DEEP LEARNING APPROACH TO FORECASTING ELECTRICITY PRICE FROM LOAD DATA

Vladimir BABUSHKIN, Gheorghe CĂPĂȚÂNĂ

State University of Moldova

The accurate forecasting of electricity price and load is essential for maintaining a stable interplay between demand and supply in the dynamic electricity market. In this work we propose a deep Convolutional Neural Network-based model for day-ahead electricity price forecasting from historical price/load data and predicted load values. The model was tested on the data for New York and New South Wales and demonstrated high prediction accuracy for both datasets.

Keywords: *Deep Learning, Machine Learning, Long-Short Term Memory Networks, Convolutional Neural Networks.*

APLICAREA INVĂȚĂRII PROFUNDE LA PROGNOZA PREȚULUI ELECTRICITĂȚII DIN DATE DE ÎNCĂRCARE

Previziunea exactă a prețului și încărcării energiei electrice este esențială pentru menținerea unei interacțiuni stabile între cerere și ofertă pe piața dinamică a energiei electrice. Articolul descrie un model profund bazat pe rețeaua neuronală convoluțională pentru prognozele viitoare ale prețului energiei electrice din datele istorice ale prețului / tarifelor și valorile prognozate ale tarifelor. Modelul a fost testat pe datele pentru New York și New South Wales și a demonstrat o precizie ridicată a predicțiilor pentru ambele seturi de date.

Cuvinte cheie: *învățare profundă, învățare automată, rețele de memorie pe termen scurt, rețele neuronale convoluționale***.**

Introduction

Accurate forecasting of the electrical energy demand is crucial for electricity production, since electrical energy belongs to those commodities that cannot be fully stored, and needs to be consumed as soon as it is produced. Inefficient production and distribution of electricity incurs additional costs. The accurate prediction of electrical energy demand allows not only to avoid these costs by scheduling the load production, but also to manipulate with electricity price to achieve desired levels of consumption and maximize the profit. Moreover, the knowledge of expected demand allows more efficient use of natural resources for industrial production of electrical energy [1].

In the traditional electricity market framework, electricity consumption was strongly aligned with the regular peak demand curve, and thus, was quite predictable. The power production was more resource-consuming and expensive, since it ignored the immediate consumption feedback. Introduction of smart grids on the electricity market resulted in more efficient load management, allowing the supplier to set the price according to the consumption profile and making power production more cost-effective and environmentally friendly. It facilitates the adjustment of hourly consumption patterns of the end-users to minimize expenditures, thus lowering the electrical load in the grid at peak hours. It also resulted in increased dynamics of the electricity market – the end-users became dependent on the price value at a particular time of the day. The knowledge of future prices is also important for traders, retailers and other entities of the market, since it allows them to optimize financial strategies. Therefore, the implementation of price prediction mechanisms is essential in the deregulated market settings.

The electrical load depends on other external factors, such as seasonal patterns, increasing with the drop in temperature. Traditional statistical, linear and non-linear models in most cases fail to deal with the high volatility in electricity price time series data. Artificial neural networks are known for learning complex non-linear dependences. Recently emerged Recurrent Neural Networks, in particular Long-Short Term Memory networks, are capable to extract long-term temporal dependences, thus providing a more suitable tool for electricity price and load prediction. Despite the growing popularity of Deep Learning, the application of its methods to time series prediction is still in its nascent state and thus requires comprehensive and thorough analysis.

In this work we present a novel CNN-based approach for day ahead electricity price prediction that takes as inputs the historical price and load data along with the predicted load to generate price forecast for the given day of interest. We validate the model on the electrical load and price data provided by the New York Independent System Operator [2] for a New York City district (N.Y.C.) and by the Australian Energy Market Operator (AEMO) [3] for New South Wales. The proposed network can be deployed for both online and offline forecasting of electricity prices.

To our best knowledge, CNN-based mapping between price and load has not been addressed in the literature. An attempt to predict electricity price by mapping from price and load historical data was reported in [4]. However, the difference of our approach and the one in [4] is that we are using deep CNN instead of backpropagation network of two hidden layers. Our algorithm takes fewer parameters than the one proposed in [4] (5 vs 9). It is done intentionally to avoid over fitting and to allow the hidden layers of CNN to infer subtle dependences from the inputs.

1. Previous research

In the past the electricity suppliers were not market-oriented and the electricity price was set by governments. However, after the commercialization of the electricity supply and demand networks, the electricity price is determined by the newly emerged electricity market. In other words, the price is set according to the fluctuations in supply and demand on the electricity market, where electricity is regarded as a commodity. Compared to the load, the electricity price is more volatile, i.e. it can significantly change over the given period of time. Moreover, the load data is relatively homogenous and demonstrates cyclic variations, while the price data lacks such homogeneity and its variations follow quite a transient cyclic pattern [4]. In these volatile settings the knowledge of the electricity price beforehand is essential for decision making in electricity market. In fact, all entities of the market will benefit from the accurate price forecasts. The *Independent Power Producers* (IPPs) will be able to adjust the generation of power according to the expected price, reducing the cost of production. Market players will devise strategies for selling and buying the power on the market. The end-users will shift their most electricity consuming activities to the times where electricity prices are low, thus reducing the load in the grid during the peak hours. The problem of price forecasting relies on several factors such as correlations between power demand and supply, generation costs, market structure, influence of other market participants, etc. These factors are used to create models, reflecting causation between historical electricity prices and future prices. Knowing this causation allows to accurately predict electricity prices in the future [26].

Since electricity is not a storable commodity, the balance between power supply and demand can only be achieved a stable power system. Therefore, in most cases, electricity price is highly volatile and easily affected by such factors as sudden changes in weather conditions, transmission problems in the power system, outages of large power plants, participants' bidding strategies, transmission congestion, fuel prices, cost of unit operations to name a few. All these factors introduce additional complexities and challenges to the price forecasting task [26], [27].

The *Artificial Neural Networks* (ANN) have become widely popular since the second half of the 80's. The Artificial Neural Networks are capable of extracting significant features from the data, inferring non-linear relationships and learning more general concepts from experience. Depending on the neuron model, ANNs architecture and learning algorithm can be *Hybrid* (*e.g. Counter-propagation*), *Supervised* (*Brain State in a Box*, *Fuzzy Cognitive Map*, *Perceptron*, *Radial Basis Function*, *Support Vector Machine*), *Unsupervised* (*Adaptive Resonance Theory*, *Self-Organizing Maps*, *Principal Component Analysis*) and with *Reinforcement* (*Award/Punishment Associative*). In comparison with conventional techniques, ANNs can detect non-linear features even using a small dataset for training, thus resulting in more robust and accurate forecast. Future models for load time series prediction will either incorporate ANNs along with conventional algorithms into hybrid models or will be solely based on ANNs [5].

Despite the effectiveness of *Artificial Neural Networks*, for quite a long time they were not as popular as the kernels methods, such as *support vector regression* (SVR). The resurrection of Neural Networks occurred in 2006, with a seminal work by Geoffrey Hinton et al. [9] who demonstrated that the *Deep Learning Networks* can perform better than conventional *Machine Learning* methods. It boosted the applications of deep learning methods in different machine learning tasks, including computer vision, natural languages processing and other domains of *Engineering and Science*. In comparison with traditional shallow learning, deep learning networks incorporate multiple hidden layers and deploy stochastic optimization to learn more complex non-linear dependences [10]. Varying number of layers allow encoding different levels of abstraction which results in improved performance. As a recently emerged branch of machine learning, deep learning networks are applicable to a broad spectrum of complex real-world learning problems. For instance, deep learning networks' implementations can be found in image classification, object detection and recognition, speech recognition, language translation, voice synthesis and many more.

The first implementation of *Deep Learning Networks* for time-series forecasting is given in [11]. Authors deployed *Deep Believe Learning Networks* (DBN) for regression and prediction of time series data by combining the outputs from all DBNs with the help of *support vector regression* (SVR). The effectiveness of the proposed method was demonstrated on three independent datasets. In [12] authors deploy various simulations to show that recurrent networks can potentially outperform conventional nonlinear kernelized methods. They also explore how the performance of the network varies with its architecture (number of hidden layers and neurons per layer), demonstrating that additional complexity does not improve performance, and the optimal performance could be achieved with low number of layers.

In another work [13], authors developed a deep learning network with contrastive divergence for natural gas load forecasting. Another example of using *Deep Believe Learning Network* with restricted Boltzmann machine (RBM) for time-series forecasting can be found in [14]. This novel approach outperformed the traditional *Multilayer Perceptron* (MLP) neural network and statistical ARIMA model. One efficient application of *Conditional Restricted Boltzmann Machine* (CRBM) and *Factored Conditional Restricted Boltzmann Machine* (FCRBM) to an individual (single-meter) residential electricity consumption forecasting is reported in [15]. The method evaluated on benchmark dataset of individual customer electricity consumption data with temporal resolution of one minute. Authors come to conclusion that FCRBM outperforms ANN, *Support Vector Machine* (SVM), *Recurrent Neural Networks* (RNN) and CRBM [15].

In general, conventional ANN based methods are vulnerable to vanishing and exploding gradient problems that occur during the backpropagation stage in multilayer architectures. Also, traditional *Artificial Neural Networks* and conventional *Machine Learning Techniques*, such as SVM, fail to capture sequential information in time-series data [16]. To overcome these issues, Recurrent Neural Networks (RNNs) were developed as a framework for capturing the non-stationary patterns in time-series as well as long-term dependences. RNNs are capable of sharing the parameters across different time steps and thus, require less training, resulting in more effective computations. These features made RNNs popular for long and short-term forecasting of electricity load. For instance, in [17] RNN with a pooling layer has been proposed to forecast the household electricity load. It is able to mitigate the overfitting by pooling different customer's load profiles into one input, thus diversifying the data. The proposed pooling-based deep RNN demonstrated higher performance in comparison with traditional technics such as ARIMA, SVR and conventional deep RNN. In [16] authors developed a hybrid RNN model, that incorporates one step-ahead concept. First RNN was trained on electricity load time series that later were passed to the hybrid RNN model to learn non-linear dependences. The proposed model demonstrated a superior performance compared to the known methods.

However, deep RNNs still suffer of vanishing and exploding gradient problems that have deteriorative effect on their performance. *Long-Short Term Memory Networks* were initially introduced by Sepp Hochreiter and Jurgen Schmidhuber [18] as variation of RNNs. Shortly, they became a popular tool for time-series prediction due to their ability to deal with the vanishing and exploding gradient problems. LSTM networks are successfully applied for time-series forecasting in a variety of domains. For example, in [19] authors successfully adapted a combination of *Deep Believe Networks*, *AutoEncoder* and LSTM for renewable energy power forecasting and demonstrated that *Deep Learning* methods perform better than ANNs. In [20] additional features such as temperature, humidity and wind speed have been fed to LSTM network with electricity load time series data. The LSTM network trained on these data was used for forecasting electricity load both for short and long forecasting horizons, varying from 24 hours to 30 days. Authors show that the LSTM based forecast is more accurate than the prediction from conventional technics.

LSTM networks can be also scaled for the individual customer consumption (single-meter) forecasting. This type of forecasting is a very challenging task due to the volatile nature of individual loads influenced by a variety of external factors, such as climate, performance of thermal systems, and residency patterns. Marino et al. [21] were first to deploy LSTM networks to a single-meter load forecast. They used the same benchmark dataset as [15] and reported a similar performance as *Factored Conditional Restricted Boltzmann Machine* (FCRBM). Another work [22] deployed LSTM network for forecasting long-term electricity load time series. The authors deployed density-based clustering to infer the inconsistent nature of residential electricity load profiles and discuss challenges associated with predicting load in residential areas. They also demonstrate that LSTM networks are capable of capturing subtle patterns in individual load profiles which makes them superior in the majority of cases.

Another class of deep learning methods, the *Convolutional Neural Networks* (CNNs), have been successfully used for object recognition due to their ability to learn hidden dependences in raw data. CNNs can also capture local trend features and scale-invariant features for closely interrelated neighboring data points, e.g. local trend patterns in nearby hours. These circumstances make CNNs widely applicable for the electricity load prediction [23]. An example of load forecasting with CNN can be found in [24], where the authors demonstrate the superiority of CNN-based approach by reporting the achievement of the smallest forecasting error among the known methods. In the proposed CNN architecture three consecutive convolutional layers were used for feature extraction. The in-between pooling layers decrease the dimension of feature map and extract the important features from the deep layers. The forecasting is done at the output layer, where the flattened values pass through fully connected structure between flatten and output layers to the inputs of sigmoid function. This algorithm allows to make prediction of electricity load for the next three days.

Amarasinghe K. et al. [25] explored the application of CNNs for electricity load forecasting on individual building level. The historical loads were passed through multiple convolutional layers and fed to the fully connected layers to perform regression and produce a prediction. As in the similar work on LSTM [21] the benchmark dataset for a single-residential customer data were used to evaluate the performance of CNN-based approach against ANN, SVM, LSTM and FCRBM. The presented CNN demonstrated comparable results to the FCRBM [15] and *Sequence to Sequence* LSTM [21] on the same dataset. All the examples above demonstrate viability of the CNNs for accurate electricity load forecasting.

Similar to the electricity load prediction, there are traditional statistical methods and recent machine-learning based approaches to electricity price forecasting. An example of statistical model based on *Autoregressive Integrated Moving Average* (ARIMA) combined with wavelet transformation is presented in [28]. The day-ahead forecast for the next 24 hours was obtained by applying inverse wavelet transform. The model has been tested on mainland Spain and California markets demonstrating quite reliable results. In [29] ARIMA has been also combined with the *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) and the wavelet transform was applied to the original time series to decompose them into two parts that can be predicted separately. To obtain the final forecast the inverse wavelet transform was applied to the prediction of both subseries. The performance of the proposed approach has been evaluated on the Spanish and PJM electricity markets and the results were compared with other forecasting methods.

The problem with conventional statistical frameworks is that they perform well on the linear data however, when it comes to non-linear time series, they demonstrate a significant decline in prediction accuracy. To deal with the extreme non-linearity of price time series data, Voronin et al [30] developed an iterative model, consisting of modules for normal price and price outliers (spikes) prediction. The module for a normal price prediction is a mixture of wavelet transform, linear ARIMA and nonlinear neural network models. The spikes were handled by compound classifiers. The combined forecasts from both normal price prediction and spike prediction modules result into the overall price prediction. The combined model outperformed other known methods for price prediction.

The ability of *Artificial Neural Networks* to learn non-linear dependences in time-series data was exploited by several authors to create successful models for electricity price prediction. In [31] *Radial Basis Function Neural Network* (RBFNN) is applied for prediction of short-term electricity load and price. The model took a day type (e.g. Sunday, Monday) as an optional parameter, and generally demonstrated better performance if the day type is supplied. Another work [4] relies on Artificial Neural Networks to extract the complex dependences between the electricity price, load and other factors. The model takes as inputs several features such as the day of the week, time slot of the day, forecasted demand, change in demand, price for one day ago, one week ago, two weeks ago, three weeks ago and a month ago. The forecast horizon can vary from hours to days and weeks. The model proved to be efficient for normal days, but demonstrated poor performance for days with spikes.

In [26] authors propose an electricity price forecasting system based on the combination of *Convolutional Neural Network* (CNN) and the *Long Short-Term Memory* (LSTM). The CNN network consists of two convolutional 1D layers, with batch normalization after the second convolutional layer. The forecasting is performed by LSTM unit with ReLU activation function. The hybrid system accepts the past 24 hours price data as its input and produces a forecast for the next hour. In comparison with the traditional machine learning approaches, the presented hybrid model demonstrates the best forecasting accuracy.

Application of deep learning framework for electricity price prediction has also been explored in [32], where authors consider four different deep learning models and compare their prediction accuracy. The deep learning approaches usually outperform conventional methods. Authors also demonstrate that machine learning approaches perform better than traditional statistical models. The conventional approaches fail to achieve high forecast accuracies for price time series data, particularly the moving average-based models. Also, the hybrid models are not superior over their simpler components. Thus, the preference should be given to simple non-linear models, e.g. deep learning networks with few neurons in each layer.

In general, deep learning methods demonstrate better performance on electricity price datasets due to their ability of learning complex non-linear dependences and patterns. Considering all the examples above we can make the following observations:

- *Long Short-Term Memory Networks* [18] are better suited for dealing with non-linear sequences and thus, provide better accuracy for electricity price prediction.
- A network with more than one layer is capable of modeling the same dependence as the single-layered network but with a smaller number of neurons that simplifies the architecture of the deep-learning network. It also improves generalization of network [32].

Therefore, deep learning approaches represent a very promising technology for electricity price prediction.

2. Problem definition

Formally, times series can be represented as sequence of vectors, depending on time t: $X(t), t = 0,1, \dots$. Components of the vector $X(t)$ are observable variables, such as temperature, speed of wind, stock price, electricity load, etc. Actually, $X(t)$ is a continuous function, however, in practice, we are dealing with discrete time intervals. $X(t)$ defined on such discrete time intervals is called *time sequences* or *time series*. In the case of continuous variable domains, such as temperature, the values are *sampled* at points through the chosen time interval (e.g. measuring the temperature once in an hour). The number of measured points at the given time interval is known as *sampling frequency*. It is an important parameter, since different sampling frequencies may affect the characteristics of the time series.

If vector $X(t)$ consists of one component only, the time series are called *univariate*. In the case of multicomponent vectors $X(t)$ we deal with *multivariate* time series. Analysis of multivariate time series focuses on identifying dependences between observable components of vector $X(t)$, for example, how the temperature affects the electricity price and utilizing these patters to predict one of the components.

Time series forecasting, i.e. predicting the future dynamics of the variable of interest, is the most popular technic in real life applications. Most of the problems involve prediction of future values of vector $X(t)$, e.g. predicting electrical energy demand to estimate the network load and schedule electricity production. Formally, the time series forecasting problem can be defined as follows [33]: given time series $X(t)$, $t = 0, 1, \dots$, find a function $f: \mathbb{R}^{k \times n + l} \to \mathbb{R}^k$, where $k = dim\big(X(t)\big)$, e.g. $k = 24$ if we are interested in predicting hourly load. n is the number of separate time series vectors to train, e.g. number of days to train the model. The ultimate goal is to obtain an estimate $\hat{X}(t+d)$ of the vector X, where d is *prediction lag* (usually, $d = 1$), given the values of $X(t)$ up to time t and also additional time-dependent exogenous features

$$\pi_i, i = 1, \dots, l: \hat{X}(t+d) = f(X(t),\ X(t-1), \dots, X(0), \pi_1, \dots,\ \pi_l)$$

In this work exogenous features $\pi_i, i = 1, \dots, l$ that can contain for example, the dynamics of population density in the district of interest, are neglected.

3. Convolutional neural networks (CNN)

Originally inspired by information processing in brain's visual cortex, the most successful of all deep learning approaches, *Convolutional Neural Networks* (CNNs) are comprised of three-dimensional layers, that are able to encode both the spatial and depth representations, depending on the number of features. This architecture reflects the organization of RGB color channels in input (*convolution*) layer and therefore, CNN is widely used for image classification. The hidden (*subsampling*) layers of CNN encode primitive shapes in lower layers, gradually progressing towards encoding more complex shapes, like loops, in higher layers. In the case of grayscale input the input layer has a depth of 1 but higher layers preserve their 3D structure. The convolution layer is termed after the *convolution operation* that relies on filter to map the activations from one layer to the next. During the convolution a 3D filter of weights is used and selection of spatial region in a layer, passed through the activation function (e.g. ReLU), defines the value of the hidden state of the next layer. In the next layer the activation preserves their spatial relationships from the previous layer. Since any activation in every hidden layer is a function of a tiny region in the previous layer, the convolutional network is sparsely connected. To compress the spatial residuals coming from previous layers, an additional subsampling layer is added, that averages the values of local regions of size 2×2 (*subsampling operation*). Traditionally, CNNs were designed for working with two-dimensional image data and they become very popular for image recognition, object localization and text processing [78].

4. DEEP CNN for electricity price prediction

The idea of multi-layer CNN for price prediction was inspired by the work of Singhal et al. [4], where authors consider a three-layer back propagation network taking as inputs time indices, forecasted demand and historical time-lagged actual price information to predict the future price value. In this work we use deep convolutional neural network with three 1D convolutional layers, three intermediary pooling layers and a dense layer as output layer (see Fig. 1). The choice of CNN is influenced by the fact that electricity price data and electrical load are correlated, even though the relationship is non-linear [34]. The price is affected by a variety of external factors, including fluctuation in end-users' consumption, seasonality, weather conditions and other social, economic and political factors and the dynamics of the market, which result in high volatility of the price data. Most of these factors are intrinsically encoded in load and historical price data. Since CNNs are capable of learning complex relationships, we anticipate that CNN with multilayer architecture will be able to learn an accurate mapping between load and price data. To our best knowledge, the deep learning CNN-based approach for predicting price from load values has not been addressed in the literature. The difference of our approach and the one proposed in [4] is that we are using deep CNN instead of backpropagation network of two hidden layers. Our algorithm takes fewer parameters than the one in [4] (5 vs 9) to avoid overfitting and to allow the hidden layers of CNN to infer subtle dependences from the inputs.

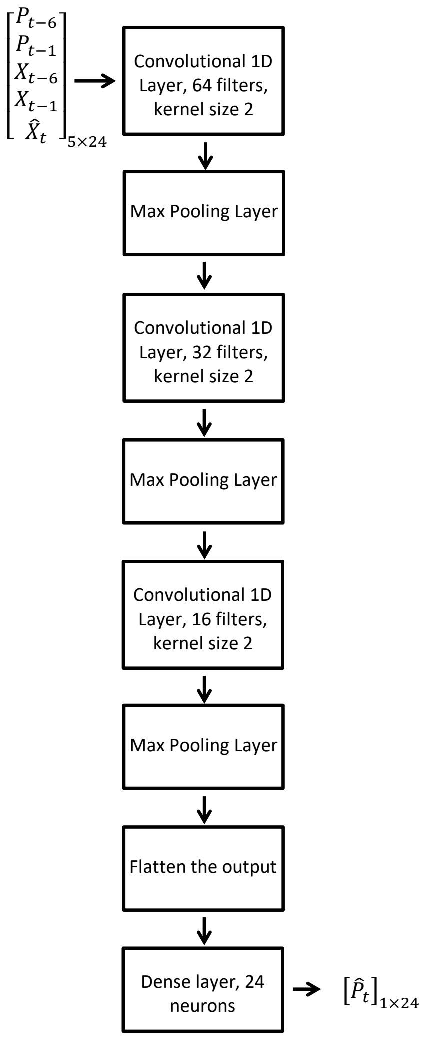

Fig. 1. Block diagram of proposed Deep Learning CNN for electricity price prediction.

Since there is no theory yet for determining an optimal number of layers and neurons in the deep learning network, we adopted a trial-and-error process for developing the architecture of multilayer deep CNN for mapping between load and price data. We assumed that the first two layers of the network would learn the general and daily trend from the load and historical price data. The third layer is responsible or capturing other external factors, such as seasonality, outages and the dynamicity of the market. Also, as it is pointed in [32], a network with more than one layers is capable of modeling the same dependence as the single-layered network, but with a smaller number of neurons. Therefore, we decided to adopt a simple multilayered architecture to avoid overfitting.

Experimenting with number of network layers demonstrated that three convolutional 1D layers for learning temporal dependencies and with three pooling layers for dimensionality reduction result in optimal performance of the model. The proposed network takes actual prices for a day before the day of interest P_{t-1} and the similar day a week ago P_{t-6} along with the actual loads for a day before the day of interest X_{t-1} and the similar day a week ago X_{t-6}. The fifth parameter is a predicted load for the day of interest $\hat{X}_t$. And the output is the predicted price $\hat{P}_t$ for the day of interest t. The network has been trained on a data for one year preceding the day t, to learn the dependencies between input $[P_{t-6}, P_{t-1}, X_{t-6}, X_{t-1}, \hat{X}_t]^T_{5\times24}$ and output vectors $[P_t]_{1\times24}$, where P_t are actual price values. The choice of load and price a week before the day of interest is justified by the presence of weekly trend in price and load variation patterns. Also, we consider load and price a day before the day of interest due to the similarity of load profile with the predicted day. The addition of more historical data as price/load two/three/four weeks ago does not produce a significant effect to the forecasting accuracy, but might result in network overfitting. We choose ReLU as activation for all layers, since it is more efficient for multilayer networks. The significant drop in loss of proposed model has been detected after 10-th epoch, however, due to the periodic spikes in validation loss we decided to increase the number of training epochs till 100. In overall, the proposed architecture demonstrated optimal performance on the price dataset for the New York City district (NYISO, [2]) and for New South Wales (AEMO) [3].

5. Results

Electricity price prediction is a difficult problem due to volatility, non-stationarity and non-homogenous structure of price time series data. In contrast to electrical load, electricity price data has a less pronounced trend (see Fig. 2 and Fig. 5). Also, price data is heavily affected by external factors. For example, price spike occurs if there is a transmission congestion detected in the system. After the spike, price usually reverts to a more reasonable level [4]. Regardless of volatile nature, electricity price is not a random variable. As it can be noticed from Fig. 2 and Fig. 5, there is a periodic structure and variations in daily trend, correlated with electrical load. Therefore, it is possible to learn hidden patterns and rules that determine price fluctuations from load and historical price data. Knowing these hidden dependences will result in more accurate electricity price forecast.

The non-linear correlation between electrical load and price is influenced by different factors. Load is mostly affected by the consumption patterns, weather conditions, social events (e.g. holidays), seasonality and the inability to store the electric power. In the case of electricity price, all the load's factors come into play along with other factors imposed by market, such as regulations of authorities,

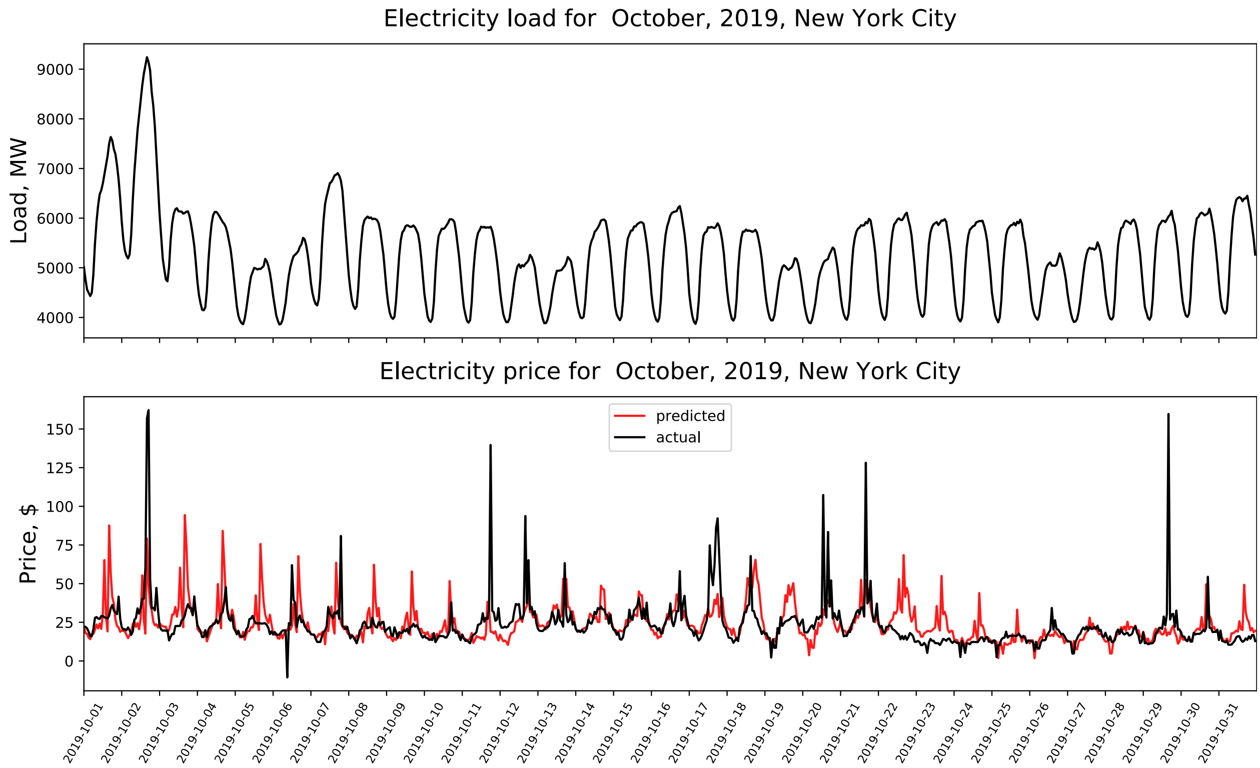

Fig. 2. Load and predicted/actual electricity price values for NYC in October, 2019.

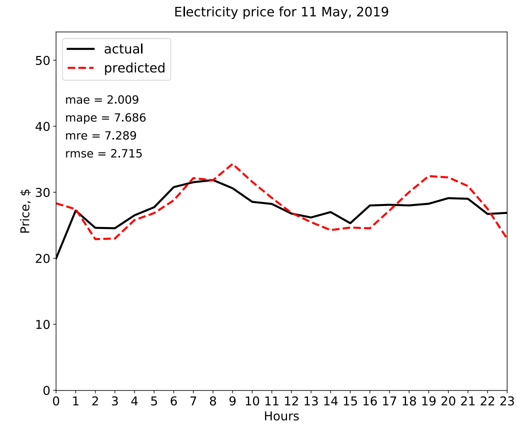

Fig. 3. Best price forecast for NYC, 2019.

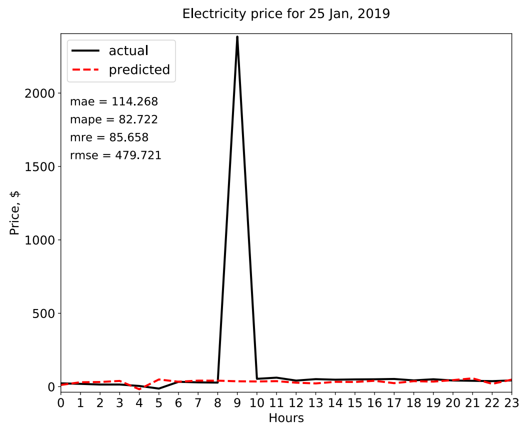

Fig. 4. Worst price forecast for NYC, 2019.

competitors' pricing schemes, and other macro- and micro-economic factors. All these factors contribute to high volatility of price data and need to be considered for more accurate mapping between load and prices.

We applied the proposed Convolutional Neural Network to establish the mapping between electricity load and price. As pricing data, we used Location-Based Marginal Pricing (LBMP) - cost to provide the next MW of Load at a Specific Location in the grid provided by the New York Independent System Operator [2] for the same location as load (N.Y.C.).

The validation of CNN-based algorithm for electricity price prediction has been conducted on the data for 2019. To train the model, the load and price data for the previous year was used. As a predicted load values, we used output from LSTM network, since it demonstrated relatively high average performance in comparison with the rest of the non-statistical models considered in this work. On the validation stage the inputs for each day were tested and after obtaining prediction the corresponding inputs were generated for predicted day's actual value and the whole network has been retrained. This procedure is known as online learning and it allows to generate predictions for the longer forecasting horizons (in this case we need to use the predicted price for the given day instead of the actual price, since we assume the actual value is unknown). In overall, the proposed algorithm was able to achieve the average MAE of 28.3 and average MRE 32.48 on NYC and average MAE of 22.35 and MRE 25.52 for NSW dataset for 2019.

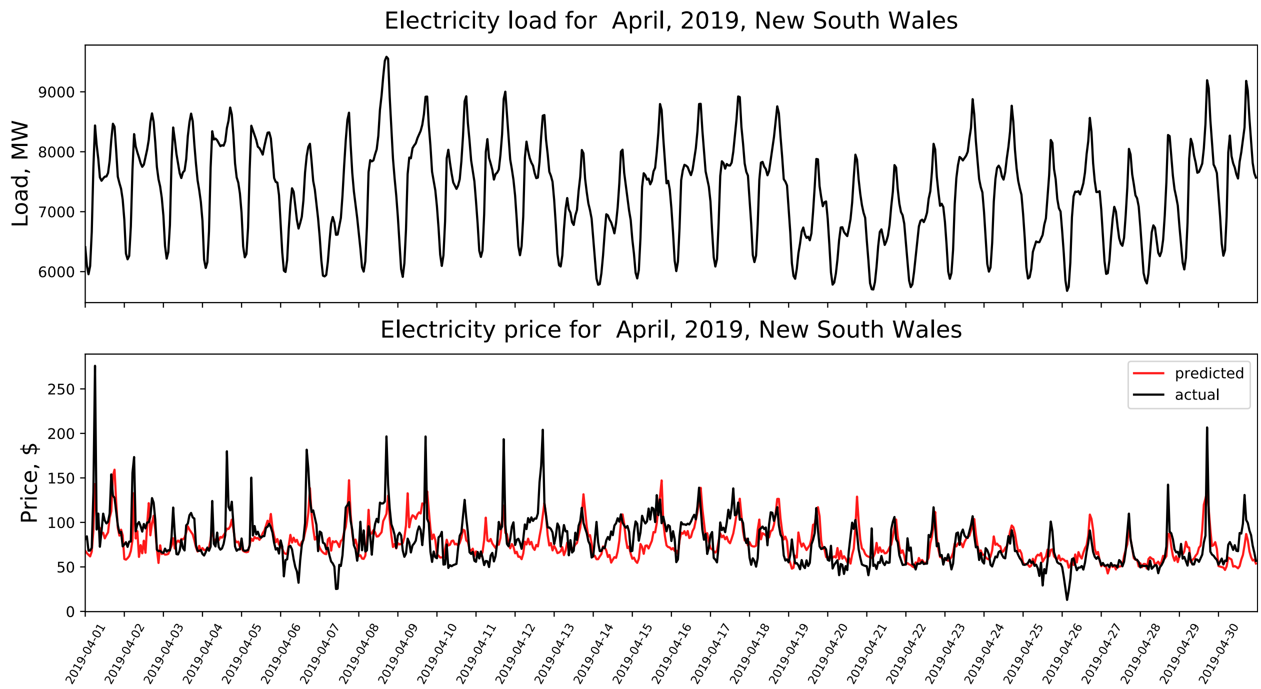

Fig. 5. Load and predicted/actual electricity price values for NSW in April, 2019.

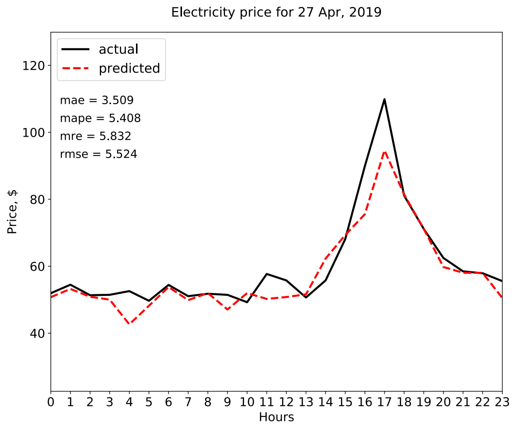

Fig. 6. Best price forecast for NSW, 2019.

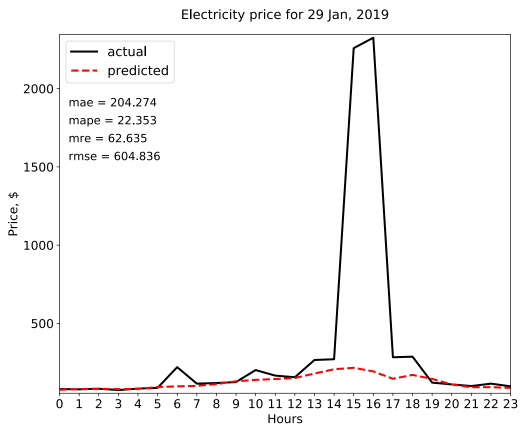

Fig. 7. Worst price forecast for NSW, 2019.

Fig. 3 and Fig. 4 show the best and worst predictions for the year of 2019 for NYC and Fig. 6 with Fig. 7 demonstrate the best and worst predictions for NSW.

As it is easy to notice, the model fails to perform well when it faces the spike in electricity price. In addition, the forecasting accuracy of the model significantly drops down after the spike. Therefore, the forecast for the next few days after the outburst cannot be considered 100% reliable. For instance, in Fig. 2 after the spike occurred on October 21, the model predicted small peaks for October 22 and 23 while the actual trend was almost flat. However, the peak for October 11 did not affect the predictions for October 12 and 13. Thus, the presence of a spike may indicate a need for revising the forecast for the next few days after the outburst.

Usually, spikes are the product of temporal discordance between electricity supply and demand. Even the discrepancies on the scale of seconds can induce significant spikes. The other factors resulting in spikes include power generation outages, congestion, behavior of market entities, manipulations on the market to name a few.

Significant volatility of price data creates challenges for accurate forecasting of the future prices. As a result, the price forecasting accuracy is usually lower than that of the electrical load forecasting. However, the requirements for price forecasting are not as stringent as for the electrical load and some degree of inaccuracy in prediction is acceptable.

Conclusion

In this study we present the CNN - based model for direct mapping between electricity load and price values. Since the price is very hard to predict due to its volatility, most of the conventional models are unable to capture all hidden dependencies resulting in unpredictable behavior of price values. The proposed CNN–based deep learning approach is capable of learning some of these dependences, which results in more-or-less adequate performance.

For future work it would be interesting to attempt learning the features inducing the peaks in price data. It can be accomplished by deploying wavelet transform to extract the deterministic and noise component and analyze them individually. It would also be interesting to explore the performance of more advanced deep learning techniques such as Generative Adversarial Networks (GANs) and Deep Reinforcement Learning for time series prediction, or hybrid models, incorporating CNN-s for feature learning and LSTM-s for prediction.

References:

[1] STOFT, S. Power system economics: designing markets for electricity, Piscataway, NJ: IEEE Press, 2002. Pages 1-46, ISBN 0-471-15040-1, Available: http://stoft.com/metaPage/lib/Stoft-2002-PSE-Ch-1-3,4,5,6.pdf [Accessed 27 07 2020].

[2] Energy Market & Operational Data, New York Independent System Operator, [Online]. Available: https://www.nyiso.com/energy-market-operational-data [Accessed 20 04 2020].

[3] Aggregated price and demand data, Australian Energy Market Operator, [Online]. Available: https://www.aemo.com.au/energy-systems/electricity/national-electricity-market-nem/data-nem/aggregated-data [Accessed 20 04 2020].
[4] SINGHAL, D., SWARUP, S. Electricity price forecasting using artificial neural networks, In: International Journal of Electrical Power & Energy Systems, vol. 33, pages 550-555, March 2011. ISSN: 0142-0615. DOI: 10.1016/j.ijepes.2010.12.009.
[5] HIPPERT, H.,S., PEDREIRA, C.,E., SOUZA, R., C. Neural Networks for Short-Term Load Forecasting: A review and Evaluation, In: IEEE Transactions on Power Systems, Vol. 16, No. 1, 2011, pp. 44-55. Print ISSN: 0885-8950. Electronic ISSN: 1558-0679. DOI: 10.1109/59.910780.
[6] EKONOMOU L., CHRISTODOULOU, C., A., MLADENOV, V. A short-term load forecasting method using artificial neural networks and wavelet analysis. In: International Journal of Power Systems, vol. 1, pp. 64-68, 2016. Available: http://www.iaras.org/iaras/journals/ijps [Accessed 27 07 2020]. ISSN: 2367-8976.
[7] REIS, A.J.R., A.P.A. da SILVA. Feature extraction via multiresolution analysis for short-term load forecasting. In: IEEE Transactions on Power Systems, Volume: 20, Issue: 1, Feb. 2005. Pages: 189–198, Available: https://ieeexplore.ieee.org/document/1388509 [Accessed 27 07 2020]. Print ISSN: 0885-8950. Electronic ISSN: 1558-0679. DOI: 10.1109/TPWRS.2004.840380.
[8] KOPRINSKA, I., RANA, M., TRONCOSO, A., MARTÍNEZ-ÁLVAREZ, F., Combining Pattern Sequence Similarity with neural networks for forecasting electricity demand time series. In: Proceedings of the International Joint Conference on Neural Networks, Dallas, Texas, USA, August 4-9, 2013, p. 940-947. ISBN 9781467361279.
[9] HINTON, G., E., OSINDERO, S, TEH, Y-W. A fast learning algorithm for deep belief nets. In: *Neural computation*, vol. 18, pp. 1527-54, 2006. ISSN: 0941-0643 (Print) 1433-3058 (Online). Available: http://www.cs.toronto.edu/~fritz/absps/ncfast.pdf [Accessed 27 07 2020].
[10] BOUKTIF, S., FIAZ, A., OUNI, A., SERHANI, M., A. Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches. In: Energies, June 2018,. 11(7), 1636. ISSN (electronic) 1996-1073. DOI: 10.3390/en11071636.
[11] QIU, X., ZHANG L., REN, Y., SUGANTHAN, P., N., AMARATUNDA G. Ensemble deep learning for regression and timeseries forecasting. In: IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL), December 2014. DOI: 10.1109/CIEL.2014.7015739
[12] Busseti, E., Osband, I., WONG, S. Deep learning for time series modeling, CS 229 Final Project Report, Stanford University, December 14th, 2012. Available: http://cs229.stanford.edu/proj2012/BussetiOsbandWong-DeepLearningForTimeSeriesModeling.pdf [Accessed 27 07 2020].
[13] MERKEL, G., D., POVINELLI, R.,J., BROWN, R., H. Short-Term Load Forecasting of Natural Gas with Deep Neural Network Regression. In: Energies, vol. 11, 2008. ISSN (electronic) 1996-1073. DOI: 10.3390/en11082008.
[14] KUREMOTO, T., KIMURA, S., KOBAYASHI, K., OBAYASHI, M. Time Series Forecasting Using Restricted Boltzmann Machine. In: Neurocomputing 137:47-56, August 2014. ISSN 0925-2312 DOI: 10.1016/j.neucom.2013.03.047.
[15] MOCANU, E., NGUYEN, H., P., GIBESCU, M., KLING, W., L. Deep Learning For Estimating Building Energy Consumption. In: Sustainable Energy, Grids and Networks, 6, 91–99. ISSN: 2352-4677. https://doi.org/10.1016/j.segan.2016.02.005.
[16] WEI, L.-Y., CHIH-HUNG, T., CHUNG, Y.-C., KUO-HSIUNG, L., Chueh H.-E., C., Lin J.-S. A Study of the Hybrid Recurrent Neural Network Model for Electricity Loads Forecasting. In: International Journal of Academic Research in Accounting, Finance and Management Sciences,

vol. 7, no. 2, April 2017, pp. 21-29. E-ISSN: 2225-8329, P-ISSN: 2308-0337. DOI: 10.6007/IJARAFMS/v7-i2/2786.
[17] SHI, H., XU, M., Li, R. Deep Learning for Household Load Forecasting - A Novel Pooling Deep RNN. In: IEEE Transactions on Smart Grid, vol. 9, Issue: 5 Sept. 2018, Pages: 5271-5280. Print ISSN: 1949-3053. Electronic ISSN: 1949-3061DOI: 10.1109/TSG.2017.2686012.
[18] HOCHREITER, S., SCHMIDHUBER, J. Long-short term memory, In: *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, November 1997. Print ISSN 0941-0643. Electronic ISSN 1433-3058. https://doi.org/10.1162/neco.1997.9.8.1735 .
[19] GENSLER, A. HENZE, J., SICK, B., RAABE, N. Deep Learning for solar power forecasting - An approach using AutoEncoder and LSTM Neural Networks. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), October 2016. DOI: 10.1109/SMC.2016.7844673.
[20] SHAHZAD, M., AFSHIN, A. Short-Term Load Forecasts Using LSTM Networks. In: ScienceDirect, 10th International Conference on Applied Energy (ICAE2018), 22-25 August 2018, Hong Kong, China. Energy Procedia, 158 (2019) 2922-2927, vol. 158, pages 2922-2927, 2019. Available online at www.sciencedirect.com
[21] MARINO, D., L., AMARASINGHE, K., MANIC, M. Building energy load forecasting using Deep Neural Networks. In: IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, 2016, pages 1-6. Electronic ISBN: 978-1-5090-3474-1. Print on Demand(PoD) ISBN: 978-1-5090-3475-8DOI: 10.1109/IECON.2016.7793413.
[22] ZHENG, J., XU, C., ZNANG Z., LI, X. Electric Load Forecasting in Smart Grids Using Long-Short-Term-Memory based Recurrent Neural Network. In: 2017 51st Annual Conference on Information Sciences and Systems (CISS). 22-24 March 2017, Baltimore, MD, USA, Electronic ISBN: 978-1-5090-4780-2. Print on Demand(PoD) ISBN: 978-1-5090-2697-5. DOI: 10.1109/CISS.2017.7926112.
[23] DU, S., LI, T., GONG, X., Yu Y., HORNG, S.-J. Traffic flow forecasting based on hybrid deep learning framework. In: 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 2017. ISBN: 9781538618301.
[24] KUO, P.-H., HUANG, C.-J. A High Precision Artificial Neural Networks Model for Short-Term Energy Load Forecasting. In: Energies, vol. 11(1): 213, January 2018, pages 1-13. ISSN (electronic) 1996-1073. DOI: 10.3390/en11010213
[25] AMARASINGHE K., A., MARINO, D. L., MILOS, M. Deep neural networks for energy load forecasting. In: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE 2017), 19-21 June 2017, Edinburgh, Scotland, United Kingdom, pages 1-6, ISBN: 9781509014132. Available: http://www.people.vcu.edu/~mmanic/papers/2017/ISIE2017_AmarMarinoManic_DLEnergyPrediction.pdf [Accessed 27 07 2020].
[26] KUO, P.-H., HUANG, C.-J. An Electricity Price Forecasting Model by Hybrid Structured Deep Neural Networks. In: Sustainability, vol. 10(4), April 2018, pages 1-17. ISSN 2071-1050. https://doi.org/10.3390/su10041280
[27] HU, L., TAYLOR, G., WAN, H.-B., IRVING, M. A review of short-term electricity price forecasting techniques in deregulated electricity markets. In: 44th International Universities Power Engineering Conference (UPEC 2009), 1-4 September 2009, Glasgow, United Kingdom, pages 1-6. ISBN: 9780947649456.
[28] CONTRERAS, J., ESPINOLA, R., NOGALES, F., G., CONEJO, A., J. ARIMA models to predict next-day electricity prices. In: IEEE Transactions on Power Systems. Volume: 18 , Issue: 3, Aug. 2003, pages: 1014-1020. ISSN 08858950, DOI: 10.1109/TPWRS.2002.804943.
[29] TAN, Z., ZHANG, J., WANG, J., XU, J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. In: Applied Energy, vol. 87, no. 11, pages: 3606-3610, 2010. ISSN: 0306-2619, https://doi.org/10.1016/j.apenergy.2010.05.012

[30] VORONIN, S., PARTANEN, J. Price forecasting in the day-ahead energy market by an iterative method with separate normal price and price spike frameworks. In: Energies, vol. 6 (11), pages: 5897-5920, November 2013. ISSN (electronic) 1996-1073. DOI: 10.3390/en6115897.

[31] SINGH, N., K., TRIPATHY, M., SINGH, A., K. A radial basis function neural network approach for multi-hour short term load-price forecasting with type of day parameter. In: 6th International Conference on Industrial and Information Systems, 16-19 August 2011, pages 316-321. Print ISSN: 2164-7011. DOI: 10.1109/ICIINFS.2011.6038087.

[32] JESUS, L., RIDDER F., D., SCHUTTER B., D. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. In: *Applied Energy*, vol. 221, pages: 386-405, July 2018. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2018.02.069.

[33] DORFFNER, G. Neural Networks for Time Series Processing, vol. 6, Neural Network World, 1996, pp. 447-468. ISSN 23364335, 12100552.

[34] NEUPANE, B., WOON, W., L., AUNG, Z. Ensemble Prediction Model with Expert Selection for Electricity Price Forecasting. In: Energies, 10(1), January, 2017. ISSN (electronic) 1996-1073. DOI: 10.3390/en10010077