G.; Petoșanu, D.-M. Designing, Developing, and Implementing a Forecasting Method for the Produced and Consumed Electricity in the Case of Small Wind Farms Situated on Quite Complex Hilly Terrain. *Energies* 2018, 11, 2623. [CrossRef]

35. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The Performance of LSTM and BiLSTM in Forecasting Time Series. In Proceedings of the —2019 IEEE International Conference on Big Data, Big Data 2019, Los Angeles, CA, USA, 9–12 December 2019; IEEE Publisher: Piscataway, NJ, USA, 2019; pp. 3285–3292.

36. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* 1997, 9, 1735–1780. [CrossRef]

37. Murad, A.; Pyun, J.-Y. Deep Recurrent Networks for Human Activity Recognition. *Sensors* 2017, 17, 2556. [CrossRef]

38. Zheng, H.; Yuan, J.; Chen, L. Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation. *Energies* 2017, 10, 1168. [CrossRef]

39. Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M. 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]

40. Huang, C.-J.; Kuo, P.-H. A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities. *Sensors* 2018, 18, 2220. [CrossRef] [PubMed]
41. Nguyen, M.-T.; Nguyen, V.-H.; Yun, S.-J.; Kim, Y.-H. Recurrent Neural Network for Partial Discharge Diagnosis in Gas-Insulated Switchgear. Energies 2018, 11, 1202. [CrossRef]
42. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. IEEE Trans. Signal Process. 1997, 45, 2673–2681. [CrossRef]
43. Baldi, P.; Brunak, S.; Frasconi, P.; Soda, G.; Pollastri, G. Exploiting the past and the future in protein secondary structure prediction. In Bioinformatics; Oxford University Press: Oxford, UK, 1999; pp. 937–946.
44. Rasifaghii, N. Predictive Analytics: Regression Analysis with LSTM, GRU and BiLSTM in TensorFlow. Available online: https://towardsdatascience.com/predictive-analysis-rnn-lstm-and-gru-to-predict-water-consumption-e6bb3c2b4b02 (accessed on 14 November 2020).
45. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. In Neural Networks; Elsevier: Amsterdam, The Netherlands, 2005; Volume 18, pp. 602–610.
46. Mohan, A.T.; Gaitonde, D.V. A deep learning based approach to reduced order modeling for turbulent flow control using LSTM neural networks. arXiv 2018, arXiv:1804.09269.
47. Kiperwasser, E.; Goldberg, Y. Simple and Accurate Dependency Parsing Using Bidirectional LSTM Feature Representations. Trans. Assoc. Comput. Linguist. 2016, 4, 313–327. [CrossRef]
48. Pollastri, G.; McElysaght, A. Porter: A new, accurate server for protein secondary structure prediction. Bioinformatics 2005, 21, 1719–1720. [CrossRef]
49. Sundermeyer, M.; Alkhoul, T.; Wuebker, J.; Ney, H. Translation modeling with bidirectional recurrent neural networks. In Proceedings of the EMNLP 2014—2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 14–25.
50. Dernoncourt, F.; Lee, J.Y.; Szolovits, P. Neuroner: An easy-to-use program for named-entity recognition based on neural networks. In Proceedings of the EMNLP 2017—Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Copenhagen, Denmark, 9–11 September 2017; pp. 97–102.
51. Liwicki, M.; Graves, A.; Bunke, H.; Schmidhuber, J. A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks. In Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR 2007), Parana, Brazil, 23–26 September 2007; pp. 367–371.
52. Graves, A.; Jaitly, N.; Mohamed, A.R. Hybrid speech recognition with Deep Bidirectional LSTM. In Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2013, Piscataway, NJ, USA, 8–13 December 2013; pp. 273–278.
53. Graves, A.; Fernández, S.; Schmidhuber, J. Bidirectional LSTM networks for improved phoneme classification and recognition; In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer International Publishing: Cham, Switzerland, 2005; pp. 799–804.
54. MathWorks Documentation. Time Series Forecasting Using Deep Learning. Available online: https://www.mathworks.com/help/deeplearning/ (accessed on 14 November 2020).
55. Ibrahim, N.K.; Abdullah, R.S.A.R.; Saripan, M.I. Artificial Neural Network Approach in Radar Target Classification. J. Comput. Sci. 2009, 5, 23–32. [CrossRef]
56. Lippmann, R.P. An Introduction to Computing with Neural Nets. IEEE ASSP Mag. 1987, 4, 4–22. [CrossRef]
57. Schmidhuber, J. Deep Learning in neural networks: An overview. Neural Netw. 2015, 61, 85–117. [CrossRef] [PubMed]
58. Levenberg, K. A method for the solution of certain non-linear problems in least squares. Q. J. Appl. Math. 1944, 2, 164–168. [CrossRef]
59. Hagan, M.T.; Menhaj, M.B. Training Feedforward Networks with the Marquardt Algorithm. IEEE Trans. Neural Networks 1994, 5, 899–903. [CrossRef] [PubMed]
60. MacKay, D.J.C. Bayesian Interpolation. Neural Comput. 1992, 4, 415–447. [CrossRef]
61. Foresee, F.D.; Hagan, M.T. Gauss-Newton approximation to bayesian learning. In Proceedings of the International Conference on Neural Networks, Houston, TX, USA, 12 June 1997; pp. 1930–1935. [CrossRef]
62. Möller, M.F. A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 1993, 6, 525–533. [CrossRef]
63. Huang, Y. Advances in Artificial Neural Networks—Methodological Development and Application. Algorithms 2009, 2, 973–1007. [CrossRef]
64. Petrosanu, D.-M.; Căruntuș, G.; Căruntuș, N.L.; Pirjan, A. A Review of the Recent Developments in Integrating Machine Learning Models with Sensor Devices in the Smart Buildings Sector with a View to Attaining Enhanced Sensing, Energy Efficiency, and Optimal Building Management. Energies 2019, 12, 4745. [CrossRef]
65. Eurostat. Energy Statistics—Electricity Prices for Domestic and Industrial Consumers, Price Components. Available online: https://ec.europa.eu/eurostat/cache/metadata/en/nrg_pc_204_esms.htm (accessed on 14 November 2020).