

## Article

# Deep Learning-Based Approaches to Optimize the Electricity Contract Capacity Problem for Commercial Customers

Rafik Nafkha, Tomasz Ząbkowski and Krzysztof Gajowniczek \* 

Department of Artificial Intelligence, Institute of Information Technology, Warsaw University of Life Sciences-SGGW, 02-776 Warsaw, Poland; rafik\_nafkha@sggw.edu.pl (R.N.); tomasz\_zabkowski@sggw.edu.pl (T.Z.)

\* Correspondence: krzysztof\_gajowniczek@sggw.edu.pl

**Abstract:** The electricity tariffs available to customers in Poland depend on the connection voltage level and contracted capacity, which reflect the customer demand profile. Therefore, before connecting to the power grid, each consumer declares the demand for maximum power. This amount, referred to as the contracted capacity, is used by the electricity provider to assign the proper connection type to the power grid, including the size of the security breaker. Maximum power is also the basis for calculating fixed charges for electricity consumption, which is controlled and metered through peak meters. If the peak demand exceeds the contracted capacity, a penalty charge is applied to the exceeded amount, which is up to ten times the basic rate. In this article, we present several solutions for entrepreneurs based on the implementation of two-stage and deep learning approaches to predict maximal load values and the moments of exceeding the contracted capacity in the short term, i.e., up to one month ahead. The forecast is further used to optimize the capacity volume to be contracted in the following month to minimize network charge for exceeding the contracted level. As confirmed experimentally with two datasets, the application of a multiple output forecast artificial neural network model and a genetic algorithm (two-stage approach) for load optimization delivers significant benefits to customers. As an alternative, the same benefit is delivered with a deep learning architecture (hybrid approach) to predict the maximal capacity demands and, simultaneously, to determine the optimal capacity contract.

**Keywords:** contracted capacity; optimization; genetic algorithm; deep learning; electricity load time series forecasting



**Citation:** Nafkha, R.; Ząbkowski, T.; Gajowniczek, K. Deep Learning-Based Approaches to Optimize the Electricity Contract Capacity Problem for Commercial Customers. *Energies* **2021**, *14*, 2181. <https://doi.org/10.3390/en14082181>

Academic Editor: Annamária R. Várkonyi-Kóczy

Received: 28 February 2021

Accepted: 11 April 2021

Published: 14 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The electricity market is unique in terms of storage and supply conditions, which makes it very demanding and difficult in comparison to other production systems [1]. Therefore, forecasting the load demand is of great importance. To address these inconvenient conditions, energy producers propose different energy tariffs and contract options for their customers. Usually, voltage level and individual contracted capacity are the main factors used to assign proper tariffs for commercial customers. This strategy ensures that fluctuations in energy demand are controlled, which provides insight into the energy quantity required to be generated and allows it to be transmitted to customers.

One of the main variables considered in the tariff structures is the capacity component, so the users are charged for the availability to use the maximum power, in line with the connection agreement, which is the maximum value of the averaged consumed power within the period of 15 min in an hour span [2]. In practice, households and small businesses are not charged for exceeding contracted capacity. On the other hand, if the declared capacity quantity is exceeded by business or industrial consumers connected to a low-, medium- or high-voltage network, a penalty charge is levied. In line with the government's regulation with regard to the specific rules for the determination and calculation of tariffs and billing in the electricity industry [2], a fee is charged for exceeding the contractual

capacity defined in the contract. The fee constitutes the product of the rate of capacity charge and the sum of ten largest quantities of the surplus consumed capacity over the contractual capacity, indicated by the measuring and clearing device or ten times the maximum amount of surplus of consumed capacity over the contracted capacity recorded during the reference period.

Generally, to avoid an overcontracted capacity amount, customers declare a level of contracted capacity that is much higher than their needs, which makes customers not use their planned capacity and pay, in fact, for unused power. Therefore, the objective of this paper is to select a suitable contract capacity for commercial customers to achieve the minimal electricity bill during the analyzed period. By using the proposed algorithms, the contract capacity can be estimated with good accuracy and thus deliver a value added to the customers.

### 1.1. Literature Review

**Time horizon selection.** Numerous papers consider load forecasting, and most of the works are mainly related to long-term optimization of the electricity purchase and distribution process by suppliers and distributors. In general, load forecasting has been investigated by utilities and electricity suppliers, where long-term load forecasts (LTLFs) are used to predict the annual peak of the power system [3–5] to manage future investments in terms of modernization and launch new units to maintain the stability of nationwide electricity demand over time periods of up to 20 years [6]. Medium-term load forecasts (MTLFs) use hourly loads to predict the weekly peak load for both power and system operation planning [7]. Short-term load forecasts (STLFs) usually aim to predict the load up to one week ahead, while very short-term load forecasts (VSTLFs) are used for a time horizon of less than 24 h. Both STLF and VSTLF have engaged the attention of most researchers since they provide necessary information for day-to-day utilities' operations [8]. These forecasts also become useful when dealing with smart grids, microgrids, peak load anticipation, and intelligent buildings [9].

**STLF techniques.** The capacity consumed is usually measured directly from the measurement systems and is accounted for in accordance with the energy policy of a given country. For example, in Poland, the maximum load is measured in 15-min allowances [2], the same way it is counted in Taiwan [10]; in France, the interval span is 10 min for industrial customers [11].

In this article, we propose the short-term load forecasting (STLF) technique to forecast the maximum load taken from the grid. The STLF technique usually aims to predict the load up to one week ahead. There are a number of papers dedicated to short-term load forecasting of commercial customers for the purpose of contract capacity optimization. Most recent works include some classical techniques, such as autoregression (AR) models [12], linear regression models [13], dynamic linear and nonlinear models, general exponential smoothing models, spectral methods, and seasonal ARIMA models, which have been used to forecast load consumption [14–16]. Unfortunately, their capability to solve time series with complex seasonality and nonlinear series is limited, in favor of artificial neural network (ANN) techniques and expert systems [17–20]. Interestingly, the load forecasting field is one of the most successful applications of ANNs in power systems. Neural networks are able to deliver better performance when dealing with highly nonlinear series resulting from, e.g., the noninteger seasonality appearing as a result of averaging ordinary and leap years (365.25 days).

**Optimization.** There are various types of hardware and software solutions to support customers with optimization. For instance, load limiters are used to automatically disconnect the power supply circuit of the electrical installation in the event of exceeding the determined value of power and are often used to prevent overruns [21]. Another, more sophisticated option includes the electric power distribution and utility monitoring system, which support large and medium-sized sites in Japan in terms of energy efficient management, preventive maintenance and obvious capacity overruns [22].

In Poland, an application prototype was recently presented [23]. An optimization module has been introduced to optimize the level of contracted power using the Microsoft Time Series algorithm, which includes the ARTXP (Autoregressive Tree) algorithm for initial power value prediction and the ARIMA algorithm used to improve the prediction accuracy. The application analyzes historical data that may come from various sources, i.e., received invoices or direct measurements from monitoring devices. The described algorithm assumes monthly settlements with the electricity supplier and minimizes the total cost of power orders along with penalties for exceeding the contracted level on a monthly and annual basis. The disadvantage of this solution is the inability to capture the maximum values of the consumed power, where averaged consumption values are used to predict future values. In addition, when reading invoice data, reports for 15-min consumption values and daily consumption values are not available.

An easy solution is presented by Chan et al. [10] where to determine the electricity contract capacity level, they formulate the problem as a linear program to optimize the cost function, which only requires polynomial time. The application of the model to two real case studies demonstrated that industrial customers who are billed on the basis of Taipower tariffs can reduce their electricity bills with respect to contracted capacity.

Another hybrid solution presented by Hong et al. [24] is dedicated to power consumers with self-owned generation units. The study consists of introducing the improved Taguchi method, which includes the traditional Taguchi method [25] and particle swarm optimization method to search for the best combination of contracted capacities and the dispatched output of self-owned generating units. The comparison of the improved method with other traditional methods, including particle swarm optimization (PSO) [26], genetic algorithm (GA) [27] or linear programming (LP) optimization [10], demonstrated that, on the basis of real data of an optoelectronics factory in Taiwan with a self-owned generator, the total expenses for electricity are reduced.

In [28], the authors proposed a methodology for the reduction of electrical costs for large customers, such as industrial companies or hospitals, whenever the electric supply tariff contains a time-of-use factor and penalizations due to excess capacity demand (which is typical in many countries). The proposed methodology uses mathematical programming and forecasting tools to deliver substantial reduction on the electric bill of a large Spanish hospital.

With respect to the retail electricity market for the Polish market, the study by [29] revealed that 81% of customers could benefit from tariff optimization or switching. This suggests that customers are not necessarily aware of the benefits due to tariff change, since the majority of the individual customers in Poland have flat tariff plans.

Finally, it is worth mentioning that modern energy systems target the convergence of infrastructure for different energy carriers, which has recently become a major focus of research and development [30]. This creates additional conditions for optimization and long-term forecasting.

## 1.2. Contribution

Motivated by the aforementioned discussions, this paper presents a solution for small- and medium-sized enterprises from the C tariff group, concerning short-term load forecasts as a basis for calculating and optimizing the capacity required to avoid any additional fees related to exceeding the contracted capacity level.

The main contributions of the article can be summarized as follows:

- We develop a new, two-stage approach to determine the optimal capacity contract based on the predicted maximal capacity demands and genetic algorithms;
- As an alternative to the two-stage approach, we create a new deep learning architecture to predict the maximal capacity demands and, simultaneously, to determine the optimal capacity contract;
- We propose the incorporation of the quantile loss function in deep learning model learning for the benefit of accurate prediction of the maximum consumption;

- Through empirical analysis, we compare and choose the best forecasting strategy, i.e., direct multistep forecasting, recursive multistep forecasting and multiple-output forecasting.

Specifically, a long short-term memory (LSTM) artificial neural network (ANN) is constructed to forecast the load values and the moments of exceeding the contracted capacity in the short-term horizon, i.e., up to one month ahead. The forecast is further used to optimize the capacity volume to be contracted in the following month for the commercial customer to minimize network charge for exceeding the contracted level.

Long short-term memory networks belong to a complex area of deep learning methods. These are types of recurrent neural networks (RNNs) capable of learning order dependence in sequence prediction problems such as time series. The reason for using recurrent networks is that they are different from traditional feed-forward neural networks, and taking into account the complexity and volatility in electricity time series, RNNs are able to identify patterns and behaviors that traditional methods cannot achieve [31]. Recurrent neural networks contain cycles that feed the network activations from a previous time step as inputs to the network to influence predictions at the current time step. These activations are stored in the network, which can hold long-term temporal contextual information. This mechanism allows ANNs to exploit a dynamically changing contextual window over the input sequence history [32].

Standard RNNs often fail to learn correctly in the presence of time lags greater than 5–10 discrete time steps between the input events and target signals. As provided in [33], the LSTM model is not affected by this problem, and it can learn to bridge minimal time lags in excess of 1000 discrete time steps by enforcing constant error flow through “constant error carousels” (CECs) within special units, called cells.

The remainder of this paper is organized as follows. Section 2 proposes a two-stage approach to optimize the electricity contract capacity problem. As an alternative to the two-stage approach, a deep learning solution is tested. Section 3 provides a detailed description of the tariff structure in Poland. Section 4 applies the models to real datasets for two commercial customers in Poland. Section 5 concludes with the comments and provides directions for future research.

## 2. Proposed Approaches to Optimize the Electricity Contract Capacity Problem

### 2.1. Two Stage Approach

#### 2.1.1. Stage One—LSTM Electricity Load Time Series Forecasting

Time series forecasting is typically considered a one-step prediction. Because electricity load forecasting is essential for both the utility and the customer, it cannot be designed with one-step prediction. Maximum power is used by the utility to provide the right amount of power for customers, whereas it is the basis for calculating, usually monthly, fixed charges for industrial and business electricity users. Predicting multiple time steps is considered a multistep forecasting and includes prediction of the load values  $[L_{+1}, \dots, L_{+h}]$  based on the historical load time series  $[L_{-1}, \dots, L_{-N}]$  composed of  $N$  load observations, where  $h > 1$  denotes the forecasting horizon. In this article, we use abbreviation  $h$  since our goal is to provide hourly forecasts.

In this paper, we analyze four strategies for electricity load forecasting: naïve forecasting, recursive multistep forecasting, direct multistep forecasting and a multiple-output strategy.

**Naïve forecasting.** The naïve forecast was considered in the following manner: for the forecasting horizon, the values observed for the same hour and same day of the four previous weeks were averaged and taken as a forecast. In practice, such approaches are considered a starting point to which other proposed methods should be compared. Once the results for those new methods are better, it is valid to conclude that it is worth putting the effort into the usage of these methods.

**Direct multistep forecasting.** This approach involves the development of separate models for each hour in the following month, which gives us up to 744 models

(31 days  $\times$  24 h). In the numerical example, we develop separate models for predicting the load required for each hour. The load prediction time series model form is given by:

$$\begin{aligned} L_{+1} &= \text{model1}([L_{-1}, \dots, L_{-N}]) \\ L_{+2} &= \text{model2}([L_{-1}, \dots, L_{-N}]) \\ &\vdots \\ L_{+744} &= \text{model774}([L_{-1}, \dots, L_{-N}]) \end{aligned} \quad (1)$$

Since separate models are used, there is no opportunity to consider the dependencies between the forecasts, e.g., the forecast at hour 2 is dependent on the prediction at hour 1, which is often the case in a time series.

**Recursive multistep forecasting.** This approach repeatedly uses a one-step model in which the prediction for the previous time interval is used as an input to make a prediction for the next time interval. Following the previous example, to predict the load for the next two months, we first developed a one-step load forecasting model. The developed model was used to predict hour 1 electricity load. The obtained value was further used as an observation input to predict the load for hour 2. The recursive multistep load prediction model is given by:

$$\begin{aligned} L_{+1} &= \text{model}([L_{-1}, \dots, L_{-N}]) \\ L_{+2} &= \text{model2}([L_{+1}, L_{-1}, \dots, L_{-N-1}]) \\ &\vdots \\ L_{+744} &= \text{model}([L_{+743}, \dots, L_{-1}]) \end{aligned} \quad (2)$$

Because forecasts are used in place of previous observations, in the case of exceeding the contract capacity value in one month, the prediction values for the next months can quickly upgrade as the prediction time horizon increases.

**Multiple-output forecasting.** In this case, forecasting involves the development of a single model that is capable of predicting the entire forecast time horizon in a one-shot approach. Therefore, to predict the load required for the next month, e.g., one month, we developed one model and used it to predict the next month as one operation. The model form would be as follows:

$$L_{[+1, \dots, +744]} = \text{model}([L_{-1}, \dots, L_{-N}]) \quad (3)$$

The model can learn the dependence structure between inputs and outputs as well as between outputs.

Specifically, for the latter three approaches, i.e., direct multistep, recursive multistep and multiple-output, LSTM artificial neural networks (for more details, please see Section 2.2) were constructed to forecast the load values and the moments of exceeding the contracted capacity in the short-term horizon, i.e., hour by hour up to one month ahead. The forecast was further used to optimize the capacity volume to be contracted in the following month for the commercial customer to minimize network charge for exceeding the contracted level. The reason for choosing the neural network algorithms is that they are directly applicable to all three approaches (with almost the same structure) without the need to convert (combine) the one-output model to a multiple-output strategy. Additionally, LSTM networks are capable of learning long-term dependencies observed in electricity load time series.

Since the standard regression models predict the average value, it might happen that most of the forecasts will be underestimated. This would have huge implications for the optimization step because the search space will be compressed (squeezed) and the optimal contract will be set too low, e.g., 50 kW instead of 55 kW. In this case, a penalty charge will be applied to the exceeded amount, which is up to ten times the basic rate (please see Formula (6)).

To overcome the above problem, we apply the modified loss function for the aforementioned strategies to predict the contract value with the maximum consumption values as

well as the maximum load at specific hours and days of the week. That is why we carried the function for 100 quantiles to check how large the maximum loads are. The quantile loss for an individual data point is defined as:

$$Q(\zeta_i|\alpha) = \begin{cases} \alpha\zeta_i, & \text{if } \zeta_i \geq 0 \\ (\alpha - 1)\zeta_i, & \text{if } \zeta_i < 0 \end{cases}, \quad (4)$$

where  $\alpha$  is the required quantile (a value between 0 and 1) and

$$\zeta_i = y_i - f(x_i), \quad (5)$$

where  $f(x)$  is the predicted (quantile) model and  $y$  is the observed value for the corresponding input  $x$ .

### 2.1.2. Stage Two—Load Forecast Optimization

In this article, we consider peaks over contracted capacity in a given month. Most customers order the same amount of power for individual months of the year. If the peak demand does not exceed the contractual capacity, a fixed capacity charge will be levied. It constitutes the product of the fixed capacity rate  $R$  [PLN/kW], where PLN stands for Polish Zloty and contracted capacity demand for month  $R_t$  in kW. To exceed the contractual capacity defined in the contract, an additional surcharge for excess demand will be added. The annual cost can be, therefore, expressed as:

$$Cost_m = \begin{cases} R_m * d_m & d_m < d_m^c \\ R_m * d_m + R_m * (d_m - d_m^c) * n_m & d_m^c < d_m \end{cases} \quad (6)$$

$$Cost_{total} = \sum_{m=1}^{year} Cost_m \quad (7)$$

where

$d_m^c$ —contracted capacity (kW) in month  $m$ ;

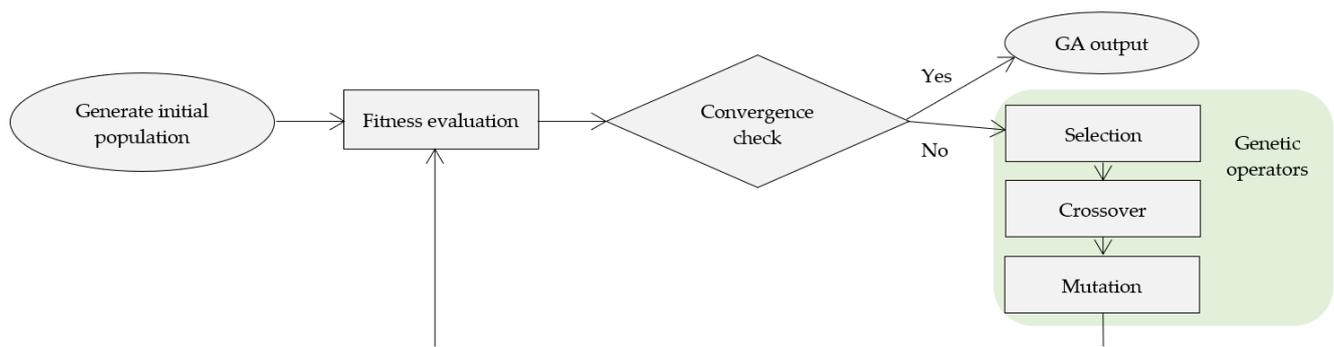
$d_m$ —maximum demand amount (kW) in month  $m$ ;

$n_m$ —the sum of up to ten largest amounts of surplus consumed capacity over the contractual capacity, indicated by the measuring;

$R_m$ —rate of contractual capacity (PLN/kW) in month  $m$ .

In this work, since 12 months of data are available, we consider the total cost over 10 months, i.e., March–December 2016, because January and February were considered for model training (including variable calculations with delays). The solution that minimizes the annual total contracted capacity cost and the penalties for excessive consumption over the fixed capacity amount can be solved using particle swarm optimization [26], the genetic algorithm (GA) [27] or even Excel's solver for linear programming unless the solution is trapped in a local minimum [10]. However, in this paper, we propose a GA that can find multiple Pareto solutions for a multiobjective optimization problem in one run.

In principle, genetic algorithms are stochastic search algorithms inspired by biological evolution and natural selection processes. The GA simulates an evolutionary process in which the strongest individuals dominate the weaker individuals, reflecting the biological mechanism of evolution, such as crossover, mutation, and selection (see Figure 1). For experiments, we used the R package for GA because it provides a set of different optimization functionalities.



**Figure 1.** Flow chart of the genetic algorithm.

Since the search space is constructed based on the received forecasts, the genotype is represented as a floating-point value of real numbers. In this representation, in the first step, the algorithm produces the initial population (as presented on the left-hand side of Figure 1). The population consists of 50 individuals randomly generated from the uniform distribution from the range  $[0,100]$ , i.e., The maximum demand across the investigated period for all datasets is 95; therefore, 100 in our opinion is reasonable. The fitness function controlling the optimization process is defined in Formula (6), i.e.,  $\arg \min_{d_m} Cost_m$ . The selection process is conducted based on fitness proportional selection with fitness linear scaling. Other genetic operators were set at local arithmetic crossover and uniform random mutation. The maximum number of iterations to run before the GA search is halted is set at 100. By default, the top 5% of individuals will survive at each iteration.

## 2.2. Hybrid Model for Optimization and Multiple Output Forecasting

The core part of the aforementioned strategies constitutes the deep neural network composed of several layers. These are fully connected dense layers [34] and LSTM layers. LSTMs are clearly designed to avoid a long-term dependency problem (e.g., vanishing gradients). Remembering knowledge for a long time is practically their default behavior. All recurrent neural networks take the form of a chain of repeating neural network modules. LSTM also has this chain-like structure, but instead of a single layer of neural network, there are four, interacting with each other in a very special way.

The key to LSTM is the so-called cell state  $C_t$ , which is presented by the horizontal line running through the top of Figure 2. The state of the cell is similar to that of a conveyor belt. It runs straight down the entire chain, with slight linear interactions. It is very easy for unchanged information to flow through it. LSTM has the ability to delete or add information to the state of the cell, precisely regulated by structures called gateways. This decision is undertaken by the “forget gate layer” built based on a sigmoid layer, denoted as  $f_t$  and with Formula (8). It checks  $h_t$  and  $x_t$  and displays a number from 0 to 1 for each number in the cell state  $C_{t-1}$ , where 1 means “keep it completely” and 0 means “getting rid of it completely”.

During the next step (composed of two parts), the network decides what new information should be maintained in the cell state. First, the sigmoid layer  $i_t$  (Formula (9)), called the “input gate layer”, determines which values should be updated. Then, the  $\tilde{C}_t$  layer (Formula (10)) creates a vector of new candidate values that can be added to the state. Both outputs are then combined to create an update to the state.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (8)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (9)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad (10)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \quad (11)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (12)$$

$$h_t = o_t * \tanh(C_t), \quad (13)$$

where  $x_t \in \mathbb{R}^d$  is the  $d$ -dimensional input vector to the LSTM unit,  $f_t \in \mathbb{R}^h$  denotes the forget gate's activation vector,  $i_t \in \mathbb{R}^h$  is the input/update gate's activation vector,  $o_t \in \mathbb{R}^h$  denotes the output gate's activation vector,  $h_t \in \mathbb{R}^h$  denotes the hidden state vector also known as the output vector of the LSTM unit,  $\tilde{C}_t \in \mathbb{R}^h$  denotes the cell input activation vector,  $C_t \in \mathbb{R}^h$  denotes the forget gate's activation vector, and  $W \in \mathbb{R}^{h \times d}$  and  $b \in \mathbb{R}^d$  are the weight matrices and bias vector parameters that need to be learned during training.

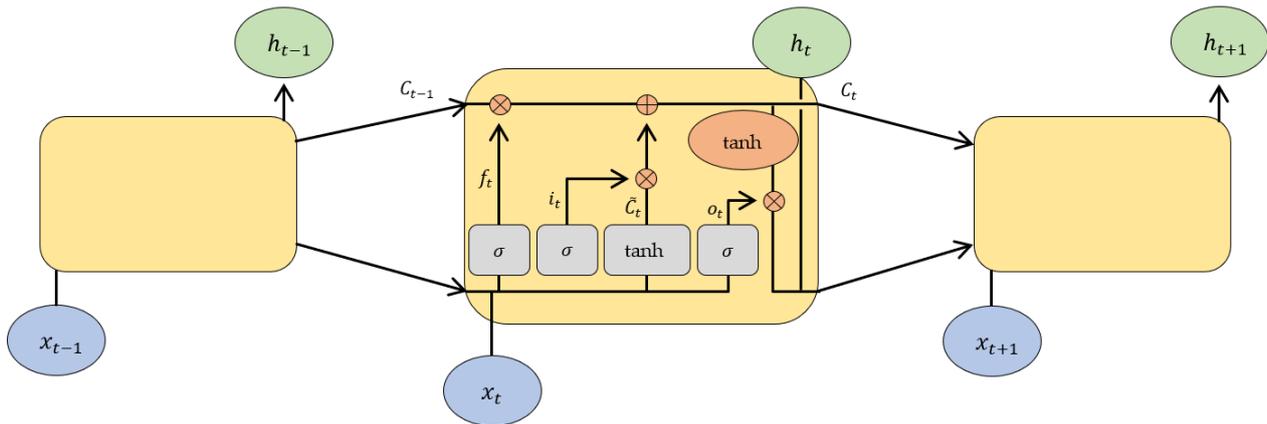
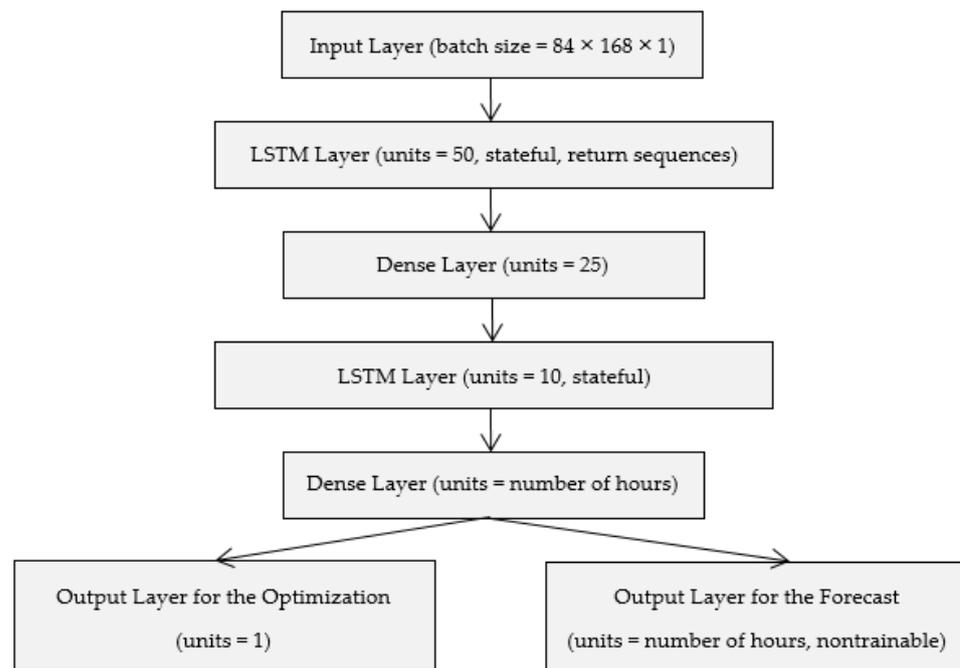


Figure 2. Long short-term memory network.

After the above operations, the old cell state,  $C_{t-1}$ , is updated to the new cell state  $C_t$ . It is done by multiplying the old state  $C_t$  by  $f_t$ , i.e., forgetting the desired information and adding the  $i_t * \tilde{C}_t$  (see Formula (11)). Ultimately, the network has to decide what output to produce. First, a sigmoid layer  $o_t$  is run deciding what parts of the cell state bring out to (Formula (12)). Next, the state of the cell  $C_t$  is activated by  $\tanh$  and multiplied through the output of the sigmoid gate  $o_t$ , producing the final output  $h_t$ .

As stated in Section 2.1.1, the reason for choosing the deep learning architecture is that it is directly applicable to all three strategies (please see Figure 3 and description below). In summary, the architecture for the direct multistep forecast and recursive multistep forecast is the same. The difference is in the preparation of the training subsets and the number of trained models (please see Formulas (1) and (2)). In the case of the multiple-output strategy, the difference is in the last output layer, having not one but multiple outputs. Finally, the newly proposed architecture consists of an additional output layer producing the optimal capacity contract.

Initially, in the deep learning architecture presented in Figure 3, there is an input layer receiving a batch of size  $84 \times 168 \times 1$ . This is because the deep network receives a three-dimensional vector (tensor) with the dimension samples  $\times$  timesteps  $\times$  features. Since for the learning we use the last three weeks, each epoch should process  $24 \times 21 = 504$  observations and because in a stateful network, we should only pass inputs with a number of samples that can be divided by the batch size, the latter one is set at 84. For lag observations, we present the feature vector of the last week  $24 \times 7 = 168$ . There are no other external variables used for learning, which is why the last dimension is 1.



**Figure 3.** Hybrid model for optimization and multiple-output forecasting.

The first long short-term memory layers consist of 50 units whose activation functions are set at hyperbolic tangent (to push the values to be between  $-1$  and  $1$ ) in the cell state, and later, the output is multiplied by the output of the sigmoid gate. Each unit returns the last output in the output sequence (instead of the full sequence), and the layer is stateful, meaning that the last state for each sample in the batch will be used as the initial state for the sample in the following batch.

Next, the first dense layer has 25 fully connected hidden neurons (with a linear activation function) constructing a  $50 \times 25$  weight matrix, which is subject to learning. All neurons are connected with their biases. This layer is then connected to the second LSTM layer having 10 units. Up to this point, all of the employed deep networks have the same structure. After that layer, there is a fully connected dense output layer with one unit (one forecasting model for each future hour) for the direct multistep and recursive forecast. When considering the multiple-output forecast, the last layer consists of a number of neurons equal to the number of forecasted hours.

The newly proposed hybrid model for optimization and multiple-output forecasting consists of two additional layers (both output layers). The first ultimate layer produces the desired forest (presented on the right in Figure 3). This layer is built based on the identity dense layer to obtain the same outputs as those provided by the previous dense layer. It is achieved by the nontrainable identity weights matrix. The second output (presented on the left in Figure 3) produces the optimal contract capacity based on the received forecasts (this layer has one unit).

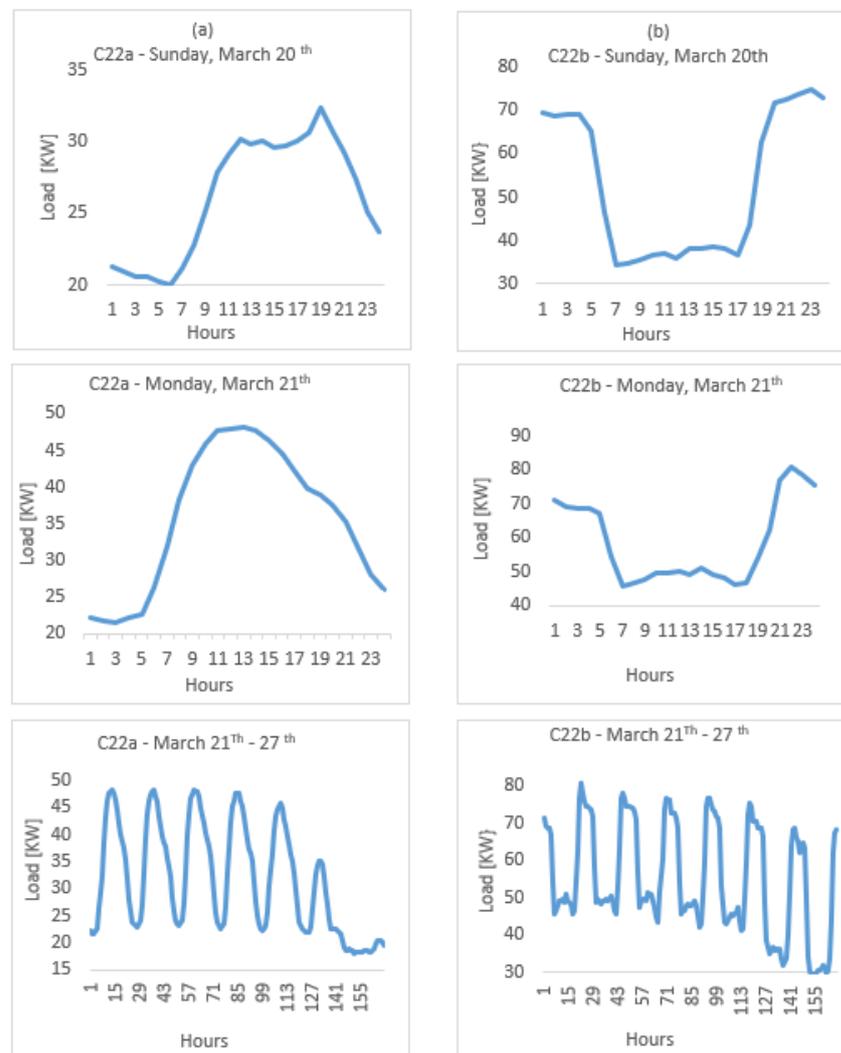
All models are trained using quantile loss defined in Formulas (4) and (5), while the additional output for optimization in the hybrid model minimizes (likewise genetic algorithms) loss defined in Formula (6). The adaptive moment estimation (Adam) [35] algorithm is used for weight updating.

### 3. Data Characteristics and Tariff Structure

There were two separate datasets for medium-sized commercial customers used in the analysis, each with the data points gathered at 15-min intervals and covering the time interval between 1 January 2016 and 31 December 2016. In total, there are 35,136 observations in each dataset. The customers belong to the C tariff group, which is applicable to small and medium-size enterprises where electricity is supplied with low voltage lines.

The group includes C2x tariffs where contracted capacity is over 40 kilowatts and the letter “x” designates the number of energy consumption zones per day. The following tariffs are available: C22a tariff with two zones (peak and off-peak), C22b tariff with two zones (day and night) and C23 tariff with three zones (morning peak, afternoon peak and off-peak).

The first dataset contains details for customers who belong to the C22a tariff. The customer is classified as a small pharmaceutical plant with a contracted capacity greater than 40 kilowatts and who is mainly using electricity during the day hours. The contracted capacity for the customer is 51 kW. Figure 4a shows lower electricity consumption during night hours, while much higher consumption is observed between 9:00 and 17:00 (except for Sundays).



**Figure 4.** Daily and weekly energy consumption structures of two different users: (a) user C22a, (b) user C22b.

The second dataset contains details for customers who belong to the C22b tariff. It is a confectionery plant that performs the majority of its activities during the night. The contracted capacity for the customer is 80 kW. Figure 4b shows lower electricity consumption in the daytime zone, i.e., between 6:00 and 21:00, and higher consumption in the night time zone, i.e., between 21:00 and 6:00 (including weekends).

Most of the users within C2x tariff groups do not possess detailed usage data to control energy consumption parameters and to ensure their optimal adjustment. As shown in Figure 5, the contracted capacity is not adequately set, as it is often exceeded in reality. On the other hand, the average load consumption does not exceed 70% of the contracted

capacity level, which translates into losses due to unused capacity. Therefore, it is crucial to determine the optimal contract capacity for each month to minimize the total cost of the electricity bills.

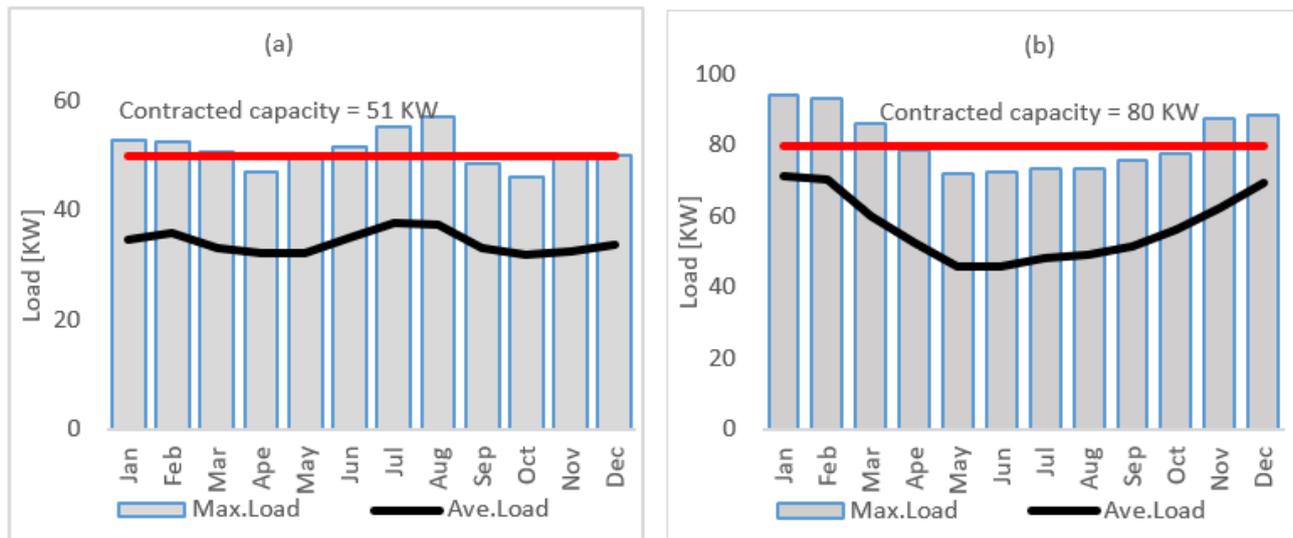


Figure 5. Consumed load and contracted capacity (a) for customer C22a and (b) customer C22b.

#### 4. Numerical Experiments

In this section, within the two-stage approach outlined previously, we used four forecasting approaches for electricity load forecasting, and then, in the second stage, we applied a genetic algorithm to optimize the user's contract capacity. Additionally, as an alternative to the two-stage approach, we tested a hybrid model to predict the maximal capacity demands and, simultaneously, to determine the optimal capacity contract.

Initially, we started with hourly forecasts for the entire year. Although the settlement with the power plant or electricity supplier was made on the basis of the maximum load in the monthly billing period, hourly forecasting is necessary for load optimization at peak times in the second stage. From the forecasted hourly short-term load values, we selected the largest daily load values. Thus, we obtain 366 values for 2016, from which a set of 12 maximum values from each month is selected. These values constitute the input data to predict and then optimize the amount of capacity required in the next period.

For the forecasting approach, we determined the following components: (1) the quantities and costs incurred on the basis of the actual load consumption and contract capacity, i.e., the constant value declared by the user at the beginning of the contract period; and (2) the optimal load amount and the cost that the user would incur. This is the case when we know, in advance, the amount of power required at the end of the billing period. (3) The optimal amount and the costs that the user would incur on the basis of the predicted load quantities using a naïve forecast and three approaches supported by LSTM neural networks. Ultimately, we determined the optimal contract capacity using a genetic algorithm.

In the following, the results of the forecasting experiments and optimization will be discussed. The following notations are used in Tables 1 and 2:

- Actual contract—the value of the customer's contracted capacity in kW;
- Actual cost—the customer's total cost of contracted capacity and the penalties exceeding the contracted capacity level in PLN;
- Above actual contract—the amount of capacity consumed over the contracted level in kW;
- Opt contract capacity—the optimal amount of consumed capacity based on the historical usage in kW;

- Opt contract cost—the optimal cost of consumed capacity based on historical usage in PLN;
- Above opt contract—the number of loads over the contracted capacity based on historical usage;
- Opt contract capacity pred—the optimal contract based on the forecast obtained by the neural network and optimized by GA in kW;
- Opt cost capacity pred—the total cost of optimal contract predicted by the network and optimized by GA in PLN;
- Above opt capacity pred—the number of loads over the contracted capacity based on the forecast obtained by the neural network and optimized by GA.

Table 1 shows the results of the analysis for the customer who belongs to the C22a tariff group, having a contracted capacity of 51 kW per month. During the June–August period, the customer consumed more capacity, and therefore, the contracted level was exceeded several times in those months, e.g., even 122 times in July, which significantly increased the cost. In total, the actual cost for the customer between March and December was 6166.37 PLN. With a retrospective analysis, based on historical usage, one could see that the optimal values for contracted capacity would vary between 46 kW and 57 kW, depending on the month, as presented in Table 1. Knowing that, the customer could benefit from lower bills, so the cost of the optimal contract would be 5093.99 PLN, which is 17.4% less than actual cost. Of course, for the customer, it is difficult to correctly specify what would be the capacity required in the following months; therefore, the optimal contract capacity should be forecasted. In our case, we used four forecasting strategies to estimate the maximum load for each hour one month ahead, and these values were further used as the input to the genetic algorithm to search for the optimal contract level so that the total cost was minimized.

The naïve approach resulted in a forecasted capacity between 47 kW and 57 kW (depending on the month), which would cause the contract to be exceeded several times in May, June, July and November (even 94 times in July). After all, the total cost of the optimal contract predicted by the naïve approach and optimized by the genetic algorithm was 6339.84 PLN which is far from the optimal cost (5093.99 PLN). In comparison to the actual cost of 6166.36 PLN, there is no benefit for the customer.

The direct multistep approach forecasted a capacity between 58 kW and 62 kW. These values were overestimated, so there would be no breach of contracted capacity, and the total cost of electricity, after genetic algorithm optimization, would be 5950.00 PLN. In comparison to the actual cost, the benefit for the customer would be 216.36 PLN (6166.36–5950.00), which is 3.5% of the actual bills.

With the recursive multistep approach, the capacity was forecasted to vary between 55 kW and 60 kW. These values were less overestimated in comparison to the direct multistep approach, so the total cost of electricity, after genetic algorithm optimization, would be 5700.00 PLN. In comparison to the actual cost, the benefit for the customer would be 466.36 PLN (6166.36–5700.00), which is 7.6% of the actual bills.

Eventually, as a result of the multiple-output strategy, the capacity was forecasted to be between 48 kW and 57 kW, depending on the month. Importantly, only once, in August would we exceed the contracted capacity. This helped to keep the total cost very low, i.e., close to the optimal cost values. Eventually, the total cost of the optimal contract predicted by the network and optimized by the genetic algorithm was 5211.32 PLN which is very close to the optimal one (5093.99 PLN). In comparison to the actual cost, the benefit for the customer is quite material and amounts to 955.04 PLN (6166.36–5211.32), which is 15.5% of the actual bills.

Finally, the application of the hybrid model for optimization and multiple-output forecasting resulted in the same capacity volume forecasts as for the multiple-output strategy within the two-stage approach; therefore, the same benefit for the customer was estimated.

Table 1. The results with Q100 approach for C22a tariff user.

Month	Actual Values			Optimal Values Based on Historical Usage			Naïve Forecast Optimal Values Based on Predicted Usage			Direct Multi Step Forecast Optimal Values Based on Predicted Usage			Recursive Multi Step Forecast Optimal Values Based on Predicted Usage			Multiple-Output Forecast/Hybrid Approach (Deep Learning) Optimal Values Based on Predicted Usage		
	Actual Contract [kW]	Actual Cost [PLN]	Opt Contract Capacity Pred [kW]	Opt Contract Capacity [kW]	Opt Contract Cost [PLN]	Above Opt Contract	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred
March	51	510.00	0	50	506.74	1	53	530.00	0	60	600.00	0	58	580.00	0	53	530.00	0
April	51	510.00	0	47	471.39	1	51	510.00	0	59	590.00	0	56	560.00	0	49	490.00	0
May	51	510.00	0	50	501.72	1	47	787.28	24	59	590.00	0	56	560.00	0	51	510.00	0
June	51	522.89	2	52	520.00	0	50	664.49	17	59	590.00	0	56	560.00	0	52	520.00	0
July	51	950.30	122	56	560.00	0	52	860.31	94	60	600.00	0	58	580.00	0	56	560.00	0
August	51	1123.16	95	57	571.31	1	56	661.85	9	62	620.00	0	60	600.00	0	57	571.31	1
September	51	510.00	0	49	490.00	0	57	570.00	0	60	600.00	0	59	590.00	0	54	540.00	0
October	51	510.00	0	46	466.28	2	49	490.00	0	58	580.00	0	55	550.00	0	48	480.00	0
November	51	510.00	0	50	500.00	0	47	759.40	54	58	580.00	0	56	560.00	0	50	500.00	0
December	51	510.00	0	50	506.52	3	50	506.52	3	60	600.00	0	56	560.00	0	51	510.00	0
Total costs		6166.36			5093.99			6339.84			5950.00			5700.00			5211.31	

Note: Multiple Output Forecast and Hybrid approach (deep learning) delivered the same capacity volume forecasts; thus, the same benefit for the customer is estimated.

Table 2. The results with Q100 approach for C22b tariff user.

Month	Actual Values			Optimal Values Based on Historical Usage			Naïve Forecast Optimal Values Based on Predicted Usage			Direct Multi Step Forecast Optimal Values Based on Predicted Usage			Recursive Multi Step Forecast Optimal Values Based on Predicted Usage			Multiple-Output Forecast/Hybrid Approach (Deep Learning) Optimal Values Based on Predicted Usage		
	Actual Contract [kW]	Actual Cost [PLN]	Opt Contract Capacity Pred [kW]	Opt Contract Capacity [kW]	Opt Contract Cost [PLN]	Above Opt Contract	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred	Opt Contract Capacity Pred [kW]	Opt Cost Capacity Pred [PLN]	Above Opt Capacity Pred
March	80	1402.72	62	85	860.27	1	93	930.00	0	99	990.00	0	99	990.00	0	88	880.00	0
April	80	800.00	0	78	786.63	1	86	860.00	0	98	980.00	0	97	970.00	0	85	850.00	0
May	80	800.00	0	72	722.82	1	79	790.00	0	98	980.00	0	93	930.00	0	77	770.00	0
June	80	800.00	0	72	724.29	1	73	730.00	0	94	940.00	0	87	870.00	0	73	730.00	0
July	80	800.00	0	73	737.82	2	73	737.82	2	94	940.00	0	87	870.00	0	74	740.00	0
August	80	800.00	0	74	740.00	0	74	740.00	0	97	970.00	0	91	910.00	0	74	740.00	0
September	80	800.00	0	76	760.00	0	74	878.72	8	96	960.00	0	87	870.00	0	76	760.00	0
October	80	800.00	0	78	780.00	0	76	923.78	11	97	970.00	0	92	920.00	0	78	780.00	0
November	80	1547.07	58	87	874.71	1	78	1727.07	98	99	990.00	0	92	920.00	0	87	874.71	1
March	80	2239.56	266	89	943.96	1	88	1199.78	5	99	990.00	0	98	980.00	0	89	943.96	1
Total costs		10,789.34			7930.48			9517.17			9710.00			9230.00			8068.66	

Note: Multiple-Output Forecast and Hybrid approach (deep learning) delivered the same capacity volume forecasts; thus, the same benefit for the customer is estimated.

In a similar manner, the analysis for the second customer was prepared. Table 2 shows the results of the analysis for the customer who belongs to the C22b tariff group, having a contracted capacity of 80 kW per month. During March, November and December, the customer consumed more capacity, and therefore, the contracted level exceeded the number of times, specifically 266 times in December, which significantly impacted the actual bills. In total, the actual cost for the customer between March and December was 10,789.34 PLN. With a retrospective analysis, based on historical usage, one could see that the optimal values for contracted capacity would vary between 72 kW and 90 kW, depending on the month, as presented in Table 2. Knowing that, the customer could benefit from lower bills, so the cost of the optimal contract would be 7930.49 PLN, which is 26.5% less than the actual cost.

The naïve approach has a forecasted capacity between 73 kW and 93 kW, which would cause the contract to exceed 98 times in July in November. After all, the total cost of the optimal contract predicted by the naïve approach and optimized by the genetic algorithm was 9517.17 PLN which is far from the optimal cost (7930.48 PLN). In comparison to the actual cost, the benefit for the customer is 1272.17 PLN (10,789.34–9517.17), which is 11.8% of the actual bills.

The direct multistep approach forecasted a capacity between 94 kW and 99 kW. These values were overestimated, so there would be no breach of contracted capacity, and the total cost of electricity, after genetic algorithm optimization, would be 9710.00 PLN. In comparison to the actual cost, the benefit for the customer would be 1079.34 PLN (10,789.34–9710.00), which is 10.0% of the actual bills.

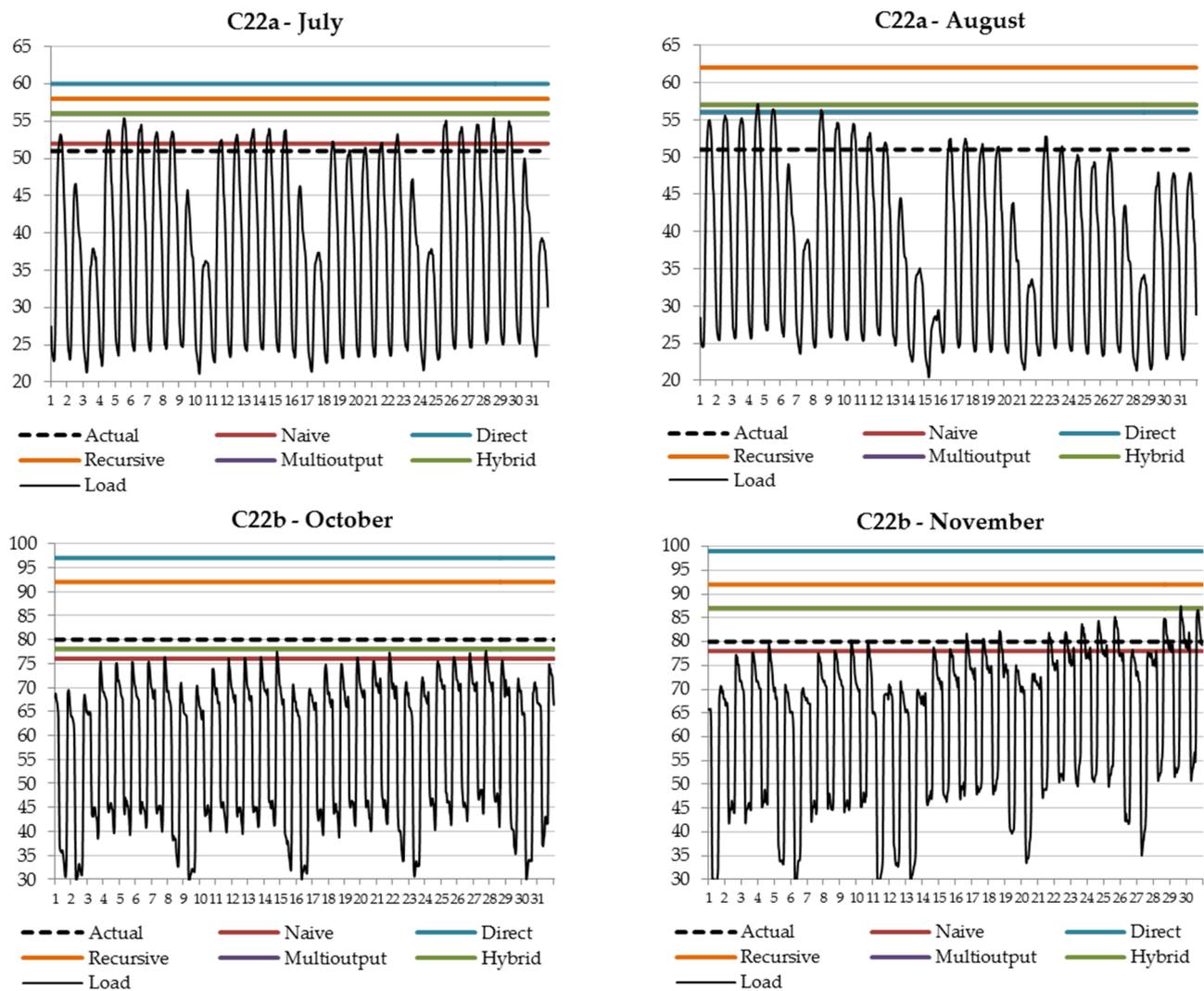
With the recursive multistep approach, the capacity was forecasted to vary between 87 kW and 99 kW, and there will be no breach of contracted capacity, which makes the total cost of electricity equal to 9230.00 PLN. In comparison to the actual cost, the benefit for the customer would be 1559.34 PLN (10,789.34–9230.00), which is 14.5% of the actual bills.

Finally, with a multiple output forecast strategy to estimate the maximum load, the forecasted capacity was between 73 kW and 89 kW, depending on the month. There were instances where the usage would exceed the contracted capacity, e.g., once in November and once in December. Finally, the total cost of the optimal contract predicted by the network and optimized by the genetic algorithm was 8068.66 PLN. In comparison to the actual cost, the benefit for the customer is material, and it amounts to 2720.68 PLN (10,789.34–8068.66), which is 25.2% of the actual bills.

As previously mentioned, the application of a hybrid model for optimization and multiple-output forecasting resulted in the same capacity volume forecasts as for the multiple output strategy within the two-stage approach; therefore, the same benefit for the customer was estimated.

The visualization of the approaches that were tested is presented in Figure 6. For this reason, the real load and the capacity contracts estimated with various approaches were provided for both customers. The upper part of the figure shows the details for the C22a customer, while the lower part of the figure depicts the details for the C22b customer.

It was observed that the multiple-output forecast strategy and hybrid approach delivered the same results, i.e., the same levels of capacity contracts were proposed in the analysis period. Please note that the hybrid approach and multiple-output method overlap in Figure 6. Both approaches were able to estimate the capacity of the contract very close to the actual load.



**Figure 6.** Exemplary consumed loads versus optimized contracted capacities based on the investigated methods for customers C22a and C22b. Note: Hybrid and multiple-output methods overlap.

## 5. Conclusions

In this paper, we presented several solutions applicable for commercial customers.

First, a two-stage approach was proposed to determine the appropriate contract capacity amount that minimizes financial losses in the case of exceeding the amount of capacity defined in the contract. The first stage was to apply four strategies to forecast hourly capacity values as the basis to determine the monthly maximum capacity required. These maximum values were used to determine the optimal monthly capacity at the second stage, so the values were provided as the input to the genetic algorithm to establish such a monthly contract capacity level that would help the user avoid charges for exceeding the contracted level.

Second, as an alternative to the two-stage approach, we created a hybrid approach, i.e., a new deep learning architecture to predict the maximal capacity demands and, simultaneously, to determine the optimal monthly capacity contract.

As shown through the experiments, the application of a two-stage approach, i.e., multiple-output forecasting with an artificial neural network model and a genetic algorithm for load optimization, delivered significant benefits to commercial customers. In comparison to the actual costs, the benefit for the customers, due to optimization, is mate-

rial. Specifically, the benefit for the C22a customer is 15.5% of the actual bills, while for the C22b customer, it is 25.2%.

For the hybrid approach, we observed that forecasts of the deep learning model resulted in the same capacity contracts proposed for consecutive months as for the two-stage approach; thus, the benefits for the customers were exactly the same. However, the advantage of this approach is that prediction and optimization are performed simultaneously, which simplifies the process.

With our analysis, we confirm that customers are not necessarily aware of the benefits due to capacity volume optimization. The reason for that might be that customers do not have means or methods to analyze and to draw conclusions from the data, so they could discover the efficiency potential.

In future work, we will continue the research towards fitting the models so that these models could potentially better deal with the seasonality of the demand on the customer end. Although this research deals with Polish tariffs, we believe it can be applied to other electricity customers in capacity cost decision making. Additionally, we intend to extend the analysis to the broader set of customers.

**Author Contributions:** R.N. wrote Sections 1–5 of the manuscript; T.Z. wrote Sections 1, 2, 4 and 5 of the manuscript; K.G. prepared the simulation and analysis and wrote Sections 1, 2 and 5 of the manuscript. All authors have read and agreed to the published version of the manuscript.

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

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

## References

1. Tostado-Véliz, M.; Icaza-Alvarez, D.; Jurado, F. A novel methodology for optimal sizing photovoltaic-battery systems in smart homes considering grid outages and demand response. *Renew. Energy* **2021**, *170*, 884–896. [CrossRef]
2. Regulation of the minister of economy of 27 april 2012 amending the regulation on the specific rules for the determination and calculation of tariffs and billing in electricity trading. *J. Laws* **2012**, *640*, 5–16.
3. Lindberg, K.; Seljom, P.; Madsen, H.; Fischer, D.; Korpås, M. Long-term electricity load forecasting: Current and future trends. *Util. Policy* **2019**, *58*, 102–119. [CrossRef]
4. Rajakovic, N.L.; Shiljkut, V.M. Long-term forecasting of annual peak load considering effects of demand-side programs. *J. Mod. Power Syst. Clean Energy* **2017**, *6*, 145–157. [CrossRef]
5. Gajowniczek, K.; Nafkha, R.; Ząbkowski, T. Electricity Peak Demand Classification with Artificial Neural Networks. In Proceedings of the 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), Prague, Czech Republic, 3–6 September 2017; IEEE: Prague, Czech Republic, 2017; pp. 307–315.
6. Mojica, J.L.; Petersen, D.; Hansen, B.; Powell, K.M.; Hedengren, J.D. Optimal combined long-term facility design and short-term operational strategy for chp capacity investments. *Energy* **2017**, *118*, 97–115. [CrossRef]
7. Torkzadeh, R.; Mirzaei, A.; Mirjalili, M.M.; Anaraki, A.S.; Sehhati, M.R.; Behdad, F. Medium Term Load Forecasting in Distribution Systems Based on Multi Linear Regression & Principal Component Analysis: A Novel Approach. In Proceedings of the 2014 19th Conference on Electrical Power Distribution Networks (EPDC), Poonak, Tehran, 6–7 May 2014; IEEE: Poonak, Teheran, 2014; pp. 66–70.
8. Hippert, H.S.; Pedreira, C.E.; Souza, R.C. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Trans. Power Syst.* **2001**, *16*, 44–55. [CrossRef]
9. Ryu, S.; Noh, J.; Kim, H. Deep neural network based demand side short term load forecasting. *Energies* **2016**, *10*, 3. [CrossRef]
10. Chen, C.-Y.; Liao, C.-J. A linear programming approach to the electricity contract capacity problem. *Appl. Math. Model.* **2011**, *35*, 4077–4082. Available online: <https://www.sciencedirect.com/science/article/pii/S0307904X11000898> (accessed on 10 February 2021). [CrossRef]
11. CRE. *Règles Tarifaires Pour L'utilisation des Réseaux Publics D'électricité*; Commission de régulation de l'énergie: Paris, France, 2012.
12. Baharudin, Z.; Kamel, N. Autoregressive Method in Short Term Load Forecast. In Proceedings of the 2008 IEEE 2nd International Power and Energy Conference, Johor Bahru, Malaysia, 1–3 December 2008; IEEE: Johor Bahru, Malaysia, 2008; pp. 1603–1608.
13. Karpio, K.; Łukasiewicz, P.; Nafkha, R. Regression technique for electricity load modeling and outlined data points explanation. In Proceedings of the International Multi-Conference on Advanced Computer Systems, Cairo, Egypt, 26–28 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 56–67.
14. Alberg, D.; Last, M. Short-term load forecasting in smart meters with sliding window-based arima algorithms. *Vietnam. J. Comput. Sci.* **2018**, *5*, 241–249. [CrossRef]
15. Mohamed, N.; Ahmad, M.H.; Ismail, Z. Improving Short Term Load Forecasting Using Double Seasonal Arima Model. *World Appl. Sci. J.* **2011**, *15*, 223–231.

16. Mado, I.; Soeprijanto, A.; Suhartono, S. Applying of double seasonal arima model for electrical power demand forecasting at pt. Pln gresik indonesia. *Int. J. Electr. Comput. Eng.* **2018**, *8*, 4892–4901. [[CrossRef](#)]
17. Kowm, D.; Kim, M.; Hong, C.; Cho, S. Artificial neural network based short term load forecasting. *Int. J. Smart Home* **2014**, *8*, 145–150. [[CrossRef](#)]
18. Feinberg, E.A.; Genethliou, D. Load forecasting. In *Applied Mathematics for Restructured Electric Power Systems*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 269–285.
19. Shahriar, S.M.; Hasan, M.K.; Abrar, S.R.A. An effective artificial neural network based power load prediction algorithm. *Int. J. Comput. Appl.* **2019**, *178*, 35–41.
20. Charytoniuk, W.; Chen, M. Very short-term load forecasting using artificial neural networks. *IEEE Trans. Power Syst.* **2000**, *15*, 263–268. [[CrossRef](#)]
21. Haniffuddin, H.; Mansor, W. Plc based load limiter. In Proceedings of the 2011 IEEE Control and System Graduate Research Colloquium, Shah Alam, Malaysia, 27–28 June 2011; pp. 185–188.
22. Hayakawa, H.; Watanabe, T.; Kimura, T. Electric Power Distribution and Utility Monitoring System for Better Energy Visualization. *Hitachi Rev.* **2014**, *63*, 427–431.
23. Fraczek, J.; Blach, D. Contracted power order optimization. *Studia Inform.* **2014**, *35*, 199–212.
24. Yang, H.-T.; Peng, P.-C.; Huang, C.-H. Improved taguchi method based contracted capacity optimization for power consumer with self-owned generating units. In Proceedings of the 6th WSEAS International Conference on Applications of Electrical Engineering, Istanbul, Turkey, 27–29 May 2007; Citeseer: Istanbul, Turkey, 2007.
25. Liu, D.; Cai, Y. Taguchi method for solving the economic dispatch problem with nonsmooth cost functions. *IEEE Trans. Power Syst.* **2005**, *20*, 2006–2014. [[CrossRef](#)]
26. Niu, D.; Dai, S. A short-term load forecasting model with a modified particle swarm optimization algorithm and least squares support vector machine based on the denoising method of empirical mode decomposition and grey relational analysis. *Energies* **2017**, *10*, 408. Available online: <https://www.mdpi.com/1996-1073/10/3/408> (accessed on 10 February 2021). [[CrossRef](#)]
27. Ali, S.; Kim, D.-H. Optimized power control methodology using genetic algorithm. *Wirel. Pers. Commun.* **2015**, *83*, 493–505. [[CrossRef](#)]
28. Fernández, M.; Zorita, A.; García-Escudero, L.; Duque, O.; Morínigo, D.; Riesco, M.; Muñoz, M. Cost optimization of electrical contracted capacity for large customers. *Int. J. Electr. Power Energy Syst.* **2013**, *46*, 123–131. [[CrossRef](#)]
29. Nafkha, R.; Gajowniczek, K.; Ząbkowski, T. Do customers choose proper tariff? Empirical analysis based on polish data using unsupervised techniques. *Energies* **2018**, *11*, 514. [[CrossRef](#)]
30. Tostado-Véliz, M.; Arévalo, P.; Jurado, F. A comprehensive electrical-gas-hydrogen microgrid model for energy management applications. *Energy Convers. Manag.* **2021**, *228*, 113726. [[CrossRef](#)]
31. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Networks* **1994**, *5*, 157–166. [[CrossRef](#)]
32. Gers, F.; Schmidhuber, U.; Cummins, F. Learning to forget: Continual prediction with lstm. In Proceedings of the 1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470), Edinburgh, UK, 7–10 September 1999.
33. Sak, H.; Senior, A.; Beaufays, F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. *arXiv* **2014**, arXiv:abs/1402.1128.
34. Gajowniczek, K.; Ząbkowski, T. Short term electricity forecasting based on user behavior from individual smart meter data. *J. Intell. Fuzzy Syst.* **2015**, *30*, 223–234. [[CrossRef](#)]
35. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.