The Dissertation Committee for Jian Xu
certifies that this is the approved version of the following dissertation:

Electricity Market Forecast using Machine Learning Approaches

Committee:

Ross Baldick, Supervisor

Hao Zhu

Nanshu Lu

Johnathan Bard

David Adelman

Electricity Market Forecast using Machine Learning Approaches

by

Jian Xu

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2019

Dedicated to my dear mother Jinhua Liu, my dear wife Yaguo Wang and my dear daughter Eilim Jin Xu.

Acknowledgments

I am so thankful for the opportunity Dr.Ross Baldick provided me to finish my PhD study in such a renowned electrical engineering program, otherwise my American Dream can not be fully completed.

I want to express thanks to the God who created my flesh and gave me life, although my flesh is sinful and weak, and although my life was sometime very difficult. In order to make me survive, He brought me a smart and strong wife, and sent many angels during those difficult situations I would otherwise had quit. I am most thankful for my daughter Eilim, which is the most precious gift I can get in my life.

I also want to give thanks to my dear committee. Everyone of you is a good example of modesty, diligence, tolerance and serving. Everyone of the committee is outstanding in his/her field and talented in teaching. It is my great honor to have Dr.John Goodenough, the Nobel Price laureate for the development of lithium-ion batteries, in my committee at his 97, even though he cannot attend my final defense because he is going to attend the Nobel Price Ceremony.

Thanks to Dr.Ross Baldick for every help during the last three and half years. Without your help I could accomplish nothing. Thanks to Dr.Hao Zhu for spending time discussing on my research problems and gave me many good guides and suggestions. Thanks to Dr.Adelman for many enlightening discussions which had brought up many good ideas. Thanks to Dr.Bard. You are just so considerate on everything. Thanks to Dr.Nanshu Lu for her kindness to join my committee as Dr.Goodenough had to go to the Nobel Price Ceremony. Thanks to my colleague Manuel Garcia for helping with the formatting of my dissertation and presentation.

Thanks to the brothers and sisters from my Bible study fellowship for spending time in the nights to help me practice my presentations. Thanks to all the known and unknown prayers that are holding me from failing in this wonderful journey. The journey is a witness to the amazing Grace.

May God bless you all!

Electricity Market Forecast using Machine Learning Approaches

Jian Xu, Ph.D.
The University of Texas at Austin, 2019

Supervisor: Ross Baldick

Electricity generation and load should always be balanced to maintain a tightly regulated system frequency in the power grid. Electricity generation and load both depend on many factors, such as the weather, temperature, and wind. These characteristics make the dynamics of electricity price very different from that of any other commodities or financial assets. The electricity price can exhibit hourly, daily, and seasonal fluctuations, as well as abrupt unanticipated spikes. Almost all electricity market participants use wind/load/price forecasting tools in their daily operations to optimize their operation plans, and bidding and hedging strategies, in order to maximize the profits and avoid price risks. However, the unreliable and inaccurate predictions with current forecasting tools have caused many serious problems, which can cause system instabilities and result in extreme prices even in the absence of scarcity. This dissertation presents an implementation of state of the art machine learning approaches into the forecasting tools to improve the reliability and accuracy of electricity price prediction.

Most existing wholesale electricity markets consist of a Day-Ahead Market and a Real-Time Market that work together to ensure the adequacy of electricity generation capacity for the Real-Time operation to secure the reliability of the grid. The two markets have different purposes, with the Day-Ahead Market serving as preparation for and hedging against variation in the Real-Time Market. Also, the Day-Ahead Market uses hourly Day-Ahead forecasting information and the Real-Time Market uses most up-to-date Real-Time information when running calculations. So the forecasting strategies of Day-Ahead

and Real-Time Markets should be different as well. The dissertation has two parts. The first part focuses on Day-Ahead price forecasting and the second part focuses on Real-Time price forecasting.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Electricity price forecasting is not a new topic, but there are still many unexplored areas. The following questions are posed in this dissertation: How to improve the forecasting accuracy based on current methods and resources? How to reduce the forecasting time so it can be used in Real-Time operation? How to forecast the peak-electricity-price more accurately? As perspective on these questions, Weron in his review paper [74] summarized all the popular electricity price forecasting methods, compared the performances and listed challenges and opportunities in electricity price forecasting research.

As observed from the review paper by Weron [74], most previous papers on the topic were focusing on improving the forecasting accuracy with a new algorithm or a new technology [45–48, 57, 61, 69, 74, 85, 88]. The typical issues for the researches on algorithm improvement is that the new algorithm is tested using very limited data, and the algorithm may only work for certain markets and particular time windows. For a new technology like machine learning, it can bring significant accuracy improvement at the very beginning, but it is difficult for the researcher to make innovations further because machine learning is like a black box which is not as easy to interpret as mathematical algorithms.

The reduction of forecasting time mainly relies on technology improvement, such as higher computing power and new technology based on improved computing. Simplifying the forecasting algorithm in order to reduce forecasting time is usually a trade off with worsening the forecasting accuracy.

Peak-electricity-price becomes more and more critical in electricity wholesale markets today, as the increasing amount of renewable energy is causing more fluctuations of the electricity prices during the peak-hours. During August 2019 the Electricity Reliability Council of Texas (ERCOT) set a new load record and average peak-hour electricity price was over $1,000/MWh for several days. Large generators can make extra million dollar profit a day if they were self-scheduled during those days. A reliable peak-electricity-price forecasting tool is very meaningful for conventional generation owners and the market, but there are many challenges for peak-electricity-price forecasting. The underlying reasons for peak-electricity-prices can vary from peak load to wind fluctuation to system congestions to outages to weather condition, and can be due to human errors too. Breakthrough of the forecasting methods/strategies are needed to address the complex causes.

An important contributor to high wholesale electricity prices is high system load. During the week of August 12, 2019, ERCOT grid hit a new load record of 74,531 MW. Level 1 of Energy Emergency Alert (EEA 1) was issued twice by ERCOT.

On their website (www.ercot.com) ERCOT describes EEA 1 as a signal that the grid is in a critical situation with risks of rotating outages. The ERCOT website describes the market condition under EEA 1 as " when operating reserves drop below 2,300 MW and are not expected to recover within 30 minutes, ERCOT can call on all available power supplies, including power from other grids, if available" [32].

Issuing an EEA 1 is a strong signal that the market is running in a very rare scarcity condition. Real-Time wholesale prices of all ERCOT regions were over $1,000/MWh for most of the peak hours during that week of August 12, 2019, and the price hit and stayed at the price cap of $9,000/MWh for many hours. Figure 1.1 is a screenshot of ERCOT Locational Marginal Price (LMP) Map of August 12, 2019, on which the 74,531 MW new load record was set.

During that peak week ERCOT market participants with generation and load could plan to self-arrange the generation and avoid the risk to buy electricity to cover the load in the Real-Time Market. Conversely, market participants with generation and load that did not plan well could have lost the opportunity to receive the high Real-Time price for generation and may even have paid a huge amount of money to cover the load. A good short-term price forecasting tool with sensitivity to peak-electricity-prices can be extremely helpful for those circumstances, especially when the price spikes become normal behaviors in the markets nowadays as the renewable generation percentage keeps growing.

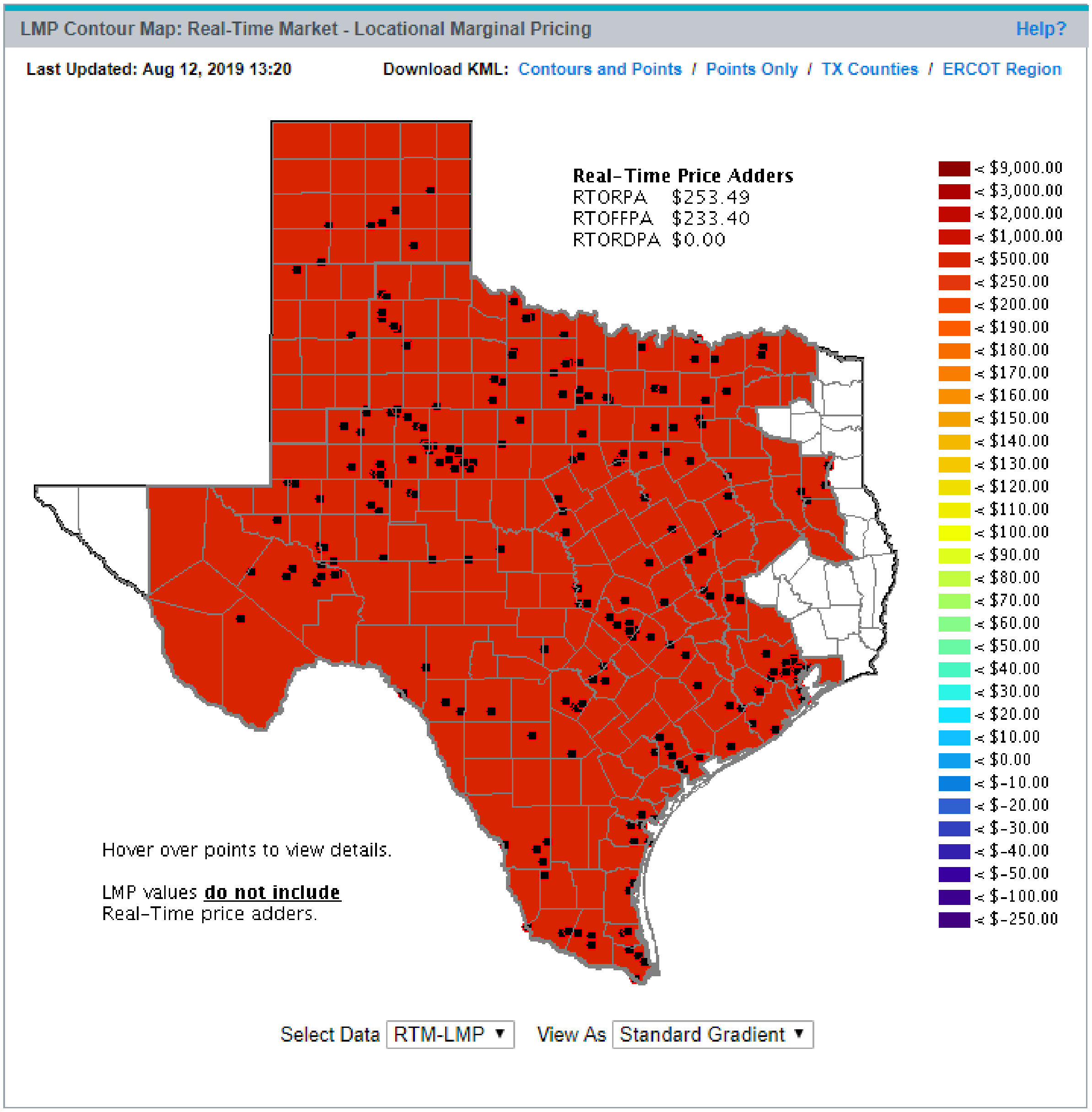

Figure 1.1: ERCOT LMP Map on August 12, 2019. Source: www.ercot.com

Having established the significance of peak-electricity-price and peak-electricity-prices forecasting, further details will be discussed in the rest of this chapter, which has seven sections. Section 1.1 gives a brief introduction of electricity price forecasting, including the usage of electricity forecasting and the main approaches to conduct electricity forecasting. Section 1.2 gives an overview of the ERCOT electricity market, especially focusing on ERCOT Nodal

Market implementations. Section 1.3 lists the possible benefits of electricity forecasting to both ERCOT and its market participants. Section 1.4 introduces state-of-the-art machine learning neural networks, and comparison between neural networks. Section 1.5 discusses the specific problem of peak price forecasting. Section 1.6 discusses how price forecasts can benefit the Real-Time Co-optimization (RTC) of wholesale markets. The last section is a literature survey on the history of electricity price forecasting, the popular methods, and how machine learning methods are improving the forecasts.

1.1 Electricity Price Forecasting

Electricity price forecasting tools and technologies are used by market participants to help optimize their market operations. In the longer term, bilateral contracts are priced based on forecasts of future Day-Ahead and Real-Time Market prices [48]. Major electricity consumers can minimize wholesale purchase costs by operating during low-price hours or periods.

The accuracy of electricity price forecasting is very important [83]. Hong [40] estimates that a 1% improvement of the short-term forecasting accuracy can result in about $0.5 million savings per year for a utility company with 1 GW peak load. For such a utility company, if the load factor is 50% and average production cost is $30/MWh, this comes to a total cost of about $0.1 billion per year, and the saving is about 0.5% of the cost. Although the saving percentage might seem to be a small amount, this is nevertheless a significant ongoing savings due to just improved forecasts and could have a significant effect on profitability.

However, due to the special characteristics of electricity such as not being storable, the necessity of balancing generation and demand to control system frequency all the time, and weather and wind dependency, the behavior of the price of electricity is different from that of any other commodity or financial asset. The forecasting of the electricity prices is very challenging due to these idiosyncratic characteristics.

Many modeling and statistical methods have been proposed during the last few decades, but it is very difficult to build a model with good prediction accuracy that can cover the characteristics of the whole system. Electricity price forecasting using classical models and statistical approaches became an outdated and saturated topic in the last few years. This is a signal that it is difficult to make further improvements using the classical approaches. However, newly developed approaches like Artificial Intelligence (AI) might make substantial improvements on electricity price forecasting.

AI can solve very complicated classification and regression problems as the technology keeps developing and computing power becomes cheaper. Machine learning is one of the most popular AI approaches and has been widely applied in many areas. Neural network-based machine learning methods are very promising in computer vision, speech recognition, and natural language processing [39]. A neural network loosely simulates how the human brain works in learning and making decisions. By training the neural networks, each neuron in the network can remember the weights of inputs and outputs.

1.2 ERCOT Nodal Market

ERCOT is the Independent System Operator (ISO) of Texas operating the electricity grid and managing the wholesale electricity market for most of the Texas region. ERCOT is one of the 9 members of the ISO/RTO Council (IRC) [7] across North America. Figure 1.2 shows the ISOs/RTOs in the US and Canada.

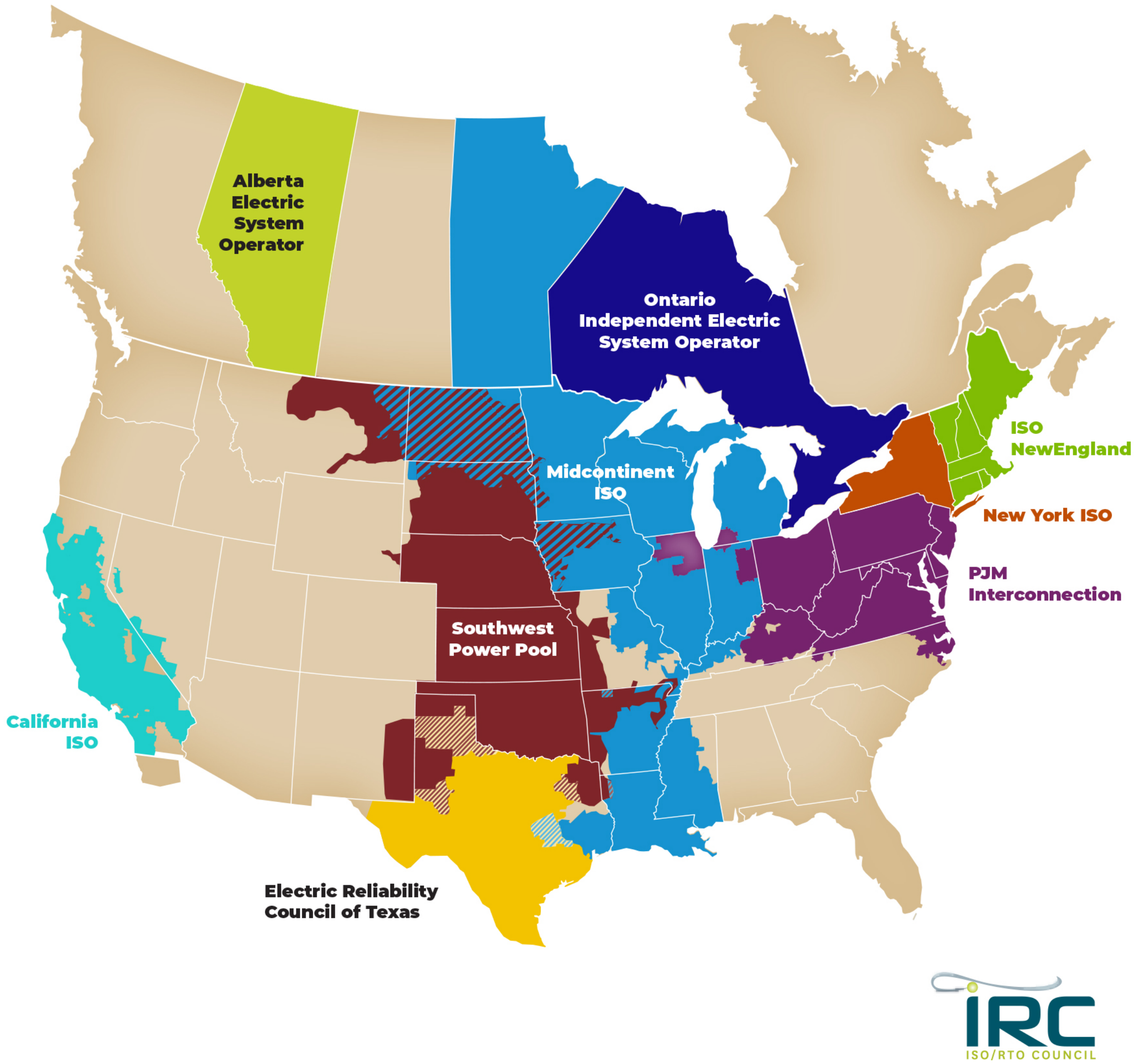

Figure 1.2: ISO/RTOs Operating Map Source: https://isorto.org

As stated in ERCOT Quick Facts [12], "ERCOT manages the flow of electricity to more than 25 million Texas customers — representing about 90 percent of the state's electric load. As the independent system operator for the region, ERCOT schedules power on an electric grid that connects more than 46,500 miles of transmission lines and 600+ generation units. It also performs financial settlement for the competitive wholesale bulk-power market and administers retail switching for seven million premises in competitive choice areas."

Also according to ERCOT Quick Facts dated July 31, 2019 [12], ERCOT has 22,051 MW wind generation capacity as of January 21, 2019, the most among any States in the United States. The wind generation set a record of 19,672 MW on January 21, 2019. ERCOT also has installed 1,858 MW of utility-scale solar capacity as of June 2019, and it is anticipated that considerably more will be installed.

In December of 2010, ERCOT successfully upgraded from the previous Zonal Market to the LMP based Nodal Market. The ERCOT Nodal Market architecture includes a Day-Ahead Market (DAM), a Real-Time Market (RTM), and also a Congestion Revenue Right (CRR) auction market. CRR owners will get charged or paid in the Day-Ahead or Real-Time Market based on the difference in LMPs.

The ERCOT Day-Ahead Market is a forward financial market cleared before the operating day, with the main purposes to schedule energy and ancillary services, to facilitate generator commitment decisions, and to provide price references for the next operating day. The Day-Ahead Market clearing process co-optimizes the energy offers and bids from market participants, ancillary services and CRRs to maximize system wide economic surplus. The market participants with generators can submit start-up cost, minimum energy cost, and offer curve above minimum energy (Three-Part-Offer) to ERCOT for the Day-Ahead Market and ERCOT will consider the choice to commit the generator in its optimization. Alternatively, generators can submit energy-only offers if they intend to self-commit without direction from ERCOT. The Day-Ahead Market co-optimization engine will give these cleared results: unit commitments of resources with Three-part-offer submitted, the awards of energy offers and bids, awards of ancillary services and awards of CRR that are taken to Real-Time Market Settlement.

The ERCOT Nodal Market also implemented the Reliability Unit Commitment (RUC) and the Security Constrained Economic Dispatch (SCED) in the Real-Time Market. The SCED in the Real-Time Market reduced the Real-Time Market clearing interval to 5 minutes from 15 minutes in the previous Zonal Market. The following paragraphs describe RUC and SCED in the Real-Time Market.

RUC is a process to ensure sufficient generation capacity is committed to cover the forecasted ERCOT demand, and to monitor the transmission system security by performing the network security analysis [3]. RUC is performed daily and hourly to check that there is enough capacity for the Real-Time operation. The ERCOT Day-Ahead Market clearing process is based on the voluntary energy offers and bids instead of the load forecast, so the energy committed in the Day-Ahead Market may not be sufficient for the actual energy and ancillary service requirements in the Real-Time operation. The RUC process will check the shortage and procure enough generation capacity to meet load forecast plus enough ancillary service capacity.

SCED in the Real-Time Market dispatches generators based on their offer curves to match the total ERCOT demand while satisfying generator ramp-rate constraints and the transmission constraints during Real-Time operation. The SCED process produces the base point and LMP for each generator [8]. The base point is the instructed target dispatch level to be achieved at the end of the upcoming 5-minute dispatch interval.

1.3 Effect of Electricity Price Forecast on ERCOT Wholesale Market

For ERCOT market participants, no matter what kind of entities they represent, a good price forecasting tool will benefit them in daily planning and market operation. ERCOT market participants may make more informed bids and offers with the advanced forecasting tool. ERCOT publishes Load and Wind forecasts on the website (www.ercot.com) but it does not publish electricity price forecast. Consequently, forecasting of price will be discussed under the assumption that load and wind forecasts are available as potential inputs to forecast prices.

A good price forecasting tool can also help market participants to optimize the ways they run the generators and cover the loads. For market participants that only represent generation and not load, they can have many strategies for Day-Ahead Market trading. They can choose to offer or not to offer the generators into the ERCOT Day-Ahead Market. In the case of generators, if the price forecasting tool predicts that the Day-Ahead Market average price is higher than Real-Time Market average price, and the Day-Ahead price is above the operation cost, then they will be more confident to offer as much as possible into the Day-Ahead Market. However, if the forecasting tool predicts that the Real-Time Market average price is higher, and the additional profit in the Real-Time Market compared with the Day-Ahead Market is more than (start-up + minimum energy) cost of the generator(s), then the market participants may choose to self-schedule the generator(s) in the Real-Time Market.

The goal of operation should not only be to cover the cost, but also to make maximum profit. As the renewable resources are getting tax credit, and extra carbon dioxide tax may be charged in the future, it will be more difficult for the fossil fuel resources to recover their investments. The optimization of running non-renewable resources becomes more and more important, as the only advantage that non-renewable resources have is that they can be dispatched upon request. For the quick-start units like gas turbine resources, they can come online and reach the upper limit within half an hour or even less time. In reality they can be planned to start up one hour ahead of the hour at which the high Real-Time price is forecasted.

Large companies typically have many types of generators, such as coal, gas, nuclear, quick-start gas turbine, hydro, wind, solar, etc. and they may also spread across one State or even multiple States. With thermal generators, fuel prices play a role. Should the gas or coal resources be used if only a few generators are needed and the rest can be offline for maintenance? If a monthly price forecast is available, maintenance for different units can be arranged more economically.

If a market participant plans to sell ancillary services in the Day-Ahead Market, they need to carry awarded ancillary services on the resources during the Real-Time operation under current ERCOT market structure. (This will change after implementation of RTC.) Under the ERCOT market design, the company has the flexibility to rearrange ancillary services amounts on any qualified generators. If they have an accurate price forecasting tool, they can let expensive resources carry more ancillary services so the cheap resources can be used to generate more energy, which can make plan for each hour more optimized and economical.

If the market participant is also a CRR account holder, they can hedge the price differences between the nodes to reduce the risk caused by transmission congestions. The CRR Market is a competitive market; however, the ability to forecast prices varies between market participants. A good price forecasting tool available to all the market participants will help them to bid and offer CRR at more reasonable prices, which will help optimize the CRR Market.

For market participants that represent load, if they can predict prices of both Day-Ahead and Real-Time Markets, then they can decide whether to purchase energy to cover load in the Day-Ahead Market or be self-scheduled in the Real-Time Market.

Another option is bilateral trading with other market participants in the market. A bilateral trade involves an agreement between two parties in the market to trade electricity or ancillary services at fixed prices. If one can predict the prices ahead, better deals can be reached for the company.

Electricity price forecast is a dynamic process. When market participants all generate at an accurately forecasted Peak-Electricity-Price hour, the Peak-Electricity-Price may disappear quickly due to the abundant generation. However, even the successful forecast of a $9000/MWh price for a 5-minute short period can bring big benefit for big capacity generators. the forecast should continuously be updated by taking inputs from most recent market conditions and track changes in the marketplace such as the growth of renewables and storage assets.

The development of solar and battery storage projects could help reduce the amount of Peak-Electricity-Prices in the market, but it will take a while based on the current development speed. As the load also keeps increasing in the ERCOT region and older generation resources retire, the price fluctuation will be a long-term issue for the ERCOT Market.

1.4 State-of-the-Art Machine Learning Approaches

Machine learning approaches have been widely used to solve problems that are difficult to model from first principles. Using a classical modeling method, researchers may have to write thousands of lines of code but still cannot cover the whole problem. Inspired by how the human brain works, neural network algorithm has brought a revolution to machine learning. Through training a neural network-based machine learning model using input data and expected output data, each neuron in the machine learning model will remember a proper weight, which will generate a hidden algorithm to solve the problem. By using non-linear activation functions and controlling gates, neural networks have strong non-linear modeling capability. Neural network models have been used to solve many difficult problems from image recognition to sound recognition to all types of forecasting. Machine learning has changed the world, but will bring much bigger revolution to the world as it keeps developing and improving.

For this proposed research, neural network-based machine learning approaches will be used. There are several popular state-of-the-art Neural Networks, including deep neural network (DNN), convolutional neural network (CNN), recurrent neural network (RNN), long-short term memory network (LSTM), and gated recurrent units (GRU). In the following sections, brief introduction of the basic principles behind each approach will be given.

1.4.1 DNN

DNN is the earliest and simplest neural network, which is a combination of layers of neurons. By increasing neurons and layers, DNN can be trained to model linear and non-linear problems. The weight at each neuron can be adjusted when different patterns are detected in each training.

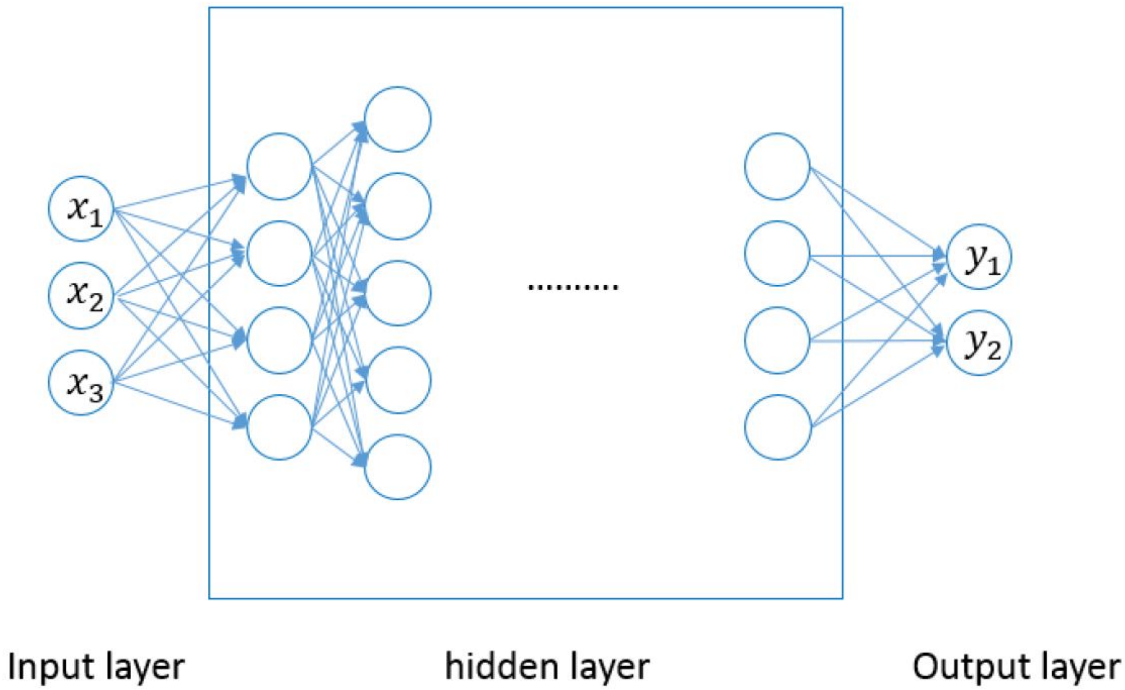

Figure 1.3: DNN structure

As shown in Figure 1.3, a DNN network has at least an input layer, a hidden layer, and an output layer. Based on the problem multiple layers can be added in the hidden layer, and each layer can have different neurons. The more neurons and layers added, the greater the ability to represent non-linearity. But complex networks having more neurons will take a long time to train.

1.4.2 CNN

CNN was invented for image recognition, but it can also be used in data prediction problems. The most common layer for CNN is the filter layer, which has a defined size scanner to scan the inputs into smaller groups. Then a max pooling layer can be added as needed to find the biggest signal of each group. Multiple filter layers can be added, and hidden layers before the output layer can be added as well to implement more complex study. CNN is very popular not only for image data processing, but also to summarize the important information for data inputs. Sometimes CNN-RNN combination is used to perform data deep learning.

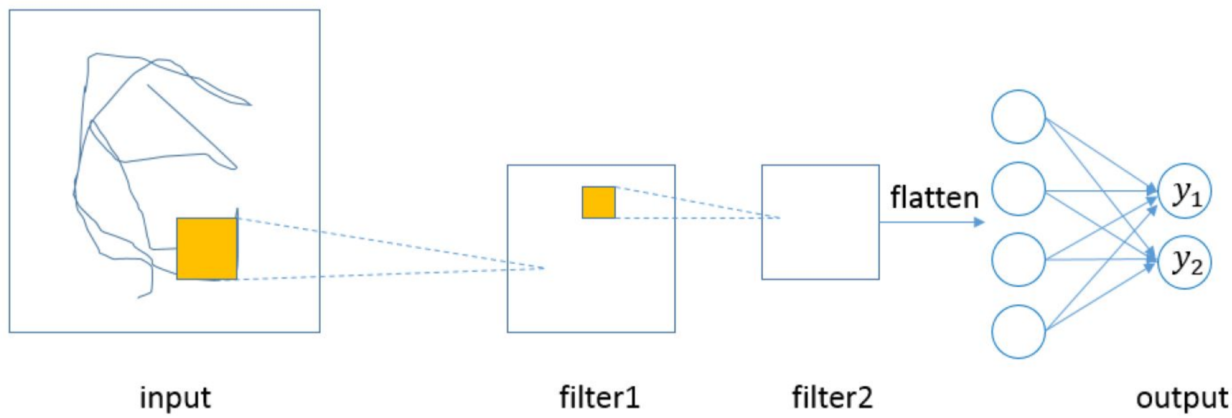

Figure 1.4: CNN structure

Figure 1.4 shows a simple CNN example. Filter1 scans the input data into summarized block data, then filter2 scans the summarized data in filter1. The multiple dimension data is flattened later and passed to the output layer.

1.4.3 RNN

RNN incorporates implementation of time sequence learning capability. Unlike DNN, which just conducts forward computation, RNN will feed previous output as an input to the current block. As a result, it can remember what happened before, and makes prediction what will happen using both current input and previous information as shown in Figure 1.5. RNN is very useful for problems that have time sequential events.

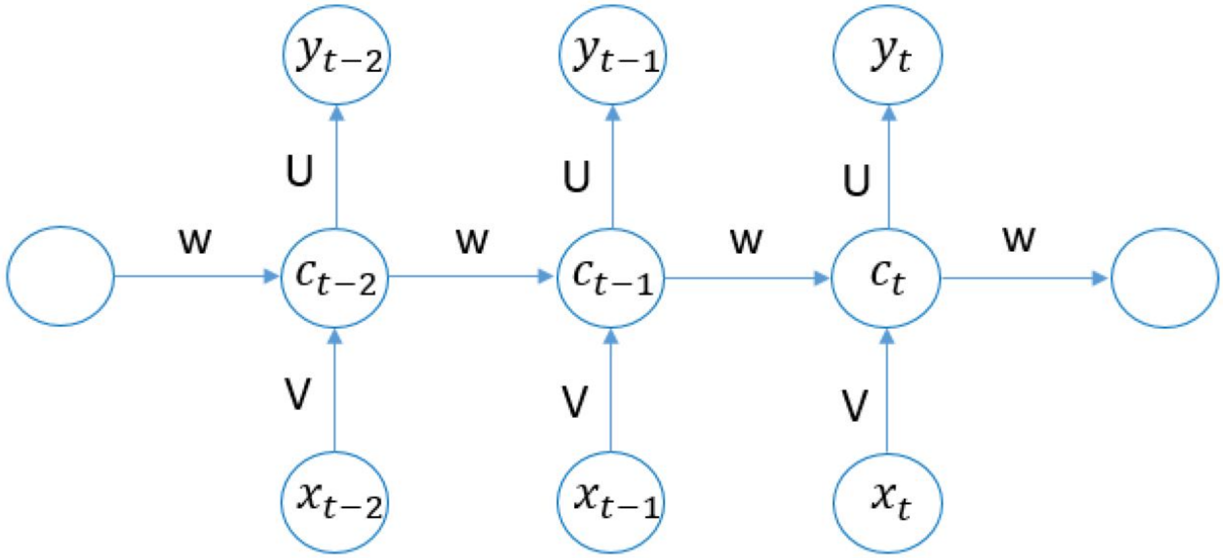

Figure 1.5: RNN structure

In Figure 1.5 w is the weight function of the sequence data passing from the previous step, V is the function of the input at each time step and U is the output function at each time step. Equations can be written as follows:

$$c_t = w \times c_{t-1} + V(x_t) \tag{1.1}$$

$$y_t = U(c_t) \tag{1.2}$$

RNN is capable of representing many time sequence problems, but for some cases it is not satisfactory because of the so-called vanishing and exploding gradient issue to be explained below. From (1.1) and (1.2), the output can be calculated:

$$\begin{aligned} y_t &= U(w \times c_{t-1} + V(x_t)) = U(w \times (w \times c_{t-2} + V(x_{T-1})) + V(x_t)) \\ &= U(w \times (w \times (w \times c_{t-3} + V(x_{t-2})) + V(x_{T-1})) + V(x_t)) = ... \end{aligned} \tag{1.3}$$

For each training RNN will predict the output in the forward direction and then do backward propagation based on (1.3) to reset the weight for each layer in order to optimize the output value. The detail of neural network training will be explained in section 1.4.6. If, as typical, w is small at the beginning then over time it will create very small gradient, which is called vanishing gradient. If w is big then over time it will create huge gradient which is called exploding gradient. The gradient is used to reset the weight, so both vanishing and exploding gradients will lead to malfunctions of RNN.

1.4.4 LSTM

LSTM was invented to solve the vanishing and exploding gradient problem of RNN. Base on the idea of RNN, LSTM added three gates in the architecture: Input Gate, Forget Gate, and Output Gate, along with a LSTM cell, as shown in Figure 1.6. The Forget Gate can make the decision to forget the previous values which can prevent the vanishing and exploding gradient issue of the regular RNN from happening [2].

The idea of LSTM is demonstrated in Figure 1.6. In a LSTM unit, the Forget Gate F_t controls how much previous cell state c_{t-1} will be passed through to the new cell state; the Input Gate I_t controls how much the new information will be stored in the new cell state c_t; the Output Gate O_t controls the output of the unit. The 3-gates structure makes LSTM extremely powerful in studying time sequential data [2].

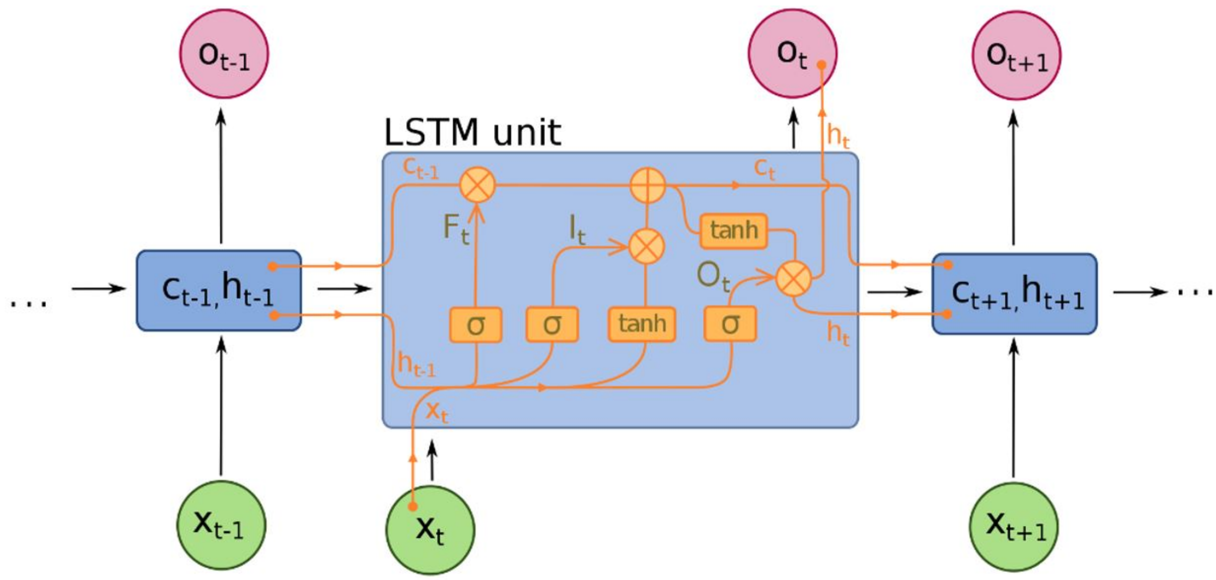

Figure 1.6: LSTM unit. Source: reproduced from figure "Long short-term memory unit" of [2].

Suppose that in the Forget Gate, weight of h_{t-1} is w_{fh} and weight of x_t is w_{fx}; In the Input Gate, weight of h_{t-1} is w_{ih} and weight of x_t is w_{ix}; In the Output Gate, weight of h_{t-1} is w_{oh} and weight of x_t is w_{ox}.

Also as shown in Figure 1.6, there is a *tanh* function to pass new information to the memory cell. For the input of the *tanh* function, the weight of h_{t-1} is w_{ch} and the weight of x_t is w_{cx}. σ in the figure means *sinh* function. The equations to calculate the output value are as follows:

$$F_t = sinh(w_{fh}h_{t-1} + w_{fx}x_t) \tag{1.4}$$

$$I_t = sinh(w_{ih}h_{t-1} + w_{ix}x_t) \tag{1.5}$$

$$c_t = F_t \times c_{t-1} + I_t \times tanh(w_{ch}h_{t-1} + w_{cx}x_t) \tag{1.6}$$

$$O_t = sinh(w_{oh}h_{t-1} + w_{ox}x_t) \tag{1.7}$$

$$y_t = O_t \times tanh(c_t) \tag{1.8}$$

In (1.6) $F_t \times c_{t-1}$ decides whether to forget the input from the previous memory cell value c_{t-1} and how much of the previous memory cell c_{t-1} can pass through.

By forgetting the previous input values, controlling current input values, and shaping output values, LSTM can solve the vanishing and exploding gradients issue of the regular RNN. LSTM now is the most popular neural network for sequential data study. LSTM has a very complicated architecture, but some simplified versions have been invented, including GRU, which is described in the next section.

1.4.5 GRU

GRU is a simplified version of LSTM, having fewer gates (Reset Gate and Update Gate) and less complicated architecture. Because of the simplified structure it can also reduce tensor calculations during training, which means faster training speed. As shown in Figure 1.7, compared to LSTM, the Reset Gate controls how much previous step information will be passed, and the Update Gate function is similar to the combination of the Input and Forget gates of LSTM [26].

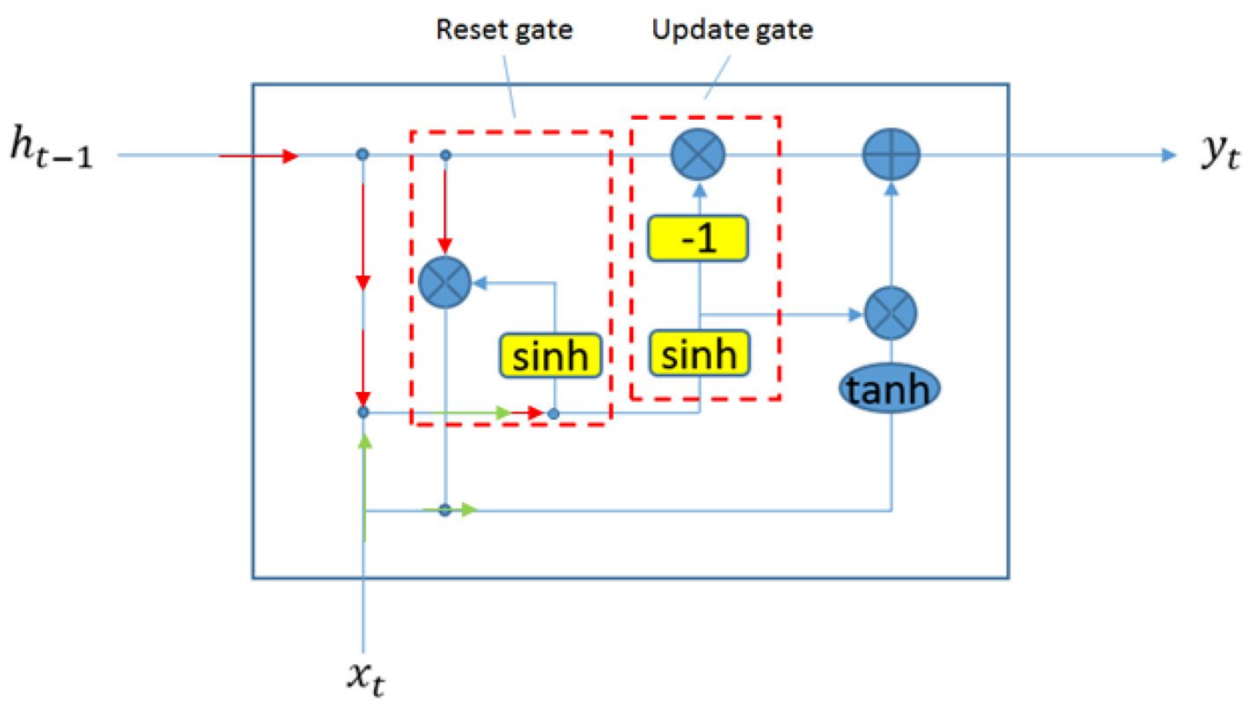

Figure 1.7: GRU structure

1.4.6 Neural Network Training

Neural networks provide easier and more flexible ways to formulate complex non-linear forecasting problems. By using the advanced neural networks, researchers and data scientists can avoid the development of large programs and complicated models, but can instead call the existing neural network packages in Python or R. They can spend more time in data mining, data processing, and data analysis. The goal of this dissertation is to use the state-of-the-art machine learning approaches to improve forecasts of electricity price, and to develop new forecasting methods based on neural networks to forecast peak-electricity-prices.

All the machine learning models have similar training process. After a machine learning model is designed, before being used to do forecasting work, it needs to be trained first, as shown in Figure 1.8.

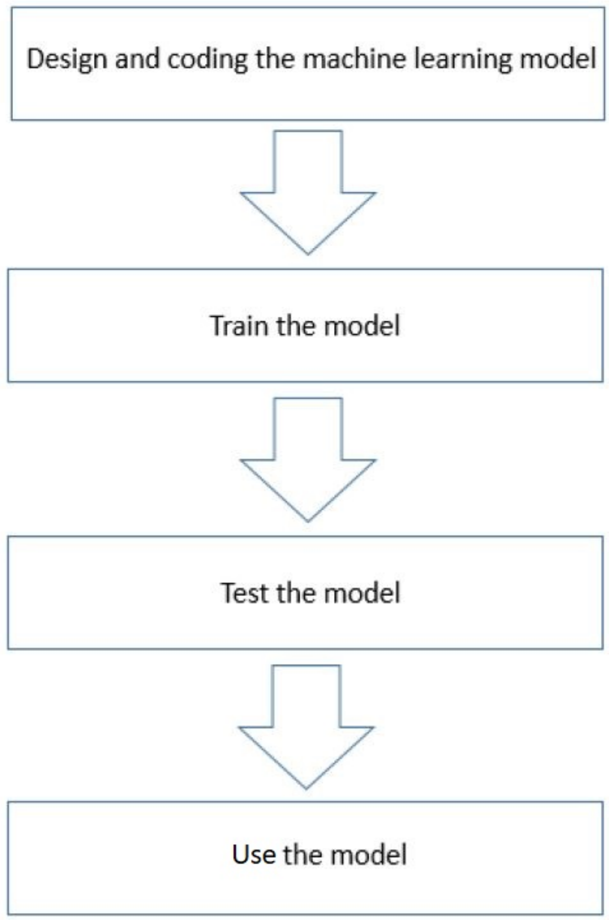

Figure 1.8: The flow of implementing a machine learning model

To train the model is to teach the model how to generate the right output based on the input. In the training process when the input data is fed to the model, it moves forward to the output layer, then the model will do backward propagation using the error calculated between the forecasted values and the actual values in the training data set, to adjust the weight value stored at each neuron.

During the training process, there is an option to give the model a validation data set to validate the training at each step. Validation is part of the training process.

The more epochs of training are given, the more adjustments will be made to the weight of each neuron. The ideal number of trainings depends on the architecture of the model and the quality of the training data. If the model does not have enough epochs of training, it will not forecast at its best; however, if the model is trained for too many epochs, then there will be an over fitting problem. It is helpful to observe the errors during the training. If the error keeps going down then more epochs of training can be given, but if the error increases it may be a signal of over fitting.

The following describes a few parameters that need to be set for machine learning model trainings:

> Epoch, which is how many times the model is trained with the complete set of data.
>
> Batch is a data group in each epoch that will be fed to the model together.
>
> Iterations is how many batches in one epoch.

For example, if the training data set has 10,000 rows of data, then one epoch means all the 10,000 rows of data passes the network once. If each batch has 100 rows of data, then the iterations of each epoch is 10,000/100=100.

1.5 Peak Price Forecasting

Peak price forecasting has been very important in financial areas, including the forecasting of big jumps in stock prices, gas prices, oil prices, metal prices, and electricity prices.

The predominant forecasting methods for peak events can be divided into three groups [59, 74]. The first group forecasts peak events together with the non-peak events [44, 45, 59, 64, 74, 76]. In this case, the models learn the pattern of peak and non-peak events together.

The second group treats peak and non-peak events separately [16,18,22,27,36,54,55,84]. The third group is aimed at forecasting the probabilities of events in different ranges, which is called probabilistic forecasting methods [41]. However, none of these methods can resolve the hard limitation that the causes of peak events can vary widely. For example, peak-electricity-prices in the wholesale electricity markets can be triggered by extreme high system load and low generation but can also be triggered by the fluctuation of wind generation in the system, and it can also be caused by human operation faults, extreme weather, software failures, and transmission congestions [41]. Even if we treat peak-electricity-prices separately to non-peak-electricity-prices, we still have hundreds of reasons for the peak-electricity-prices themselves.

The dissertation will introduce some methods developed focusing on peak-electricity-price forecasting. The methods are trying to improve the accuracy of forecasting the peak-electricity-price through improving the forecasting strategy, input features and results processing.

1.6 Electricity Market Real-Time Co-optimization

ERCOT is planning a Real-Time Co-optimization (RTC) project, which will co-optimize electricity generation and ancillary services every 5 minutes in Real-Time operation [6].

Compared to the current ERCOT SCED structure, there will be additional AS offer and AS demand curves included as inputs to SCED. AS awards and AS prices will be new outputs of SCED under RTC. The change of SCED data flow is as shown in Figure 1.9. SCED will need two more inputs and have two more outputs, as circled.

Electricity wholesale prices based on current optimization algorithms will be affected by RTC where the Real-Time frequency responsive capacity of the market will play a key role. RTC will bring more challenges and opportunities to the market, and Real-Time electricity price forecasting will be more valuable under RTC.

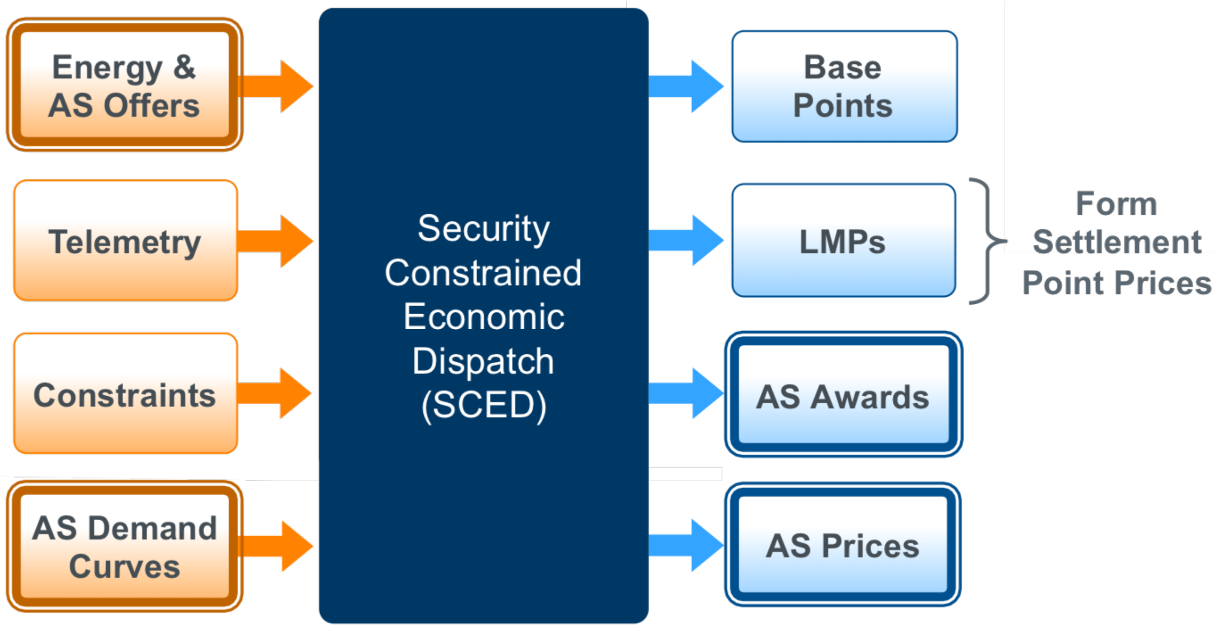

Figure 1.9: ERCOT SCED in RTC. Source: ERCOT Real-Time Co-optimization Task Force [13]

1.7 Literature Survey

Based on the review paper [74], the research of electricity price forecasting started in late 1990s, and the main methods of electricity price forecasting can be categorized into five groups: multi-agent, fundamental, reduced-form, statistical, and machine learning models.

Multi-agent methods simulate the operation of the system and calculate the price by matching supply and demand in a model of economic competition [63, 88]. Fundamental methods study the physical and economic factors that impact the electricity prices, then predict prices based on the study [23, 29]. Reduced-form methods study the statistical characteristics of the electricity price over time, which include spot price models and forward price models. Markov regime-switching and jump diffusion are the most well-known reduced-form spot price models [17, 73].

Statistical methods predict electricity price using a mathematical combination of historical prices and all the related data. Statistical methods are widely used since it is easy to configure inputs and outputs, very easy to understand and operate, and the forecasting results are quite accurate. The most widely used statistical approaches are:

1. Find a similar day in the history and make adjustments based on differences [44, 46];
2. Mathematical regression [69];
3. Time series models [45, 75].

During the last 10 years many machine learning methods were used for electricity price forecasting. Early pioneers foresaw the potential revolution that machine learning would bring to forecasting in electricity markets, and tried simple neural networks in electricity price forecasting [24, 47, 57, 61, 85]. Even though the accuracy at that time cannot compete with advanced neural networks nowadays, nevertheless the results were already very impressive at that time.

As neural networks have kept improving and computing power has been growing rapidly while the cost drops significantly during the last 10 years, machine learning has become a very popular approach for electricity related forecasts [28, 33, 34, 52, 71, 86], including wind speed and load forecasts. By using neural network modes, wind speed forecasting accuracy has been improved by 30% compared to the previous best forecasting methods based on the study in [71]. DNN and CNN were used in the studies, but for electricity price forecasting, time sequence is one of the key factors to fully understand the price trend.

After RNN was initially proposed for time sequence study, later the improved version LSTM and GRU have shown significant memory ability. Now a machine learning approach can model complex nonlinear time series forecasting problem with satisfactory accuracy [53, 68]. LSTM and related networks are becoming the best tools for solving any time sequence problem. However, electricity grid and market systems are more complicated than most other systems, which has caused difficulties for a single recurrent neural network to fully understand the underlying factors behind the electricity market data. Peak-electricity-price forecasting is a big challenge for recurrent neural networks because the event is rare and could be caused by very different reasons, but it is very important information to the market. The dissertation will discuss the approaches to solve these problems.

Xu and Baldick tried to forecast Day-Ahead electricity prices of the ERCOT wholesale market using different neural networks with and without including the peak-electricity-prices [77], as will be discussed in Chapter 2. It drew the conclusion that the peak-electricity-price needs to be studied separately from typical price evolution because there are some underlying special market conditions behind the peak-electricity-price.

Nowotarski in his review paper [59] summarized the forecasting methods that have sensitivity to peak-electricity-prices, and he especially recommended the probabilistic forecasting methods that can forecast the probability of different price ranges; however, the probabilistic methods would have problems effectively dealing with the big range of prices in real electricity markets. For example, in the ERCOT market the wholesale electricity price in Real-Time market can move from negative to $9,000/MWh within a few hours.

Xu and Baldick then pointed out in their papers [76,78,79] that it is more practical to forecast the time of the peak-electricity-price, rather the price value itself. Xu and Baldick introduced several neural network based methods to forecast time of peak-electricity-prices in three papers [76, 78, 79]. The details will be discussed in the following chapters.

In the remainder on the dissertation, Part I will cover the neural networks developed to perform ERCOT wholesale Day-Ahead electricity price forecasting, and the analysis of the forecasting results. Part II will cover the forecasting of peak-electricity-prices in the ERCOT wholesale Real-Time Market. Part III will give conclusions and future work plans.

Part I

Day-Ahead Market Wholesale Electricity Price Forecasting

Chapter 2

ERCOT Day-Ahead Price Forecast[1]

Electricity price forecasting tools and technologies are used by market participants to help optimize market operations. In the longer term, bilateral contracts are priced based on forecasts of future Day-Ahead and Real-Time Market prices [48]. Large electricity consumers can minimize wholesale purchase costs by operating during low price hours. As stated in Chapter 1, the accuracy of electricity price forecasting is very important [40, 83]. However, due to the special characteristics of electricity, the forecasting of the electricity prices is very challenging. Many modeling and statistical methods have been proposed during the last few decades, but it is very difficult to build a model with good prediction accuracy that can cover the characteristics of the whole system. It is difficult to make progress in electricity price forecasting using classical models and statistical approaches. However, as stated in Chapter 1, newly developed approaches like machine learning, especially neural network methods, might make substantial improvements in electricity price forecasting. Advanced recurrent neural networks like LSTM and GRU have proven strength in handling complexity and non-linearity as the technology keeps developing and computing power becomes cheaper, and are already applied in energy forecasting researches [14,24,25,28,35,47,60,61,65,71,74,80,82,85]. This chapter focuses on exploring the forecasting capability of neural network models for the ERCOT Day-Ahead Market.

Chapter 2 has 4 sections. Section 2.1 explains how Day-Ahead Market price forecasting will benefit the market. Section 2.2 discusses the input data for machine learning

[1]This Chapter is based on the paper "Day-Ahead Price Forecasting in ERCOT Market Using Neural Network Approaches" in Proceedings of the Tenth ACM International Conference on Future Energy Systems (e-Energy 19), June 2019, by Jian Xu and Ross Baldick. Jian Xu is the first author of the paper. Dr.Ross Baldick is the co-author and supervised the work.

models and how to test the models. Section 2.3 gives analysis of the results. The last section concludes the studies.

2.1 ERCOT Day-Ahead Market Price Forecasting

According to the 2017 State of the Market Report for the ERCOT Electricity markets by the Independent Market Monitor (IMM) [4], the ERCOT Day-Ahead Market covers about 90% of the total capacity of the next operating day. This means most market participants will carry all their Day-Ahead plans to the Real-Time operation, and their revenues and profits depend primarily on Day-Ahead prices. So the forecast of Day-Ahead hourly electricity price is especially useful for both ERCOT and the market participants to optimize their plans and market operations.

Day-Ahead Market price forecasting is one of the most popular topics among electricity forecasts for the following reasons:

1. The Day-Ahead Market is very important, since, as in ERCOT, it covers most of the capacity of the whole system;

2. The Day-Ahead price is settled hourly, with less fluctuation, which is much easier to forecast compared to the Real-Time price which is calculated by SCED every 5 minutes and settled every 15 minutes. The Real-Time price has much greater and more uncertain fluctuations than the Day-Ahead price;

3. Forecast of the Day-Ahead price is practical and extremely useful for ERCOT and market participants to make plans one day ahead. Unlike the Real-Time Market for which everything needs to be decided within minutes or seconds, so that studies that run longer than 5 minutes may be useless, the Day-Ahead Market gives a long time window to make decisions on submitting bids and offers. Most ERCOT market participants spend more time

and efforts on Day-Ahead Market planning than Real-Time Market planning. The Real-Time Market is more like execution of the Day-Ahead Market plans.

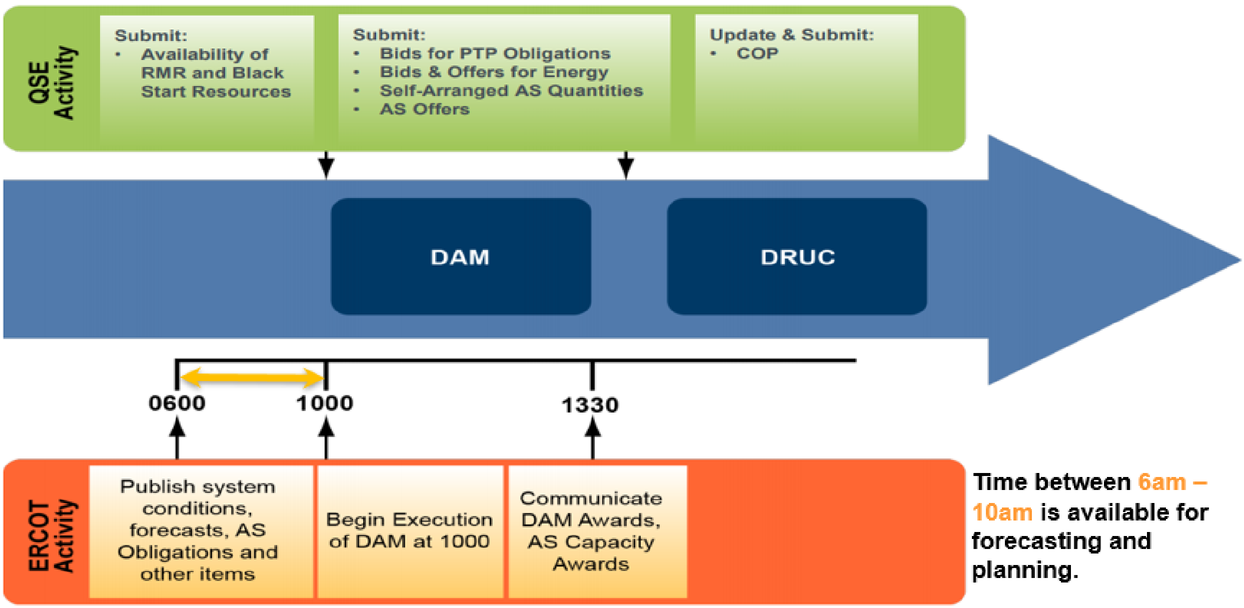

Figure 2.1: ERCOT Day-Ahead Market timeline. Source: ERCOT Basic Training [11]

During operation of the ERCOT Day-Ahead Market, market participants submit their bids and offers to ERCOT before 10 AM of the day before the relevant operating day, and the ERCOT Day-Ahead Market engine will run an optimization trying to maximize total surplus of the market (total revenue minus total cost). The load (bids) and generation (offers) sizes, wind speed, weather and temperature, transmission conditions will all affect the clearing price of the Day-Ahead Market. The ERCOT Day-Ahead Market timeline is shown in Figure 2.1.

The specific application of machine learning models in this chapter is to forecast the Day-Ahead wholesale electricity price in one specific ERCOT Load Zone. One of the key things to build a good machine learning model is to get very organized data input which can capture as much information as possible that will affect the price. The following sections will explain how to prepare input data, and how different models perform in forecasting the Day-Ahead price. The machine learning forecasting approach is demonstrated in Figure 2.2. The right part is the ERCOT optimization engine and the left part is the machine learning forecasting model. The forecast tries to match the future outcomes of the optimization engine based on the existing available information.

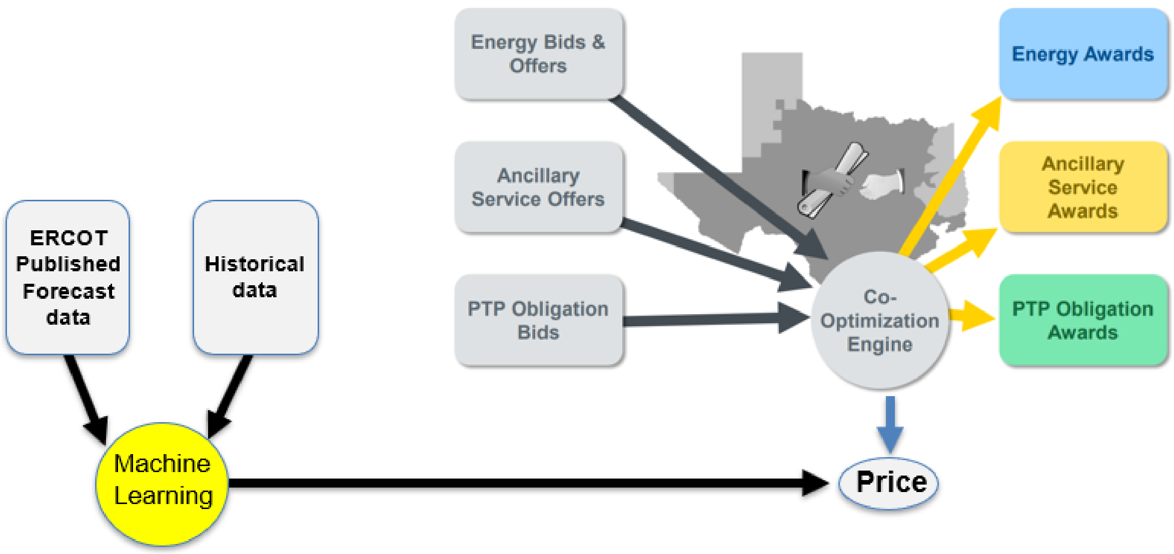

Figure 2.2: ERCOT Day-Ahead Forecasting using Machine learning Approach. Source: Created based on ERCOT Basic Training. [11]

2.2 Model Description and Preliminary Results

2.2.1 Data Input

The following information is available and can be used for Day-Ahead price forecasting in the ERCOT Market: Day-Ahead cleared prices of previous days, wind and load long-term forecasting information from ERCOT, transmission information from ERCOT, and weather and temperature forecasts. Weather and temperature information is not used explicitly in this chapter. Instead the models use month/day/hour information, which not only implicitly represents temperature information (since the month and day will reflect average temperature range and seasonal information) but also represents peak and non-peak information into the neural network models. In particular, the hour input feature will reflect peak and non-peak hours of the day, with ERCOT peak hours being 7:00 - 22:00 of weekdays, and non-peak hours being 22:00 - 7:00 next day and weekends. Based on the data availability, the selected input features for this chapter are as in Table 2.1.

Table 2.1: Selected input features

Selected Input	Comments
Load forecast	Day-Ahead Market uses load forecast to run unit commitment.
Wind forecast	Day-Ahead Market uses load and wind forecast to calculate net load.
Previous Day-Ahead prices	Provide price and system condition reference.
Month, Day	Provide seasonal information, temperature range.
Hour	Provide peak and off-peak information, and temperature information.

ERCOT can be roughly divided into four zones: West, North, South, and Houston. A testing dataset from the ERCOT South Load Zone was used to train, test, and validate the neural network models. The data set is partitioned into training, testing and validating subsets. An example of input data for Day-Ahead price forecasting is as in Table 2.2. Each column of the table corresponds to an input feature of the neural network model.

Table 2.2: Example of input data for Day-Ahead price forecasting

Time	Day-Ahead Price of Day d-1 ($/MWh)	Wind forecast of Day d (MW)	Load of Day d-1 (MW)	Month	Day	Hour
01/01/2018 01:00	24	1,000	32,000	1	1	0
01/01/2018 02:00	25	900	32,400	1	1	1
01/01/2018 03:00	22	900	32,800	1	1	2
01/01/2018 04:00	25	920	33,780	1	1	3

2.2.2 Model Architectures

Models using all of the neural networks described in section 1.4, namely Deep Neural Network (DNN), Convolutional Neural Network (CNN), Long-short Term Memory Network (LSTM), and Gated Recurrent Units (GRU), were set up for testing and comparison.

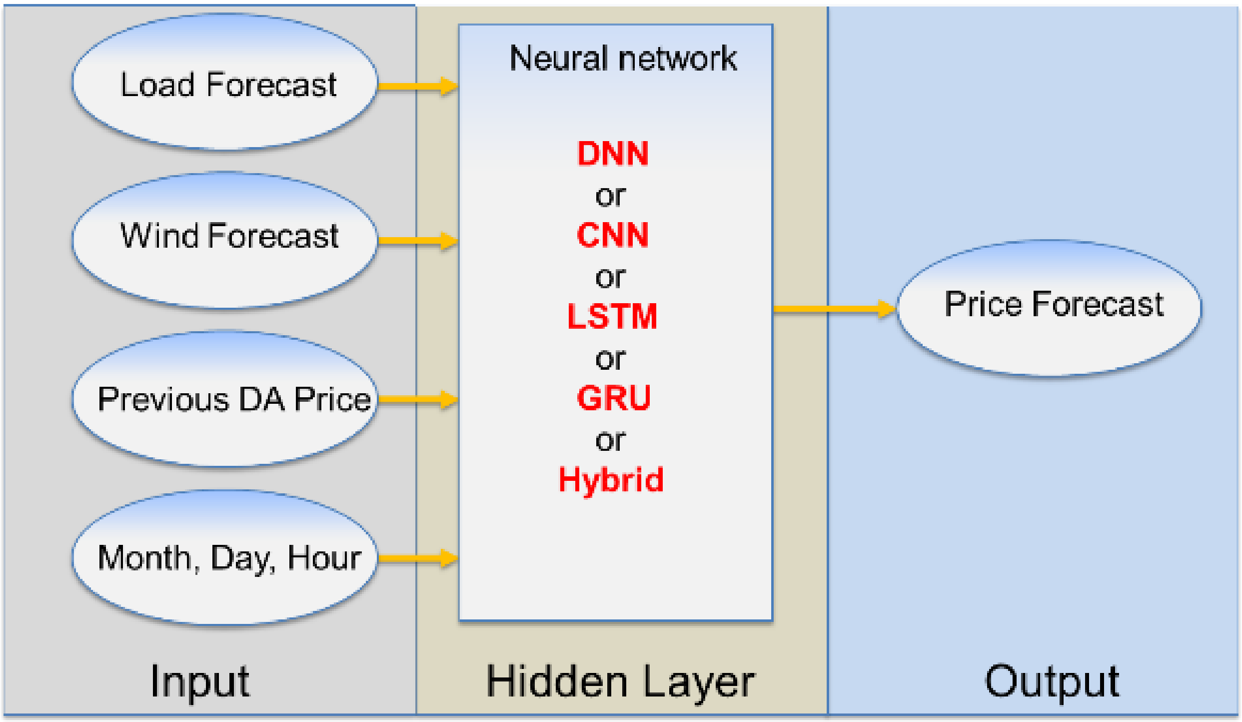

Figure 2.3: Machine learning data flow

As shown in Figure 2.3, all the model architectures can be simplified into three parts: input, hidden layers, and output. The design of hidden layers is the core, which will determine how the model adapts the trainings, and the consistency between training and testing.

The neural networks introduced in Section 1.4 can be divided into two groups: feed-forward neural networks and recurrent neural networks (RNN). DNN and CNN belong to the first group while LSTM and GRU belong to the second. Figures 2.4 and 2.5 show examples of the two kinds of neural networks. The structure on an RNN, as shown in Figure 2.5, matches better with time sequential data. For this reason, an RNN has better capability in studying the time sequential influences from the input features compared to feed-forward neural networks illustrated in Figure 2.4.

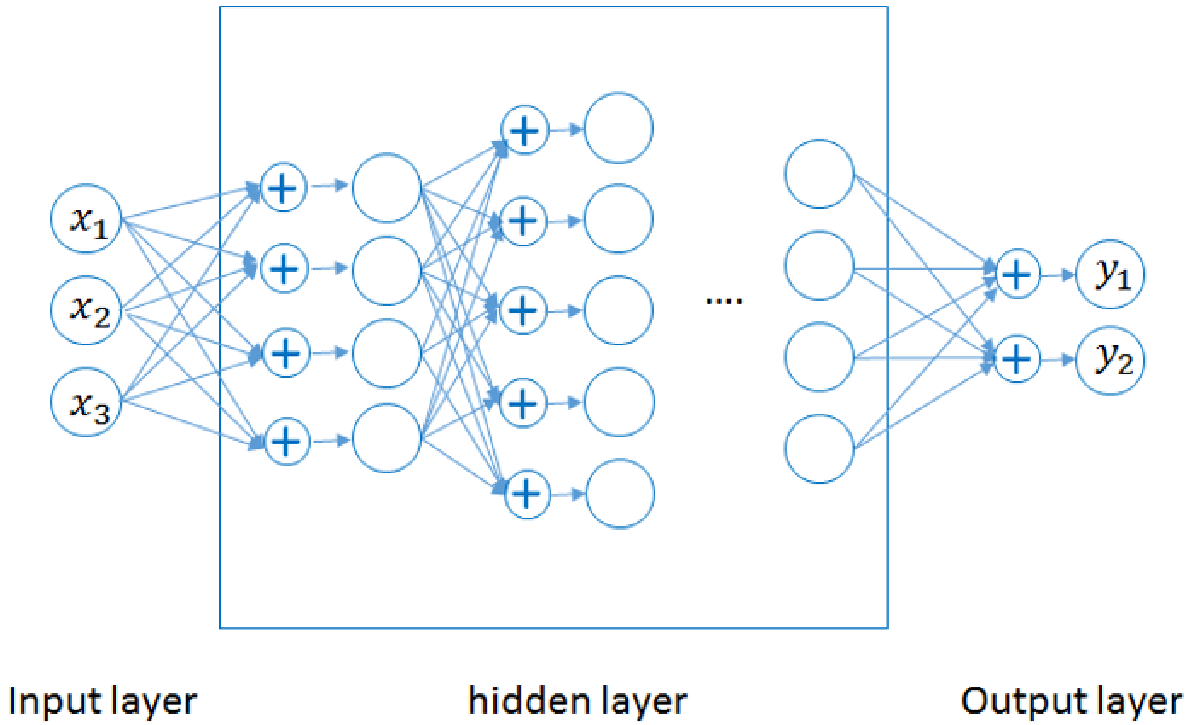

Figure 2.4: Feed-forward neural network (DNN)

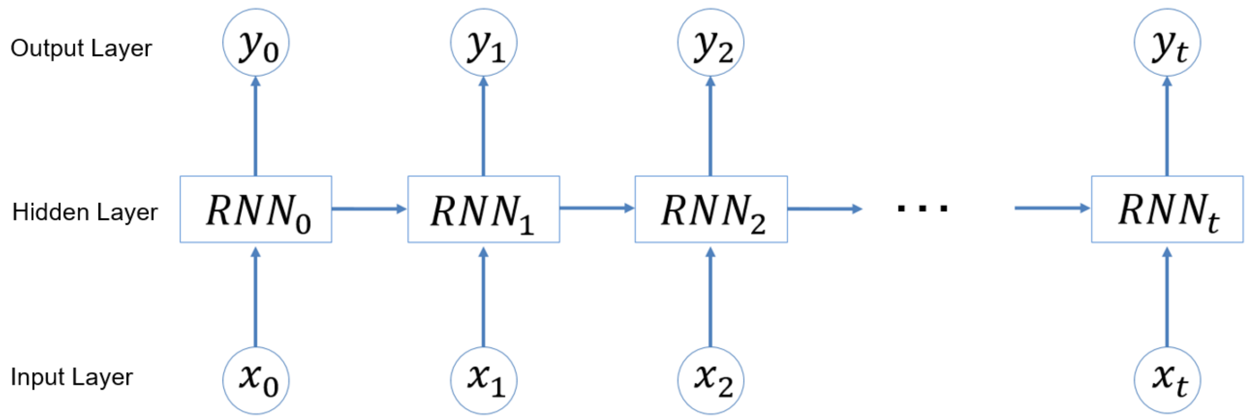

Figure 2.5: Recurrent neural network

Hidden layers of a LSTM model is one of the most complex structures because of the recurrent structure of the LSTM cells [38]. Figure 2.6 shows the data flow structure of a simple three hidden layers, three nodes per layer LSTM model.

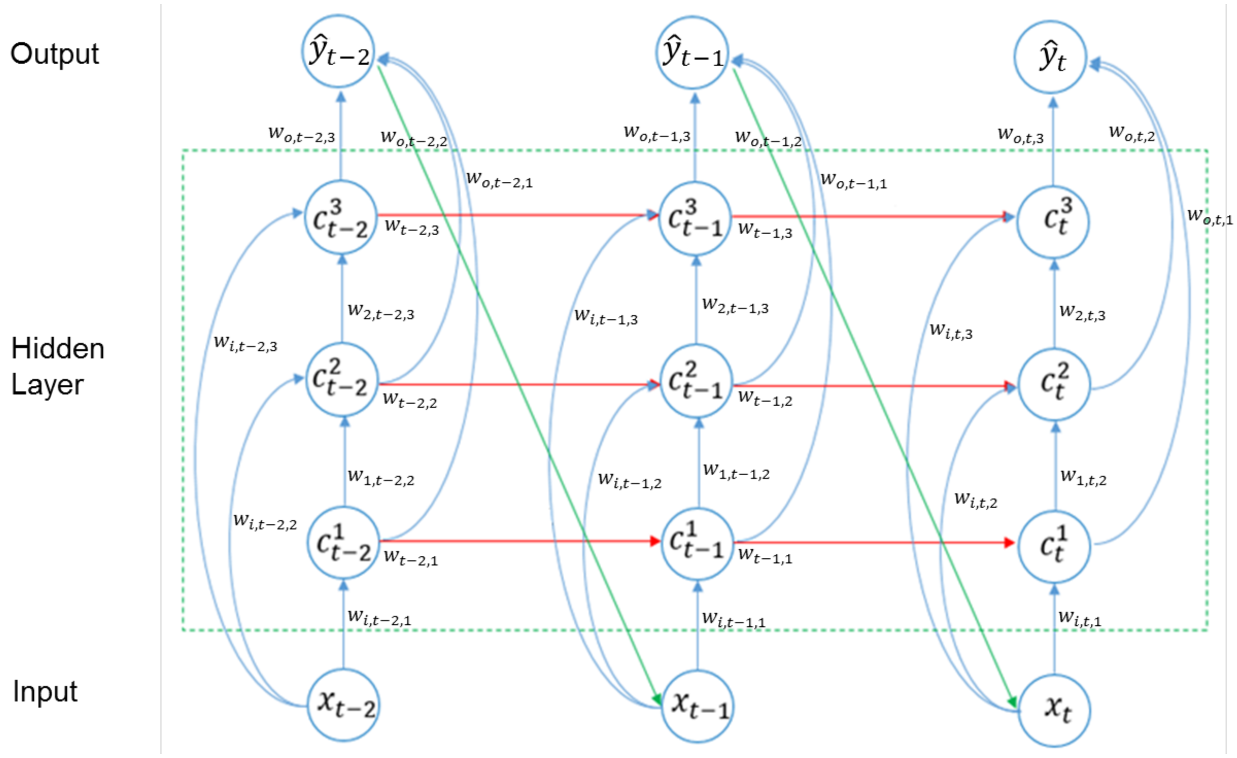

Figure 2.6: A simple LSTM data flow

The input sequential vectors are sent to the hidden layers and the hidden sequential vectors will be calculated, and then the output vectors will be calculated. Each output vector $\hat{y}_t$ will be passed to the next input vector x_{t+1} for parameterizing a predictive distribution of the input [38]. The value at the cell at time sequence t of the layer n is:

$$c_t^n = H(w_{t-1,n} \times c_{t-1}^n + w_{i,t,n} \times x_t + w_{n-1,t,n} \times c_t^{n-1}) \tag{2.1}$$

where H is the hidden layer function and the w terms are weights as shown in Figure 2.6. The forecast, $\hat{y}$ is the output at time t and is determined by the output layer function using all the layers:

$$\hat{y} = \sum_{k=1}^{n} w_{o,t,k} \times c_t^k \tag{2.2}$$

The LSTM model used in this chapter has more than three layers, and each layer has over 100 nodes. Moreover, the inputs are vectors instead of scalars as illustrated in the simple model in Figure 2.6. Vector inputs require huge calculations for every training. Figure 2.7 shows an example of the time sequential data flow of LSTM.

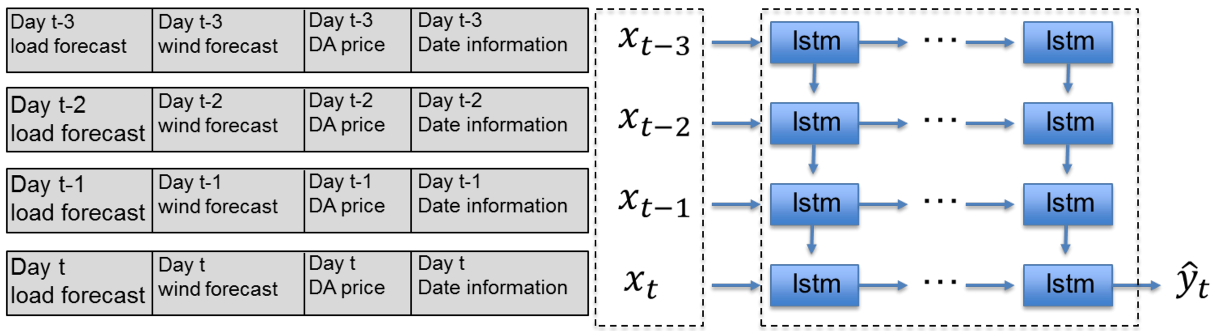

Figure 2.7: Time sequence data flow of LSTM

2.2.3 Testing of the Models

All the four neural network models, DNN, CNN, LSTM, and GRU, will be trained and validated using the equivalent settings and the same set of data. Once the model has been trained, a set of (different) testing data will be fed to the trained model to predict the Day-Ahead prices. The error is calculated by statistically comparing the forecasted Day-Ahead prices to the actual Day-Ahead prices during the same time window. Mean Absolute Error (MAE) was used to evaluate the results in the tests. If the testing data set has T hours in total, y_t is the actual Day-Ahead price at hour t, and $\hat{y}_t$ is the forecasted Day-Ahead price at hour t, then MAE can be calculated as follows:

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |(y_t - \hat{y}_t)| \tag{2.3}$$

2.3 Result and Discussion

2.3.1 Performance Testing of Different Neural Networks

Performance of each neural network has been tested for comparison. Hybrid models combining feedforward neural network and recurrent neural network have also been tested to see how the hybrid models can improve the forecasting capability. Testing result of each neural network is shown in Figure 2.8. The result shows recurrent neural networks LSTM and GRU are more accurate than feed forward neural networks DNN and CNN in predicting the prices.

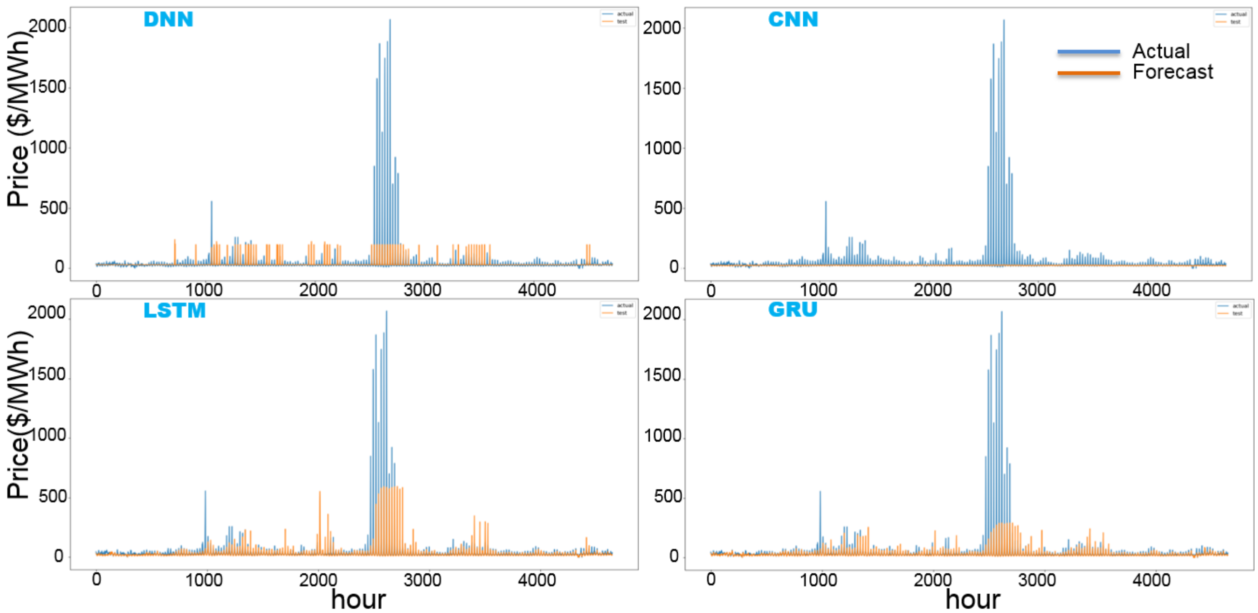

Figure 2.8: Performance of each neural network

A hybrid model can be constructed by using the outputs of a feed-forward neural network as the inputs for a recurrent neural network. Hybrid models studied in this chapter are LSTM-DNN and LSTM-CNN. Figure 2.9 shows how the structures of the two neural networks for the case of LSTM-DNN. LSTM-CNN has an analogous structure.

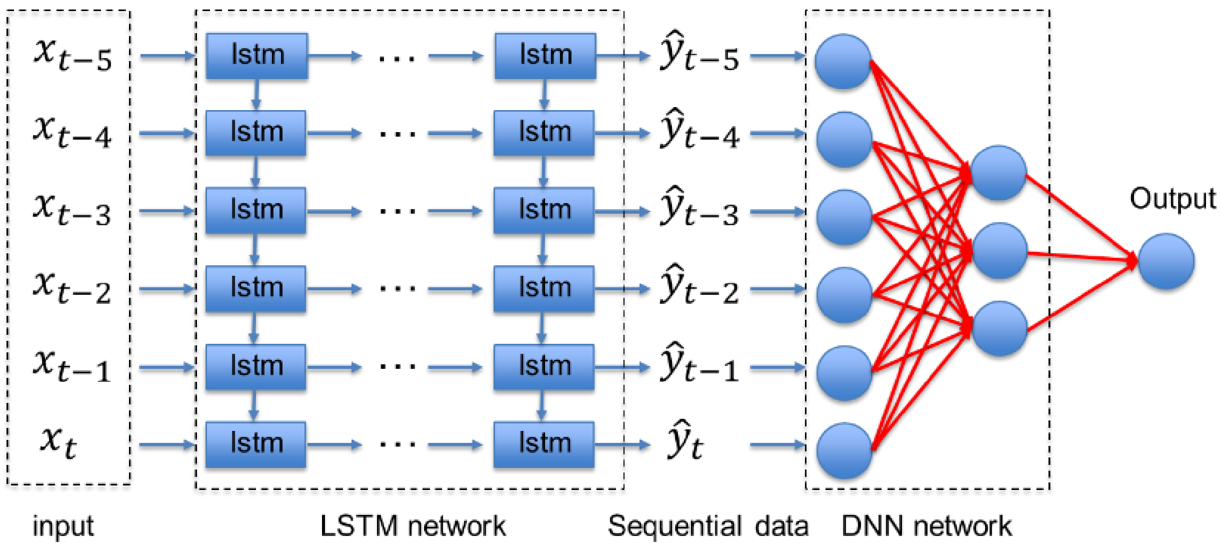

Figure 2.9: Hybrid model architecture

The testing results of the hybrid neural networks are shown in Figure 2.10, the results show that LSTM-CNN is more accurate than LSTM-DNN in forecasting prices. All the testing data are summarized in Table 2.3 for comparison.

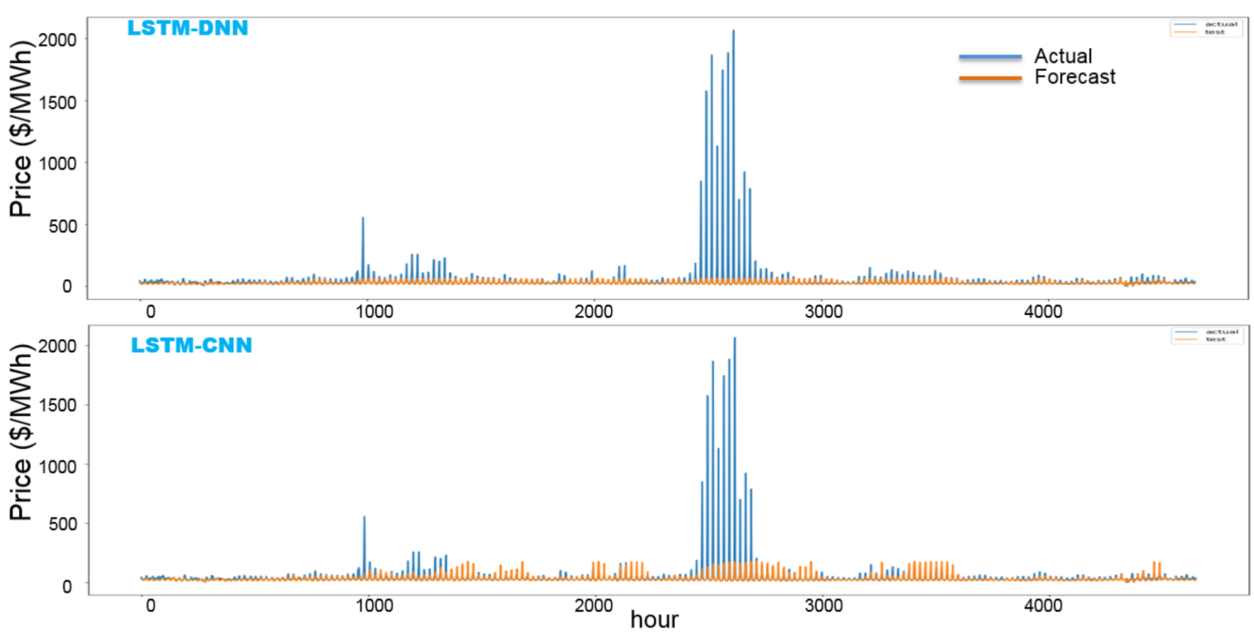

Figure 2.10: Performances of hybrid neural networks

Table 2.3: Summary of testing results for different neural networks

Neural Network	MAE	CPU Time (minutes)
DNN	19.328	202
CNN	16.532	607
LSTM	15.838	2,487
GRU	14.863	2,316
LSTM-DNN	15.394	3,015
LSTM-CNN	16.092	2,482

From the results, GRU performs best among all the neural network models that were tested, which is consistent with the statement in [49] that GRU has better performance than LSTM in many cases. Hybrid models LSTM-DNN and LSTM-CNN do not improve forecasting accuracy much compared to LSTM alone and are not superior to GRU.

2.3.2 Study of the Impact of Peak-electricity-prices on the Forecasting

As discussed in Chapter 1, sudden "spike" prices are challenging to predict, both from the perspective of timing and from the perspective of the peak-electricity-price reached. The previous section utilized all data, including both "normal" prices and price spikes. This section analyzes how removing peak-electricity-prices will affect forecasting accuracy of normal prices. Price spikes over \$100/MWh occur about 200 times a year in the ERCOT South Load Zone, which is less than 2% of the total hours. It is believed that the sporadic high price events will affect the ability of neural networks to forecast normal prices if the high prices are included in the training and validation data set. Figure 2.11 shows the price distribution in each month during 2013-2018, omitting price spikes above \$100/MWh.

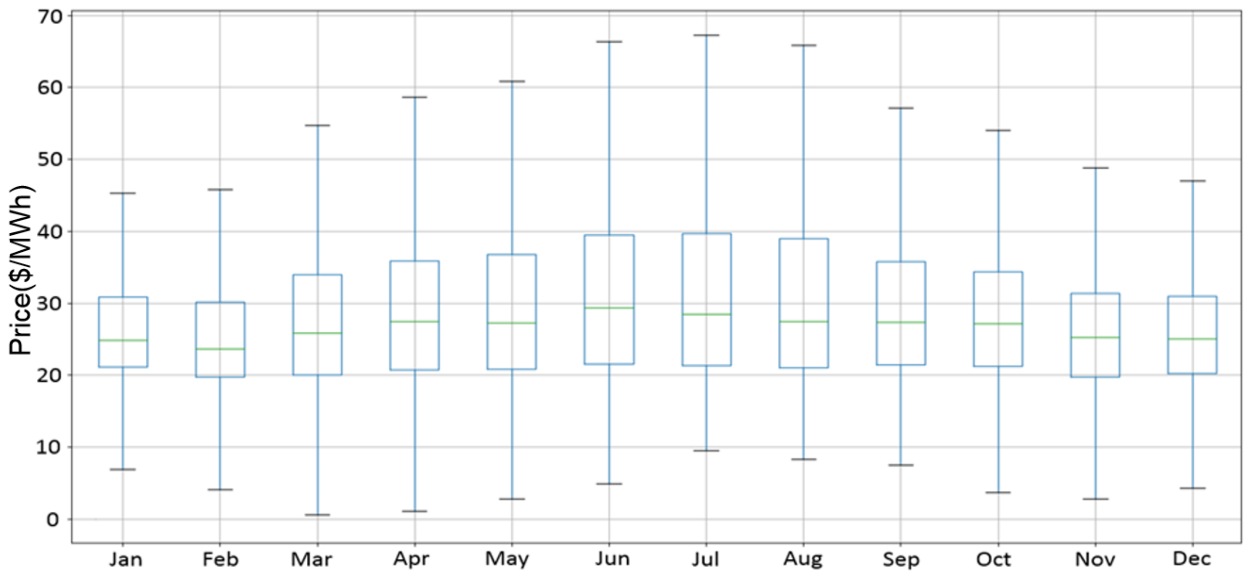

Figure 2.11: Boxplot omits price outliers (2013 - 2018)

Figure 2.12 shows the counts of price spikes in different price ranges during 2013-2018 in the ERCOT Day-Ahead Market. Prediction of timing and magnitude of price spikes apparently needs to be done separately to prediction of normal prices, and needs to consider more inputs like transmission conditions, system conditions, and wind fluctuations, which are not necessary inputs for accurate forecasting of normal price occurrences. If the main goal is to predict prices during the 98% of time when they are normal, then the price spikes can be excluded and only basic data is need for accurate forecasts. Part II will address the important issue of predicting timing and magnitude of price spikes in the context of wholesale Real-Time electricity prices.

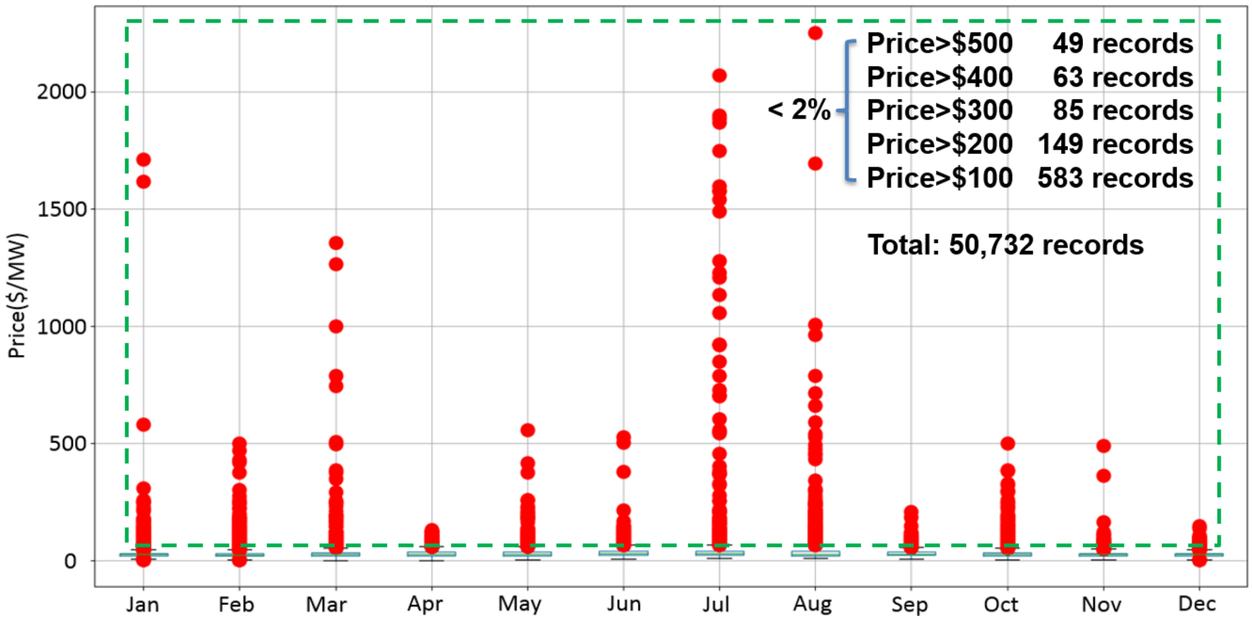

Figure 2.12: Boxplot includes price outliers (2013 - 2018)

Some specific testing results using GRU omitting price spikes are shown in Figure 2.13. Since recurrent neural networks are suitable for time sequential price forecasting, GRU, which is a popular recurrent neural network, is selected for testing and comparison in this section. The results are also summarized in Table 2.4.

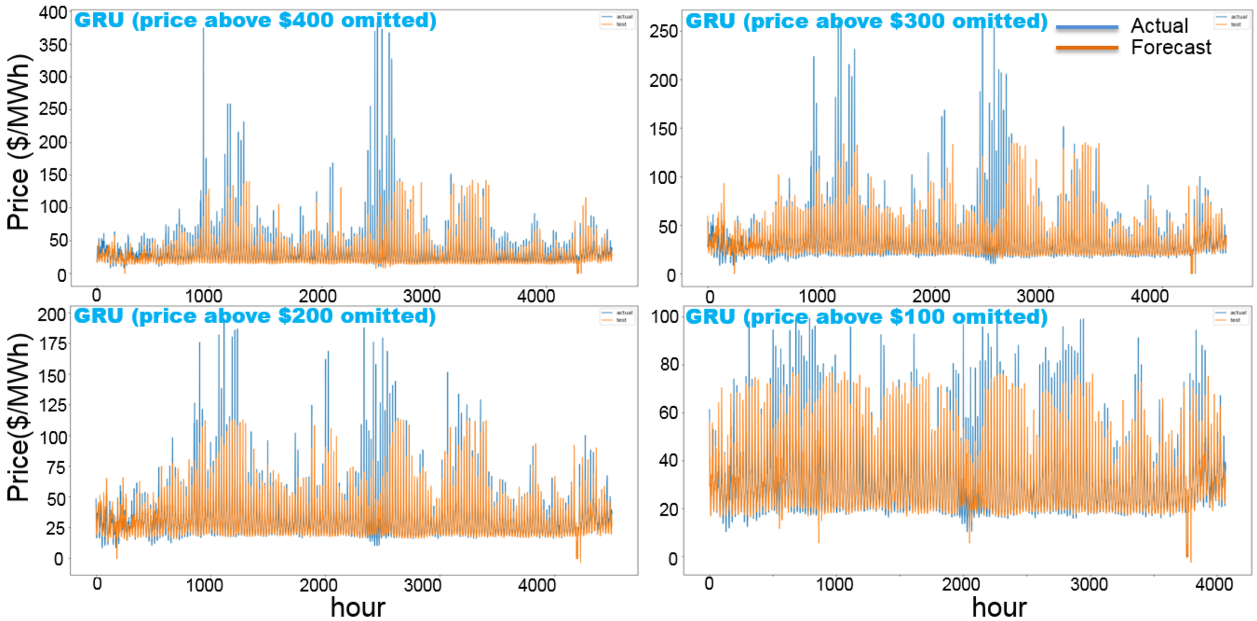

Figure 2.13: Testing result of GRU omitting peak-electricity-prices

Table 2.4: Summary of testing results for GRU omitting peak-electricity-prices

Neural Network	Filter	MAE of testing
GRU	Price over $500 omitted	13.750
	Price over $400 omitted	8.324
	Price over $300 omitted	5.979
	Price over $200 omitted	5.146
	Price over $100 omitted	4.528

The testing results have shown that the prediction of the normal conditions can be improved in accuracy if the price spikes are excluded. In practice, there can be two sets of forecasting algorithms: one focuses on forecasting the normal conditions, and another can focus on forecasting the peak-electricity-prices. Since the peak-electricity-prices themselves show great variability, it is easier to predict the time when price spikes will occur than it is to forecast the actual peak-electricity-prices. As mentioned previously, this will be explored in more detail in Part II.

2.3.3 Neural Networks Compared to Other State of the Art methods

This section compares state-of-the-art classical forecasting methods to the advanced neural network methods. Data used in this testing omits records with price above $100/MWh and so is comparable to the ANN results from Section 2.3.2 corresponding to the last row in Table 2.4. The state of the art classical forecasting methods for comparison are: Autoregressive Integrated Moving Average (ARIMA), Holt Winter's Exponential Smoothing (HWES), Autoregressive Moving Average (ARMA) and Persistence Algorithm (the "naive" forecast).

Table 2.5: Summary of testing results for other state-of-the-art methods

Method	MAE
ANN(GRU)	4.528
ARIMA	5.408
HWES	6.367
ARMA	10.719
NAIVE	5.756

The results for the statistical methods are shown in Table 2.5. From the results we can see GRU has MAE 4.528 for the same data set, which is an approximately 10% improvement over other state-of-the-art forecasting methods. Interestingly, most of the other models cannot outperform a naive persistence forecast.

2.4 Chapter Summary

The overall accuracy of neural networks has outperformed classical forecasting approaches in several contexts [68]. This chapter shows that the same is true for electricity price forecasting. However, different neural networks have different performances, and recurrent neural networks perform the best among all the neural networks, and are better than state of the art statistical methods in forecasting ERCOT Day-Ahead Market prices.

The empirical performance of the hybrid models was not significantly improved compared to the best non-hybrid model. Using more inputs, such as transmission conditions, may help improve the forecasting accuracy further and it is possible that the hybrid models might show better performance when there is an even greater variety of input data to be used in the forecast.

If the models are not designed to forecast the 2% of time when there are peak-electricity-prices, then the removal of peak-electricity-prices as noise from the training data set can improve the accuracy of the forecasting of those prices that are under \$100/MWh. Peak-electricity-prices will be forecasted separately using special methods and strategies in the later parts of the dissertation.

Compared to Part I, which has addressed on Day-Ahead Market price forecasting, Part II will focus on the Real-Time Market price forecasting. Part II is also going to address the critical peak-electricity-price forecasting problem.

Part II

Real-Time Market Wholesale Electricity Price Forecasting

Part I has discussed Day-Ahead Market forecasting problems. Part II will focus on Real-Time Market and peak-electricity-price forecasting. Part II has three chapters: Chapter 3, Chapter 4 and Chapter 5.

Chapter 3 introduces a new forecasting method called Multiple-Run Time-of-Peak Forecasting (MRTPF), which combines the different outcomes of neural networks with different configurations to forecast the time of peak-electricity-price.

Different from the strategy of Chapter 3, Chapter 4 tries to improve the forecasting of the time of peak-electricity-price by translating the long-range raw price values into binary data 0 and 1.

Finally, Chapter 5 introduces a forecasting method called Three-step Real-Time Electricity Price Forecasting (TREPF), which covers the forecasting periods from Day-Ahead Market to Hour-Ahead and 5-Minute-Ahead Real-Time operations.

Chapter 3

A Time of Peak Forecasting Method based on Artificial Neural Network [1]

Electricity markets are subject to significant variation in prices. Market participants are typically concerned about peak-electricity-price events because they can cause huge benefits or risks compared to non-peak events, despite being occasional or rare. In other fields, occasional or rare peaks or other outliers may also have particular significance. However, as discussed in Chapter 2, it is very challenging for an artificial neural network (ANN) to forecast accurately about peak or other unusual events if the input features are limited and the historical peak events are rare. In some contexts, including electricity markets, the timing of the peak event may be the most critical issue. That is, we care more about when the peak events will occur, rather than how high the peaks will be, in order to plan ahead for these peaks. This chapter introduces a new method which will focus on forecasting when the peak events will happen. Based on the study of this chapter, the statistical combination of forecasts of several ANNs that are configured differently can forecast a greater fraction of the occurrences of peaks than any single ANN can, with acceptable levels of false positives. The method proposed in the Chapter, called Multiple-Run Time-of-Peak Forecasting (MRTPF), uses a combination of ANNs having different numbers of neurons and layers in order to make a significant contribution to time of peak forecasting. As with the forecasting experiments reported in Chapter 2, the MRTPF method is tested using data from ERCOT wholesale electricity market for performance verification.

[1] This Chapter is based on the paper "A New Time of Peak Forecasting Method based on Artificial Neural Network" submitted to the Journal of Modern Power Systems and Clean Energy, November 2019, by Jian Xu and Ross Baldick. Jian Xu is the first author of the paper. Dr.Ross Baldick is the co-author and supervised the work.

Chapter 3 has 5 sections. Section 3.1 introduces peak-electricity-price forecasting and the MRTPF method. Section 3.2 discusses the design of MRTPF. Section 3.3 analyzes the basic approach of the new method. Section 3.4 tests the method using the real market data from ERCOT. The last section gives conclusions about the studies.

3.1 Introduction

As discussed in Section 1.5, peak-electricity-prices have a big financial impact on the market, and while there are already many methods developed to forecast the peak-electricity-prices, none of them can resolve the hard limitation that the causes of peak events can vary widely. As mentioned in Chapter 1, peak-electricity-prices can be triggered by extreme high load and low generation but can also be triggered by the fluctuation of wind generation in the system, and it can also be caused by human operation faults, extreme weather, software failures, and transmission congestion. Even if we treat peak-electricity-prices separately to non-peak-electricity-prices, we still have hundreds of reasons for the peak-electricity-prices themselves. From this observation, a new method, called Multiple-Run Time-of-Peak Forecasting (MRTPF) is proposed. MRTPF relies on the characteristic of ANNs that when several individual ANNs are setup with different neuron configurations, they can perform very differently [19, 67]. By adjusting the neurons and layers to form a range of models with strengths in representing different causes, the statistical combination of all the results can cover the peak-electricity-price possibilities better than any single method.

3.2 Design of MRTPF

The MRTPF method can be based on any collection of forecasting models that have heterogeneous forecasting capabilities, but here the ANN will be used for demonstration in this chapter. Figure 3.1 shows the ANN configuration that will be used to construct the MRTPF model using a family of differently specified configurations of the ANN.

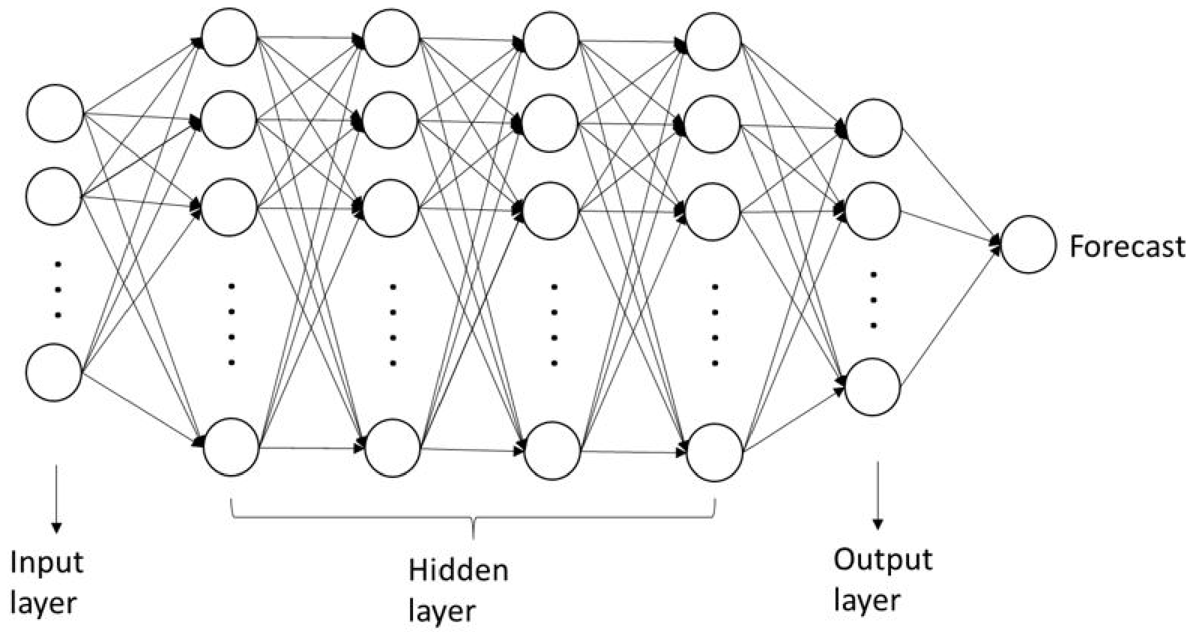

Figure 3.1: Artificial neural network used to demonstrate MRTPF.

The architecture of the MRTPF method is shown in Figure 3.2. The configurations of ANNs are chosen to have high sensitivity to various peak events. Once the ANNs are chosen they will be used to generate forecasts. The group results are combined statistically further into a single forecast as the final result. The same data set which was divided into two groups for training and testing will be used to tune the neural network configurations, to provide inputs to the method, and to evaluate the results.

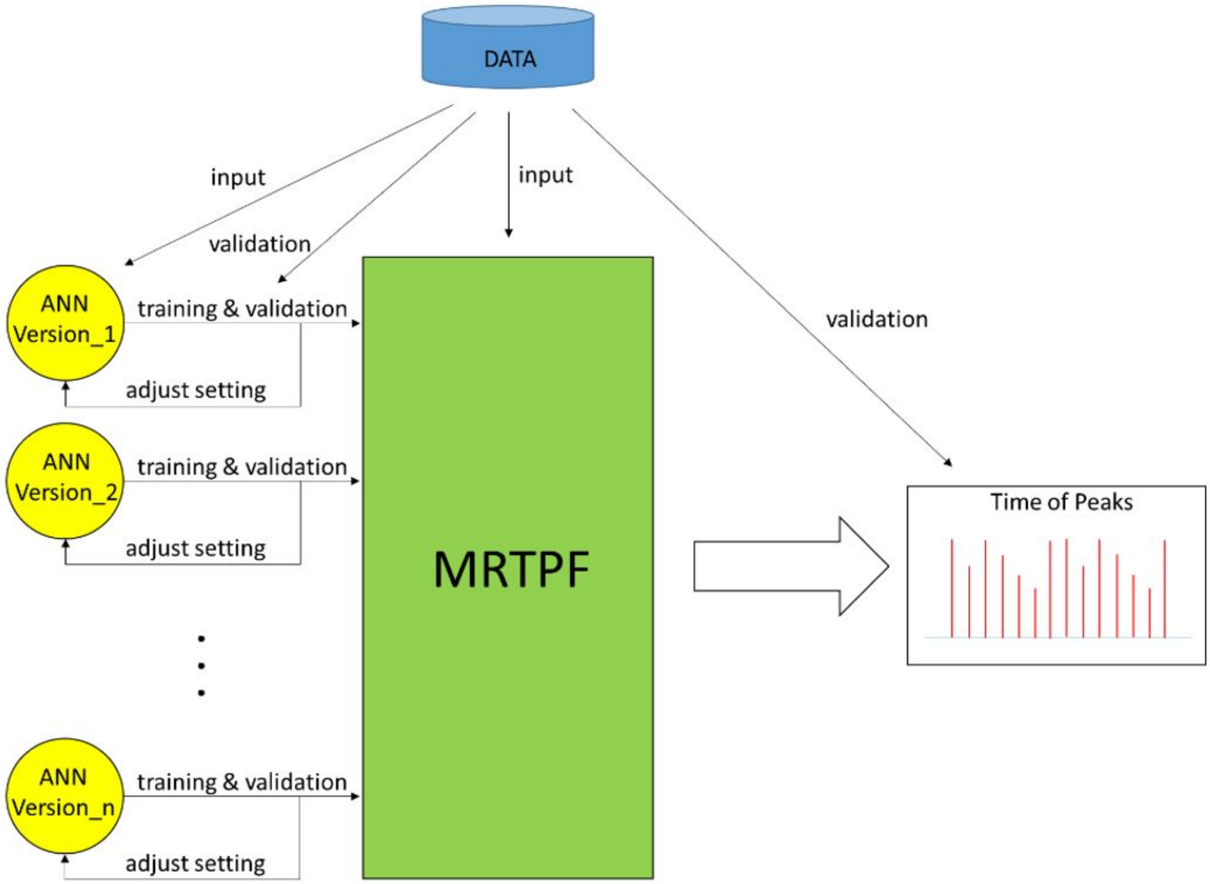

Figure 3.2: Architecture of MRTPF. A combination of pre-tuned neural networks are sent to MRTPF as parts of the machine, and MRTPF will run these nueral networks and statistically generate a single forecast from the results.

The group of forecasts generated by the chosen ANNs will be combined into the final forecasting result as shown in Figure 3.3. The next section describes the basis of this approach in more detail.

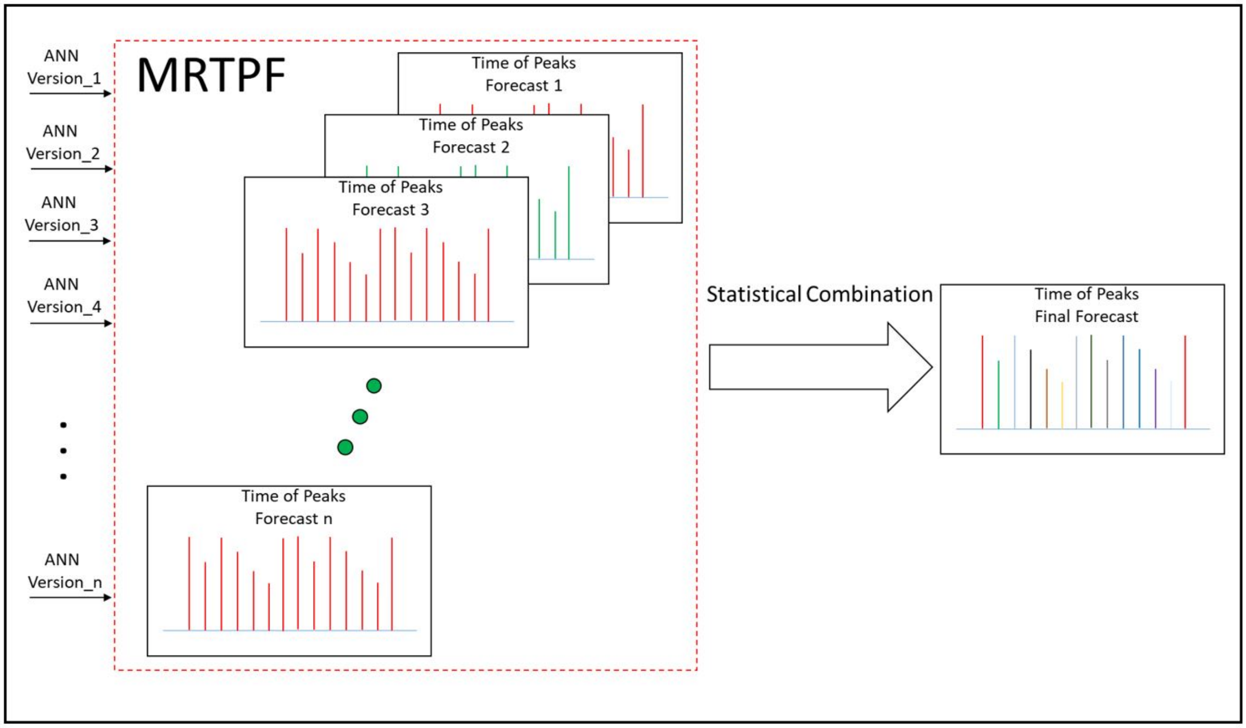

Figure 3.3: Generation of MRTPF forecast. The results from the neural networks within MRTPF are combined using some mathematical rules to generate a single forecast. In this Chapter a simple maximization algorithm is used for demonstration.

3.3 Basis of Approach

A simple neural network example is used to demonstrate the idea of MRTPF. In Figure 3.4, there is only one single layer with one neuron. The activation of the hidden layer is $ReLU$. $ReLU$ is a popular non-linear activation function widely used in neural networks which returns the maximum of 0 and the input value [30], as shown in (3.1) and Figure 3.5.

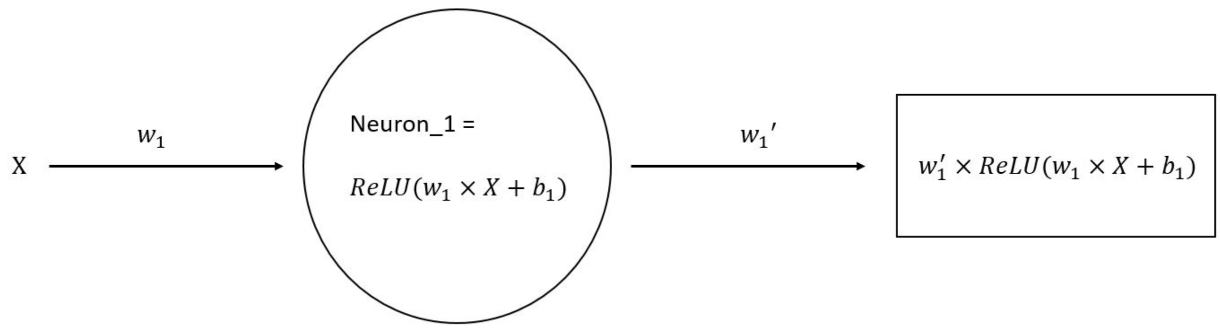

Figure 3.4: Data flow of a single layer and single neuron neural network.

$$ReLU(x) = max(0, x) \tag{3.1}$$

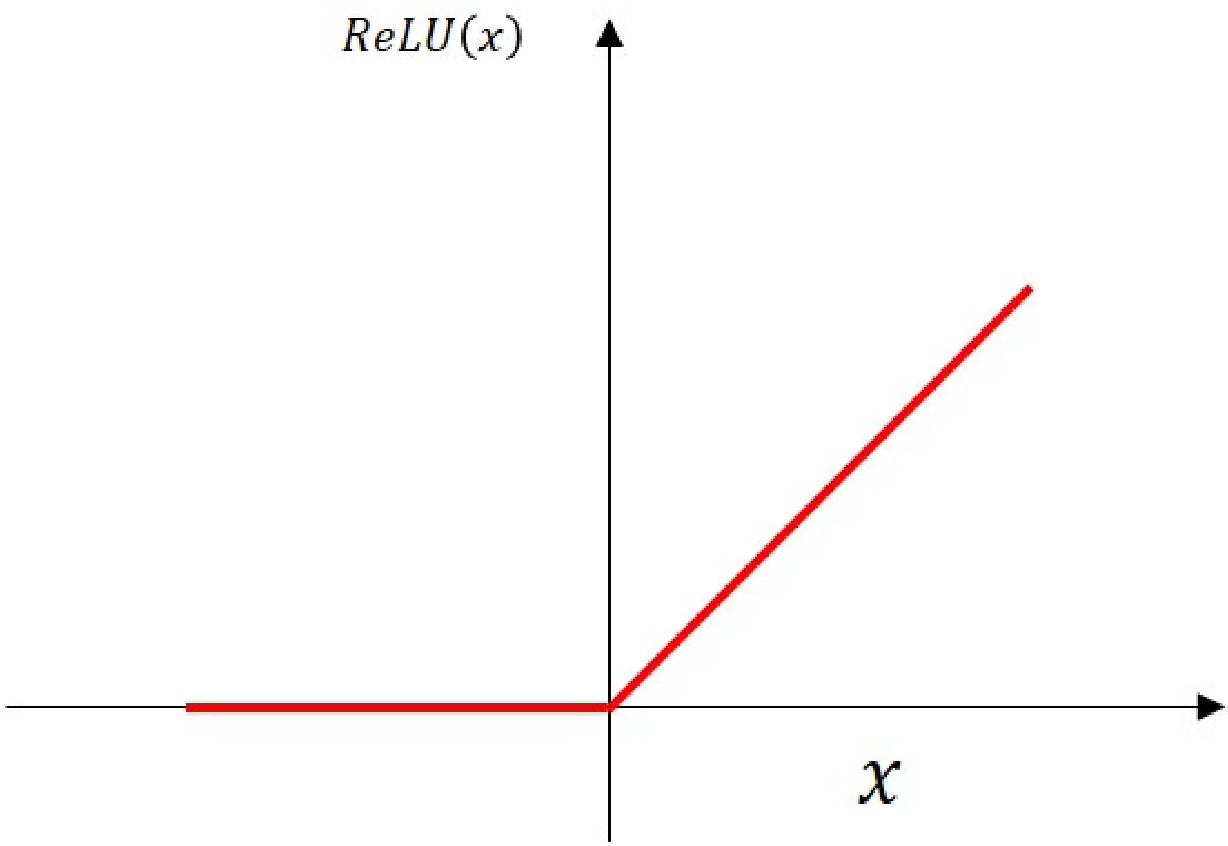

Figure 3.5: ReLU activation.It filters out the negative values from the input.

After the model is trained, the data input-output relationship is illustrated by the red solid line in Figure 3.6. Given the input, the estimated output function will be:

$$\hat{y} = w_1{}' \times ReLU(w_1 \times X + b_1) \tag{3.2}$$

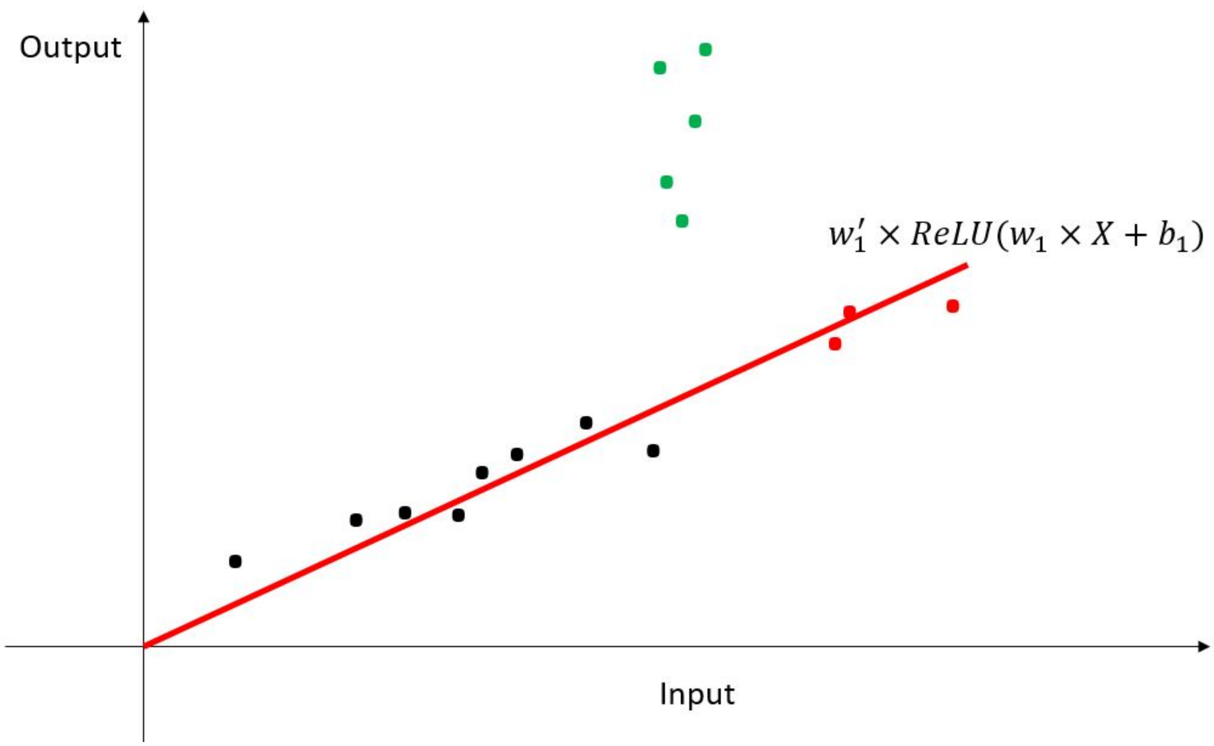

Figure 3.6: Single layer and single neuron neural network forecasting demonstration. Real values are distributed sparsely, and the red line is the trained output function.

If one more neuron is added, the neural network will be like in Figure 3.7.

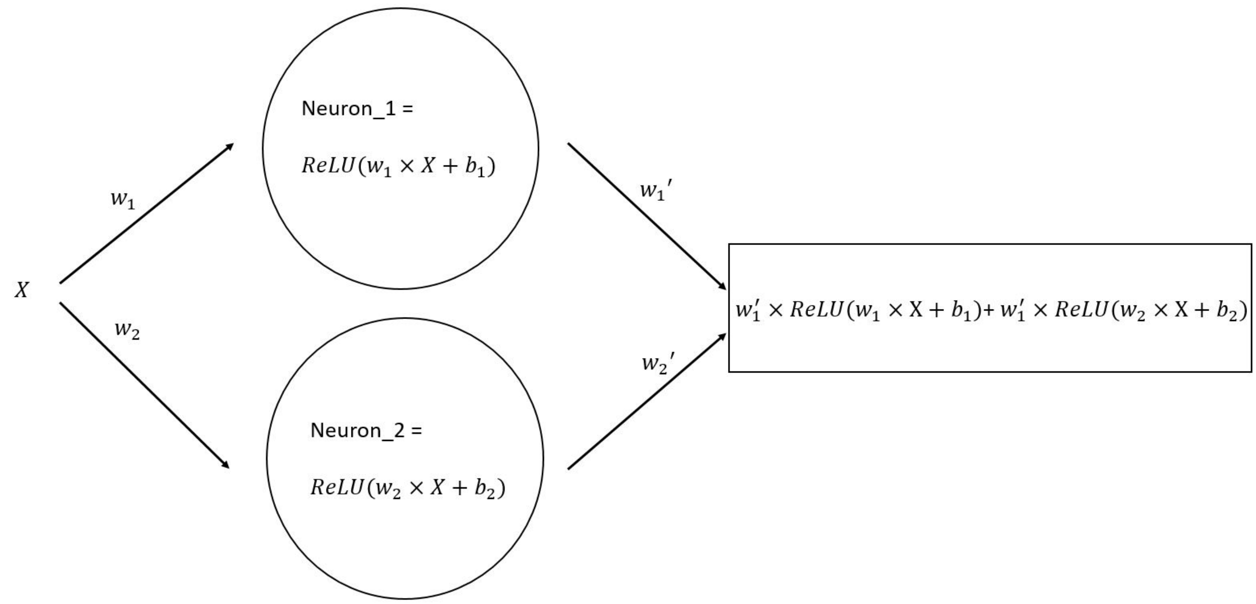

Figure 3.7: Data flow of a single layer and two neurons neural network.

Given the input, the estimated output function will be:

$$\hat{y} = w_1{}' \times ReLU(w_1 \times X + b_1) + w_2{}' \times ReLU(w_2 \times X + b_2) \tag{3.3}$$

We assume, for example, that this response is shown by the blue line in Figure 3.8.

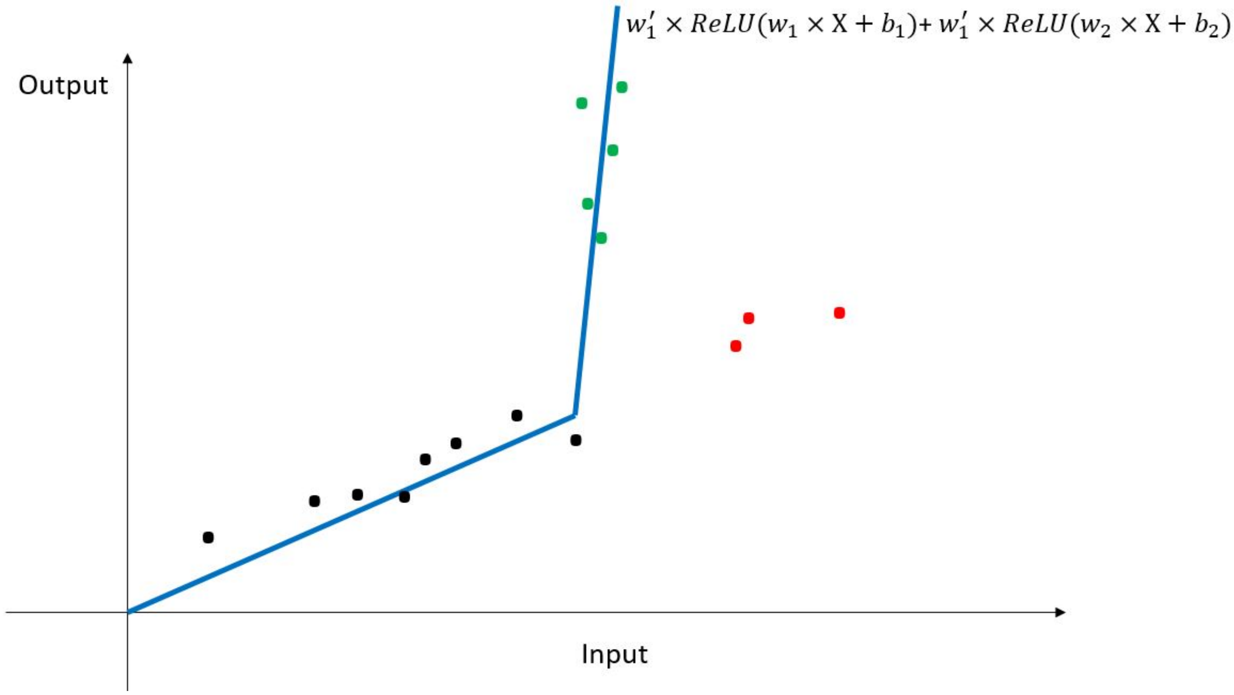

Figure 3.8: Single layer and two neurons neural network forecasting. The blue line is the trained output function which can cover dots but will miss the red dots.

From Figure 3.7 and Figure 3.8 it can be see that even though the improved neural network can reduce the total error of the estimated output function by adding extra flexibility, it will still miss the red dots which can be very extreme peak events. If we do not rely on either of them alone but instead utilize information from the two estimated output functions together, which is the idea of MRTPF method, then the combined estimated output function will be as shown in Figure 3.9. In Figure 3.9 both green and red dots are considered and covered.

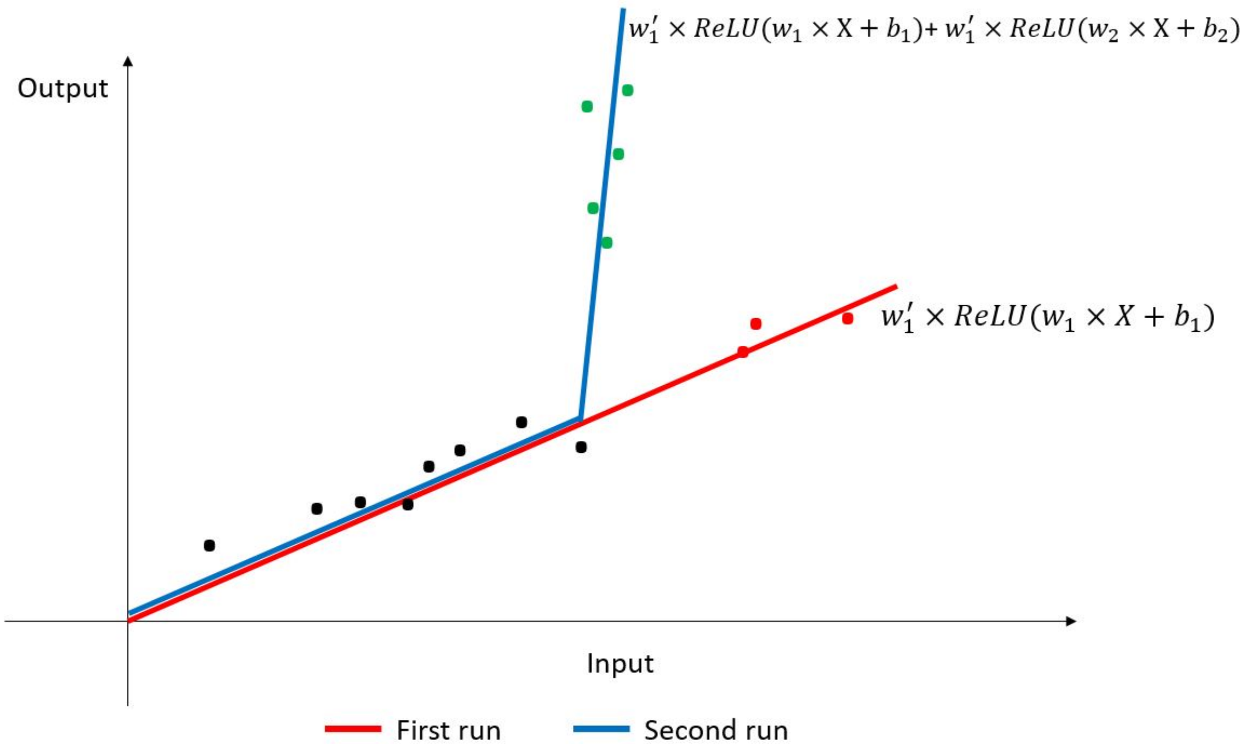

Figure 3.9: MRTPF forecasting. This simple MRTPF model contains two neural networks and trained output functions from both (red and blue lines) can cover most dots, which performs better than any single of the two neural networks.

The above discussion of MRTPF, suggests the following hypothesis: In a complicated system no single neural network model can forecast all the peak events, but by adjusting neuron numbers the combination of the models may be able to forecast more peak events than can any single model. The basic idea of the MRTPF method is to combine the forecast of peak events from a collection of heterogeneous models. This approach is akin to "boosting" [15, 66] in that the results of multiple forecasts are combined. However, standard boosting involves combining forecast values, typically as a weighted average, whereas here the approach is to combine the forecasting of the times of peak.

The MRTPF method has significant potential advantages, particularly in the context of peak event forecasting methods. For example, in electricity markets an indication that a peak-electricity-price is likely to occur is extremely helpful, even without any precise probability associated with it. This is because there are generally significant downsides to missing a peak-electricity-price event, but relatively lower downsides for a false positive. The MRTPF approach is well suited to detecting conditions of peak-electricity-prices as will be demonstrated in the case study in the next section.

3.4 Testing Using ERCOT Data

As in Chapter 2, ERCOT historical market data will be used for a case study of the MRTPF method. Different from Chapter 2 that uses Day-Ahead information, this chapter will use hourly averaged information (Price, Net Load, and Wind) for study.

Electricity price is time sequential. For ERCOT Real-Time operation, the LMP at each electricity bus is calculated every 5 minutes and the market is settled based on 15-minute averages of the 5-minute prices. It is useful to forecast at the 15-minute timescale, but for practical purpose, generation owners in the market normally plan over hours, matching the slower time constants of generator start-up and shut-down. Hourly data is used for demonstration of the idea, and 15-minute forecast will be tested in the future research.

There is not a clear definition of a boundary between peak-electricity-prices and non-peak-electricity-prices in wholesale electricity markets, but most generators' costs are below \$40/MWh without considering the startup cost (startup cost can range from hundreds to thousands of dollars). In this chapter we define a price over \$100/MWh is peak-electricity-price.

Although the peak-electricity-price events are only occasional, about 400 times out of 8,760 price intervals per year since 2015, there is therefore still a rich set of peak-electricity-price events. Moreover, the consequences of peak-electricity-price events can be quite severe, justifying effort to forecast the times of peaks. As mentioned in Chapter 1, ERCOT Real-Time prices hit the \$9,000/MWh price-cap continuously and stayed at the price-cap for hours during the week of August 12, 2019. Considering the average price in 2017-2019 in ERCOT Real-Time Market is less than \$30/MWh, that week can have more impact to ERCOT market participants than the total of the rest of 2019.

The rest of this section is organized as follows. Section 3.4.1 discusses input features selection for the ERCOT market data, while Section 3.4.2 describes the formulation of the MRTPF model and the results for this specific data set. Section 3.4.3 provides a discussion of using the predictions of the occurrence of peak-electricity-price periods.

3.4.1 Input Features Selection

Inputs selection is very important for peak events forecasting. There is no possibility to forecast peak events precisely if the key input features are not selected.

It is widely believed that wind generation fluctuations have a big impact on the peak-electricity-prices in the ERCOT Market [21, 37]. As wind installation grows each year in ERCOT Market, the wholesale electricity prices will continue to fluctuate and the effect of wind on price fluctuations will increase.

Another important input feature is the Net Load of the system. From basic microeconomic principles, there is an increasing relationship between price and Net Load in the electricity market, all else equal. Table 3.1 shows an example of the data from ERCOT website. The last two columns are calculated based on the Net Load and Wind Generation. These columns show, respectively the percentage changes of Net Load and Wind Generation compared to the previous hour.

Table 3.1: Input features for MRTPF. Each column corresponds to an input feature.

Time	HUB Price ($/MWh)	Net Load (MW)	Wind (MW)	% change of Net Load	% change of Wind
01-Jan-16 03:00:00	14.98	28368	4665	-1.21	4.63
01-Jan-16 04:00:00	15.15	28402	4691	0.12	0.56
01-Jan-16 05:00:00	15.53	29063	4678	2.32	-0.28

If we use the raw data of Net Load and Wind Generation as the input features to forecast the price, the forecast versus actual is as shown in Figure 3.10. Interestingly, if we use the percentage change of load and wind compared to the previous hour as the input features to forecast the price, the forecast versus actual is as shown in Figure 3.11. The forecasting performance improved significantly compare to that in Figure 3.10 in sensing the price changes.

Time	Price	Net Load	Wind	% change_N	% change_W
01-Jan-16 03:00:00	14.98	**28369**	**4665**	-1.21	4.63
01-Jan-16 04:00:00	15.15	**28402**	**4691**	0.12	0.56

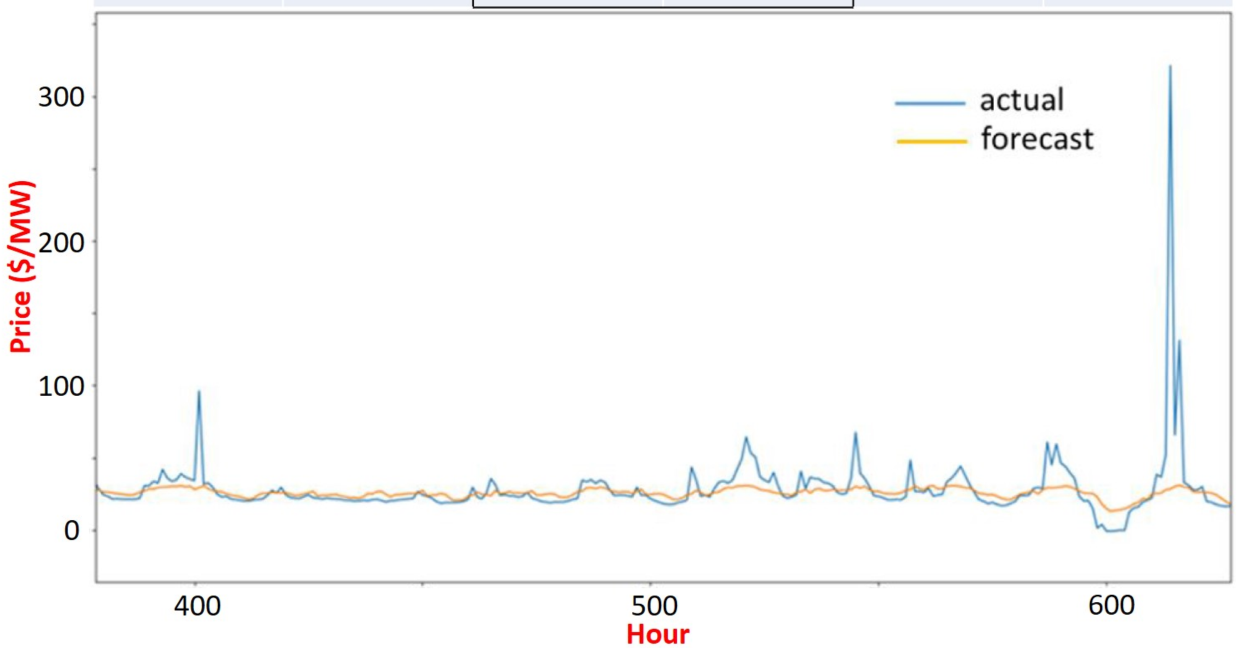

Figure 3.10: Forecast using net load and wind. The forecast has smooth turnings everywhere.

Time	Price	Net Load	Wind	% change_N	% change_W
01-Jan-16 03:00:00	14.98	28369	4665	-1.21	4.63
01-Jan-16 04:00:00	15.15	28402	4691	0.12	0.56

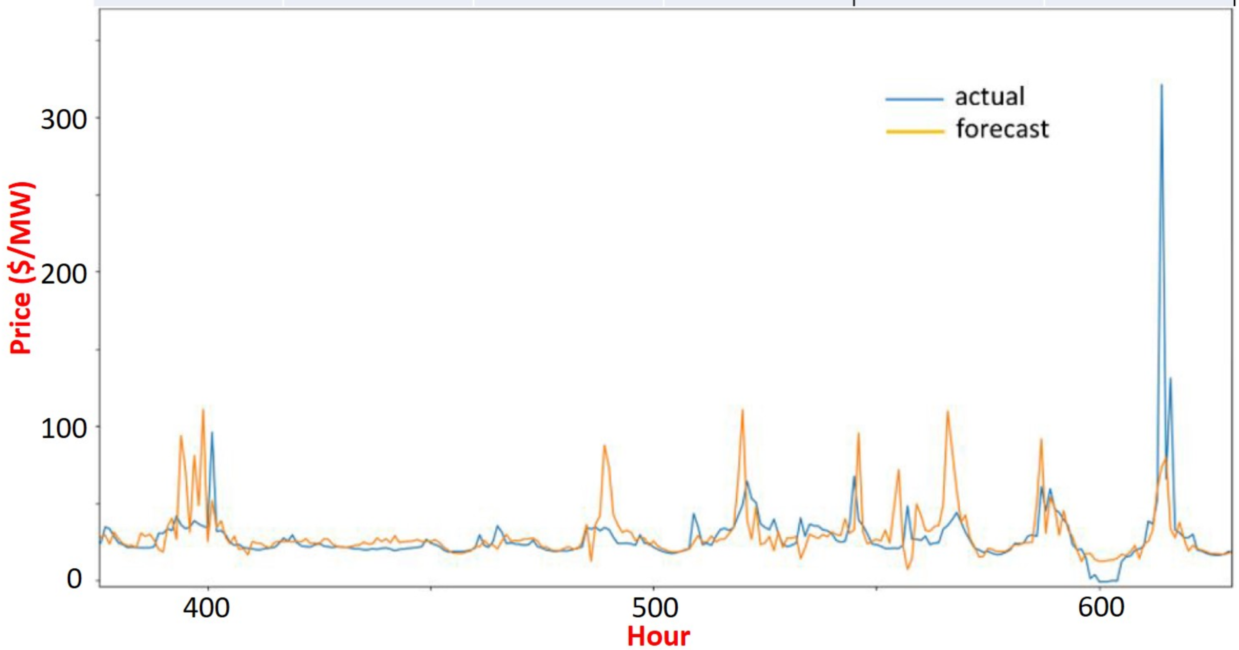

Figure 3.11: Forecast using change percentages of Net Load and Wind.The forecast has cuspate turnings which track the peak-electricity-prices better.

From the above results we can see the percentage-change data can forecast peak-electricity-prices better than the raw Net Load and Wind Generation data. We can say that peak-electricity-prices are mostly driven by the percentage changes of Net Load and Wind in the system. This can also be interpreted as suggesting that there are many factors, in addition to the level of Net Load, that affect prices and that considering changes is allowing the estimation to adapt to these varying conditions without explicitly modeling them.

3.4.2 Formulation of the MRTPF Model

In this testing, a simple ANN like in Figure 3.1 is used, with an input layer, 4-hidden-layers, and an output layer to generate output. The arrows show the occurrence of peak-electricity-prices that are forecasted by the models. The two graphs in Figure 3.12 show the forecast vs actual when each hidden layer has 100 neurons and 600 neurons, respectively. The two different ANN configurations cause the models to have sensitivity to different occurrences of peak-electricity-prices. That is, the ANNs of different configurations can track the peak-electricity-prices caused by different phenomena.

The hourly data from the years 2016-2018 from the ERCOT website (www.ercot.com) will be used for study. A total of 1,300 hours during the end of 2018 are used to test the performance, while the previous hours are used to train and validate the ANN models.

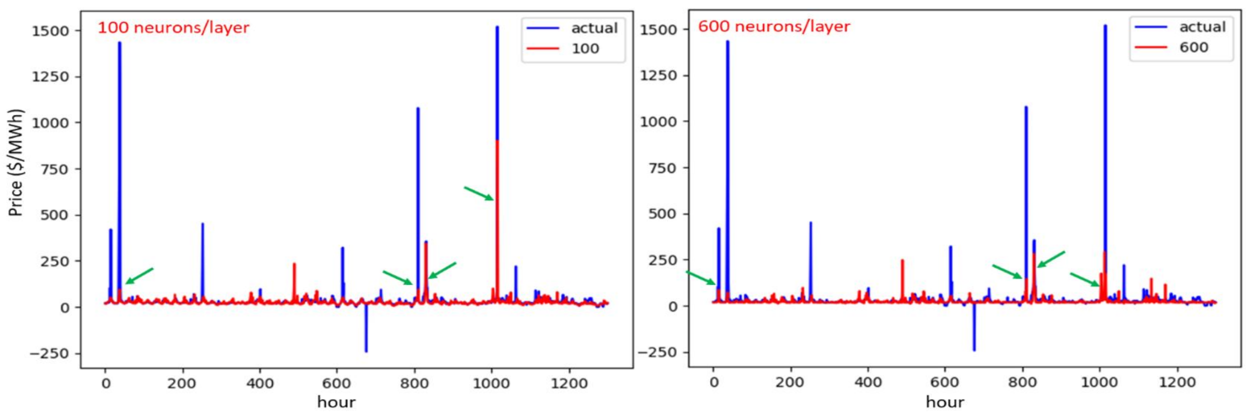

Figure 3.12: Results of Four-hidden-layers ANNs with 100 neurons (left) and 600 neurons (right) at each layer. The green arrows track how the neural networks capture the peak-electricity-prices. The two neural networks can track different peaks.

Four different ANN configurations were tested, and the results are shown in Figure 3.13. Each configuration can forecast several times of peak-electricity-prices but none of them can predict all of the times of peak-electricity-prices. Note that increasing the number of neurons does not monotonically increase the number of peaks detected. Moreover, different peaks are detected by different numbers of neurons. Consequently, the combination of the forecasts can cover more times of peak-electricity-prices than is possible with any single ANN.

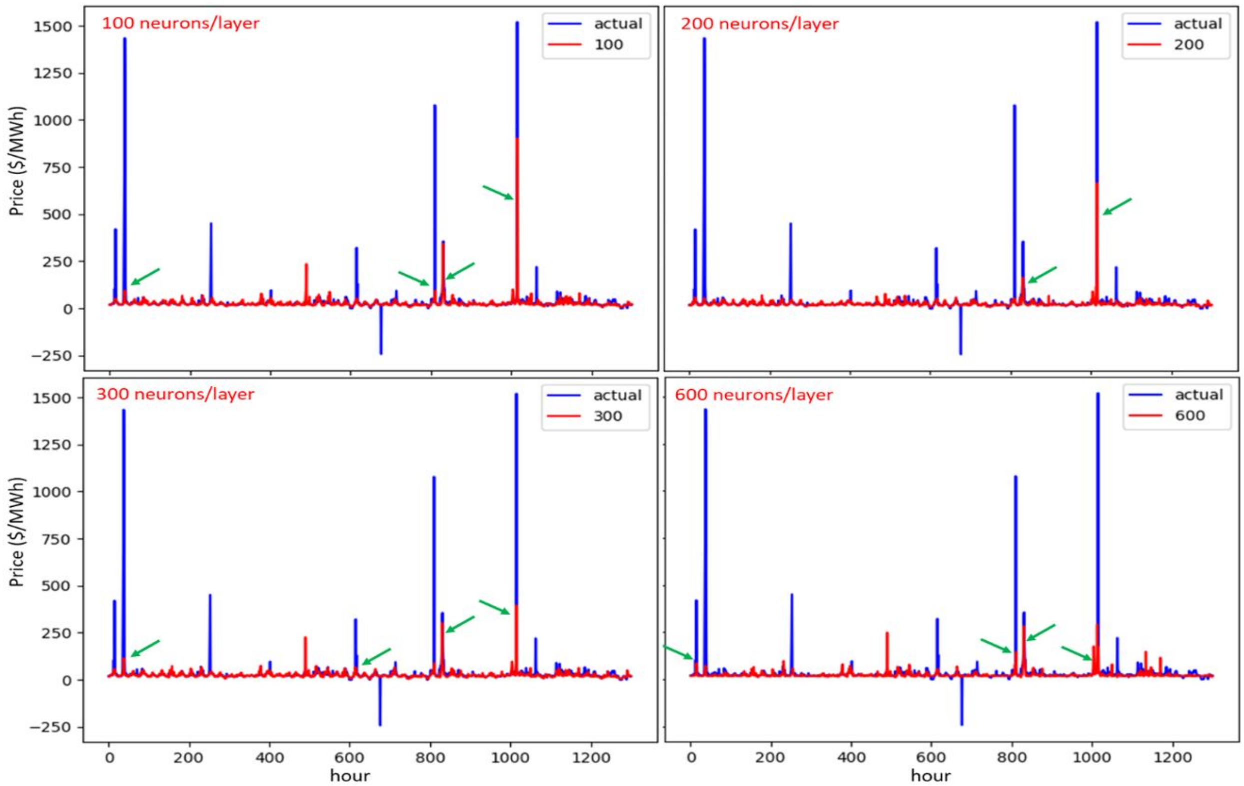

Figure 3.13: Results of four ANNs with different neuron configurations. The performance does not keep improving as neuron number grows, but the combination of them is better than any single neural network in capturing the peak-electricity-prices.

Figure 3.14 shows the results of 16 configurations of ANN models compared to the actual values.

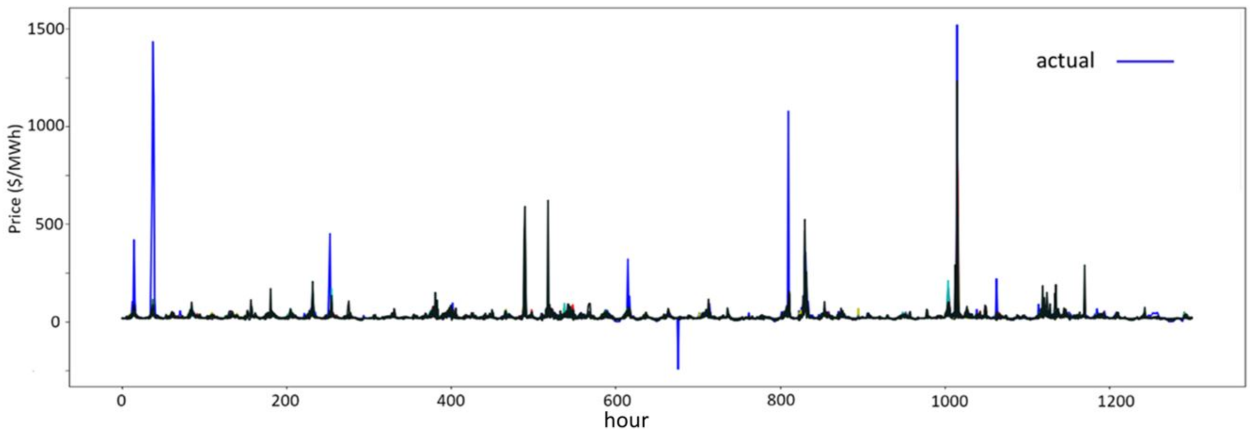

Figure 3.14: Results of ANNs with different neuron configurations. The blue color is the actual values while other colors are forecasts from different ANNs. If any color covers a big part of the blue color at the peak hour then the peak is forecasted successfully.

If the maximum forecasting result of the 16 models is selected for each hour, then the result will as shown in Figure 3.15, which is used as the final result of the MRTPF model. Figure 3.15 shows that almost all times of peak-electricity-prices are captured by the ensemble of ANNs. Only the visible blue peaks not covered by any red were not predicted. This includes the two blue peaks during hour 616 and 1063.

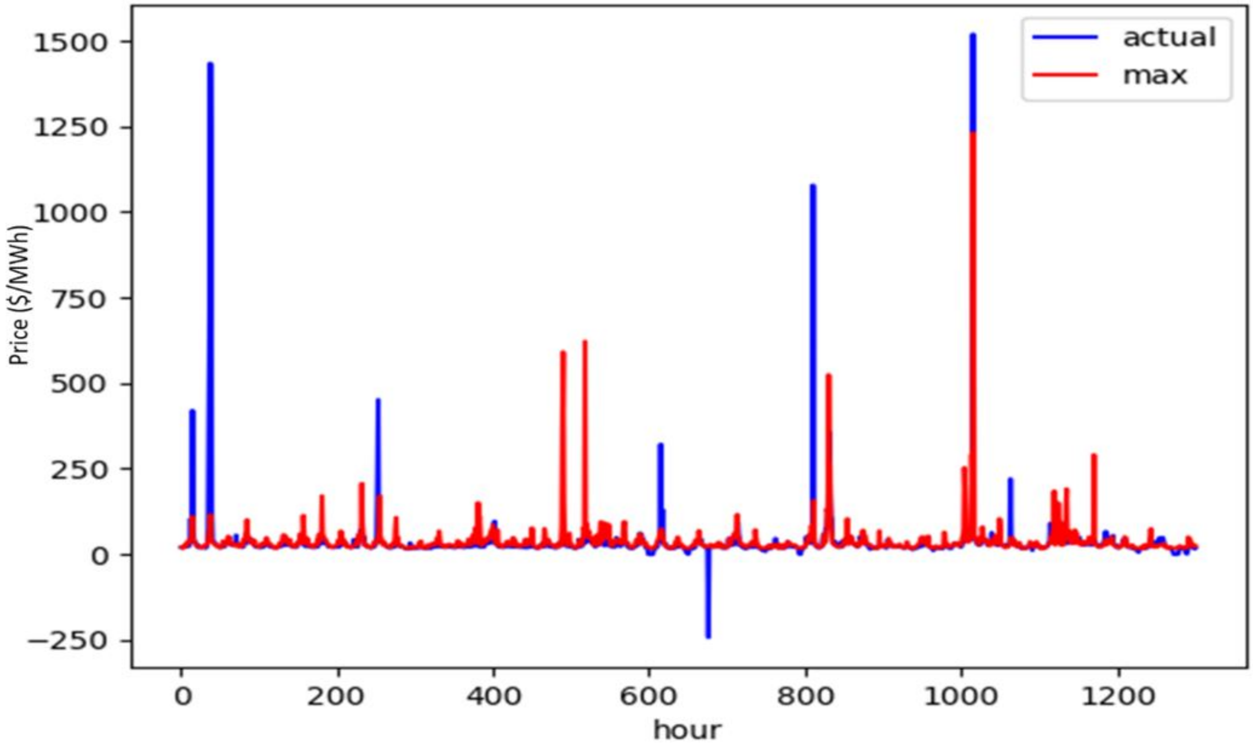

Figure 3.15: Final result of MRTPF to forecast the time of the peak prices. The red color is the maximum of all the results from the ANNs of the MRTPF model.

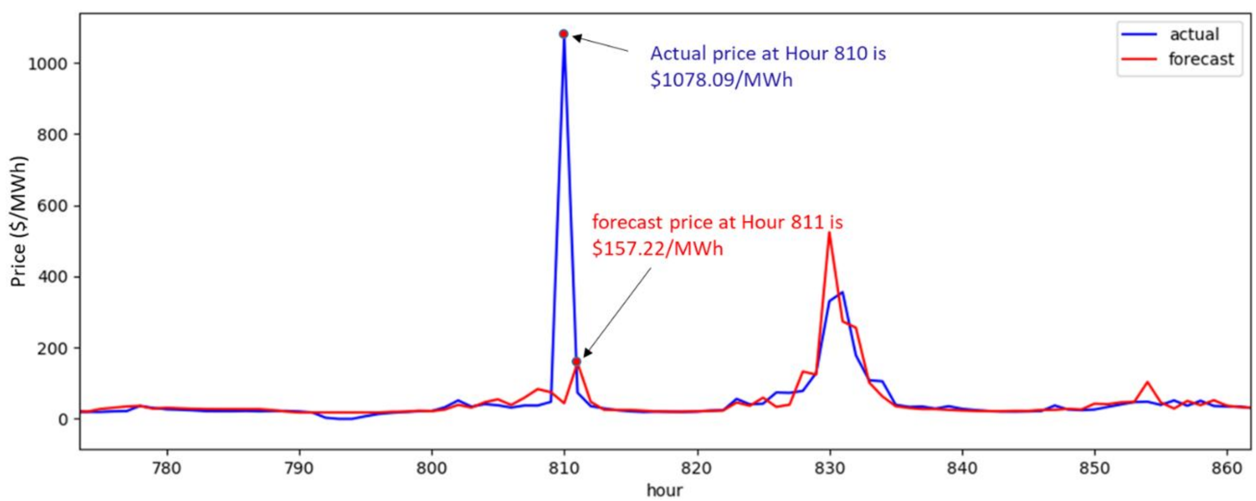

Figure 3.16: Peak-electricity-prices (blue) and the forecast (red). Peak-electricity-prices can last for minutes or a few hours. It is acceptable to forecast a part of the peak event, or a time that is very close to the peak event.

Since generators are usually committed for a few hours because there is a startup cost and it takes time (few minutes to hours depending on the kind of the generator) to startup and shutdown, if the forecasted peak hour is very close to the actual peak hour then it can be counted as a successful forecast. In this Chapter it will be counted as a success if the forecasted peak hour is within 3 hours of the closest actual peak hour. Figure 3.16 shows that the MRTPF model forecasted a peak-electricity-price at hour 811 while the actual peak-electricity-price is at hour 810. In this case the peak hour 810 can still be considered as successfully forecasted.

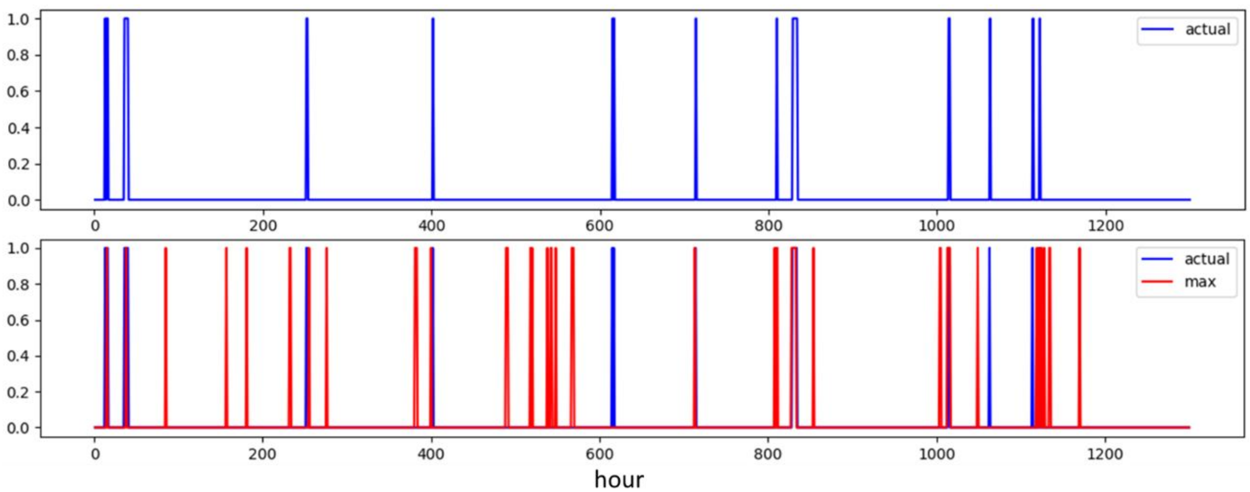

Figure 3.17: Times of peak-electricity-rices (blue) and the forecast (red). This is generated from Figure 3.16 which just shows the time of the peak-electricity-prices and how the MRTPF forecast captures it.

Figure 3.17 shows actual and forecasted time of peak-electricity-prices, which eliminated the magnitude of the prices. There are some false forecasts (22 false positive forecasts and 2 false negative forecasts total in 1,300 testing hours), and many of them are during hour 500 - 600 (9 false positive forecasts) and 1100 - 1200 (2 false positive forecasts). The actual price distributions of the false positive forecasted hours are shown in the box plot (Figure 3.18), in comparison with the whole price distribution during all the tested hours. Most prices of the false positive hours are above \$40/MWh, which is well above the overall average price in ERCOT of about \$24/MWh. For generation planning purposes, most types of generators can make a profit if they run during the false positive forecasted hours. Also, the prices during the two periods are at a very high unstable level, which could have been easily triggered into peak-electricity-prices under some circumstances, so it is prudent to trust the forecast and treat the hours as peak-electricity-price hours for generation planning.

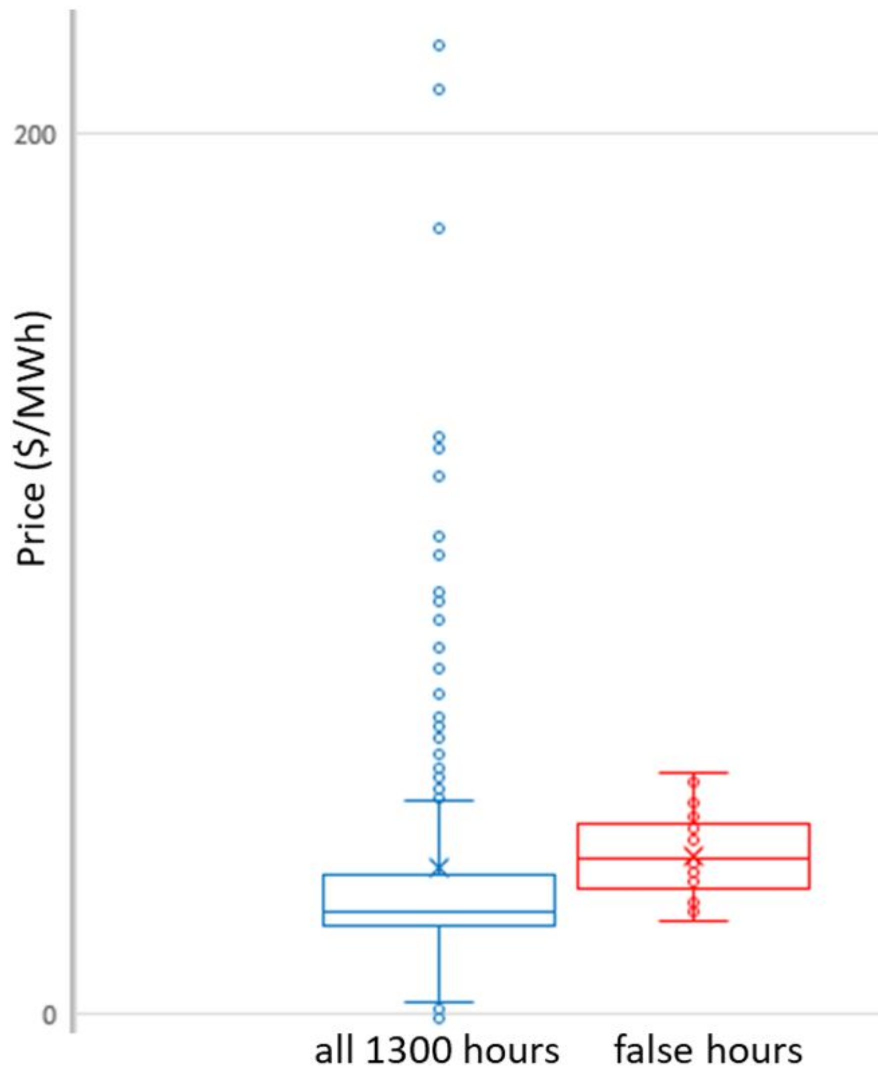

Figure 3.18: Price distributions of all the hours and false positive forecasted hours.

MRTPF is focusing on forecasting when the peak-electricity-price (>$100/MWh) will happen, rather than the price itself. This means if the actual price is over $100/MWh when a peak-electricity-price is forecasted, then the forecast approach is a success. For both ISOs and market participants, false positives for peak-electricity-price forecasts are not particularly harmful. As it happens, however, there are very few false positive forecasts based on the result from MRTPF, and they only show up when the prices are relatively high (>$50/MWh). Slightly different conditions might possibly have developed into higher prices, so from the risk perspective it is better to mark them out as potential peak-electricity-price hours.

3.4.3 Use of the Forecast of MRTPF

This section describes a hypothetical application of forecasting peak-electricity-price periods. Assume a market participant has a 50MW capacity gas turbine generator with $30/MWh cost available through the testing 1,300 hours period. Startup cost is ignored in this example. If the generator generates through all the 1,300 hours, then the total profit over the cost will be $205,219. Yet if the generator only generates during the forecasted 50 peak-electricity-price hours (including 22 false positive forecasts), and is shut down during the rest of the hours, then the total profit will be $233,815, which is even more than the former profit because it has avoided producing during low price periods. The real situation will be more complicated than this simple example, and so this example just demonstrates the benefit the method can bring to generation planning.

The definition of the value of Peak-Electricity-Price will affect the forecasting accuracy and number of false positives. The numerical experiment was rerun with different thresholds for the Peak-Electricity-Price. Table 3.2 shows the summary of total profit over the cost for different threshold settings. All the profits are better than the profit of running through all the 1,300 hours. Varying the threshold results in the performance tracing out the typical "receiver operating characteristic curve" trade-off between false positive and false negative rates.

Table 3.2: Summary of profits of different thresholds.

Threshold of the peak price ($/MWh)	Total forecated hours (including falses)	Total actual peak hours	Correctly forecasted hours	False Positives	False Negatives	Total profit to run at forecasted hours ($)
70	74	30	28	46	2	243,977
60	108	39	32	76	7	316,139
50	166	62	54	112	8	429,170

3.5 Chapter Summary

MRTPF is an advanced idea in forecasting the time of peak events, which is very applicable and easy to understand. It has advantages over the existing single-ANN based forecasting methods, and it provides more certain information than probabilistic forecasting methods on peak-electricity-price forecasting and can be evaluated easily.

From the test, the MRTPF method can forecast the time of peak-electricity-prices well. Most times of peaks were forecasted and acceptable false positives and negatives were generated.

This Chapter focused on a demonstration of the idea of MRTPF, but not on accuracy achievement. By adding more inputs, and replacing the simple ANN with more advanced neural network models like LSTM, or other machine learning models, the performance of MRTPF can be improved further.

The following Chapter 4 will introduce another time-of-peak-price forecasting method, which transfers the long-range raw price data into simple binary values 0 and 1 in order to improve the forecasting performance. It presents another strategy to improve forecasting, that is through data study and processing rather than improving the models.

Chapter 4

Forecasting the Time of Peak of Electricity Market Price Using Artificial Neural Network with Binary Representation of Input Data[1]

As discussed in Chapter 1, many peak-electricity-price forecasting methods have been discussed in the literature review Section 1.7. New approaches have also been proposed in the previous chapters. But all of them used the raw price values in the forecasting, which can spread over a big range in electricity markets from negative prices up to around $9,000/MWh. Chapter 3 brought the idea to forecast the time of peak-electricity-price rather than the price itself, and this Chapter will try to explore this further through pre-processing the price data to make it more understandable to neural network models.

Chapter 4 has 4 sections. Section 4.1 introduces the new binary forecasting method. Section 4.2 discusses the design of the model. Section 4.3 tests the method using the real market data from ERCOT. The last section gives conclusions about the studies.

4.1 Introduction

The method proposed in this chapter is called Binary Time of Peak-Electricity-Price Forecasting (BTOPEPF). It will translate the raw price values into simple binary data, 0 and 1, depending on whether the price is below or above a predefined threshold, which will significantly reduce the complexity of the price data, which otherwise can range from

[1]This Chapter is based on the paper "Binary Forecasting Method of the Time of Peak Electricity Price based on Artificial Neural Network" accepted to the 1st IEEE Sustainable Power and Energy Conference (iSPEC), November 2019, by Jian Xu and Ross Baldick. Jian Xu is the first author of the paper. Dr.Ross Baldick is the co-author and supervised the work.

negative to thousands of dollars per MWh. The method is aimed at forecasting the time of peak-electricity-price rather than the peak-electricity-price itself. The usefulness of this new forecasting method is that, as discussed in Chapter 3, electricity market operations are concerned more about the time when the peak-electricity-price will happen rather than about the exact magnitude of the peak-electricity-price. This method is intended to be very easy to apply, practical to use, and very easy to understand and evaluate.

4.2 Model of Binary Forecasting of Peak-electricity-prices

The method will use a simple DNN model as in Figure 3.1 from the previous chapter to demonstrate the idea, which has an input layer, a 4-layers hidden layer, an output layer, and a forecasting value from the output layer at each time step. As in Chapter 3, hourly averaged ERCOT Real-Time Market data will be used in this chapter for demonstration of the idea.

For this research a price that is over \$50/MWh is defined as peak-electricity-price, which occurs quite often in the market, providing for a rich data set. In contrast, if the peak-electricity-price is defined to be over \$100/MWh then this only occurs about 400 hours every year in the recent past, which as disscussed in Chapter 3, is a limitation to train the neural network model. It is acknowledged that choice of boundary is application and market-specific, so that a different boundary may be suitable, for example, for the California Market.

The binary peak-electricity-price forecasting idea is demonstrated in Figure 4.1. The raw price data is translated into binary data for every hour, and this binary data is used as the input instead of the raw price data to train the model and test the forecasting accuracy. The forecasting output from the binary forecasting model can be interpreted as a proxy of the probability of the occurrence of the peak-electricity-price (1 means there will be a peak-electricity-price, 0 means there will not be a peak-electricity-price, 0.5 means there will be 50% chance to have a peak-electricity-price, at the hour).

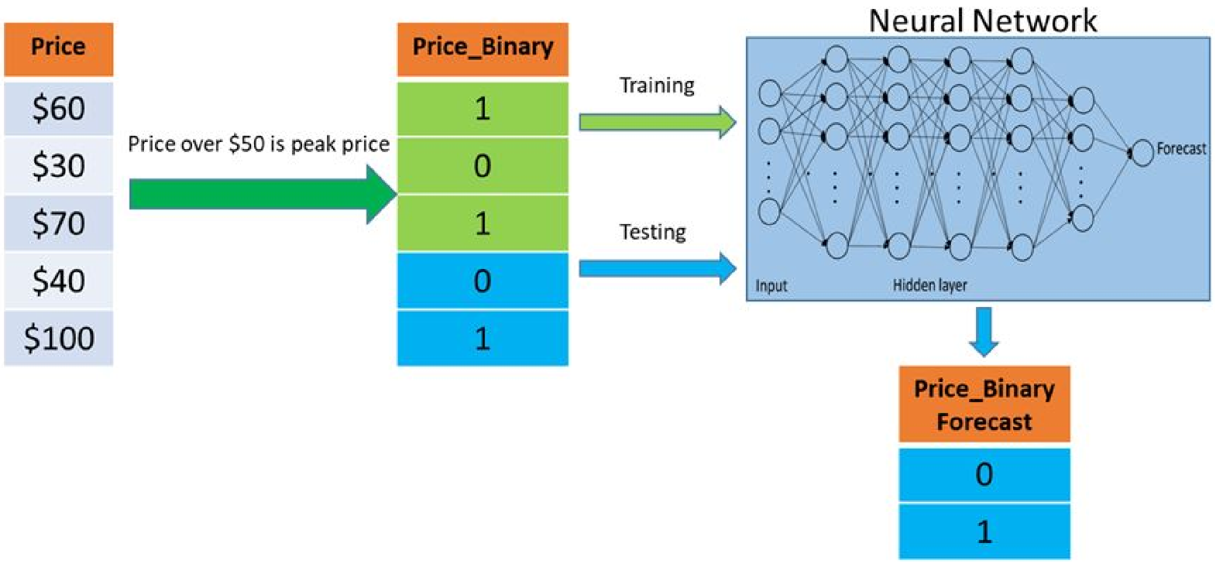

Figure 4.1: Data flow in the binary forecasting model.

4.3 Test of the Binary Forecasting Model

ERCOT Market hourly historical data of HUB SOUTH during 2015 - 2018 from the ERCOT website (www.ercot.com) is used to train and test the model, which is the same data as used in Chapter 3 except that price data has been converted into binary values. In this study a simple DNN with 4-layers hidden layer of 200 neurons each, and an output layer of 100 neurons is used. The training epoch is set to 500. The data structure is shown in Table 4.1. Net Load and Wind generations are both input features while the price is the output feature.

Table 4.1: Example of the data structure from ERCOT Market

Time	Price ($/MWh)	Binary Price	Net Load (MW)	Wind Generation (MW)
04-Jan-16 05:00:00	21.81	0	35195	2431
04-Jan-16 06:00:00	66.80	1	4691	39742
04-Jan-16 07:00:00	24.46	0	44197	1347

However, for the BTOPEPF method, the classical evaluation method such as MAE alone is not completely useful, since the result is a binary estimate. An easy way to evaluate the BTOPEPF method is to count how many times it forecasted the time of peak-electricity-price correctly, and a very close forecast should also be counted as a success. For example, the situation illustrated in Figure 4.2 around hours 336 to 339 should be considered a successful forecast. As discussed in Section 3.4.2 in Chapter 3, some peak-electricity-prices can last for hours while others only last for few minutes and so a forecast peak time within three hours of the actual peak will again be counted as a success. A generator that is committed based on the forecast will likely be operational during the high price period, even if the exact time of the peak-electricity-price is displaced from the prediction.

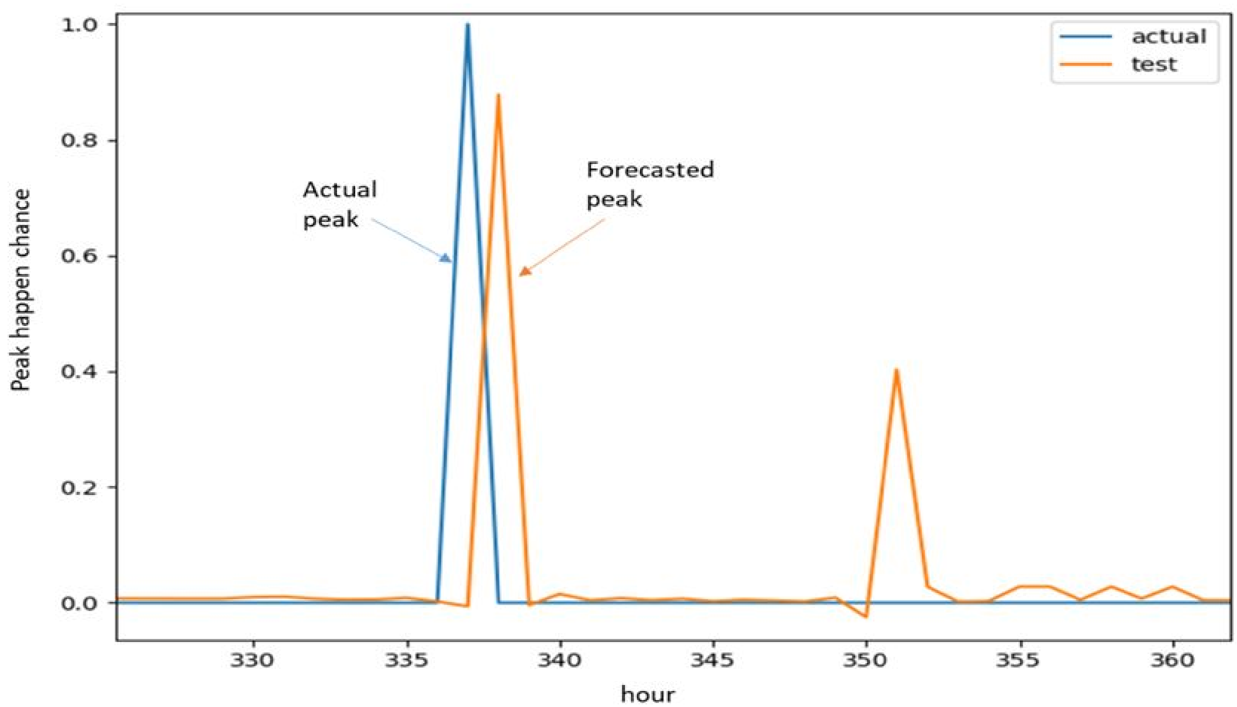

Figure 4.2: Acceptable forecasting error.

The training and validation curves of the BTOPEPF method, and the classical method that uses the same data format as the testing in Chapter 3, are shown in Figure 4.3. The curves at the top are of the classical method that uses raw price data for training, while the curves at the bottom are of the BTOPEPF method that uses the transformed binary data for training. From the comparison, the BTOPEPE method has smoother training and validation curves, which are signals that the DNN models can understand the binary price data better, compared to the raw price data during the model training. The better the DNN model is trained, the more accurate the forecast will be.

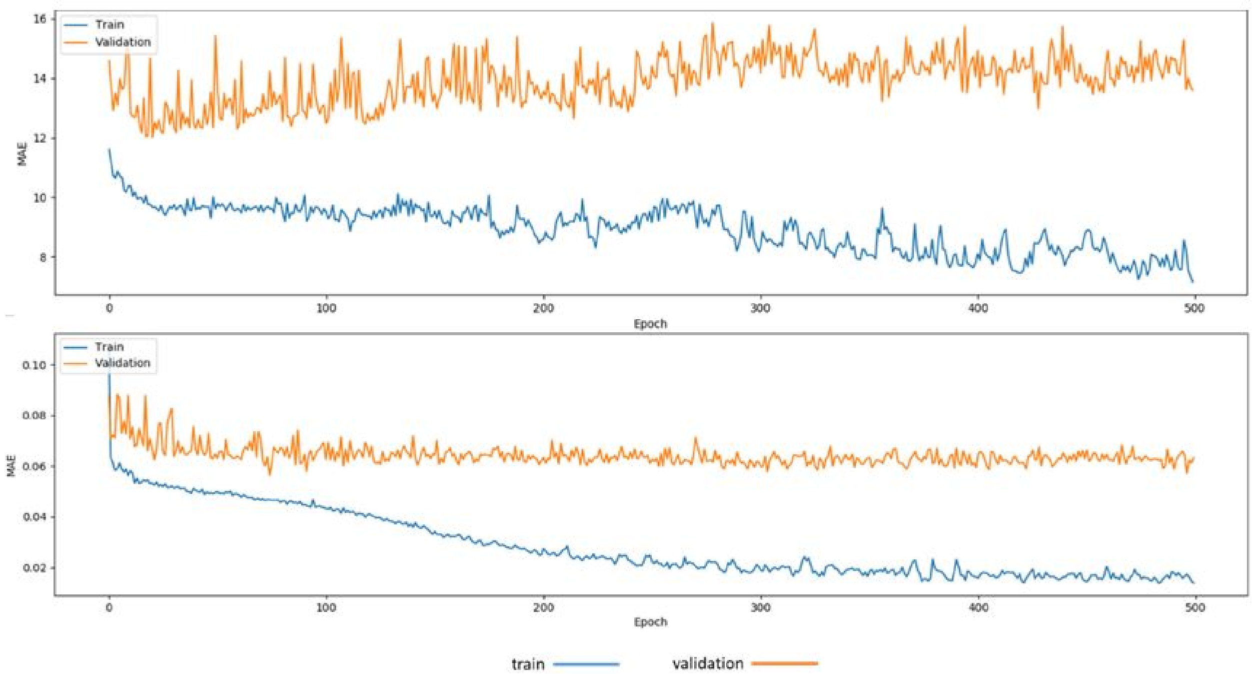

Figure 4.3: Training and validation curves of the two methods. The top graph shows the method using raw price data while the bottom graph shows the method using binary price data.

The forecasting results of the classical and the BTOPEPF methods are shown in Figure 4.4 and Figure 4.5 respectively. During the 700 testing hours there are 13 peak-electricity-prices that are greater than \$50/MWh. From Figure 4.4 it can be observed that the classical method captured the time of peak-electricity-prices 4 and 5 (when forecasts are greater than \$50/MWh), resulting in a peak forecasting accuracy of 2/13, defined to be the ratio of detected times of peak-electricity-prices to the actual total number of times of peak-electricity-prices.

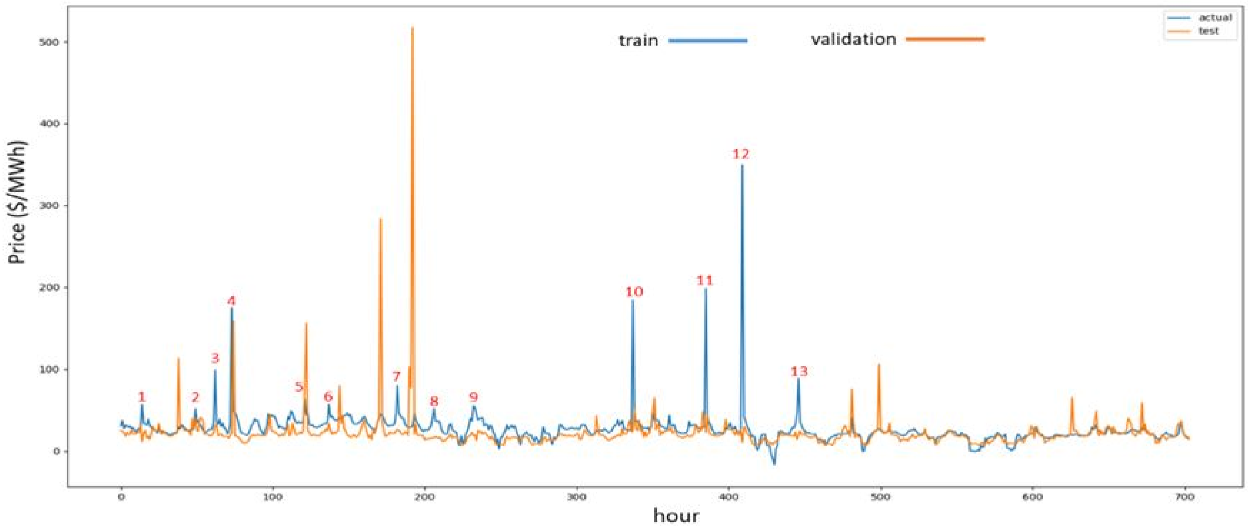

Figure 4.4: Forecast using raw price data.

From Figure 4.5 the BTOPEPF method captured time of peak-electricity-prices 3, 7, 10, 11, 12 and 13, resulting in a peak forecasting accuracy of 6/13, which is a significant improvement compared to the classical method.

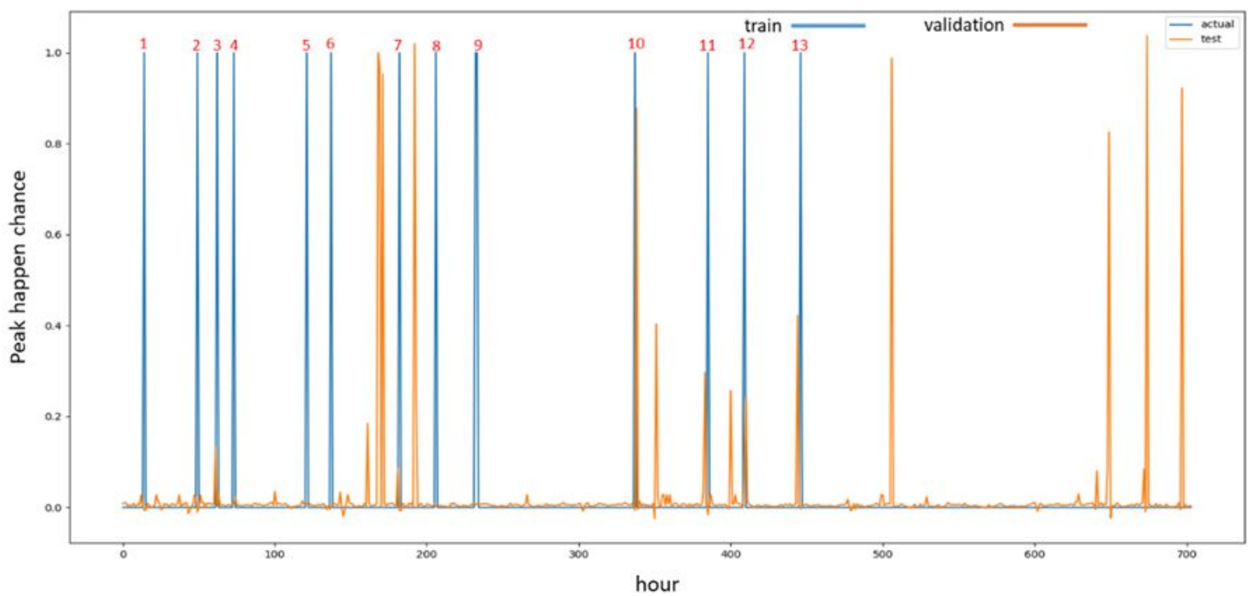

Figure 4.5: Forecast using binary price data.

As in Figure 4.5, at time of peak-electricity-price 3, the forecasted value is about 0.2. As discussed earlier, 0.2 can be interpreted as 20% chance the peak-electricity-price will happen. In real situation, even 5% chance of peak-electricity-price is worthwhile to predict because if the price goes to \$9,000/MWh, a 500 MW capacity generator can make 4.5 million dollars in one hour.

In peak-electricity-price forecasting, the false-positives are much less costly than false-negatives as the generator would not lose much money by running at a false-positive hour even if the electricity price is very low. However, they can miss a million-dollars opportunity by not running during a high peak-electricity-price (can be as high as \$9,000/MWh in ERCOT Market).

Table 4.2 shows the detail comparison of the classical method and the BTOPEPF method in the performances of forecasting time of peak-electricity-prices. Overall the BTOPEPF method is superior to the classical method in forecasting the time of peak-electricity-prices. From the test using ERCOT market data, the BTOPEPF method has two more false-positive forecasts than the classical methods using the inputs as specified in Chapter 3; however, this can be neglected considering the huge improvement in detecting the time of peak-electricity-prices. Also, the false-positives are much less costly than false negatives, as stated earlier.

Table 4.2: Comparison between the classical and the BTOPEPF methods

Method	Accuracy	False Positives	False Negatives
Classical	2/13	8	11
BTOPEPF	6/13	10	7

The forecasting accuracy and number of false positives also depend on the setting of the threshold for the forecast. If the threshold moves from 0.1 to 0.2, the forecasting accuracy will drop from 6/13 to 4/13 and false positives will drop from 10 to 7 while false negatives increasing from 7 to 9, which demonstrates the typical trade-off in receiver operating characteristic curves. The selection of the threshold depends on the experience and user preference.

4.4 Use of BTOPEPF in Market Operation

The BTOPEPF forecasting method can be used in the planning of generations in real market operations. The method's strength in predicting the occurrences of peak-electricity-prices will especially benefit Day-Ahead Market planning, to make sure the generators will run as much as possible during peak-electricity-prices for tomorrow. As the renewable generations keep growing in the power grids, overall wholesale electricity prices are being pushed lower. Therefore, the operation strategy for conventional generators becomes more important than ever before.

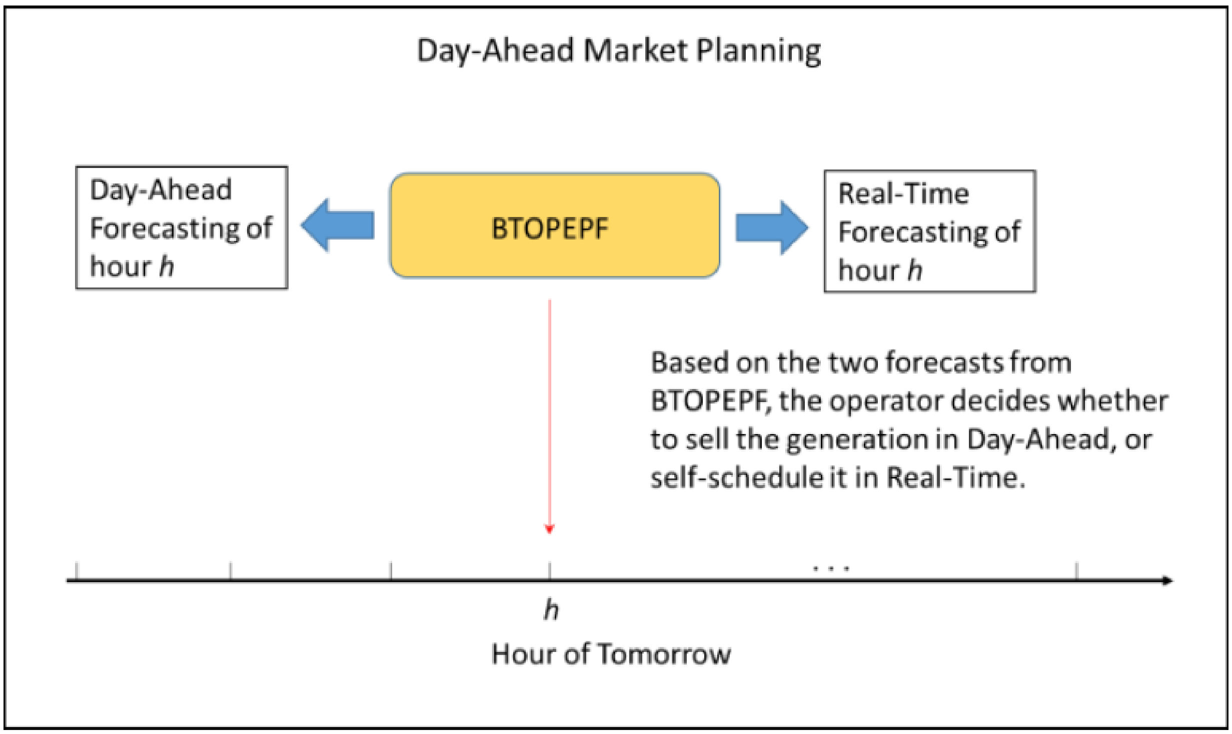

Figure 4.6: Use of the BTOPEPF method in Day-Ahead Market planning.

An application of the BTOPEPF forecasting method is shown in Figure 4.6. For the hour h of tomorrow, the operation day, BTOPEPF will use the Day-Ahead information to forecast the probability of peak-electricity-price at the generation Bus in the Day-Ahead Market and it will use the Real-Time forecasting information to forecast the probability of peak-electricity-price at the same Bus in the Real-Time Market. The two price values are compared to decide whether to sell the generation in Day-Ahead Market today, or to self-schedule the generation in Real-Time Market tomorrow at hour h. Later Chapter 5 will introduce a new forecasting method focusing on Real-Time Market.

4.5 Chapter Summary

The new forecasting method BTOPEPF is superior in forecasting the occurrences of peak-electricity-prices compared to the classical forecasting methods using the raw price data, based on the testing results using real ERCOT market data. The simplification of the price data can help the neural network model learn better about the underlying information behind peak-electricity-prices.

The following chapter will introduce a three-step peak-electricity-price forecasting method which starts from Day-Ahead Market into Real-Time operation. By considering each SCED interval, this method will be very useful for wholesale electricity markets with implementation of RTC, which need to co-optimize energy and ancillary services in every SCED interval during Real-Time operation.

Chapter 5

Three-Step Real-Time Electricity Price Forecast using Recurrent Neural Network[1]

In addition to the Day-Ahead electricity price, the Real-Time electricity price is key information for electricity market participants. As discussed in the previous chapters, accurate forecasting of the Real-Time electricity price can help market participants plan their generation scheduling and trades more economically. Methods of Real-Time electricity price forecasting are discussed in this chapter, and a new Real-Time price forecasting strategy that can enlarge the forecasting window to one day ahead and forecast at a resolution of 5-minute intervals is proposed and tested using market data from ERCOT. Unlike in Section 4.4, where the forecast of Real-Time electricity price was being made before the closure of the Day-Ahead Market, in this chapter the forecast of Real-Time electricity price is being made after the results of the Day-Ahead Market are announced. Although the methods proposed in Chapter 3 and Chapter 4 were applied to Real-Time price forecasting, they can be used for both Day-Ahead and Real-Time forecasting. However, the method proposed in this chapter is designed specifically to forecast Real-Time Market prices.

Chapter 5 has 4 sections. Section 5.1 introduces the new forecasting method to forecast Real-Time prices called TREPF. Section 5.2 discusses the design of model of TREPF. Section 5.3 presents a case study of the method using the real market data from ERCOT. The last section gives conclusions about the studies.

[1]This Chapter is based on the paper "Three-Step Real-Time Electricity Price Forecast using Recurrent Neural Network" submitted to the IEEE Transactions on Power Systems, November 2019, by Jian Xu and Ross Baldick. Jian Xu is the first author of the paper. Dr.Ross Baldick is the co-author and supervised the work.

5.1 Introduction

As discussed in Chapter 1, Real-Time electricity prices depend on load, wind generation and fluctuation, system conditions, including generation status, and human operating habits, making it very difficult to forecast. During the last decade, big data and advanced machine learning technologies, such as recurrent neural networks that have been introduced in the previous chapters, have made it possible to forecast the Real-Time electricity price at a promising accuracy. As the computer power grows, machine learning can run a study every few seconds/minutes to capture the most current changes.

Typical discussions of Real-Time forecasting focus on forecasting in the next hour; however, electricity market participants need to prepare fuel and make operation/trade plans no later than the day before the operating day, so it is very valuable to forecast the next day before the day starts. Updated forecasts on the operating day can facilitate shorter-term planning and operations. This chapter will therefore create a method with three steps that roll forward in use of available data. The first step, using the earliest information, is to forecast hourly averages of Real-Time prices of the next day using available Day-Ahead Market cleared prices and Day-Ahead forecasting information. This method can set the base hourly Real-Time price forecast for the next day, which can be used for Day-Ahead planning. During the operating day itself, Real-Time data becomes available. The second step of the Real-Time price forecasting process then uses the most current hourly Real-Time information to update the price forecast of step one for the upcoming hour of Real-Time operations. There is also a third step of the process that forecasts at the finer timescale of the 5-minute SCED intervals to shape details within the hourly forecast from step two. (The hourly Real-Time price is the average of the SCED interval prices.) This Three-step Real-Time Electricity Price Forecasting (TREPF) method is proposed in this chapter, and it is evaluated using historical data from ERCOT website, as in the previous chapters.

5.2 Design of TREPF

When market participants prepare for future Real-Time Market operations, they prefer to estimate the financial plans days or weeks ahead, in order to purchase fuel at favorable prices, to prepare work shift plans, to prepare for holidays and maintenance, and to make long-term and short-term trading plans. Good planning can help a utility company make improvements in its profits. Step 1, involving rough Real-Time price prediction made in advance of the operating day, is very important because many important decisions must also be taken in advance of the operating day. In order to demonstrate the idea of TREPF clearly, the example in this chapter will focus on one day ahead Real-Time price forecasting, but the method can apply to further ahead forecasting by adjusting forecasting parameters. Furthermore, with increasingly large net load ramping, hourly resolution Real-Time planning is not sufficient anymore. It is worthwhile to explore the forecast of each SCED interval. To cover all the needs, TREPF is therefore composed of three successive steps, Step 1, Step 2, Step 3, as shown in Figure 5.1.

Figure 5.1: Time sequence of TREPF. Step 1 happens in Day-Ahead planning, Step 2 happens hourly ahead and Step 3 happens 5-minute ahead.

Step 1 will give a rough estimation of the next day's hourly average Real-Time prices from hour ending 1 to hour ending 24, using the available information, which is the Day-Ahead Market cleared price (after the Day-Ahead Market clears) and Day-Ahead forecasts of load and wind. With the ERCOT market as an example, if it is desired to use Day-Ahead prices to forecast the Real-Time prices then the forecast must be run after 13:00 Central Time when the Day-Ahead Market is cleared and the Day-Ahead prices become available. Also, all the forecasts (load, wind, weather) for the next day should be available at that time too. All of the information mentioned above can be used to forecast the Real-Time hourly prices of the next operation day.

Day-Ahead Market cleared hourly price and forecasts of wind are used as input features in this chapter. Suppose that a forecast for hour h of the next day is desired. Furthermore, assume that the input time-steps of LSTM is chosen to be 4 hours, which includes the 3 hours previous to the forecast hour and the forecast hour itself ($h-3$, $h-2$, $h-1$, h). The Day-Ahead hourly prices and Day-Ahead hourly wind forecasts for these hours are used as input, while the output is the forecasted average Real-Time price of hour h. After 13:00, for the Step 1 Forecast, all the forecasting information for hour h of the next day are available, so the input should include the forecasts of hour h to forecast the average Real-Time price of hour h.

The Step 2 Forecast will have higher expected accuracy for the same Real-Time hour compared to Step 1 because it uses the most current Real-Time information available on the operating day itself, which is very close to the actual situation in hour h. While the Step 1 Forecast serves as general guidance for future planning, Step 2 Forecast serves as the hourly Real-Time specific operation guide on how to run the generators more economically in the Real-Time Market at the hourly level. For example, how to arrange ancillary services [5] on generation resources, how to respond to Reliability Unit Commitment (RUC) [10] instructions (a generator can choose to opt out of RUC in ERCOT market by self-committing so as to take the opportunity of high price but to give up the make-whole guarantee [9]), how to respond to trading requests from other market participants, etc.

The hours of input features of Step 2 are slightly different from Step 1. Step 2 uses Real-Time information before the next operation hour, but has no Real-Time information about hour h available. The closest hour that has available Real-Time information is $h-1$. The big advantage of the input features in Step 2 is that actual Real-Time information from earlier in the day is available rather than only forecasted information. The actual information can be expected to represent actual conditions much better. The coordination of Step 1 and Step 2 is as shown in Figure 5.2.

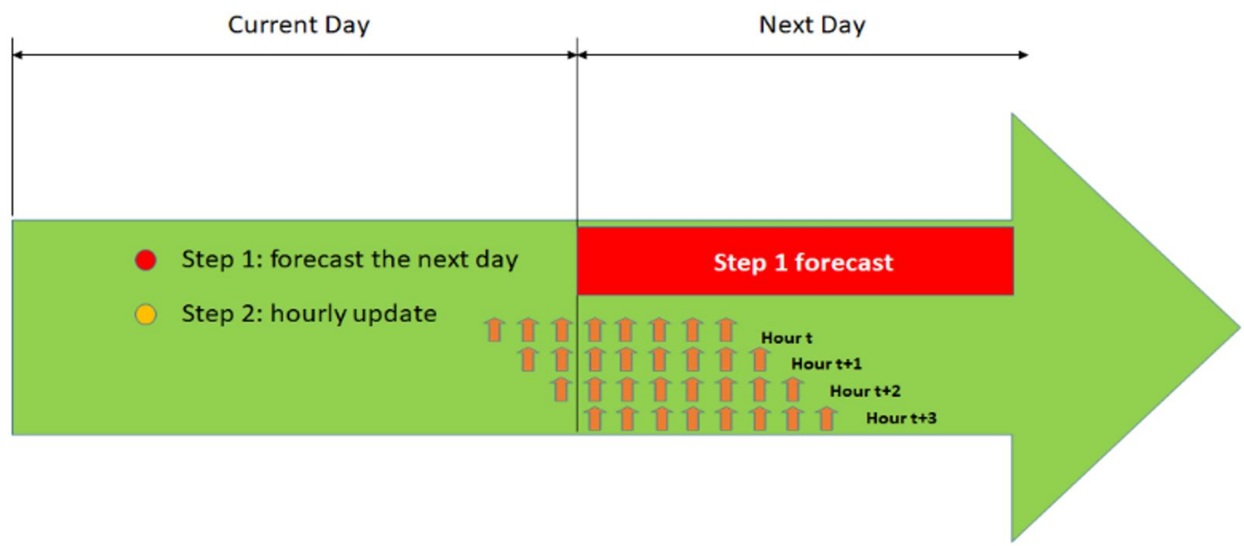

Figure 5.2: Step 1 and Step 2 of TREPF. Step 2 which gets most current Real-Time data inputs will make improvements over the hourly forecast from Step 1 on the future hour.

One of the challenges for Step 2 is that it needs to run very often and to finish within limited time, which has the potential risk to fail. However, if Step 2 fails during Real-Time operation, the results of Step 1 Forecast can still be used as a temporary guide for Real-Time hourly operations, even though its accuracy may not be as high. This is an advantage of TREPF in having a backup step.

Step 3, as shown in Figure 5.3, will forecast at a finer temporal scale, predicting the price in each 5-minute SCED interval, and providing information for the next interval planning. The information is very helpful if the market has implemented Real-Time Co-Optimization [6] between energy and Ancillary Services during each SCED 5-minute interval. The forecasted prices can be used to plan the Ancillary Services on generators during each of the future intervals. The input features of Step 3 are the previous SCED intervals cleared prices and 5-minute wind and/or load information. Step 3 can use the results of both Step 1 and Step 2 as backup if it fails.

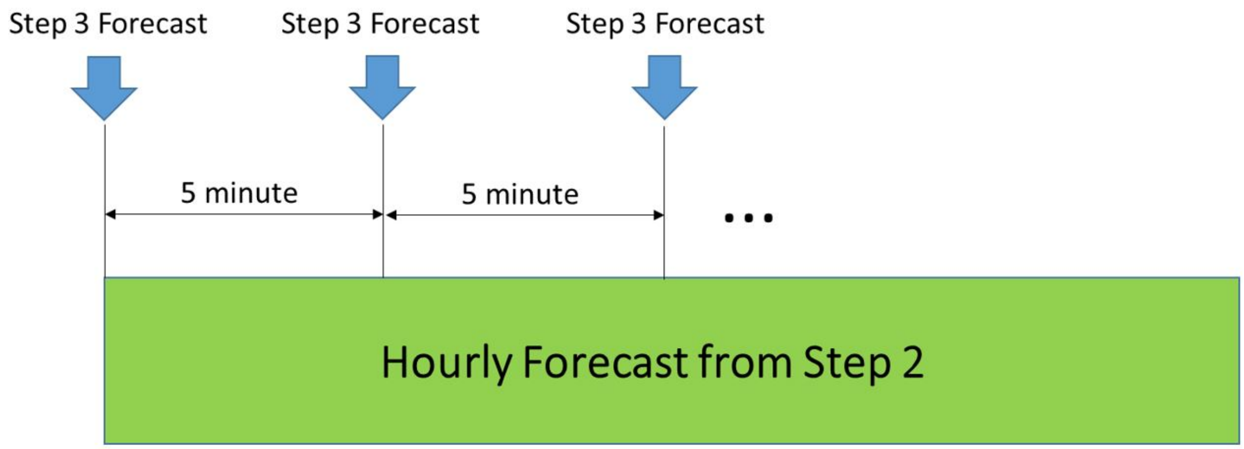

Figure 5.3: Demonstration of Step 3 Forecast of TREPF. Step 2 will make improvements into SCED 5-minute intervals over the hourly forecast from Step 2.

The data flow of all the three steps of TREPF is shown in Figure 5.4. Based on the data availability at the forecasting point, different input features can be selected for each step. As the most import information that will affect prices, the wind, load, historical prices, and weather are the top candidates for input features. But when historical prices, wind, and load are all available at the same time, there can be some overlapping effects since historical prices can reflect information of historical wind and load. This is akin to multi-collinearity in linear models and details about this issue will be discussed later.

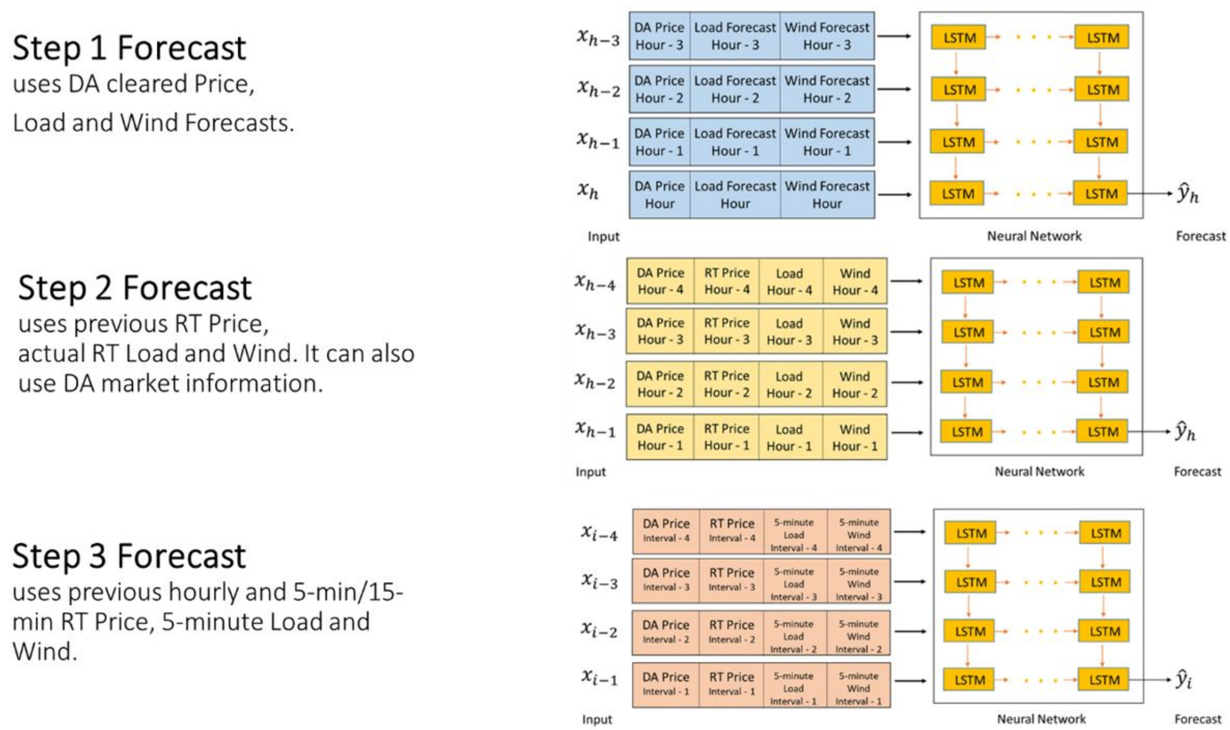

Figure 5.4: Data flow example of TREPF. Step 1 uses input data available in the Day-Ahead planning, Step 2 uses Real-Time data available one hour ahead and Step 3 uses Real-Time data 5-minute SCED interval ahead.

5.3 Case Study of TREPF

ERCOT historical data on SOUTH HUB during 2015-2018 is used for the case study. All the available 5-minute, 15-minute, and hourly wind, load, and price data were collected for this period, which has over 140,000 rows for LSTM to study. Compared to an advanced neural network like LSTM, classical statistical forecasting methods like ARIMA have limitation on understanding non-linearity and stationarity in the large data set and have overall less accuracy of forecasting when there are nonlinear relations between the inputs and outputs [50, 70].

Both $RMSE$ (5.1) and MAE (5.2) are used to evaluate forecasting accuracy:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}|(y_t - \hat{y}_t)^2} \tag{5.1}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|(y_t - \hat{y}_t)| \tag{5.2}$$

Step 1 Forecast is to forecast the Real-Time hourly prices of the next operation day, from hour 1 to hour 24. Theoretically both Day-Ahead load forecast and wind forecast can be used for study. But ERCOT does not publish Day-Ahead load forecasts and does publish the Day-Ahead wind forecast on their website. There are some commercial Day-Ahead load forecasting services in the market, but this chapter only uses ERCOT published data to ensure that results are reproducible. More input features, including non-public sources, could potentially improve price forecast performance, but this would require careful evaluation. For example, ERCOT does publish a 7-day long-term load forecast but it can be significantly different to the Day-Ahead Market conditions because of the big gap of time between the time of forecast and the operating day. To summarize, load forecast is not used as an input feature in the case study.

The ERCOT Day-Ahead Market opens at 10:00 and the Day-Ahead Market cleared prices are published around 13:00. So, if the Step 1 Forecast is made after 13:00, the Day-Ahead Market cleared prices can be used as input feature for Step 1 Forecast. That is, if the forecast is made after 13:00 the input features are Day-Ahead Market cleared prices and Day-Ahead wind forecast.

Step 2 Forecast is trying to forecast Real-Time price of the next hour during Real-Time operation. To forecast the next hour Real-Time price, the previous Real-Time prices, actual load, and wind can be used as input features. To match the input features of Step 1, hourly average Real-Time price and actual hourly average wind are used as input features so as to compare the forecasting accuracy between Step 1 and Step 2. The forecasting results over 500 hours of Step 1 and Step 2 are shown in Figure 5.5.

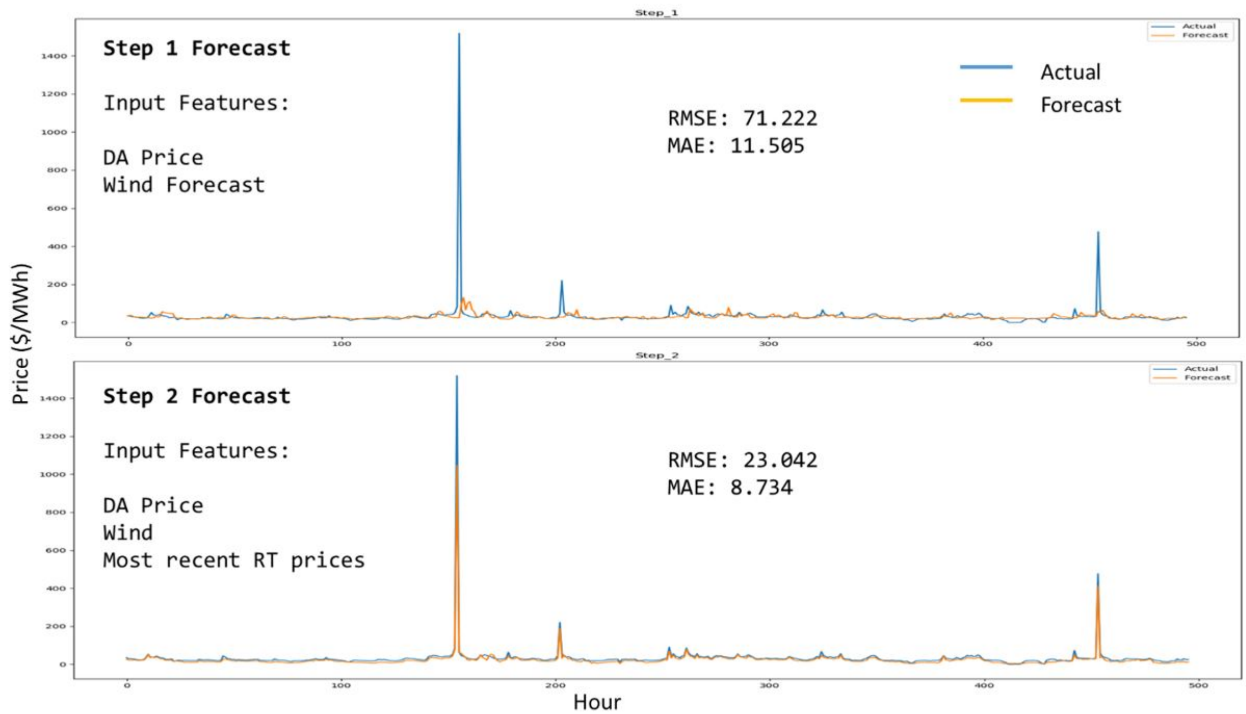

Figure 5.5: Step 1 and Step 2 forecasting results. RMSE of Step 1 is 71.222 and MAE of Step 1 is 11.505; RMSE of Step 2 is 23.042 and MAE of Step 2 is 8.732.

As mentioned earlier, among the input features Real-Time hourly average price can represent hourly load information because price depends on load and other values through the SCED calculation. Different input feature combinations were tested, with the results listed in Table 5.2. The input feature combination of wind, Day-Ahead price, and Real-Time price is superior to the combination of wind, load and Real-Time price information.

Figure 5.6 shows a more detailed comparison of Step 1 and Step 2 forecasts by zooming into a smaller window (50 hours). From the figure, Step 2 appears to have overall better accuracy, but Step 1 can also capture some important price information, such as the general price ranges at different time periods, which is sufficient for Day-Ahead planning purposes.

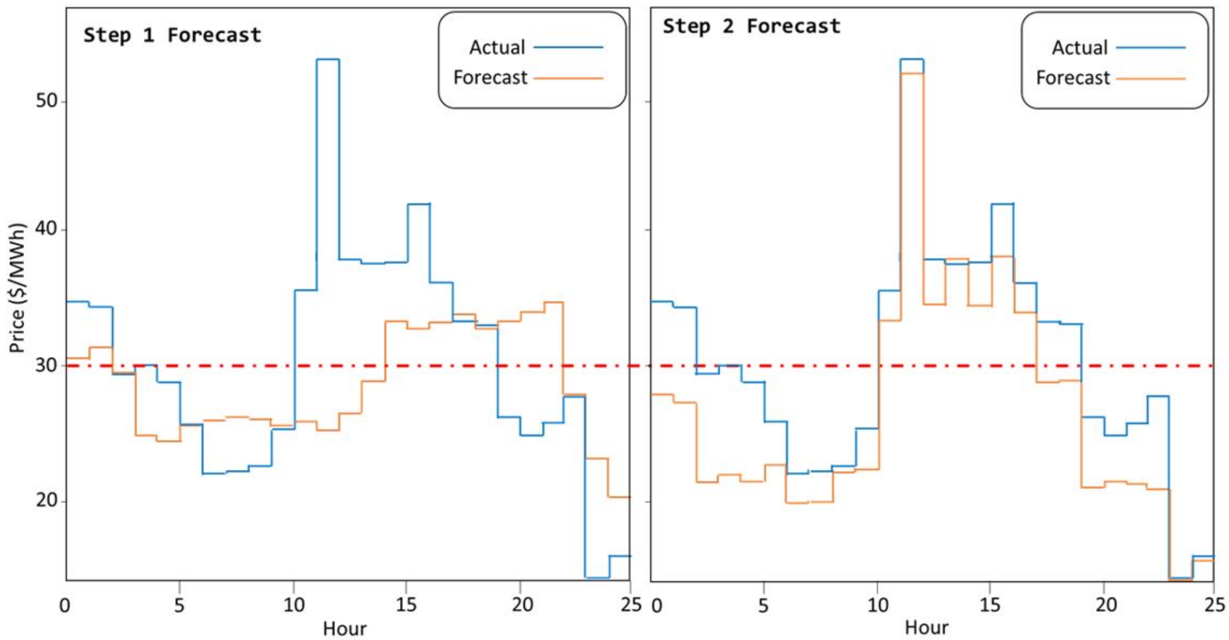

Figure 5.6: Step 1 and Step 2 forecasting results in a day-window. Step 2 has higher accuracy but Step 1 can also capture some important price information over the hours of the selected day.

Step 3 Forecast is trying to forecast prices at each 5-minute SCED interval. ERCOT publishes 5-minute nodal prices for each bus, but for a load zone such as SOUTH HUB, ERCOT only calculates the 15-minute price by time-weighted averaging of the nodal price of all the 345kV buses within SOUTH HUB zone every 15 minutes. To generate 5-minute HUB SOUTH prices, every 15-minute SOUTH HUB price is copied to three 5-minute intervals within the 15-minute period, as in the example in Table 5.1. Besides 5-minute Real-Time prices, Step 3 also uses the 5-minute wind generation of previous intervals as input features to forecast the next interval price.

Table 5.1: Example of generating 5-minute prices from 15-minute prices

Time	5-minute Price ($/MWh)	5-minute Wind (MW)
1/1/2018 0:05:00	26.09	46.2536
1/1/2018 0:10:00	26.09	46.2264
1/1/2018 0:15:00	26.1	46.5273
1/1/2018 0:20:00	26.1	46.9278
1/1/2018 0:25:00	26.1	46.6757
1/1/2018 0:30:00	25.89	46.1765
1/1/2018 0:35:00	25.89	46.0463
1/1/2018 0:40:00	25.89	46.1068
1/1/2018 0:45:00	25.97	46.6595
1/1/2018 0:50:00	25.97	46.6669
1/1/2018 0:55:00	25.97	46.3398
1/1/2018 1:00:00	26.21	45.9845
1/1/2018 1:05:00	26.21	45.2988

The Step 3 forecasting result is shown in Figure 5.7. There are 12 values to forecast in each hour compared to just one per hour in Step 1 and 2.

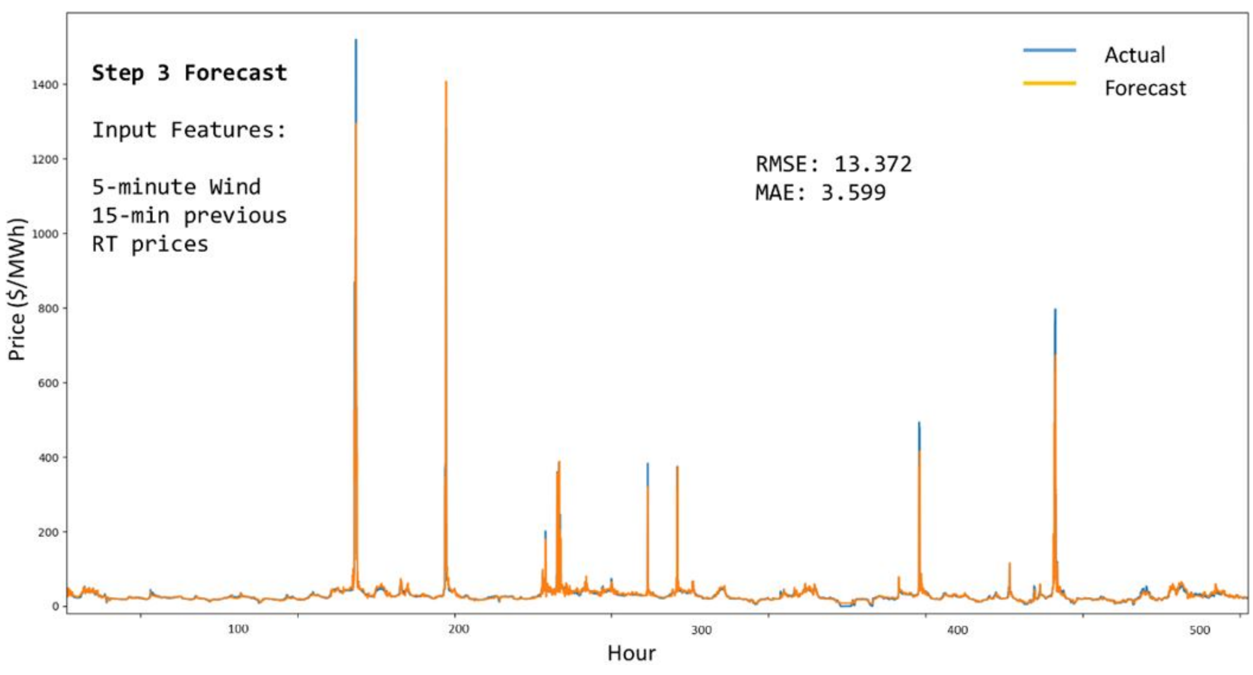

Figure 5.7: Step 3 forecasting result. RMSE of Step 3 Forecast is 13.372 and MAE of Step 3 Forecast is 3.599, which are better than those of Step 1 and Step 2.

As shown in Figure 5.8, Step 3 Forecast creates curves within the hour, to reflect the changes between the 5-minute SCED intervals. The missing information of hourly forecasting can be captured for 5-minute SCED interval planning.

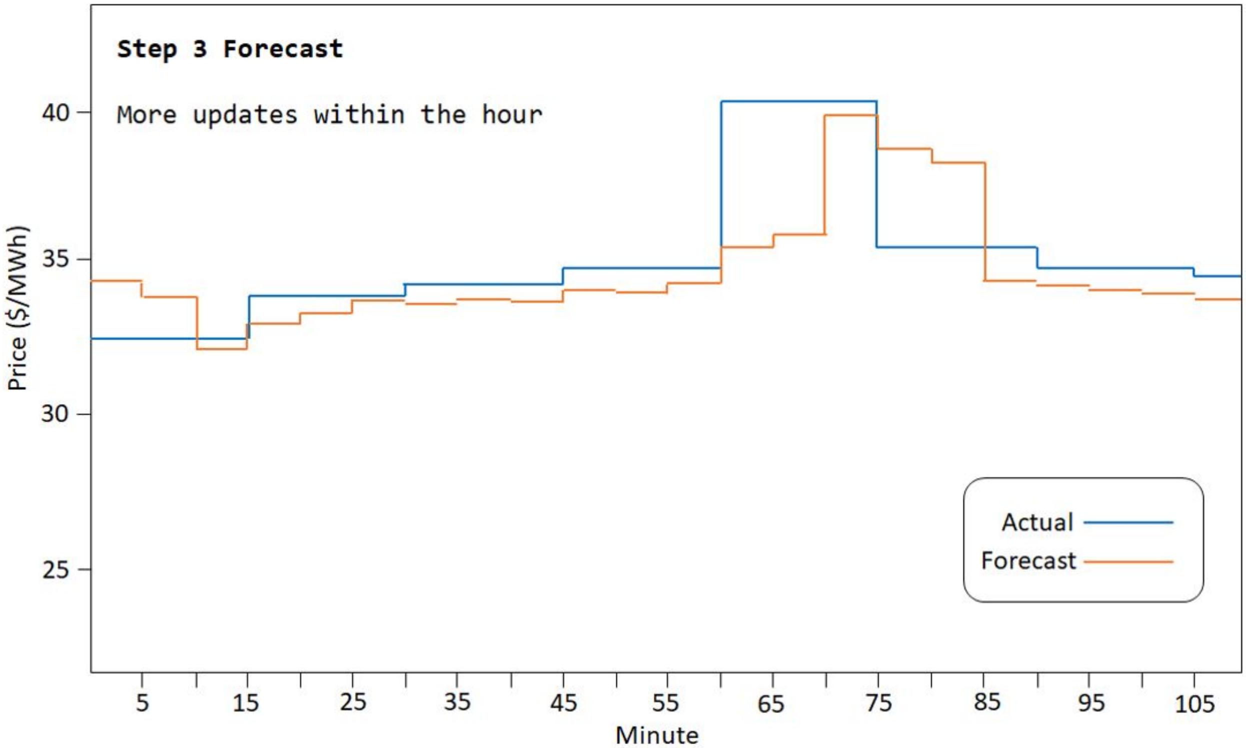

Figure 5.8: Step 3 forecasting result in a small selected time window. It forecasts into every 5-minute SCED interval.

All the forecasting results of TREPF are summarized in Table 5.2. From the table, and as expected, the forecasting accuracy improves from Step 1 to Step 2 to Step 3. When more input features are available, the accuracy of each step could be improved further.

Table 5.2: Results summary of Step 1, 2 and 3

Input Features	RMSE	MAE
Step 1		
Hourly Wind generation Forecast and Day-Ahead prices	71.22	11.51
Step 2		
Hourly Wind generation and Real-Time prices	74.48	24.01
Hourly Wind generation, Load and Real-Time prices	45.91	14.03
Hourly Wind generation, Day-Ahead prices and Real-Time prices	23.04	8.73
Step 3		
5-minute Wind generation and Real-Time prices	30.65	5.10

5.4 Chapter Summary

TREPF designed with 3-step forecasting strategy can cover the operation needs at Day-Ahead planning, hourly, and 5-minute Real-Time operations. From the ERCOT case study, Step 1 Forecast has less accuracy in forecasting Real-Time prices, but it can capture important overall Real-Time price variations for Day-Ahead planning purposes; Step 2 Forecast can be used to improve Step 1 Forecast when it comes to hourly Real-Time operation, and Step 3 Forecast is introduced to improve the Step 2 Forecast for the 5-minute SCED intervals that would especially benefit Ancillary Services planning in a market with RTC, as is planed for ERCOT in the coming year.

Recurrent Neural Network LSTM, as discussed in Chapter 1 can represent and learn the underlying data pattern very well despite the large amount of data. Based on even a limited availability of input features the forecasting accuracy is promising. More and different input features can potentially lead to different forecasting accuracies. There are some overlapping information represented by the input features, since, for example, the price, which is used as an input feature, is calculated by ERCOT using input load, but the advantage of the neural network is that it can represent the relationships to reach the best prediction.

Classical forecasting methods, like ARIMA, can also be used to implement TREPF.

However, as discussed earlier, advanced neural networks have better capability in learning the non-linearity in the underlying data pattern. During non-scarcity hours ARIMA may be able to match the performance of neural networks in forecasting the electricity prices, but will not perform as well during the scarcity events. In the ERCOT Market, the scarcity electricity prices which has a $9,000/MWh price-cap, can cause huge benefit or loss for market participants in a single hour. In order to forecast as many scarcity prices, advanced neural networks like LSTM are favorable for implementing TREPF.

Part III

Conclusions and Future work

Chapter 6

Conclusions and Future Work Plans

6.1 Conclusions

The dissertation presents several related approaches to electricity price forecasting. It has covered some emerging topics in electricity price forecasting on how to improve the accuracy of electricity price forecasting and how to forecast the peak-electricity-price in the Real-Time Market. The forecasts were based on neural networks and tried to make innovations and contributions by developing new forecasting methods, selecting the most effective input features, and processing the output data. Real ERCOT market data from the ERCOT website was used for all the studies.

Chapter 2 explored the forecasting capability of neural networks for the ERCOT Day-Ahead Market. By studying each neural network, and comparing them to the classical statistical methods, it highlighted the advantages of recurrent neural networks in Day-Ahead price forecasting.

Chapter 3 introduced a new method, called Multiple-Run Time-of-Peak Forecasting (MRTPF). By adjusting the neurons and layers of ANN to form a range of models with strengths in representing different causes of peak-electricity-prices, MRTPF statistically combines all the results to cover the peak-electricity-price possibilities. The result is better than any single method. Compared to gradient boosting forecasting methods that statistically combine different forecasts [15, 66], MRTPF uses multiple ANNs to predict the time of peak.

Chapter 4 introduced a method called Binary Time of Peak-Electricity-Price Forecasting (BTOPEPF). BTOPEPF translates the raw price values into simple binary data 0 and 1, which will significantly reduce the complexity of the price data. This method is very easy to apply, very practical to use, and very easy to understand and evaluate. Even though BTOPEPF has less accuracy compared to MRTPF, it needs much less time for ANN configuration selection and model training. However, after both models are trained, the forecasting speeds will be very similar.

Chapter 5 introduced a three-step Real-Time electricity price forecasting (TREPF) method. The method has three steps that roll forward in use of available data. The first step, using the earliest information, is to forecast hourly averages of Real-Time prices of the next day using available Day-Ahead Market cleared prices and Day-Ahead forecasting information. This method can set the base hourly Real-Time price forecast for the next day, which can be used for day ahead planning. During the operating day itself, Real-Time data becomes available. The second step of the Real-Time price forecasting process then uses the most current hourly Real-Time information, to update the price forecast of step one for every hour of Real-Time operation purposes. There is also a third step process that forecasts at the finer timescale of the 5-minute SCED intervals to shape details within the hourly forecast from step two.

The dissertation has presented several new methods in wholesale electricity price forecasting in both Day-Ahead and Real-Time Markets of ERCOT. Through exploring the use of artificial neural works, the accuracy of the forecasting is improved compared to classical statistical forecasting methods in the case studies. Analysis of input features selection is also an important contribution of the dissertation. By comparing the forecasting results, the most efficient and effective input features were chosen for case studies.

In addition, the dissertation contributes to peak-electricity-price forecasting in the ERCOT Real-Time Market, by proposing several new ideas and methods. Actual ERCOT market data is used to test the ideas and methods, and the results are consistent and promising. The methods presented can be potentially used to forecast negative prices too. The

ERCOT Market only experiences fairly limited levels of negative prices. However, other markets, such as the Australian wholesale electricity market, has more negative fluctuations, and the approach in this dissertation could be applied to forecast negative price peaks.

6.2 Future Work Plans

Future research will focus on improving the forecasting models and collecting and testing more input features in order to improve the forecasting performance in both accuracy and speed. The existing methods will be tried in other major wholesale electricity markets such as MISO, CAISO, and PJM to compare the results and confirm that the the approaches are generally applicable. In addition, efforts will be made to find possible applications of the forecasting methods in other aspects of electricity markets, such as how to use the forecasting methods to help plan the daily charge and discharge of battery storage.

As renewable generation capacity, especially wind generation capacity, is growing in almost every market, battery storage may be the only practical solution to shift the generation to cover the scarcity conditions during peak load hours. An extra 2,000 MW battery storage in ERCOT Market could have helped to avoid the EEA 1 alerts on August 12, 2019, which were mentioned in Chapter 1. Due to different characteristics of batteries compared to conventional generations, planning of charge and discharge of batteries in the grid are important. A good price forecasting tool with sensitivity to peak-electricity-prices can help determine the best schedules to charge and discharge each battery in the grid to maximize the social welfare.

In future work, more input features will be collected and tested to investigate if this indeed does improve forecast accuracy. Weather, temperature, and transmission outage information will be collected and tested to see how they can affect the forecasting results.

Bibliography

[1] Recurrent neural network, Wikipedia. *Wikipedia*. URL: https://en.wikipedia.org/wiki/Recurrent_neural_network.

[2] Understanding LSTM Networks. *Colah's blog*. URL: https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[3] MMS Reliability Unit Commitment (RUC) Requirements. 2007. URL: www.ercot.com/content/meetings/tptf/keydocs/.../14c1_mms_ruc_req_b2_v1_01.doc.

[4] 2017 State of the market report for the ERCOT Electricity markets by IMM. 10, 2017. URL: https://www.potomaceconomics.com/wp-content/uploads/2018/05/2017-State-of-the-Market-Report.pdf.

[5] Ancillary Service Assignment. *ERCOT Nodal Protocols*, 2018. URL: http://www.ercot.com/mktrules/nprotocols/current.

[6] ERCOT Stakeholders Dig into Real-time Co-optimization. *RTO Insider*, 2018. URL: https://rtoinsider.com/ercot-real-time-co-optimization-112685/.

[7] IRC Home page. 2018. URL: https://isorto.org/.

[8] Real-Time or Security Constrained Economic Dispatch (SCED). 2018. URL: http://www.ercot.org/about/wc/rt.html.

[9] Settlement for RUC Process. *ERCOT Nodal Protocols*, 2018. URL: http://www.ercot.com/mktrules/nprotocols/current.

[10] Transmission Security Analysis and Reliability Unit Commitment (RUC). *ERCOT Nodal Protocols*, 2018. URL: http://www.ercot.com/mktrules/nprotocols/current.

[11] ERCOT Basic Training. 2019. URL: http://www.ercot.com/content/wcm/training_courses/52/BTP201M2_August2016.pdf.

[12] ERCOT Quick Facts. 2019. URL: `http://www.ercot.com/content/wcm/lists/172484/ERCOT_Quick_Facts_7.31.19.pdf`.

[13] ERCOT Real-Time Co-optimization Task Force. 2019. URL: `http://ercot.com/calendar/2019/4/22/179704-RTCTF`.

[14] Apostolos N. Adamakos and Michalis K. Titsias. Short-Term Load Forecasting using a Cluster of Neural Networks for the Greek Energy Market. In *SETN '16 Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, page 15, 2016.

[15] A.S.Khwaja, X.Zhang, A.Anpalagan, and B.Venkatesh. Boosted neural networks for improved short-term electric load forecasting. *Electric Power Systems Research*, 143:431–437, 2017.

[16] A. Bello, Javier Reneses, and Antonio Islas Munoz. Medium-term probabilistic forecasting of extremely low prices in electricity markets: application to the Spanish case. *Energies*, 9(3), 2016.

[17] Fred Espen Benth, Rüdiger Kiesel, and Anna Nazarova. A critical empirical study of three electricity spot price models. *Energy Economics*, 34(5):1589–1616, 2012.

[18] Rui Bo and Fangxing Li. Probabilistic LMP Forecasting Considering Load Uncertainty. *IEEE Transactions on Power Systems*, 24(3):1279–1289, 2009.

[19] Nirmal K. Bose and Amulya K. Garga. Neural network design using Voronoi diagrams. *IEEE Transactions on Neural Networks*, 4(5):778–787, 1993.

[20] Salah Bouktif, Ali Fiaz, Ali Ouni, and Mohamed Adel Serhani. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy*, 11(7):1636–1656, 2018.

[21] Christine. Brandstatt and Gert. Brunekreeft. How to deal with negative power price spikes?—Flexible voluntary curtailment agreements for large-scale integration of wind. *Am Stat Assoc*, 39(6):3732–3740, 2011.

[22] Glenn W. Brier. Verification of forecasts expressed in terms of probability. *Mon Weather Rev*, 78(1):1–3, 1950.

[23] René Carmona and Michael Coulon. "A survey of commodity markets and structural models for electricity prices. *Quantitative Energy Finance*, pages 41–83, 2014.

[24] J.P.S. Catalão, S.J.P.S. Mariano, V.M.F. Mendes, and L.A.F.M. Ferreira. Short-term Electricity Prices Forecasting in a Competitive Market: A Neural Network Approach. *Electric Power Systems Research*, 77(10):1297–1304, 2014.

[25] Hao Chen, Keli Xiao, Jinwen Sun, and Song Wu. A Double-Layer Neural Network Framework for Highfrequency Forecasting. *ACM Trans. Manage. Inf. Syst*, 7(4), January 2017.

[26] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.

[27] Timothy M. Christensen, A. Stan Hurn, and Kenneth A. Lindsay. Forecasting spikes in electricity prices. *International Journal of Forecasting*, 28(2):400–411, 2012.

[28] Igor M. Coelho, Vitor N. Coelho, Eduardo J. da S. Luz, Luiz S. Ochi, Frederico G. Guimarães, and Eyder Rios. A GPU Deep Learning Metaheuristic based Model for Time Series Forecasting. *Applied Energy*, 201(1):412–418, 2017.

[29] Michael Coulon and Sam Howison. Stochastic Behaviour of the Electricity Bid Stack: From Fundamental Drivers to Power Prices. *The Journal of Energy Markets*, 2:29–69, 2009.

[30] George E. Dahl, Tara N. Sainath, and Geoffrey E. Hinton. Improving deep neural networks for LVCSR using rectified linear units and dropout. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.

[31] Pengwei Du. Ensemble Machine Learning Based Wind Forecasting to Combine NWP output with Data from Weather Stations. *IEEE Transactions on Sustainable Energy*, 2018.

[32] ERCOT. ERCOT's use of Energy Emergency Alerts . URL: `http://www.ercot.com/content/wcm/lists/164134/EEA_OnePager_FINAL.PDF`.

[33] Cheng Fan, Fu Xiao, and Yang Zhao. A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195(1):222–233, 2017.

[34] Cong Feng, Mingjian Cui, Bri-Mathias Hodge, and Jie Zhang. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. *Applied Energy*, 190(15):1245–1257, 2017.

[35] Tiago A. E. Ferreira, Germano C. Vasconcelos, and Paulo J. L. Adeodato. A New Evolutionary Method for Time Series Forecasting. In *GECCO '05 Proceedings of the 7th annual conference on Genetic and evolutionary computation*, pages 2221–2222, 2005.

[36] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Am Stat Assoc*, 102(477):359–378, 2012.

[37] Tilmann Gneiting and Adrian E. Raftery. Do Renewable Portfolio Standards Deliver? *EPIC Working Paper*, NO. 2019-62, 2019. URL: `https://bfi.uchicago.edu/wp-content/uploads/BFIEPIC\_WP\_201962\_v4.pdf`.

[38] Alex Graves. Generating Sequences with Recurrent Neural Networks. *arXiv preprint arXiv*, 1308(0850), 2014.

[39] Klaus Greff, Rupesh Kumar Srivastava, Jan Koutník, Bas R. Steunebrink, and Jürgen Schmidhuber. LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28:2222–2232, 2017.

[40] Tao Hong. Crystal Ball Lessons in Predictive Analytics. *EnergyBiz*, pages 35–37, 2015.

[41] Tao Hong, Pierre Pinson, Shu Fan, Hamidreza Zareipour, Alberto Troccoli, and Rob J.Hyndman. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *International Journal of Forecasting*, 32(3):896–913, 2014.

[42] Tianyu Hu, Qinglai Guo, Zhengshuo Li, Xinwei Shen, and Hongbin Sun. Distribution-Free Probability Density Forecast Through Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–14, 2019.

[43] Joanna Janczura, Stefan Trück, Rafał Weron, and Rodney C.Wolff. Identifying spikes and seasonal components in electricity spot price data: a guide to robust modeling. *Energy Economics*, 38:96–110, 2013.

[44] Joanna Janczura and Rafał Weron. An empirical comparison of alternate regime-switching models for electricity spot prices. *Energy Economics*, 32(5):1059–1073, 2010.

[45] Antonio J.Conejo, Javier Contreras, Rosa Espínola, and Miguel A.Plazas. Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting*, 21(3):435–462, 2005.

[46] Tryggvi Jonsson, Pierre Pinson, Henrik Aalborg Nielsen, Henrik Madsen, and Torben Skov Nielsen. Forecasting Electricity Spot Prices Accounting for Wind Power Predictions. *IEEE Transactions on Sustainable Energy*, 4(1):210–218, 2013.

[47] J.P.S.Catalão, S.J.P.S.Mariano, V.M.F.Mendes, and L.A.F.M.Ferreira. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Transactions on power systems*, 21(2):887–896, 2006.

[48] Dogan Keles, Jonathan Scelle, Florentina Paraschiv, and Wolf Fichtner. Extended Forecast Methods for Day-Ahead Electricity Spot Prices applying Artificial Neural Networks. *Applied Energy*, 162:218–230, 2016.

[49] Shubham Khandelwal, Benjamin Lecouteux, and Laurent Besacier. Comparing GRU and LSTM for Automatic Speech Recognition. *Research Report*, 2016.

[50] Nowrouz Kohzadi, Milton S.Boyd, Bahman Kermanshahi, and Iebeling Kaastra. A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing*, 10(2):169–181, 1996.

[51] Weicong Kong, Zhaoyang Dong, Youwei Jia, David J. Hill, Yan Xu, and Yuan Zhang. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *IEEE Transactions on Smart Grid*, 10(1):841–851, 2019.

[52] Xiangrui Kong, Xiaoyuan Xu, Zheng Yan, Sijie Chen, Huoming Yang, and Dong Han. Deep learning hybrid method for islanding detection in distributed generation. *Applied Energy*, 210(15):776–785, 2018.

[53] Ping-Huan Kuo and Chiou-Jye Huang. An Electricity Price Forecasting Model by Hybrid Structured Deep Neural Networks. *Applied Energy*, 10, 2018.

[54] Fangxing Li. Continuous Locational Marginal Pricing (CLMP). *IEEE Transactions on Power Systems*, 22(4):1638–1646, 2007.

[55] Fangxing Li and Rui Bo. Congestion and Price Prediction Under Load Variation. *IEEE Transactions on Power Systems*, 24(2):911–922, 2009.

[56] Hui Liu, Xiwei Mi, and Yanfei Li. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy*, 159(1):54–64, 2018.

[57] Paras Mandal, Tomonobu Senjyu, and Toshihisa Funabashi. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. *Energy Conversion and Management*, 47(15):2128–2142, 2006.

[58] Jakub Nowotarski, Jakub Tomczyk, and Rafał Weron. Robust estimation and forecasting of the long-term seasonal component of electricity spot prices. *Energy Economics*, 39:13–27, 2013.

[59] Jakub Nowotarski and Rafał Weron. Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, 81(1):1548–1568, 2018.

[60] Cielito C. Olegario, Andrei D. Coronel, Ruji P. Medina, and Bobby D. Gerardo. A Hybrid Approach towards Improved Artificial Neural Network Training for Short-Term Load Forecasting. In *DSIT '18 Proceedings of the 2018 International Conference on Data Science and Information Technology*, pages 53–58, 2018.

[61] N. M. Pindoriya, S. N. Singh, and S. K. Singh. An Adaptive Wavelet Neural Network-Based Energy Price Forecasting in Electricity Markets. *IEEE Transactions on power systems*, 23(3):1423–1432, 2008.

[62] Xiangyun Qing and Yugang Niu. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. *Energy*, 148(1):461–468, 2018.

[63] Miadreza Shafie-khah and João P. S. Catalão. A Stochastic Multi-Layer Agent-Based Model to Study Electricity Market Participants Behavior. *IEEE Transactions on Power Systems*, 30(2):867–881, 2015.

[64] Mohammad Shahidehpour, Hatim Yamin, and Zuyi Li. Market Operations in Electric Power Systems: Forecasting, Scheduling, and Risk Management. *Wiley*, 2002.

[65] Vinayak Sharma, Umit Cali, Veit Hagenmeyer, Ralf Mikut, and Jorge Angel González Ordiano. Numerical Weather Prediction Data Free Solar Power Forecasting with Neural Networks. In *e-Energy '18 Proceedings of the Ninth International Conference on Future Energy Systems*, pages 604–609, New York, NY, USA, 2018. ACM.

[66] Souhaib Ben Taieb and Rob J Hyndman. A gradient boosting approach to the Kaggle load forecasting competition. *International Journal of Forecasting*, 30(2):382—-394, 2014.

[67] S. Tamura and M. Tateishi. Capabilities of a four-layered feedforward neural network: four layers versus three. *IEEE Transactions on Neural Networks*, 8(2):251–255, 1997.

[68] Umut Ugurlu, Ilkay Oksuz, and Oktay Tas. Electricity Price Forecasting Using Recurrent Neural Networks. *Energies*, 11(5):1255, 2018.

[69] Nektaria V.Karakatsani and Derek W.Bunn. Forecasting electricity prices: The impact of fundamentals and time- varying coefficients. *IEEE Transactions on Sustainable Energy*, 24(4):764–785, 2008.

[70] Eleni I. Vlahogianni and Matthew G.Karlaftis. Testing and Comparing Neural Network and Statistical Approaches for Predicting Transportation Time Series. *Transportation research record*, 2399(1):9–22, 2013.

[71] H.Z. Wang, G.B. Wang, G.Q.Li, J.C.Peng, and Y.T. Liu. Deep Belief Network based Deterministic and Probabilistic Wind Speed Forecasting Approach. *Applied Energy*, 182(15):80–93, 2016.

[72] Rafał Weron. Modeling and forecasting electricity loads and prices: a statistical approach. *Chichester: John Wiley & Sons*, 2006.

[73] Rafał Weron. Market price of risk implied by Asian-style electricity options and futures. *Energy Economics*, 30(3):1098–1115, 2008.

[74] Rafał Weron. Electricity Price Forecasting: A Review of the State-of-the-art with a Look into the Future. *International Journal of Forecasting*, 30(4):1030–1081, 2014.

[75] Rafał Weron and Adam Misiorek. Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4):744–763, 2008.

[76] Jian Xu and Ross Baldick. Binary Forecasting Method of the Time of Peak Electricity Price based on Artificial Neural Network. *Accepted to The 1st IEEE Sustainable Power & Energy Conference (iSPEC)*, 2019.

[77] Jian Xu and Ross Baldick. Day-Ahead Price Forecasting in ERCOT Market Using Neural Network Approaches. *Proceedings of the Tenth ACM International Conference on Future Energy Systems (e-Energy '19)*, June 2019.

[78] Jian Xu and Ross Baldick. Three-Step Real-Time Electricity Price Forecast using Recurrent Neural Network. *IEEE Transactions on Power Systems*, Re-submitted in November 2019.

[79] Jian Xu and Ross Baldick. A New Time of Peak Forecasting Method based on Artificial Neural Network. *Journal of Modern Power Systems and Clean Energy*, Submitted in September 2019.

[80] Binbin Yong, Zijian Xu, Jun Shen, Huaming Chen, Yanshan Tian, and Qingguo Zhou. Neural Network Model with Monte Carlo Algorithm for Electricity Demand Forecasting in Queensland. In *ACSW '17 Proceedings of the Australasian Computer Science Week Multiconference*, 2017.

[81] Maheen Zahid, Fahad Ahmed, Nadeem Javaid, Raza Abid Abbasi, Hafiza Syeda Zainab Kazmi, Atia Javaid, Muhammad Bilal, Mariam Akbar, and Manzoor Ilahi. Electricity Price and Load Forecasting using Enhanced Convolutional Neural Network and Enhanced Support Vector Regression in Smart Grids. *Electronics*, 8, 2019.

[82] Mohammadsaleh Zakerinia and Seyed Farid Ghaderi. Short Term Wind Power Forecasting using Time Series Neural Networks. In *EAIA '11 Proceedings of the 2011 Emerging M&S Applications in Industry and Academia Symposium*, pages 17–22, 2011.

[83] Hamidreza Zareipour, Claudio A. Canizares, and Kankar Bhattacharya. Economic Impact of Electricity Market Price Forecasting Errors: A Demand-Side Analysis. *IEEE Transactions on Power Systems*, 25:254–262, 2007.

[84] Hamidreza Zareipour, Arya Janjani, Henry Leung, Amir Motamedi, and Antony Schellenberg. Classification of future electricity market prices. *IEEE Trans Power Systems*, 26(1):165–173, 2011.

[85] Jun Zhang and Chuntian Cheng. Day-ahead electricity price forecasting using artificial intelligence. *2008 IEEE Canada Electric Power Conference*, 2008.

[86] Huai zhi Wang, Gang qiang Li, Gui bin Wang, Jian chun Peng, Hui Jiang, and Yi tao Liu. Deep learning based ensemble approach for probabilistic wind power forecasting. *Applied Energy*, 188(15):56–70, 2016.

[87] Lidong Zhou, Bo Wang, Zheng Wang, Fei Wang, and Minghui Yang. Seasonal classification and RBF adaptive weight based parallel combined method for day-ahead electricity price forecasting. *2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2018.

[88] Florian Ziel and Rick Steinert. Electricity price forecasting using sale and purchase curves: The X-Model. *Energy Economics*, 59:435–454, 2016.

Vita

Jian Xu was born in Yancheng, China. He received the Bachelor of Science degree in Electrical Engineering from the Southeast University in Nanjing, China, and the Master of Science degree from Purdue University in West Lafayette, Indiana. He worked as an Electrical Engineer in China before he came to West Lafayette, Indiana in 2009 right after marriage. He joined Mid-Continent Independent System Operator (MISO) in 2012, and moved to Austin, Texas in May 2013 with his wife Yaguo Wang. He worked for Electricity Reliability Council of Texas (ERCOT) since November 2013. He separated from ERCOT in 2015 and worked as an Energy Market Analyst for Austin Energy. He left Austin Energy in 2017 and worked as a Data Scientist in the Lower Colorado River Authority (LCRA). In April 2019 he joined Texas Reliability Entity (Texas RE) with the title State Reliability Engineer, Sr.

Email address: jz20000cn@gmail.com

This dissertation was typeset with LaTeX† by the author.

†LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.